ISSN 1913-1844 (Print)
ISSN 1913-1852 (Online)

MODERN APPLIED SCIENCE

Vol. 9, No. 8 August 2015

CANADIAN CENTER OF SCIENCE AND EDUCATION

Editorial Board

Contents

Contents

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Urban Sprawl Monitoring

Sassan Mohammady[1] & Mahmoud Reza Delavar[2]

[1] GIS division, Dept. Of Surveying and Geomatics Eng., College of Enginnering, University of Tehran, Tehran, Iran

[2] Center of Excellence in Geomatics Eng. in Disaster Management, Dept. of Surveying and Geomatics Eng., College of Eng., University of Tehran, Tehran, Iran

Correspondence: Sassan Mohamamdy GIS division, Dept. Of Surveying and Geomatics Eng., College of Enginnering, University of Tehran, North Karegar St, Tehran, Iran. Tel: 98-936-371-2267. E-mail: Sassanmohammady@ut.ac.ir

Received: December 18, 2014 Accepted: January 11, 2015 Online Published: February 28, 2015

doi:10.5539/mas.v9n8p1 URL: http://dx.doi.org/10.5539/mas.v9n8p1

Abstract

Urban sprawl is a common phenomenon in developed and developing countries. Population growth and immigration to cities are the most important reason to such urban expansions. Out layer growths which in the most cases results sprawl has so many attractions such as low rate of crime, lower costs of living and clean air. These are probably the reasons to rapid increase of this phenomenon. Urban sprawl is a kind of growth in cities which have derived so many negative impacts such as agriculture and natural land loss, environmental pollution, high rate of travel time and costs, high rate of energy consuming and etc.,. Thus, analyzing and monitoring urban area is a key task for urban planning to inform about urban growth. Numerous researches have attempted to characterize and explain urban sprawl. In this study, we implemented Shannon Entropy for assessments of urban sprawl. The case study is Tehran Metropolis which has experienced so fast urbanization and population growth in the recent decades.

Keywords: urban, sprawl assessment, Tehran metropolis

1. Introduction

1.1 Introduce the Problem

Nowadays, a large number of cities in the developing and developed countries are witnessing urban sprawl as a common phenomenon which causes so many issues. This phenomenon in particular is different from expansion (Galster et al., 2001, Inostroza et al., 2013). In fact, urban sprawl is defined as a specific form of urban dispersed development with low-density, auto-dependent with environmentally and socially impacts (Hasse and Lathrop 2003).

The most important negative impacts of sprawl is that this kind of development causes the increase of car-dependency for transportation (Torrens and Alberti, 2000), the loss of agricultural, green spaces and natural land, increase in energy consumption, the need for more infrastructure (Bruekner, 2001) and increase in infrastructure costs, longer travel times, increasing automobile trips and costs of travel and, therefore, greater environmental pollution, the degradation of periurban ecosystems (Johnson, 2001, Li et al., 2006). Generally, sprawl is defined as a relatively wasteful method of urbanization, characterized by uniform low densities and this phenomenon is often uncoordinated and extends along the fringes of metropolitan areas. It invades upon prime agricultural and resource land in the process (Alberti and Torrens, 2000). This form of urban growth is a common phenomenon in the developed countries and also in the developing countries which has big undesirable environmental and socioeconomic effects such as soil recourses and weather pollution (Zhang 2007, Pourahmad et al 2007), prolific farming lands loss around the cities (Brabec and Smith, 2002, Zhang 2000), and land use fragmentation and loss of biodiversity (Alberti 2005), high ecological footprint (Muniz and Galindo 2005), the decrease of even access to urban services and amenities (Burton 2001), increased traffic and demand for mobility (Ewing et al 2002), destruction of forest cover (Rudel and McDonald 2005), high automobile dependency and as a result exacerbate global warming (Hamin and Guran 2009).

Rapid urbanization is quite alarming, especially in developing countries (Kumar et al 2007). This rapid urbanization in so many cases lead to urban sprawl. Urban sprawl has been identified as an undesired trend in

many countries (Hayek et al 2011). Iran as a developing country has experienced urban sprawl in many of large, medium-size or even small cities (Roshan et al 2009). Cities in the developing countries lack such type of policies in most of the cases and they grow with all freedoms (Bhatta 2009). In developing countries due to the lack of timely information of the urbanization process and its long-term ecological impacts, urban planners have not been able to analyze consistently and restore the ecosystems in peri-urban areas (Dutta 2012). In Iran, population growth and urban sprawl have always played an important role in environmental degradation with a strong pressure on local, regional and global conditions (Akbari Motlaq and Abbaszadeh 2012).

Sprawl pattern and land use changes could be detected and analyzed cost effectively and efficiently using remote sensing data and geospatial information systems (GIS) techniques (Bhatta 2009, Barnes et al 2001). Monitoring of urban growth and understanding landscape characterization through historical and timely information provides the necessary information to evaluate environmental impacts of land use change, to determine future infrastructure requirements and to delineate urban growth and municipal service areas (Kennedy 2007).

Remote sensing is considered to be an appropriate source of urban data to support urban studies (Donnay et al., 2001). Urban expansion can be quantified by measuring the built-up area change between two dates (Jensen, 2005). Remote sensing as a reliable source of data is critical for areas with rapid land use changes and especially where the updating of information is tedious and time-consuming. In urban studies, monitoring of urban development is mainly to find out the type, amount, and location of land conversion (Yeh and Li, 1999).

In this study we implemented Shannon Entropy method for analyzing and measuring urban sprawl between 1988 and 2010. In the second step, we introduced an indicator for evaluating sprawl for Tehran Metropolis.

2. Method

Spatial metrics as a valuable tool have been used for the analyzing, monitoring, and tracking of changes in land use patterns and shapes in many research studies (Aguilera et al 2011, Kim and Ellis 2009). In this study after introducing the study area, we implemented Shannon Entropy for analyzing urban sprawl for Tehran Metropolis. The Shannon's entropy reflects the dispersion of the spatial variable in a specified area (Yeh and Li 2001). The Shannon's Entropy can be calculated from the remotely sensed data to efficiently identify and characterize the urban sprawl (Torrens and Alberti 2000, Yeh and Li 2001, Lata et al 2001, Sudhira et al 2004).

2.1 Study Area

The case study in this research is the city of Tehran, Iran. Tehran as the capital of Iran has experienced rapid expansion due to the rapid growth of population and besides this city has become the main target of immigration due to its great urbanization attractions (Roshan et al 2009). Figure 1 shows the position of this city in Iran. Tehran has experienced rapid urban population growth in the last few decades. Figure 2 shows rapid population growth in Tehran Metropolis in the last few decades.

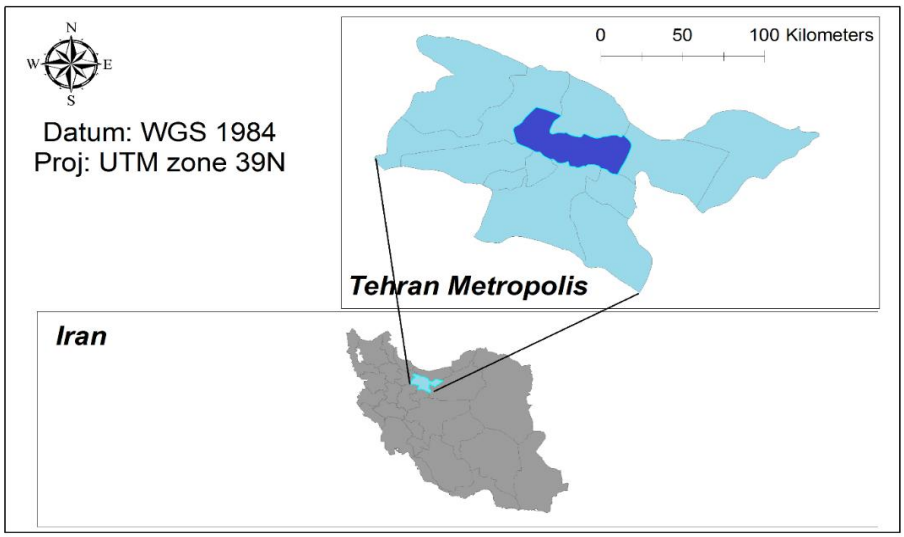


Figure 1. The study area

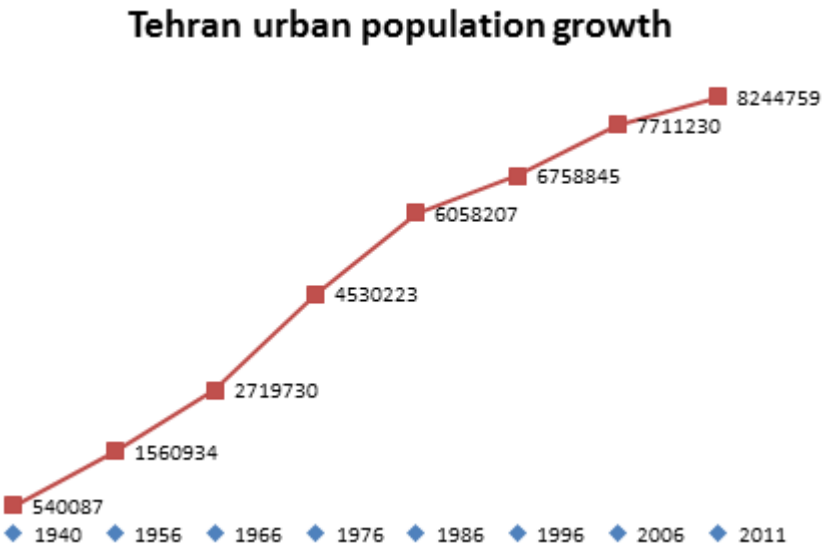


Figure 2. Urban population growth in Tehran

2.2 Pre Processing

Landsat imageries acquired at 1988, 1999 and 2010 with ground pixel size 28.5 Meter have been used (Table. 1). All of remote sensing processes have done in ENVI 4.8. The related imageries are classified based on Anderson Level 1 (Anderson et al 1976) with Support Vector Machine (SVM) classification. Support vector machines (SVM) due to its ability to handle the nonlinear classifier problems are applied in so many researches (Aghababaee et al 2012). Classification accuracies are presented in Table 2.

Table 1. Satellite imageries

Date	Sensor	Ground Pixel Size (m)	Satellite	Datum	Projection System
Sep.27, 1988	TM	28.5	Landsat-5	WGS-1984	UTM, Zone 39N
Oct.7, 1999	ETM^+	28.5	Landsat-7	WGS-1984	UTM, Zone 39N
Oct. 17, 2010	ETM^+	28.5	Landsat-7	WGS-1984	UTM, Zone 39N

Table 2. Classification assessments

Data	Overall Accuracy (%)	Kappa Statistics (%)
1988	89.43	82.22
1999	87.12	72.73
2010	91.33	88.67

According to (Thomlinson et al 1999) all of the overall accuracies are acceptable. According to (Pijanowski et al 2005), 1988 and 2010 classification resulted are considered as excellent and 1999 classification result is considered as very good.

With increased accessibility and improved quality of spatio-temporal remote sensing data as well as new analytical techniques, it enables us to monitor and analyze urban growth and land use change in a cost-effective and timely manner (Alberti 2005, Herold et al 2003). Table 3 shows urban and non-urban area in Tehran Metropolis in 1988, 1999 and 2010. Figure 3 presents Tehran urban area in 1988, 1999 and 2010.

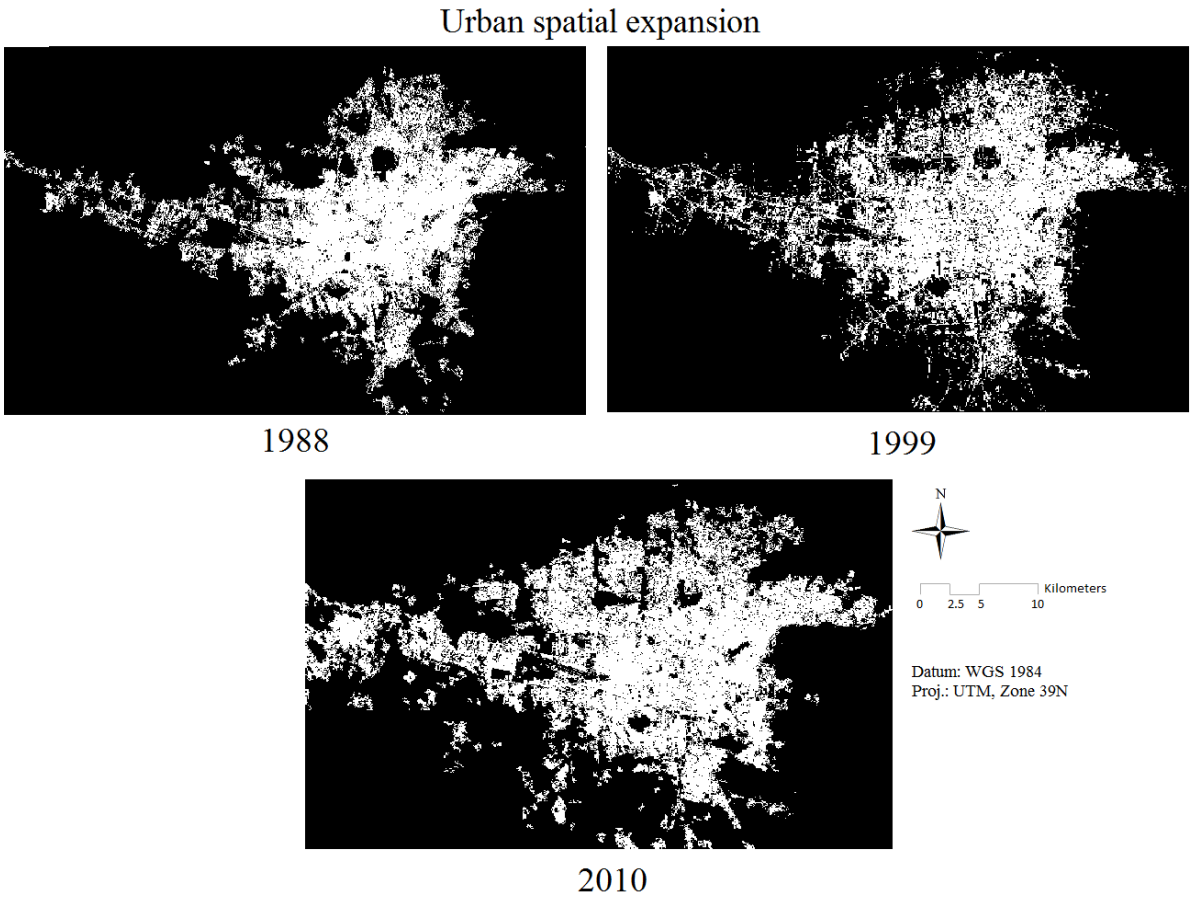


Figure 3. Tehran urban area in 1988, 1999 and 2010

Table 3. Urban and non-urban area in 1988, 1999 and 2010

Data	Urban Area (Km^2)	Urban Area (%)	Non-Urban Area (Km^2)	Non-Urban Area (%)
1988	298.813	18.02	1359.477	81.98
1999	345.9779	20.86	1312.312	79.14
2010	428.0549	25.81	1230.235	74.19

3. Results

In this study, we implemented Shannon Entropy for analyzing urban sprawl in Tehran Metropolis during 1988 to 2010. Shannon's entropy is a well-accepted method for determining the sprawled urban pattern (Li and Yeh, 2004; Sudhira et al., 2004; Yeh and Li, 2001). Shannon Entropy can be used to analysis urban sprawl in temporal span or zone wise. It means it can be used to clarify sprawl for a period of time (temporal span) and for each zone separately (zone wise). In this study we defined the zones according to distance to center. In other words, the distance between each urban pixel and center is calculated and categorized using 2500 meter interval. Each different zone has a different level of compactness leading to different patterns of growth. Therefore, a single policy for the entire city never works with equal degrees of effectiveness for all (Bhatta 2009). The implemented Shannon Entropy is defined below:

$$H_i = -\sum_{j=1}^{m} p_j \log_e(p_j) \quad (1)$$

where,

H_i = Shannon Entropy for each temporal span,

p_j = proportion of built up area in j zone to total built up area.

j is the number of zones (m=13).

i is temporal span (1, 2).

This value ranges from 0 to $\log_e(m)$, values closer to 0 indicate very compact distribution and the value closer to $\log_e(m)$ indicates that the distribution is much dispersed. In other words, larger value of entropy indicates the occurrence of urban sprawl. Table 4 presents Shannon Entropy for each temporal span. According to Bhatta et al (2009), half of the $\log_e(m)$ is a threshold for determining urban sprawl. In this study due to m=13, this value is 1.2825. It means if Shannon Entropy in any of them is larger than this value, it can safety said this region is sprawled and if become less than this value, the region is non-sprawling. According to Table 4, in both temporal spans (1988 to 1999 and 1999 to 2010) this city has experienced sprawl and Shannon Entropy for this city become bigger. It means, this city becoming more sprawling. Figure 4 shows urban cells in each zone. The urban expansion between 1988 and 2010 covered greater areas especially in west and north directions and this expansion occurred at distance from the center (Figure 4).

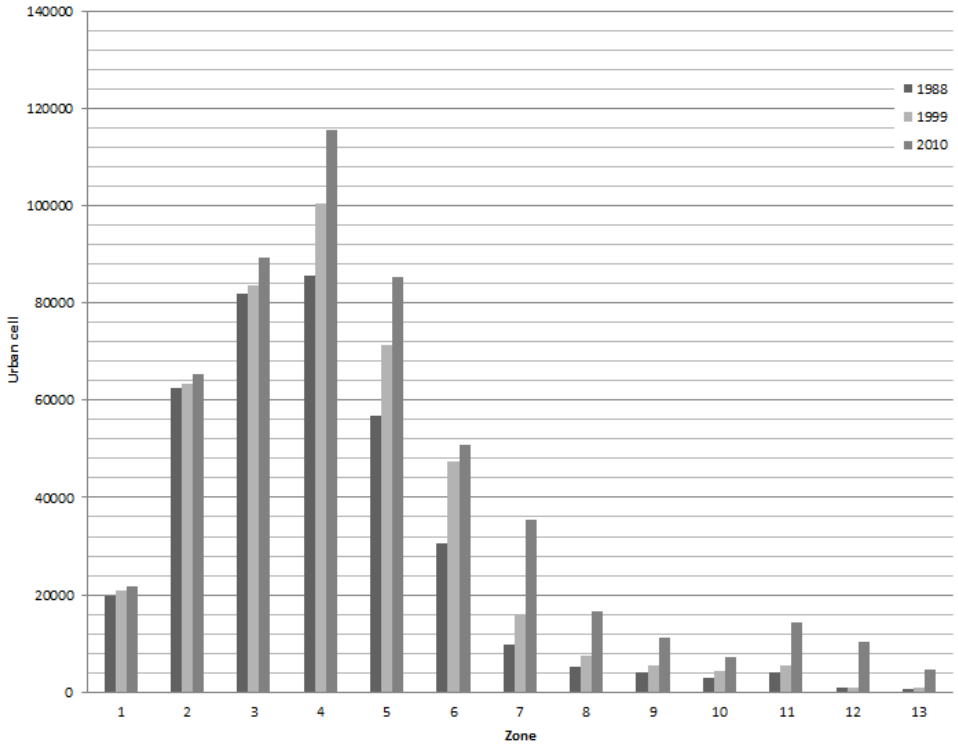


Figure 4. Urban cells in each zone

Table 4. Shannon Entropy during 1988 to 1999 and 1999 to 2010

Temporal Span	1988-1999	1999-2010
Shannon Entropy	1.9002	2.3088

Distance to center is a key factor in urban expansion. As this factor growth, urban sprawl and dispersed expansion will happen. When distance urban cell in a city from center increase, travel time, trip costs, pollution, and agricultural land lost and so many negative impacts will increase, too. This factor clarifies the importance of vertical expansion in the cities. Figure 5 shows the dispersion of the urban cells in 1988, 1999 and 2010. The number of urban cells in larger distance to center in Tehran Metropolis during the time has increased. According to Figure 5, number of the urban cell in during 1988 to 1999 with distances larger than 9 Km has increased and this expansion in more distant area has been so fast in 1999 to 2010.

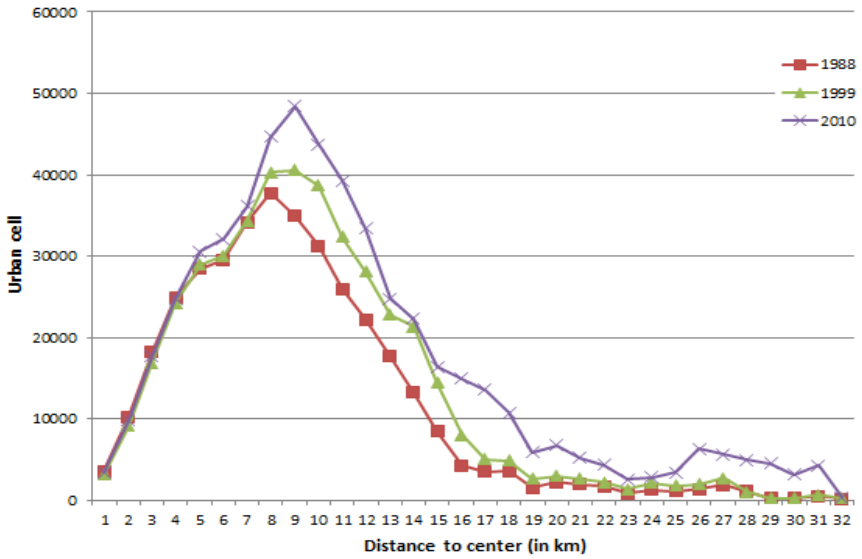


Figure 5. Dispersion of the urban cells in 1988, 1999 and 2010

3.2 Sprawl Index

Geospatial Information Systems (GIS) are widely used to represent, analyze, and display various spatial data such as topographic maps, satellite imageries and Digital Elevation Model (DEM). Many researchers use GIS as a reliable science and technology to find the spatiotemporal characteristics of landscape (Xie and Cho 2007). Using remote sensing data with GIS techniques provides a powerful tool to analyze, model and project environmental change. Therefore, integration of remote sensing and GIS have been recognized as powerful and effective methods in monitoring environmental change, especially in detecting the land use/land cover change (LUCC) (Herold et al 2003, Gao et al 2006). In the second step in this paper, the Sprawl Index (SI) is calculated, by means of specific GIS modeling, using thematic degrees of propensity to the phenomenon that depend on a set of territorial (morphological-urban development) characteristics: slope, elevation, aspect, distance to main road, distance to center, agriculture consistency, green spaces consistency, planning consistency and distance to commercial center (Table 5). The Sprawl Index (SI) is calculated through the following equation that contains a set of degrees of settlement location sensitivity to the aforementioned territorial and morphological characteristics. Figure 6 shows the sprawl parameters.

Table 5. Important parameters for obtaining SI map

Parameters	Refers
Distance to Centers (DC)	Ewing et al, 2002
Slope	Romano 2004
Elevation	Romano 2004
Aspect	Romano 2004
Distance to Main Roads (DM)	Ewing et al, 2002, Romano 2004
Planning Consistency (PC)	Jiang et al, 2007
Open Spaces Consistency (OSC)	Jiang et al, 2007, Sim & Mesev, 2011
Agriculture Consistency (AC)	Jiang et al, 2007
Distance to Commercial Centers (DCC)	Ewing et al, 2002, Romano 2004

Sprawl Index is obtained from implementation of 9 parameters. Sprawl Index is defined,

Sprawl Index (SI) =DC+ Slope+ Elevation+ Aspect+ DM+ PC+ OSC+ AC+ DCC.

Figures 6 and 7 show all of 9 parameters and the resulted SI map, respectively.

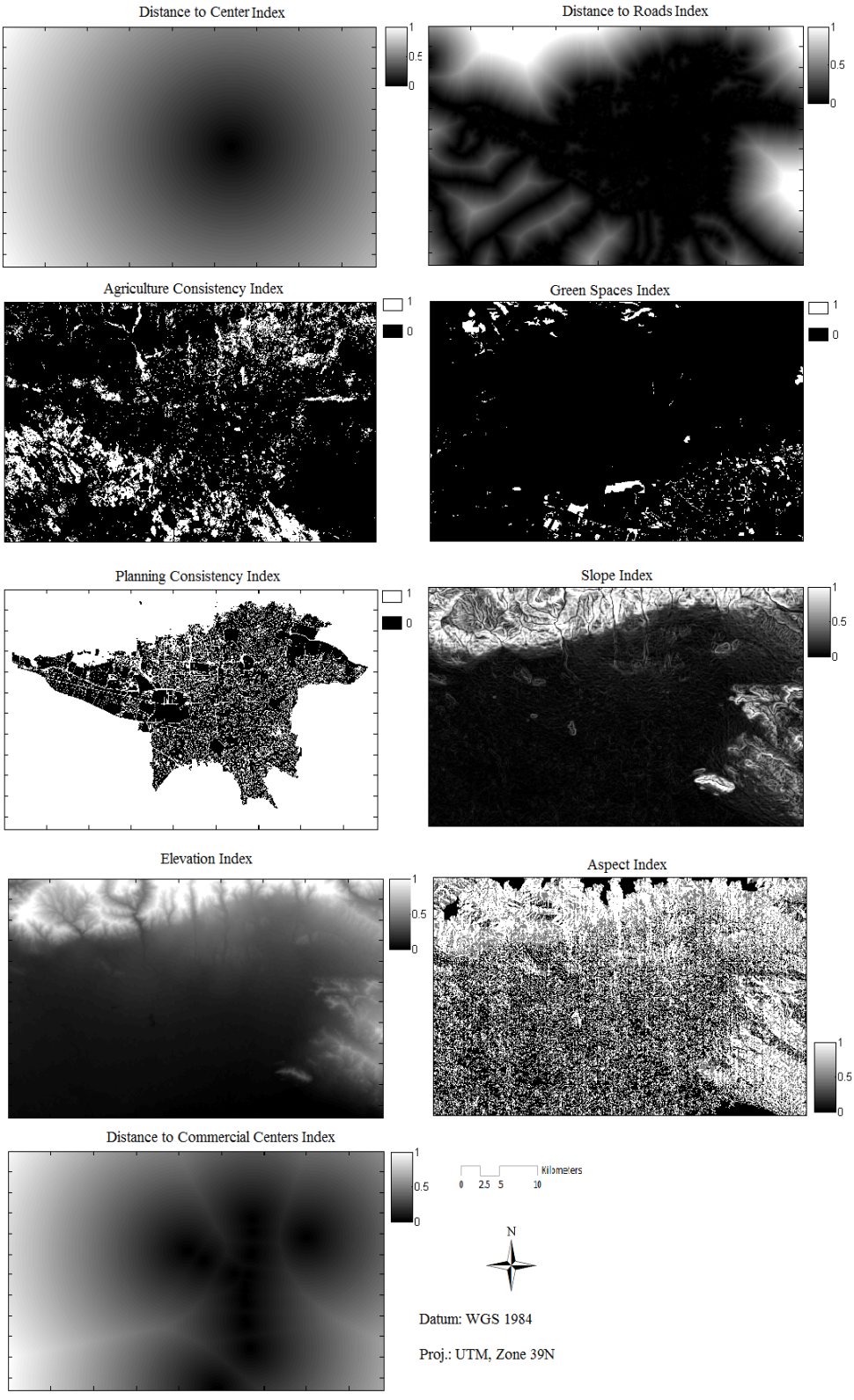


Figure 6. Important parameters for determining SI

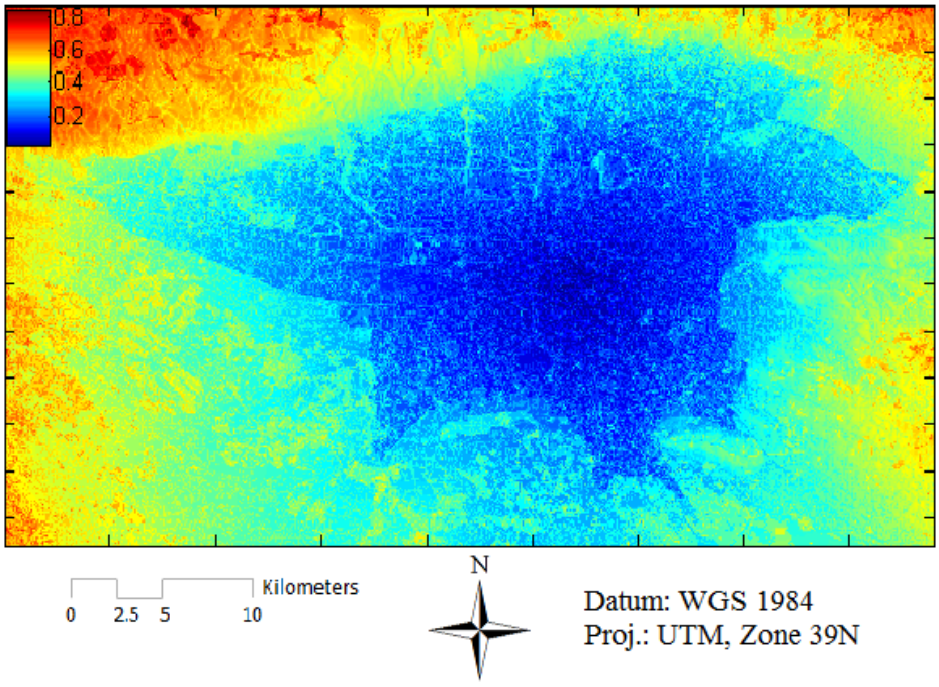


Figure 7. The SI resulted map

According to Figure 7, many part of Tehran Metropolis areas are developed in regions with approximately moderate or high level degrees of sprawl. This means these region is considered as a sprawl region. In Figure 8, we categorize SI range value between 0 to 1 in 5 classes.

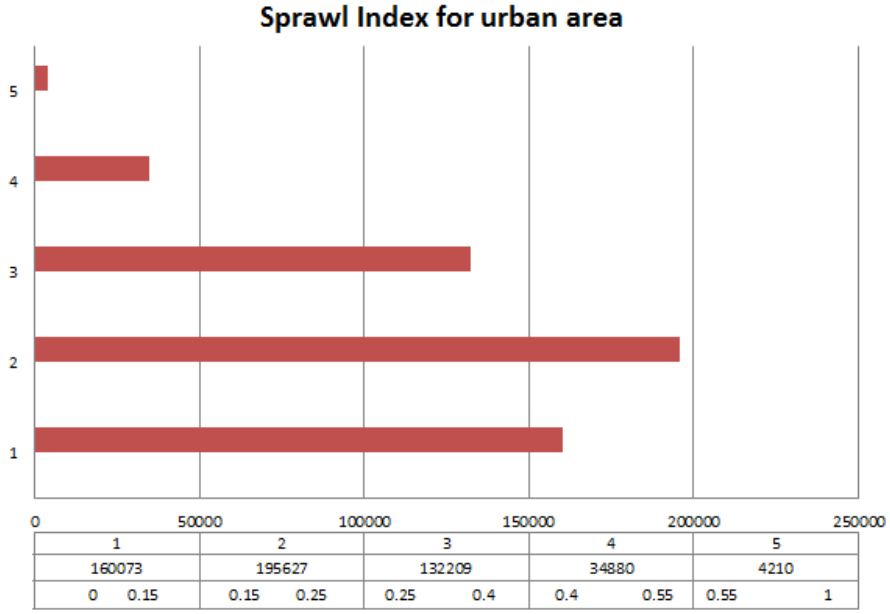


Figure 8. Categorized SI map

As the results are shown in Figure 8, the number of urban cell with SI bigger than 0.4 are 39090 cells. It means huge areas in the city are developed in area with high degree of sprawl. At the other side, some part of area in the class number 3 has approximately the big value of SI and these regions have the same problems.

In this research we have used three Landsat imageries with ground pixel size 28.5m resolutions. Number and date of these satellite imageries is an important issue in studies like urban sprawl. Because as number of satellite

imageries increase and the interval period of the imageries become even, the research precision and reliability will increase. In this research due to lack of comprehensive data (from number view) we have used three imageries. Also, resolution of the imageries is another important issue. Because as ground resolution of the imageries decrease, obviously, classification error in ground scale will decrease and this make the calculations accuracy increase.

4. Discussion & Conclusion

Since the 1988, the most important urban expansion has taken place in the North and Northwestern region of the city. These regions probably due to their appropriate elevation and slope situation and fast transportation networks growth have experienced such great expansions.

According to Table 4, Shannon Entropy for temporal span 1988-1999 is calculated 0.9667. The half of $\log_e(m)$ is 0.6931. Thus, according to earlier works (Bhatta et al 2010), it can be say due to bigger than the half of $\log_e(m)$, in this period; this city is considered as sprawled. For the second temporal span 1999-2010, Shannon entropy is calculated 1.0422 and this criterion has increased from 0.9667 in 1988-1999 to 1.0422 in 1999-2010. It seems the tendency to sprawl in this city has increased. This result obviously confirms that this city is becoming more sprawling.

As the result is shown in Figure 7, many area of this city have the SI indexes more than 0.4. Construction over the planning schemes, construction without considering urban planning criterion and so many reasons are causes to this sprawl condition in this city. A part of this inconsistency and uneven growth is derived from lack of temporal, sufficient and reliable data about history of the city.

This research can be considered as a useful method for analyzing urban growth in any cities. It should be mentioned that there are so many sprawl assessments factor for monitoring urban sprawl, but due to lack of sufficient and comprehensive data for analyzing urban and it landscape's and at the other side, practical methods which have been used in this research, it can be assumed as a powerful and great step for analyzing and monitoring urban growth in many cities. Integration of GIS, remote sensing data and statistical methods has been so productive in so many fields especially in urban planning.

According to the finding in Table 4, Tehran is sprawled and this city is becoming more sprawled. Urban growth in this city due to population growth, attraction of the city as the capital of Iran and social, economic potential of the city have increased so fast. These factors have resulted urban land and urban population growth so fast in this city.

Cities unavoidably will expand, but how to control the urban development while avoiding sprawl is an issue of urban planning. One of the solutions for this uneven dispersion is implementation of some policies that aim to restrict urban expansion in the cities to make them more compact. In other words, the policies should cover vertical growth and infilling growths of the city instead of edge expansion and out-filling expansion.

According to Figure 8, approximately all of expansion has occurred in area with distance larger than 4 km. A great part of these expansions have occurred in the area with distances larger than 9 km especially during 1999 to 2010. Trend to settlements in the distant area from center has increased especially during 1999 to 2010. Lower cost of living, cleaner air and so many reasons such as personal preferences, safety neighborhood and low range of crimes are probably the reasons for this huge expansion instead of vertical growth of the main parts of the city. But as we know it causes a lot of problems such as increasing numbers of travels and travel costs, agriculture and natural and open space loss, insufficient infrastructure and high infrastructure costs.

This research can be considered as an alarming to urban planners to inform them about situations of this city. Combination of GIS and remote sensing data can be used as a powerful method for analyzing urban growths. This should encourage urban managers and policy makers to implement such methods for determining development scenario.

Three satellite imageries have been used to analysis urban sprawl for Tehran Metropolis. The greater number of satellite imageries and using imageries with smaller ground pixel size definitely increase accuracy and reliability of results. But it should be mentioned for a developing country like Iran which historical urban data and land use map is not stored properly or even existed, this research using free and reliable satellite imageries data which approximately is the single source of data for these regions is a practical and scientific method for analyzing urban growth.

One of the major issues in cities intelligent management is the lack of proper and scientific development and as a result urban development in high slope and elevations, destruction of agricultural land, increased infrastructure and utilities costs, environmental deteriorate and natural hazards and the lack of optimum use of land have been

encountered (Mohammady et al, 2013).

An important finding of this study is the fact that this city in both the periods (1988-1999 and 1999-2010) has experienced sprawl and unfortunately, this trend is becoming larger.

References

Aghababaee, H., Amini, J., & Tzeng, Y. C. (2013). Contextual PolSAR image classification using fractal dimension and support vector machines. *European Journal of Remote Sensing, 46*, 317-332. http://dx.doi.org/10.5721/EuJRS20134618

Aguilera, F., Valenzuela, L., & Botequilha-Leitão, A. (2011). Landscape metrics in the analysis of urban land use patterns: A case study in a Spanish metropolitan area, *Landscape and Urban Planning, 99*, 226–238.

Akbari, M. M., & Abbaszadeh, G. (2013). The Physical Development of Mashhad City and Its Environmental Impacts. *Environment and Urbanization Asia, 3*(1), 79-91.

Alberti, M. (2005). The effects of urban patterns on ecosystem function. *International Regional Science Review, 28*(2), 168–192.

Anderson, J. R, Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data. *US Geological Survey, Professional Paper, 964*(28), Reston, VA/

Barnes, K. B., Morgan, J. M., Roberge, M. C., & Lowe, S. (2001). Sprawl development: Its patterns, consequences, and measurement, A white paper, Towson University. Retrieved from http://chesapeake.towson.edu/landscape/urbansprawl/download/Sprawl_white_paper.pdf

Bhatta, B. (2009). Analysis of urban growth pattern using remote sensing and GIS: a case study of Kolkata, India. *International Journal of Remote Sensing, 30*(18), 4733-4746.

Bhatta, B., Saraswati, S., & Bandyopadhyay, D. (2010). Quantifying the degree-of-freedom, degree-of-sprawl, and degree-of-goodness of urban growth from remote sensing data. *Applied Geography, 30*, 96–111.

Brabec, E., & Smith, C. (2002). Agricultural land fragmentation: the spatial effects of three land protection strategies in the eastern United States. *Landscape and Urban Planning, 58*, 255–268.

Bruekner, J. (2001). Urban Sprawl: Lessons from Urban Economics, *Brookings-Wharton Papers on Urban Affairs, 1*, 65-97.

Burton, E. (2001). *The Compact City and Social Justice*, A paper presented to the Housing Studies Association Spring Conference, Housing, Environment and Sustain ability, University of York.

Donnay, J. P., Barnsley, M. J., & Longley, P. A. (2001). Remote Sensing and Urban Analysis. In DONNAY J.P., BARNSLEY M.J., LONGLEY P.A. (Eds.), *Remote Sensing and Urban Analysis* (pp. 3-18). Taylor & Francis, London.

Dutta, V. (2012). Land Use Dynamics and Peri-urban Growth Characteristics: Reflections on Master Plan and Urban Suitability from a Sprawling North Indian City. *Environment and Urbanization Asia, 3*(2), 277-301.

Ewing, R., Pendall, R., & Chen, D. (2002). *Measuring Sprawl and its Impact*, Smart Growth America, Washington D.C.

Galster, G., Hanson, R., Ratcliffe, M. R., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept. *Housing Policy Debate, 12*(4), 681-717.

Gao, J., Liu, Y. S., & Chen, Y. F. (2006). Land cover changes during agrarian restructuring in Northeast China. *Appl. Geogr, 26*, 312–322.

Ghanghermeh, A., Roshan, G., Orosa, A., Calvo-Rolle, J. L., & Costa M. Á. (2013). New Climatic Indicators for Improving Urban Sprawl: A Case Study of Tehran City. *Entropy, 15*, 999-1013.

Hamin, E. M., & Gurran, N. (2009). Urban form and climate change: Balancing adaptation and mitigation in the U.S. and Australia, *Habitat International, 33*, 238-245.

Hasse, J. E., & Lathrop, R. G. (2003). Land Resource Impact Indicators of Urban Sprawl. *Journal of Applied Geography, 23*, 159–175.

Hayek, U., Jaeger, J. A. G., Schwick, C., Jarne, A., & Schuler, M. (2011). Measuring and assessing urban sprawl: What are the remaining options for future settlement development in Switzerland for 2030? *Applied Spatial Analysis and Policy, 4*(4), 249-279.

Herold, M., Goldstein, N. C., & Clarke, K. C. (2003). The spatiotemporal form of urban growth: Measurement,

analysis and modeling. *Remote Sensing of Environment, 86*(3), 286–302.

Inostroza, L., Baur, R., & Csaplovics, E. (2013). Urban Sprawl and Fragmentation in Latin America: A Dynamic Quantification and Characterization of Spatial Patterns. *Journal of Environmental Management, 115*, 87-97.

Jensen, J. R. (2005). Introductory Digital Image Processing: A Remote Sensing Perspective, Prentice- Hall, Upper Saddle River.

Jiang, F., Liu, S., Yuan, H., & Zhang, Q. (2007). Measuring urban sprawl in Bejing with geo-spatial indices. *Journal of Geographical Sciences, 17*(4), 469-478. http://dx.doi.org/10.1007/s11442-007-0469-z

Johnson, M. P. (2001). Environmental impacts of urban sprawl: A survey of the literature and proposed research agenda. *Environmental Planning A, 33*, 717–735.

Kennedy, L. (2007). Regional industrial policies driving peri-urban dynamics in Hyderabad, India. *Cities, 24*(2), 95–109.

Kim, J., & Ellis, C. (2009). Determining the Effects of Local Development Regulations on Landscape Structure: Comparison of the Woodlands and North Houston, TX. *Landscape and Urban Planning, 92*, 293–303.

Kumar, J. A. V., Pathan, S. K., & Bhanderi, R. J. (2007). Spatio-temporal analysis for monitoring urban growth—a case study of Indore city. *Journal of Indian Society of Remote Sensing, 35*, 11–20.

Lata, K. M., Rao, C. H. S., Prasad, V. K., Badarianth, K. V. S, Rahgavasamy, V., & Rao, C. H. S. (2001). Measuring urban sprawl: A case study of Hyderabad. *GIS Development, 5*, 26–29.

Li, X., & Yeh, A. G. (2004). Analysing Spatial Restructuring of Land Use Patterns in a Fast Growing Region Using Remote Sensing and GIS. *Landscape Urban Plan, 69*, 335-354.

Li, Y., Zhao, S., Zhao, K., Xie, P., & Fang, J. (2006). Land-cover Changes in an Urban Lake Watershed in a Mega-city, Central China. *Environmental Monitoring and Assessment, 115*, 349-359.

Mohammady, S., Delavar, M. R., & Pijanowski, B. C. (2013). Urban growth modeling using anfis algorithm: a case study for Sanandaj city, Iran, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W3, 493-498. Retrieved from http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/ XL-1-W3/493/2013 /isprsarchives-XL-1-W3-493-2013.html

Muniz, T., & Galindo, A. (2005). Urban form and the ecological footprint of commuting. The case of Barcelona, *Ecol. Econs., 55*, 499– 514.

Pijanowski, B. C., Pithadia, S., Shellito, B. A., & Alexandridis, K. (2005). Calibrating a neural network-based urban change model for two metropolitan areas of Upper Midwest of the United States. *Int. J. Geogr. Inf. Syst., 19*(2), 197–215.

Pourahmad, A., Baghvand, A., Zangenehe, S., & Givehchi, S. (2007). The Impact of Urban Sprawl up on Air Pollution. *Int. J. Environ. Res., 1*(3), 252-257.

Romano, B. (2004). *Environmental Fragmentation Tendency*. The Sprawl Index. ERSA - Porto, Portugal.

Roshan, G., Rousta, I., & Ramesh, M. (2009). Studying the effects of urban sprawl of metropolis on tourism - climate index oscillation: A case study of Tehran city. *Journal of Geography and Regional Planning, 2*(12), 310-321.

Rudel, T. K., & Macdonald, K. (2005). Sprawl and forest cover: what is the relationship? *Appl. Geogr., 25*, 67–79.

Sim, S., & Mesev, V. (2011). Measuring urban sprawl and compactness: Case study Orlando, USA. International Cartographic Conference, Paris, France.

Sudhira, H. S., Ramachandra, T. V., & Jagadish, K. S. (2004). Urban sprawl: metrics, dynamics and modeling using GIS. *International Journal of Applied Earth Observations and Geoinformation, 5*(1), 29-39.

Thomlinson, J. R., Bolstad, P. V., & Cohen, W. B. (1999). Coordinating methodologies for scaling land cover classifications from site-specific to global: Steps toward validating global map products. *Remote Sensing of Environment, 70*, 16– 28.

Torrens, P. M., & Alberti, M. (2000). *Measuring sprawl*. Working paper no. 27, Centre for Advanced Spatial Analysis, University College London.

Xie, J. Y., & Cho, N. N. (2007). Spatial and temporal dynamics of urban sprawl along two urban-rural transects: a case study of Guangzhou, China. *Landscape Urban Plan, 79*(15), 96–109.

Yeh, A. G. O., & Li, X. (2001). Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogramm Eng Remote Sens, 67*(1), 83–90.

Yeh, A. G., & Li, X. (1999). Measurement of Urban Sprawl in a Rapid Growing Region Using Entropy. *Towards Digital Earth* - Proceedings of the International Symposium on Digital Earth, Science Press.

Zhang, T. (2000). Land market forces and government's role in sprawl. The case of China. *Cities, 17*(2), 123–135.

Zhang, X., Chen, J., Tan, M., & Sun, Y. (2007). Assessing the impact of urban sprawl on soil resources of Nanjing city using satellite images and digital soil databases. *Catena, 69*, 16–30.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Isothermal Pneumatc Molding of Dome-Shaped Parts of Anisotropic Material in Short-Time Creeping Mode

Sergey Sergeevich Yakovlev[1], Sergey Nikolaevich Larin[1] & Valeriy Ivanovich Platonov[1]

[1] Tula state university, Tula, Russian Federation

Correspondence: Sergey Nikolaevich Larin, Tula state university, Tula, Russian Federation.

Received: October 16, 2014 Accepted: November 1, 2014 Online Published: June 29, 2015

doi:10.5539/mas.v9n8p13 URL: http://dx.doi.org/10.5539/mas.v9n8p13

Abstract

Studying isothermal straining of dome-shaped shells with the aim of evaluating kinematics of the material flow, stress and strained conditions, force conditions, geometrical dimensions of shells and possibility frontier of straining. The theoretical study of the isothermal straining processes has been performed on the basis of the theory of short-time creeping of anisotropic material. There has been revealed the influence of technological parameters, conditions of loading, and geometry of working tools on the kinematics of material flow produced by friction conditions on the blank and tool contact surfaces, stress and strain states, force conditions, possibility frontiers of form-shaping of operations of isothermal deforming of dome-shaped shells made out of high-strength anisotropic materials. Conclusions. The equations obtained can be used for theoretical analysis of operations of isothermal straining of dome-shaped shells in the mode of short-time creeping.

Keywords: anisotropy, dome-shaped shells, high-strength materials, stress, strain, isothermal straining, viscosity, damaging, destruction

1. Introduction

Dome-shaped parts have found their wide usage in various industries. The traditional methods of their production by means of pressing are rather labor-consuming and problematic for securing requisite geometric precision, due to residual stress, which causes deformation of the contour and hence an extensive scope of fitting and component-adjusting operations to adjust the parts to the requisite dimensions.Residual stress is largely caused by the inbuilt anisotropy of the mechanical properties in the sheet under deformation and by the uneven character of the deformations. Isothermal changing of the shape of dome-shaped parts with gas of highly strong aluminum and titan alloys is considerably advantageous versus the conventional processing methods, and is quite promising for industrial use [1-6].

Sheet material, subjected to deforming processes, as a rule, possesses certain anisotropy of mechanical properties, determined by the material grade, by technological modes of its production, which can produce a positive or a negative effect upon stable procedure of technological processes of metal treatment by pressure under various temperature-speed modes of deformation [5-15].

Single- and multilayered cellular structures (waffle-grid structure) are used for the shells of the product vessels. At a high load-bearing capacity, these structures possess relatively light weight and strength balance while loading. The existing technological processes of manufacturing cellular structures include machining operations (milling) and electroerosion. Technological principles of a hot slow-speed forming through the excess gas pressure can be used in the production of cellular structures out of aluminum and titanium alloys.

2. Methodology

The theoretical study of processes of isothermal straining is carried out on the basis of the theory of short-time creeping of anisotropic material. The object of study was stressing of an anisotropic material under conditions of short-time creeping. Short-time creeping shall be assumed as a slow straining under conditions of viscous or ductile flow. The elastic components of strain are ignored [1]. The analysis of the stress and strain conditions of the blank, of the force conditions of isothermal pneumatic molding of dome-shaped parts has been performed numerically by the method of finite-difference mutual relations, use being made of computers, for joint solution of differential balance equations, equations of state and kinematic correlations under postulated initial and boundary conditions [10]. The possibility frontiers of deforming have been set on the basis of use of

phenomenological criteria of destruction (stress and strain) of anisotropic material associated with accumulation of microdamages.

Main part. We have considered straining of a round sheet blank with radius R_0 and thickness h_0 by free bulging in the mode of viscous flow of material when exposed to the excess gas pressure $p = p_0 + a_p t^{n_p}$ into a spherical matrix (Fig. 1). Here p_0, a_p, n_p are stressing constants.

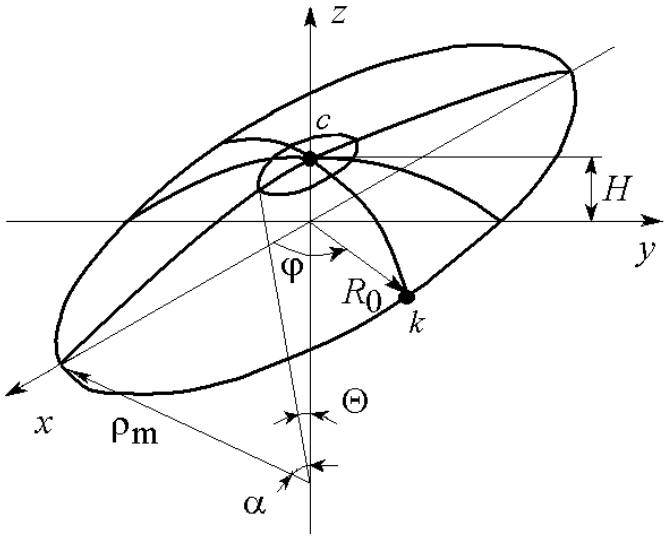


Figure 1. Scheme to the calculation of strained state of the blank middle surface in the meridian plane

The blank is fixed along the outer contour. The blank material is taken as transversally isotropic with anisotropy factor R; the stressed state of the shell is two-dimensional, i.e., the stress perpendicular to the sheet plane equals zero ([sigma]$_z$=0). We consider stressing in the meridian plane of the shell as a membrane. Due to the symmetry of the mechanical properties of the material versus the blank axis, and due to the nature of the external forces behavior, stresses and velocities of straining, which are meridian, circumferential and normal to the median surface of the blank, are the chief ones.

The median surface of the blank remains a part of the spherical surface at every stage of the straining. A radial flow of material relative to the new center at every stage of the straining takes place at any meridian section of the shell.

According to the adopted allowances, the curvature radii of the meridian section [rho]$_m$ of the median surface, and those of the shell section by the conical surface perpendicular to the meridian arch, [rho]$_t$ are equal, i.e.,

$$\rho_m = \rho_t = \rho = \frac{H^2 + R_0^2}{2H} \tag{1}$$

where H - is the height of the dome at the given time instant of straining.

It is assumed that the stresses are distributed evenly over the thickness of the element cut out of the membrane by meridian and conical surfaces in the vicinity of the point under consideration. As distinct from the available solutions, we make no limitation as to changing the thickness of the shell along the circumference arch in the meridian section.

The straining velocities in the meridian [ksi]^c_m, circumferential direction [ksi]^c_t and along the thickness [ksi]^c_z of the shell are determined accordingly by the following formulas

$$\xi^c_m = \left(\frac{\sin\theta}{\theta\sin\alpha} - ctg\alpha\right)\dot{\alpha}; \qquad \xi^c_t = \left(\frac{\cos\theta}{\sin\alpha} - ctg\alpha\right)\dot{\alpha}; \qquad \xi^c_z = \frac{\dot{h}}{h} \tag{2}$$

Here [teta] – is the actual angle between the blank vertical line of symmetry and the radius-vector determining the position of the point in the median surface section by a diagonal plane; $\dot{\alpha} = d\alpha/dt$; $\dot{h} = dh/dt$

It was assumed for the shell strain that at every step of straining there takes place a radial flow of the median

surface point in the meridian plane relative to the new center at time instant $t+dt$, i.e., in direction [teta]+*d*[teta].

The shell thickness in the dome of the middle surface of the shell ([teta]=0) is found by the following expression

$$h = h_0 \Big/ \left(1+\frac{H^2}{R_0^2}\right)^2 \tag{3}$$

Change of the shell thickness depending on straining time t at the point of fixation ([teta=alfa]) is found by the formula

$$h = h_0 \frac{H}{R_0\left(1+\frac{H^2}{R_0^2}\right) arctg\frac{H}{R_0}} \tag{4}$$

The values of the meridian $[sigma]_m$ and tangential $[sigma]_t$ stresses in the process of deforming with $[rho_m=rho_t]$, are found by the formula [10]:

$$\sigma_m = \sigma_t = \frac{p\rho}{2h} \tag{5}$$

The equivalent velocity of straining $[ksi]^c_e$ and stress $[sigma]_e$ at the dome apex (point "c") and at the fixation point of the shell in the contour (point "k") are calculated for anisotropic material respectively after formulas [5, 6]:

$$\xi^c_{ec} = \frac{2}{\sqrt{3}}\sqrt{2+R}\,\xi^c_{mc}$$

$$\xi^c_{ec} = \frac{2}{\sqrt{3}}\sqrt{2+R}\,\xi^c_{mc}; \qquad \sigma^c_{ec} = \frac{\sqrt{3}}{2\sqrt{2+R}}\sigma_{mc}; \tag{6}$$

$$\xi^c_{ek} = \left\{\frac{2}{3}\frac{(2+R)(R+1)}{2R+1}\right\}^{1/2}\xi_{mk}; \quad \sigma_{ek} = \left\{\frac{3}{2}\frac{2R+1}{(2+R)(R+1)}\right\}^{1/2}\sigma_{mk} \tag{7}$$

There has been considered a slow isothermal straining of a shell made out of a material, for which work the equations of state for the energy theory of creeping and damaging [6]:

$$\xi^c_e = B\left(\sigma_e/\sigma_{e_0}\right)^n/\left(1-\omega^c_A\right)^m; \quad \dot{\omega}^c_A = \sigma_e\,\xi^c_e/A^c_{np}, \tag{8}$$

Where B, n, m - are material constants depending on the test temperature; $[omega]^c_A$ – is damageability of material for viscous deformation after energy fracture model; A^c_{np} - specific fracture work for viscous flow of material; $\dot{\omega}^c_A = d\,\omega^c_A/d\,t$; $[ksi]^c_e$ and $[sigma]_e$ – are equivalent velocities of straining and stress; $[sigma]_{e0}$ – is equivalent stress dividing the viscous and ductile flows of material.

Since the pressure value p is distributed evenly at every moment of deformation over the surface of the shell, its value was found at the apex of the shell dome (point "*c*").

By substituting in the first equation of the material state (8) its values of $[sigma]_e$ and $[ksi]^c_e$, are to be found after formulas (6) and (7), taking account of correlations (3), (4), (5), we obtain

$$p^n dt = \frac{\sigma^n_{e0}\left(1-\omega^c_{Ac}\right)^m 2^{2n+2}\left(2+R\right)^{\frac{n+1}{2}} H^{n+1}h^n dH}{3^{\frac{n+1}{2}} B\left(H^2+R_0^2\right)^{n+1}} \tag{9}$$

Shell thickness h is found after formula (3).

The value of the accumulated damage is found after:

$$\dot{\omega}^c_{Ac} = \frac{p\left(1+\frac{H^2}{R_0^2}\right)^2}{h_0 A^c_{np}}\dot{H} \tag{10}$$

Where $\dot{H} = dH/dt$.

Destruction time t_* is found from $[\text{omega}]^c_{Ac}=1$.

Carried out in a similar manner was the study of stress and strain states of a blank in the apex of the dome shell (point "*c*") and at the point of the shell fixation (point "k"), as well as received were basic equations and proportions for solving the set task, with the assumption that material behavior is subject to the equations of the kinetic theory of creeping and damaging under the well-known law of pressure and time dependence $p = p\,(t)$ and under permanent equivalent velocity of straining in the blank dome $[\text{ksi}]_{e1}$

There has been developed the algorithm for calculation of force and straining parameters of the analyzed technological process and the software.

There have been evaluated the stress and strained states, the kinematic of the material flow, force conditions and the possibility frontier of straining process under investigation as relating to accumulation of microdamages, as depending on the anisotropy of the mechanical properties of the basic material, depending on the law of loading, and on the geometric dimensions of the blank and the product.

Calculation was made for titanium alloys BT6C at $T = 860^{\circ}C$, which behavior is described by the energy theory of creeping and damaging, and for titanium alloy BT14 at $T = 950^{\circ}C$, which behavior is subject to the kinetic theory of creeping and damaging. The mechanical characteristics of these materials under straining in conditions of viscous flow of material are shown in the works [5, 6].

The characteristic curves for changing values of gas pressure p, relative values of blank thickness in the dome $\bar{h}_c = h_c/h_0$ and at the place of its fixation $\bar{h}_k = h_k/h_0$, dependence of the height of the dome-shaped blank $\bar{H} = H/R_0$ on straining time instant t for titanium alloy BT6C ($T = 860^{\circ}C$) at a permanent equivalent straining velocity in the blank dome $[\text{ksi}]_{e1}$ are shown in Fig. 2. The experimental data are shown here with dots.

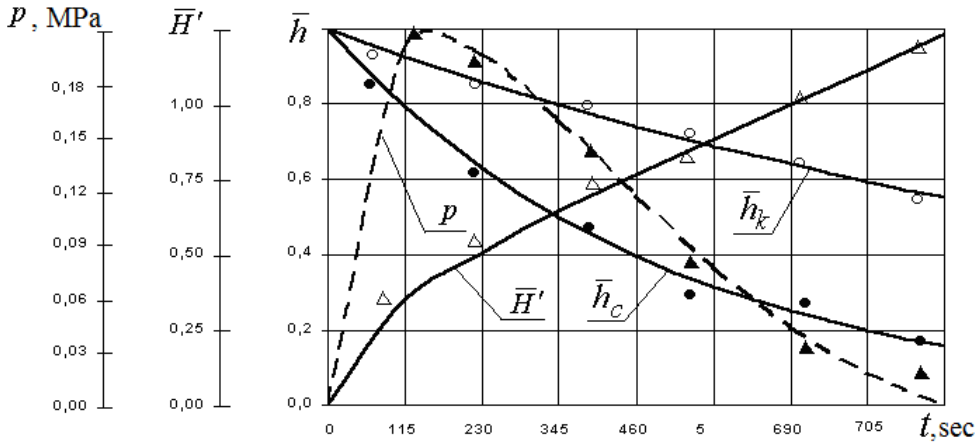


Figure 2. Dependences of p, $\bar{H}$ and $\bar{h}$ changing in the discussed points of a blank from t for titanium alloy BT6C ($\bar{R}_0 = 300$; $\xi_{e1} = 0,002\ 1/c$)

It follows from an analysis of the results of calculations and characteristic curves that as straining time t grows to a certain limit, there takes place a steep increase in the blank relative height in dome $\bar{h}_c$ and in its place of fixation $\bar{h}_k$. Further growth of straining time t results in a smooth change of the analyzed values. At time instant t, which is close to the blank destruction, there takes place an abrupt change of the relative values $\bar{H}$, $\bar{h}_c$ и $\bar{h}_k$. This is caused by an intensive growth of accumulated microdamages at the concluding stage of the process.

It has been established that the changing of the relative thickness of the blank's dome $\bar{h}_c$ takes place in a more intensive manner as versus the changing of its relative thickness at its place of fixation $\bar{h}_k$. The difference grows together with straining time t and may reach 50 %.

It has been shown that for securing a stable equivalent speed of deformation in the blank dome, the law of changing pressure p during straining time t is of a complex nature. During the initial moment of the shape-changing we witness a steep change in the pressure p, since there takes place a substantial change of the radius of the half-sphere $[rho]_m$. Subsequent growth of the straining time t is accompanied by a decrease in the magnitude of the gas pressure p.

Comparing theoretical and experimental data on the relative thickness in the blank dome $\bar{h}_c$ and at the place of its fixation $\bar{h}_k$, as well as on the relative height of the blank $\bar{H}$, shows their satisfactory coincidence (up to 10 %) [6].

It has been established that the destruction of the blank during the isothermal straining takes place in the dome of the part, where happens the maximum thinning-out of the blank.

Figure 3 shows the changed time of destruction t_*, the changed relative height $\bar{H}_*$, and the changed thickness of the blank dome $\bar{h}_*$ at the moment of destruction, as determined by the magnitude of the accumulated microdamages with $[omega]^c_A=1$, -- as having functional connections to the value of the permanent equivalent velocity of straining in the blank dome $[ksi]_{el}$. It is shown that increased parameters of the law of loading a_p, n_p and increased magnitude of the value of the permanent equivalent straining velocity in the blank dome $[ksi]_{el}$ result in a shorter time of destruction t_* and in a smaller relative height $\bar{H}_*$, as well as in an increased relative thickness of the blank dome $\bar{h}_*$.

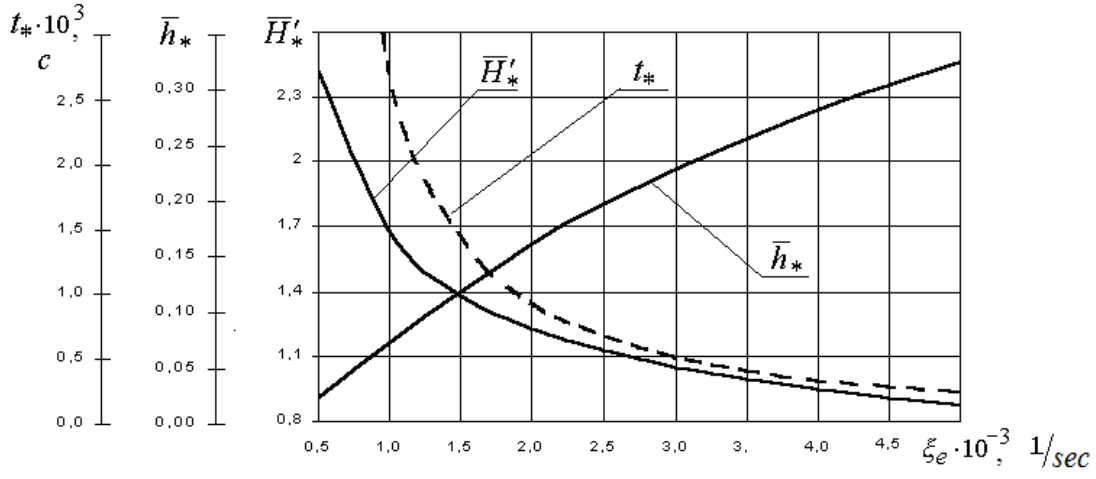


Figure 3. Dependences of changing t_* and $\bar{H}_*$, $\bar{h}_*$ in the blank dome as relative to $[ksi]_e$ for titanium alloy BT6C ($\bar{R}_0 = 300$)

It is determined that normal anisotropy factor R influences considerably the magnitude of the destruction time and the relative values $\bar{H}_*$, $\bar{h}_*$. The calculations have been made at $\bar{R}_0 = 300$; $p_0 = 0,013\ МПа$; $a_p = 4 \cdot 10^{-3}\ МПа\,/c^{n_p}$; $n_p = 0,6$ (the kinetic theory). As there grows the anisotropy factor R, the relative value of $\bar{h}_*$ steeply increases, while destruction time t_* and the relative blank height fall abruptly. It has been established that the failure of considering the anisotropy of the mechanical properties of the blank when analyzing the process of isothermal changing of the spherical shell shape – makes us err in destruction time t_* by about 35 %, while in the relative height $\bar{H}_*$ and in the thickness of the blank dome $\bar{h}_*$ at the moment of destruction – we err by 15 %.

Analysis of the calculation results shows that the limit possibilities of shape-changing in the mode of viscous

flow of material, which conduct is subject to the kinetic theory of creeping and damaging (alloy BT14), do not depend on the conditions of the blank loading. Shown is the substantial dependence of destruction time t_* on loading parameters a_p, n_p and on the value of the permanent equivalent straining velocity $[\text{ksi}]_{el}$. Increasing the loading parameters a_p from $0.2 \cdot 10^{-3}$ $МПа/c^{n_p}$ to $1.4 \cdot 10^{-3}$ $МПа/c^{n_p}$, and those of n_p from 0.5 to 0.8 with unchanged other parameters, results in reduction of the destruction time t_* by 1.8 times.

There has been shown a substantial influence of the geometric dimensions of the blank upon the destruction time magnitude t_*. It has been established that an increase in the relative size of the blank radius $\overline{R}_0 = R_0 / h_0$ from 200 to 800 results in a shorter destruction time by more than 4 times.

3. Discussion

Isothermal gas forming of the domed components made of high-strength aluminum and titanium sheet alloys has significant advantages over traditional methods of treatment and is very promising when it is used within the industrial sector. It is less time-consuming and problematic in terms of providing the necessary geometric accuracy due to the presence of residual stresses; it does not cause edge deformation and does not require a large volume of thereof associated mechanical and finishing works on parts fitting to the specified dimensions [11 - 20].

The suggested mathematic model for the isothermal straining of domed shells made of anisotropic high-strength materials in the mode of creeping, allows us to evaluate the influence of technological parameters, loading conditions, geometry of the working tool, friction conditions on the contact surface of the workpiece and tool on the kinematics of material flow, stress and strain states, force conditions, possibility frontier of forming the process in question, taking into account the anisotropy of mechanical properties, which distinguishes this work out of many other alike studies. (Authors of article offer mathematical model of isothermal deformation of dome-shaped covers by superfluous pressure of gas from anisotropic high-strength materials in a creep mode. The given model allows to estimate kinematics of a current of the material, the intense and deformed conditions of a cover, power modes, the geometrical sizes of made dome-shaped covers and limiting possibilities of deformation. The offered mathematical model can be used as base by working out of technological processes of isothermal deformation of dome-shaped covers).

4. Conclusions

Result analysis of the considered process allows drawing the following conclusions:

On the basis of the developed mathematical model the analysis of isothermal deformation of dome-shaped covers with a view of an estimation of kinematics of a current of the material strained and deformed conditions, power modes, the geometrical sizes of covers and limiting possibilities of deformation, in which result has been carried out:

- We have established the influence of technological parameters, the loading conditions, tools geometry, conditions of friction on the contact surfaces of the blank and the tool as affecting the kinematics of the material flow, the stress and strain states, force conditions, the possibility frontier in forming the operations of isothermal straining of dome shells made out of highly strong anisotropic materials.

- It has been shown that for the purpose of providing for permanent equivalent straining velocities in the blank dome, the law of pressure change during straining has a complicated character. It has been determined that changing the relative thickness of the blank dome happens in a more intensive manner by comparison to changing the relative thickness at the main points. As the straining time is increased, the difference may grow to reach 50 %.

- It has been established that in the process of isothermal straining of domed blanks, the destruction of the blank occurs in the dome of the blank where there happens a maximal thinning-out of the blank. Increase of the parameters of loading and equivalent straining velocity results in reduction of the relative limit height of the blank and of the destruction time for materials, which conduct is described by the energy theory of creeping and damaging. The possibility frontier of shape-forming during isothermal straining of anisotropic materials, which behavior is described by the energy theory of creeping and damaging, do not depend on the blank loading conditions.

- It has been established that lack of consideration for anisotropy of the blank mechanical properties when analyzing the process of isothermal straining of dome-shaped shells, gives us an error in evaluating the destruction time of about 45%, and of 30% for the relative height and thickness in the blank dome at the moment of destruction.

Acknowledgement

This work has been performed within the frameworks of the basic part of the State Task #2014/227 for performing research and development works of the Ministry of Education and Science of the Russian Federation for the years 2014-2020 and of the RFFI grant # 14-08-00066 a.

References

Abbasia, M., Saeed, A. A., & Naderi, M. (2012). The effect of strain rate and deformation temperature on the characteristics of isothermally hot compressed boron-alloyed steel. *Materials Science and Engineering: A. 538*, 356–363. http://dx.doi.org/10.1016/j.msea.2012.01.060

Alwin, S., Volker, U., Christoph, E., Rainer, K., Alfred, K., Monteroe, M. C., Roland, R., Wolfgang, S., Domenico, S., & Dominique, V. (2008). Opportunities and challenges of spray forming high-alloyed steels. *Materials Science and Engineering: A. 477*(1–2), 69–79. http://dx.doi.org/10.1016/j.msea.2007.08.082

Chana, K. C., Wanga, G. F., Wanga, C. L., & Zhang, K. F. (2005). Low temperature superplastic gas pressure forming of electrodeposited Ni/SiCp nanocomposites. *Materials Science and Engineering: A., 404*(1–2), 108–116. http://dx.doi.org/10.1016/j.msea.2005.05.042

Chunga, S. W., Higashia, K., & Kim, W. J. (2004). Superplastic gas pressure forming of fine-grained AZ61 magnesium alloy sheet. *Materials Science and Engineering: A., 372*(1–2), 15–20. http://dx.doi.org/10.1016/j.msea.2003.08.125

DONG, W. P., & Fea, J. C. (2008). Simulation of 4A11 piston skirt isothermal forging process. *Transactions of Nonferrous Metals Society of China, 18*(5), 1196–1200. http://dx.doi.org/10.1016/S1003-6326(08)60204-6

Eva-Lis Odenbergera et al. (2008). Thermo-mechanical material response and hot sheet metal forming of Ti-6242. *Materials Science and Engineering: A., 489*(1–2), 158–168. http://dx.doi.org/10.1007/s12289-012-1094-7

F Grechnikov, F. V. (1998). Deformation of Anisotropic Materials. M. Mashinostroyenie, p. 446.

Fuxiaoa, Y., zhong, C. J., Ranganathanb, S., & Dwarakadasab, E. S. (2010). Fundamental differences between spray forming and other semisolid processes. *Materials Science and Engineering: A., 304–306*, 621–626. http://dx.doi.org/10.1016/S0921-5093(00)01547-1

Golenkov, V. A., Yakovlev, S. P., Golovin, S. A., Yakovlev, S. S., & Kukhar, V. D. (2009). Theory of Metal Processing with Pressure. University text-book. M, p. 442.

Li, L. A., & Zhang, X. M. (2011). Hot compression deformation behavior and processing parameters of a cast Mg–Gd–Y–Zr alloy. *Materials Science and Engineering: A., 528*(3), 1396–1401. http://dx.doi.org/10.1016/j.msea.2014.07.025

Michael, J., O'Briena, F. von, Bremenb, H., Minoru, F., Zenji, H., & Terence, G. L. (2007). A finite element analysis of the superplastic forming of an aluminum alloy processed by ECAP. *Materials Science and Engineering: A., 456*(1–2), 236–242. http://dx.doi.org/ 10.1016/j.msea.2006.11.116

Naderia, M., Durrenbergerb, L., Molinarib, A., & Blecka, W. (2008). Constitutive relationships for 22MnB5 boron steel deformed isothermally at high temperatures. *Materials Science and Engineering: A., 478*(1–2), 130–139. http://dx.doi.org/10.1016/j.msea.2007.05.094

Puertas, I., Luis-Pérez, C. J., Salcedo, D., León, J., Luri, R., & Fuertes, J. P. (2013). Isothermal Upset Forging of AA5083 after Severe Plastic Deformation by ECAE. *Procedia CIRP., 12*, 288–293. http://dx.doi.org/10.1016/j.matdes.2013.05.089

Rusz, S., Sinczak, J., & Lapkowski, W. (1997). Isothermal plastic forming of high-carbon steel. *Materials Science and Engineering: A., 234–236*(30), 430-433.

Sinczak, J., Lapkowski, W., & Rusz, S. (1997). Isothermal plastic forming of high melting temperature alloys. *Journal of Materials Processing Technology, 72*(3), 429-433. http://dx.doi.org/10.1016/S0924-0136(97)00206-9

Yakovlev, S. P., Chudin, V. N., Sobolev, Y. A., Yakovlev, S. S., Tregubov, V. I., & Larin, S. N. (2009). Isothermal Pneumatic Moulding of Anisotropic High-Strength Sheet Materials. M. Mashinostroyenie, p. 352.

Yakovlev, S. P., Chudin, V. N., Yakovlev, S. S., & Sobolev, Y. A. (2004). Isothermal Straining of High-Strength Anisotropic Materials. M. Mashinostroyenie, p. 427.

Yakovlev, S. P., Yakovlev, S. S., & Andreichenko, V. A. (1997). Processing Anisotropic Materials with Pressure.

Kishinev, Quant, pp: 332.

Yakovlev, S. S., Kukhar, V. D., & Tregubov, V. I. (2012). Theory and Technology of Stamping Anisotropic Materials. M. Mashinostroyenie, p. 400.

Zhao, W. J., Cao, F. Y., Gu, X. L., Ning, Z. L., Han, Y., & Sun, J. Y. (2013). Isothermal straining of spray formed Al–Zn–Mg–Cu alloy. *Mechanics of Materials, 56*, 95–105. http://dx.doi.org/10.4028/www.scientific.net/MSF.788.565.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Regression Modeling for Spur Gear Condition Monitoring Through Oil Film Thickness Based on Acoustic Emission Signal

Yasir Hassan Ali[1], Roslan Abd Rahman[2] & Raja Ishak Raja Hamzah[3]

[1] Department of Applied Mechanics and Design, Faculty of Mechanical Engineering, University Technology Malaysia, Johor, Malaysia

Correspondence: Raja Ishak Raja Hamzah, Department of Applied Mechanics and Design, Faculty of Mechanical Engineering, University Technology Malaysia, 81310 UTM Skudai, Johor Bahru, Johor, Malaysia. Tel: 60-197-511-951. E-mail: rishak@fkm.utm.my/yha2006@gmail.com/rosln@fkm.utm.my

Received: January 28, 2015 Accepted: February 12, 2015 Online Published: July 6, 2015

doi:10.5539/mas.v9n8p21 URL: http://dx.doi.org/10.5539/mas.v9n8p21

Abstract

The main purpose of a gear lubricant is to provide adequate oil film thickness to reduce and prevent gear tooth surface failures .Until now, there is no study in the literature related to the estimation of oil film thickness through Acoustic emission signals. In this study, for spur gear condition monitoring a new approach based on mathematical model was presented for oil film regimes detection.

This study focuse on the ability of regression model to find whether the gearbox is running in elastohydrodynamic, mixed wear or severe wear lubrication regime. Then, predaction the accuracy of the model is measured by examining the error that been produced by using Mean Squared Error and Mean Absolute Error. In this paper a mathematical model for time-series prediction was considered and the results shows the ability of the regression model to predict oil film regime.

Keywords: regression, acoustic emission, condition monitoring, oil film thickness, spur geaer

1. Introduction

A gear system is not only a crucial component of a machine, but also represents an essential functional module. Any unpredictable failures in a gear system often result in circumstances that are dangerous and can have severe financial consequences (Ali, Rahman et al. 2014, Ali 2014). Thus gears need to be inspected periodically for sound, vibration and development of cracks (if any), teeth, and bearings failures. Gear failures are generally due to bending, fatigue, contact fatigue, wear and scuffing. These faults in gear can be detected by monitoring vibration, torque, temperature, acoustic and lubrication film thickness continuously or online. Due to the damage on teeth surface vibration, torque, temperature, acoustic emission of gear shaft increases and condition of oil film thickness change to mixed wear, severe wear or elastohydrodynamic (EH) (Sreepradha, Kumari et al. 2014).

Condition monitoring of machines is gaining importance in industry because of the need to increase reliability and to decrease possible loss of production. To achieve this, it is necessary to establish a maintenance monitoring program to regularly assess the condition of a machine. There are a number of different methods that can be employed for machine condition monitoring to support maintenance decisions. The use of oil analysis, vibration and acoustic emission (AE) signals is quite common in the field of condition monitoring of rotating machinery (Peng and Kessissoglou 2003; Samanta, Al-Balushi et al. 2003).

Estimation or measurement of lubricating oil film thickness has been carried out by many researchers to predict healthy or unhealthy condition of the machine. Major purpose of a gear lubricant is to provide adequate oil film thickness to reduce and prevent gear tooth surface wear, to predict and overcome wear related damage progression in gear transmission systems. Various condition monitoring techniques have been developed in the past, which includes vibration, acoustic emission, oil/wear and sound analysis (Sung, Tai et al. 2000; Amarnath, Sujatha et al. 2009). The experimental results from the tests clearly show a direct correlation between AE RMS and the specific film thickness (λ) it was reported that the measured AE RMS was very sensitive to changes in specific film thickness (Hamzah, Al-Balushi et al. 2008; Raja Hamzah and Mba 2009; Hamel, Addali et al. 2014).

Mathematical modeling and statistical analysis is used here to characterize the causes of the gear failure in terms

of classifying oil film thickness with acoustic emission signal. Regression model was used to estimate/classify oil film thickness for condition monitoring of satisfactory gear operation and thereby resolving the faults diagnosing issues. Thus, to carry out the failure analysis of gear, AE RMS are used to predict the oil film thickness. The Regression modeling is one of the statistical approaches to derive the mathematical model for machine condition monitoring (ZHANG and NIE 2009). The regression based model is considered as the simplest approach to build a normal behavior model and may be used for monitoring (Schlechtingen and Ferreira Santos 2011).Time series regression models attempt to predict a future response based on the response history. In recent times there have been widespread interests in the use of multiple models for pattern classification and regression in statistics and neural network community (Subasi 2007).

Classification program was designed and used to classify the oil film thickness to Hydrodynamic lubrication (HL), Elastohydrodynamic lubrication (EHL) and Boundary lubrication (BL) based on the specific film thickness magnitude.

In this paper, we try to predict specific film thickness through AE signal because AE technique is a successful method in machinery condition monitoring and fault diagnosis due to its high sensitivity on locating micro cracks in high frequency domain. Acoustic emission waves are sent from an emission source and transferred to the surface by the transmission medium. The low displacement or high frequency mechanical waves can be picked up as electronic signals. The signal strength can be increased by using a pre-amplifier before the data are interpreted by the AE equipment (Ali, Rahman et al. 2014; Ali 2014). The data used in this research were taken from earlier research (Hamzah, Al-Balushi et al. 2008; Raja Hamzah and Mba 2009).

2. Methodology

2.1 Experimental Setup and Data Acquisition System

Figure1 illustrates the test rig used for the experiments which is a standard back-to-back gearbox with oil-bath lubrication. A summary is listed in table 1, which describes the test gear description. Using Mobil MOBILGEAR 636 as the lubricant, the tests were conducted under high pressure. Table 2 summarise the selection of the lubricant to decrease the frequency of natural pitting and wear, and the lubricant characteristics.

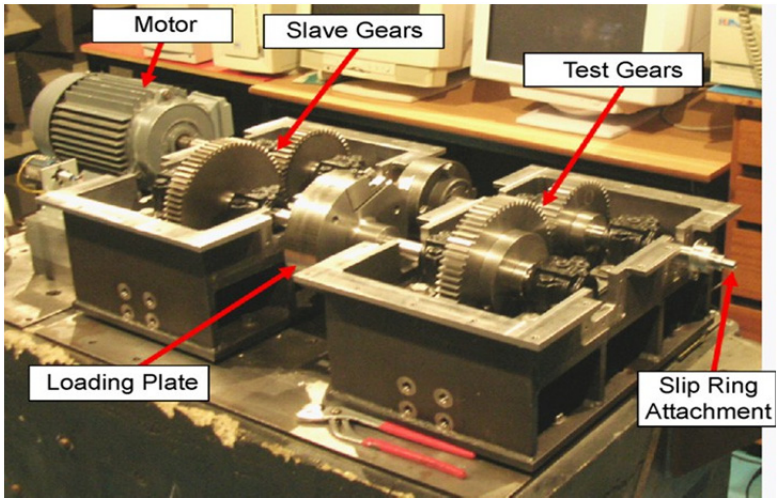


Figure 1. Back-to-back test gearbox arrangement. (Raja Hamzah and Mba 2009)

Table 1. Test gears specifications

Number of teeth, pinion: gear	49:65
Base diameter, pinion: gear (mm)	138.13:183.24
Pitch diameter, pinion: gear (mm)	147.00:195.00
Tip diameter, pinion: gear (mm)	153.00:201.00
Contact ratio	1.33
Module (mm)	3.00
Addendum modification coefficient	0
Surface roughness, Ra (mm)	2.00
Face width (mm)	30.00

Pressure angle (^{0})	2.00
Modulus of elasticity (GPa)	228.00

Table 2. Lubricant properties

Viscosity	
40 ^{0}C (cSt)	680.0
100 ^{0}C (cSt)	39.2
Density at 15.6 ^{0}C (kg/l)	0.91
Viscosity index	90.0
Pour point (^{0}C)	-9.0
Flash point (^{0}C)	285.0
Pressure viscosity coefficient, α (mm2/N)	2.2 x 10-8

The set-up included a Model WD wideband AE sensor capable of picking up relatively flat responses in the range 100 kHz–1 MHz (Physical Acoustic Corp.).To a commercial data acquisition system, the silver contact slip rings transmitted the AE signals from the rotating test pinion to a commercial data acquisition system (figure1).At 20 and 40 dB, sensor signals were preamplified throughout the testing. With an analogue-to-digital converter (ADC), software interfaced captured the gearbox test-rig was set to constantly record AE RMS signals. The torque loading parameters were 60, 120, and 250 Nm and the gearbox was run at 700 rpm. At this speed, one complete revolution of the pinion was signified by the selected speed approximately. Using the accumulated squared ADC values, the RMS could then be calculated. An anti-aliasing filter was used before signal sampling at the ADC. The AE waveform sampling rate was 10 MHz and the digital filtering range was 100–1200 kHz(Hamzah, Al-Balushi et al. 2008, Raja Hamzah and Mba 2009).

2.2 Mathematical Principles

Oil lubrication prevents the gear teeth coming into direct contact and decreases friction, vibration, heat build-up and corrosion. (Hamel, Addali et al. 2014). λ ratio is defined as the oil film thickness (h) divided by the composite surface roughness (σ_{rms}).

$$h = \frac{k\left(\eta_o u\right)^{0.7} R^{0.43}}{w^{0.13}} um \tag{1}$$

Where $k = 1.6\alpha^{0.6} E^{0.03}$

$$\lambda = \frac{h}{\sigma_{rms}} \tag{2}$$

ηo = Dynamic viscosity in Pa.s

μ = Entraining velocity in m/s

R = Equivalent radius in m

w = Load applied along the line of contact in N/m

α = Pressure viscosity coefficient in mm2/N

E = Modulus of elasticity of the gear in Pa

The film thickness is indicative of the lubrication regime between two rough surfaces. The film is affected by high pressure contact and sliding, which causes heat generation and changes in physical properties. Typical operating conditions cause the lubricant to become thin, reducing protection against rubbing at the surfaces and resulting in lubricant failure. The characteristics of the lubricant are therefore crucial to maintain the minimum film thickness under specific operating conditions and this would require a sufficiently large λ ratio (Sreepradha, Kumari et al. 2014).

Three different lubrication regimes can be differentiated depending on the lubricant film thickness which are Hydrodynamic lubrication (HL), Elastohydrodynamic lubrication (EHL) and Boundary lubrication (BL). Mixed lubrication is an intermediate regime among elastohydrodynamic and boundary lubrication. Full hydrodynamic lubrication would normally occur at $\lambda > 10$. This condition would create minimal friction and wear with no direct surface interaction. Pressure transmitted through the lubricant would characterize the only likely fatigue mechanism. Gears generally operate with the elastohydrodynamic regime ($2 < \lambda < 10$) and at a lower film

thickness ratio there is possibility of interaction on the moving surfaces from asperities in the lubricant producing wear. The effect of thermal and pressure conditions on the surfaces can effect in larger asperities in the lubricant and lead to film failure, which bases severe wear, such as scuffing. Boundary and dry lubrication would classically arise at $\lambda < 2$. Figure 2 shows the 'Stribeck Curve' illustrating the λ values in several lubrication regimes (Hamel, Addali et al. 2014).

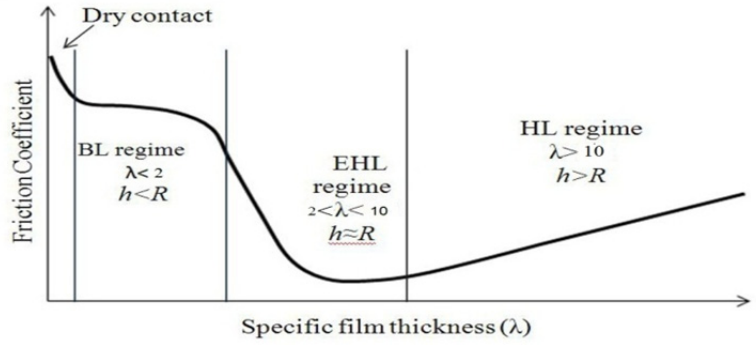


Figure 2. The Stribeck curve and specific film thickness (λ) (Hamel, Addali et al. 2014)

2.3 Statistical Error

The performance of the mathematical models is evaluated by statistical error analyses which identify the most suitable model. Two types of error analysis were used to evaluate the models.

2.3.1 Mean Squared Error (MSE)

An error is the difference between the predicted and actual. The MSE finds the average of the squares of the predicted errors, which corresponds to the risk factor the model represented, because of quadratic or squarederror loss. The cause of the difference is either due to the random approach or certain information has not been processed during the prediction, which would have produced a more accurate estimate. If $\acute{Y}i$ is a vector of n predictions, and Yi is the vector of the actual values, then the (estimated) MSE of the predictor is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\acute{Y}_i - Y_i)^2 \tag{3}$$

2.3.2 Mean absolute error (MAE)

The MAE is a quantity is commonly used to predict error when analysing time series, which comparing the closeness of predicted and actual values given by the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\acute{Y}_i - Y_i\right| \tag{4}$$

2.4 Prediction by Regression Model

Regression analysis is a statistical tool for the investigation of relationships between variables.

$$Y = \alpha + \beta X + \mu \tag{5}$$

Regression models quantitatively describe the variability among the observations by partitioning an observation into two parts (Li, Wunsch et al. 2001). The first part of this decomposition is the predicted portion having the characteristic that can be desecribed to all the observations considered as a group in a parametric framework. The remaining portion, called the residual, is the difference between the observed and the predicted values and must be ascribed to unknown sources. In order to develop relationship between Lambda and RMS, set of equations are be developed for speed S such that a total of 3 set of equations is established such as:

Generalized form of equation for S speed:

$$\lambda_{SnLn} = \alpha + \beta\,(RMS)_n + \mu \tag{6}$$

Where λ stands for specific film thickness of oil (Lambda)

S is speed of machine

L is load condition of S speed

Where n=1, 2, 3…n

α is the intercept and is taken as constant

β is the slope coefficient of RMS

RMS is root mean squaredof acoustic emission signal.

And μ is the stochastic error term which absorb the effect of all those variables which are not included in the equation or omitted and which can affect the relationship between λ and RMS

A machine with speed S will have the following set of equations:

$$\lambda_{SL1} = \alpha + \beta (RMS)_n + \mu$$

$$\lambda_{SL2} = \alpha + \beta (RMS)_n + \mu$$

$$\lambda_{SL3} = \alpha + \beta (RMS)_n + \mu$$

First, the Regression model was built by using the acoustic emission signals and specific film thickness, where the acoustic emission signals used to build the regression model is shown in Figure 3.

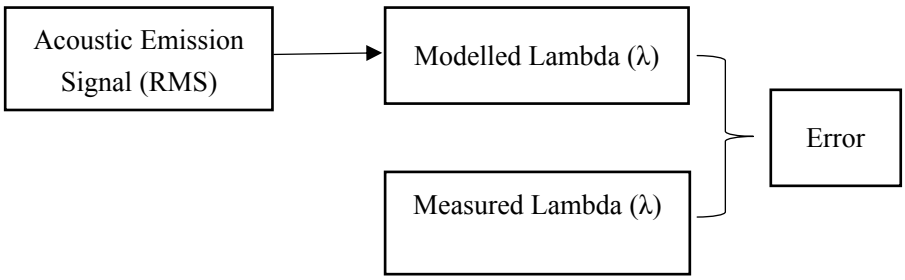


Figure 3. Regression model schematic (Schlechtingen and Ferreira Santos 2011)

3. Result and Discussion

The whole model test for speed and load conditions and all the result are listed in this section.

Table 3. Estimated elasticities for a machine with speed S (700rpm) and load condition L1 (60Nm) using the Least Squaredapproach

Variable	Coefficient	Std. Error	Prob.
Constant	11,4	0.0844	0.0000
RMS	-112.106	1.208	0.0000

R-squared= 81.2% ; Adjusted R-squared=81.2%.

Table 3 shows the result for a machine having speed S (700rpm) and a load condition of L1 (60Nm) depicts a negative relationship between Lambda and RMS. The coefficient of RMS indicates that in every second an increase in the Lambda will cause a decrease in the acoustic emission signal (RMS). The overall model is statistically significant and robust as revealed by the result from the table. The value of R-Squaredand Adjusted R-Squaredsuggests that the model is best fit showing that overall 81.2% of variation in the RMS is caused by Lambda.

Table 4. Estimated elasticity's for a machine with speed S (700rpm) and load condition L2 (120Nm) using the Least Squaredapproach

Variable	Coefficient	Std. Error	Prob.
Constant	12.2500	0.0915	0.0000
RMS	-85.4427	0.8630	0.0000

R-squared= 83.1%; Adjusted R-squared== 83.1%.

Table 4 shows an inverse relationship between thickness of oil film (Lambda) and acoustic emission signal (RMS). It reveals that when a machine shift to L2 load condition while with the same speed S1 still there is a negative association between Lambda and RMS .Thus with the increase in the thickness of oil film the acoustic emission signal tends to get weakened or reduced. The equation/ model developed was found to be robust where, the R-squaredand Adjusted R-squaredvalues shows a strong relationship between lambda and RMS as 83.1% of variation in the RMS occurs due to the changes in lambda.

Table 5. Estimated elasticity's for a machine with speed S (700 rpm) and load condition L3 (250Nm) using the Least Squaredapproach

Variable	Coefficient	Std. Error	Prob.
Constant	17.8429	0.1500	0.0000
RMS	-63.3236	0.6621	0.0000

R-squared= 82.1%; Adjusted R-squared=82.1%.

The thickness of oil film (lambda) inversely affects the acoustic emissionl signal (RMS) when the machine with S speed (700 rpm) has a load condition of L3 (250Nm). The R-squaredand Adjusted R-squaredvalues shows a strong relationship between lambda and RMS as 82.1% of variation in the RMS occurs due the changes in lambda. As shown in the tables above, the regression model established from the acoustic emission data collected through spur gear is significant in predicting oil film thickness. The parameter estimated and the R2 values in tables tell that RMS is a significant factors in explaining Lambda.

The prediction error for the full regression model for the speed and three load condition is illustrated in figure 4 which represents the output error for the regression models. The output error denotes the differences between the target (lambda) and the modelled output from the regression model.

Output error = target –modelled output.

From this figure it can be seen that apart from the few outliers (×), the regression model is capable of predicting the specific film thickness in the range of ±4. The remaining outliers are caused by extreme transient situations during the experiment, the oil mist generated as a result of the churning of the oil formed a 'frozen mist' as the oil temperature approached 0ºC. This frozen oil mist was thrown into the gear mesh by the rotating wheel resulting in the generation of AE activity. In this sense normal operational temperature profiles are modeled well. The problem can be either overcome by implementation of separate models for transient situations or by simple filtering. Filtering of non-operational periods is possible, since determination of the component condition based on online data is most reasonable if the gear is operating. By the use of the value of RMS and lambda in EHL regime we can get better prediction and the prediction error will be lower.

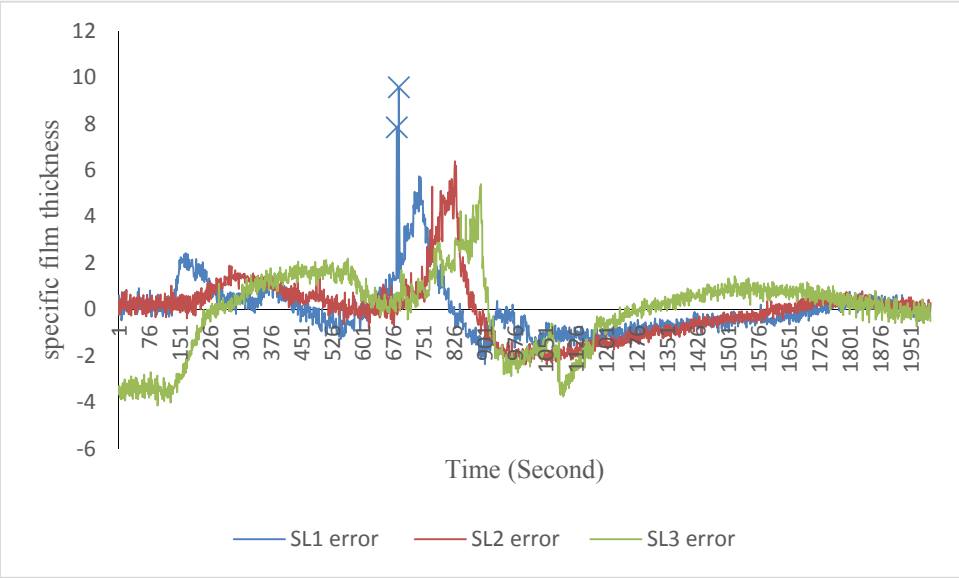


Figure 4. Regression model error

This work is important because it shows that the regression model can be used for predicting the oil film thickness through acoustic emission signal. For that purpose, acoustic emission signal (RMS) used to predict the oil film thickness in spur gear and these data are used to build the regression model. Two statistical errors have been used to assess the performance of the model. The result in Table 6 indicates that the performance of the model for the three load and speed conditions. The model represents good results for the first load and speed conditions (SL1) where the MSE and MAE values are small compared to the other conditions. At the same time, the value of the statistical errors measurement for load and speed conditions (SL3) is higher than the other conditions because the performance of the model for (SL3) is less than the others.

Table 6. Regression model performance and classification success results

Load and Speed Conditions	MSE	MAE	Regression Model Classification
SL1	1.21553	0.75694	72.90%
SL2	1.52712	0.86500	85%
SL3	2.59517	1.22352	97%

Classification program is designed and used to classify the oil film regime in gear, the program used with estimated data. When the regression models were tested with acoustic emission data sets and oil film thickness, the models yielded a predicted oil film thickness. This value was used in the program and the results is shown in Table 6. The result of this program demonstrates that the regression model have the ability to give a good performances of classification accuracy.

4. Conclusion

This study proposes a regression model to improve the accuracy of oil film thickness prediction for spur gear based on acoustic emission signal. The results showed that the regression models were can be used in prediction and this suggested technique attained 85% success in prediction.

This study shows the overall model developed and is found to be statistically significant. The statistical errors measurement which have been used to assess the performance of the models indicates that the performance of the regression model is good. Moreover the R-squared and Adjusted R-squared values shows a strong relationship between Lambda and RMS as over 80% of variation in the RMS occurs due the changes in Lambda. For this reasons the proposed method can be used to monitor the spur gear.

References

Ali, Y. H., Rahman, R. A., & Hamzah, R. I. R. (2014). Acoustic Emission Signal Analysis and Artificial Intelligence Techniques in Machine Condition Monitoring and Fault Diagnosis: A Review. *Jurnal Teknologi, 69*(2). http://dx.doi.org/10.11113/jt.v69.3121

Ali, Y. H. M. H. O., Rahman, R. A., & Hamzah, R. I. R. (2014). Acoustic emission technique in condition monitoring and fault diagnosis of gears and bearings. *International Journal of Academic Research Part A 6*(5), 6.

Amarnath, M., Sujatha, C., & Swarnamani, S. (2009). Experimental studies on the effects of reduction in gear tooth stiffness and lubricant film thickness in a spur geared system. *Tribology international, 42*(2), 340-352. http://dx.doi.org/10.1016/j.triboint.2008.07.008

Hamel, M., Addali, A., & Mba, D. (2014). Investigation of the influence of oil film thickness on helical gear defect detection using Acoustic Emission. *Applied Acoustics, 79,* 42-46. http://dx.doi.org/10.1016/j.apacoust.2013.12.005

Hamel, M., Addali, A., & Mba, D. (2014). Monitoring oil film regimes with acoustic emission. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 228*(2), 223-231. http://dx.doi.org/10.1177/1350650113503631

Hamzah, R. R., Al-Balushi, K. R., & Mba, D. (2008). Observations of acoustic emission under conditions of varying specific film thickness for meshing spur and helical gears. *Journal of Tribology, 130*(2), 021506. http://dx.doi.org/10.1115/1.2908915

Li, S., Wunsch, D. C., O'Hair, E., & Giesselmann, M. G. (2001). Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation. *Journal of Solar Energy Engineering 123*(4), 327-332. http://dx.doi.org/10.1115/1.1413216

Peng, Z., & Kessissoglou, N. (2003). An integrated approach to fault diagnosis of machinery using wear debris and vibration analysis. *Wear, 255*(7), 1221-1232. http://dx.doi.org/10.1016/S0043-1648(03)00098-X

Raja Hamzah, R., & Mba, D. (2009). The influence of operating condition on acoustic emission (AE) generation during meshing of helical and spur gear. *Tribology International, 42*(1), 3-14. http://dx.doi.org/10.1016/j.triboint.2008.06.003

Samanta, B., Al-Balushi, K., & Al-Araimi, S. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence, 16*(7), 657-665. http://dx.doi.org/10.1016/j.engappai.2003.09.006

Schlechtingen, M., & Santos, I. F. (2011). Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical Systems and Signal Processing, 25*(5), 1849-1875. http://dx.doi.org/10.1016/j.ymssp.2010.12.007

Sreepradha, C., Kumari, A. K., Perumal, A. E., Panda, R. C., Harshabardhan, K., & Aribalagan, M. (2014). Neural network model for condition monitoring of wear and film thickness in a gearbox. *Neural Computing and Applications, 24*(7-8), 1943-1952. http://dx.doi.org/10.1007/s00521-013-1427-6

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications, 32*(4), 1084-1093. http://dx.doi.org/10.1016/j.eswa.2006.02.005

Sung, C., Tai, H., & Chen, C. (2000). Locating defects of a gear system by the technique of wavelet transform. *Mechanism and Machine Theory, 35*(8), 1169-1182. http://dx.doi.org/10.1016/S0094-114X(99)00045-2

Zhang, J., & Nie, H. (2009). Experimental study and logistic regression modeling for machine condition monitoring through microcontroller-based data acquisition system. *Journal of Advanced Manufacturing Systems, 8*(2), 177-192. http://dx.doi.org/10.1142/S0219686709001742

Copyrights

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The Nexus between Female Labor Force Participation and Female Total Fertility Rate in Selected ASEAN Countries: Panel Cointegration Approach

Nira Hariyatie Hartani[1], Nor Aznin Abu Bakar[2] & Muhammad Haseeb[2]

[1] School of Government, Law and International Studies (COLGIS), Universiti Utara Malaysia, Sintok, Kedah, Malaysia

[2] School of Economics, Banking and Finance (SEFB), College of Business (COB), Universiti Utara Malaysia, Sintok, Kedah, Malaysia

Correspondence: Muhammad Haseeb, School of Economics, Banking and Finance (SEFB), College of Business (COB), Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia. Tel: 60-172-957-409. E-mail: scholar_economist@yahoo.co.uk

Received: January 20, 2015 Accepted: February 1, 2015 Online Published: July 6, 2015

doi:10.5539/mas.v9n8p29 URL: http://dx.doi.org/10.5539/mas.v9n8p29

Abstract

The main objective of this study is to empirically investigate the relationship between female labor force participation (FLFP) and female total fertility rate (FTFR) for the ASEAN-6 countries from the period 1995 to 2013. The Fully Modify OLS (FMOLS) has been applied to explore the cointegration and causality between the suggested variables. The contegration results confirm that the female labor force participation rate and total female fertility rate are cointegrated for the panel of ASEAN-6 countries. Whereas, long-run Granger causality confirm the causality run from the female total fertility rate to the female labor force participation rate. Moreover, the results show that 1percent increase in the female total fertility rate causes in a 0.44 percent decrease in the female labor force participation rate for the ASEAN-6 countries. The FTFR highest negative effect observed in Indonesia and smallest observed in Thailand. The results of FMOLS confirm the long run panel relationship between female labor force and total female fertility rate.

Keywords: FLFP, FTFR, FMOLS, ASEAN-6 countries

1. Introduction

The relationship between female labor force participation (FLFP) and female total fertility (FTFR) got considerable attention from the researchers of economics and demography. The FLFP rate generally falls around childbirth; mothers who have young children have traditionally been considered as having low labor force attachment. The study of Kenjoh (2005) investigates that among the OECD countries only Scandinavian countries are showing that majority females worked continuously over the life. Nevertheless, more recently, this situation has started to change and there are now other OECD countries where women are working continuously throughout their lives or with only a short interruption at the time of childbirth. As (Kenjoh, 2005) described it, *"one could say that the increase in the labor force participation rate of mothers is one of the most prominent developments of the recent labor market in OECD countries"*.

The nexus of female labor force participation rate (FLFP) and the female total fertility rate (FTFR) is a general problem for the developed and less develop economies. The existing studies only discussed the correlation between female labor force participation and total fertility rate. While the current study deal with correlation as well as causation between them. Conversely, it is possible that two variables are correlated with but not cause each other. Similarly, model based on correlation presume a one period proportional stationary framework, whereas the effect of FTFR and FLFP on each other is unlikely to be immediate and this reality have led to FLFP and FTFR to be model led in a dynamic manner and also as an autoregressive procedure. In the current economic literature question of causality "what causes what?" has received attention. The main intention of the current study is to inspect closely and thoroughly the causal relationship for a panel of ASEAN-6 countries over the period 1995–2013. Furthermore, for the purpose to confirm causality study apply unique combination of econometric approaches in three different directions.

In the first step study applied panel unit root for the purpose to confirm stationarity and level of integration. For the purpose to examine cointegration, Kao panel cointegration technique has been used. In the first step ultimate goal of the study is to confirm the causality, where stationarity and Kao cointegration leads study towards causality because these are the necessary steps before apply granger causality. So, after confirming the stationarity and cointegration Granger causality approach has been applied. Second, the long-run relationship is examined through FMOLS. Third, the study used the most recent data from ASEAN-6 countries.

1.1 ASEAN-6 Country's Comparison between FTFR and FLFP

1.1.1 Indonesia

The total Indonesian population 234,181,400, is recorded and ranked 4th as most populated countries in the world. Among ASEAN-6 countries, Indonesia was ranked 3rd with the average FTFR 2.347 per woman and ranked 4th with the average FLFP 51 percent during the period of 1995-2013. In 1995 FTFR start from 2.699 and continued decreasing up till 2.163 in 2013, While, FLFP begin with 49 percent and increased up to 51 percent.

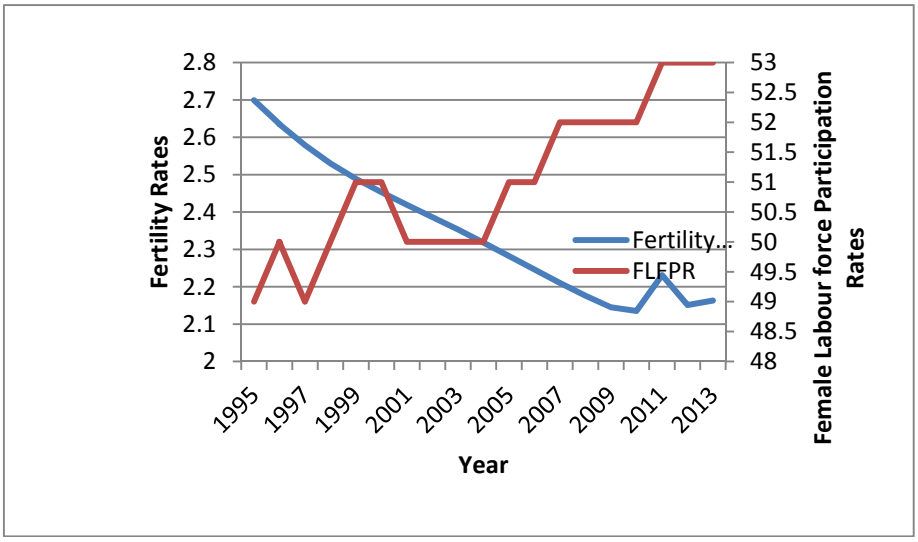


Figure 1. FTFR and FLFP comparison of INDONESIA

1.1.2 Malaysia

The total Malaysian population 28,306,700 is recorded and ranked 44th as most populated country in the world. Among ASEAN-6 countries, Malaysia was ranked 2nd with the average FTFR 2.891 per woman and ranked 6th with the average FLFP 43 percent during the period of 1995-2013. In 1995 FTFR start from 3.301and continuously decreasing up till 2.555 in 2013, While, FLFP begin with 43 percent and decreased in 1997 up to 42 percent but 2000 it is continuously increasing up to 45 percent.

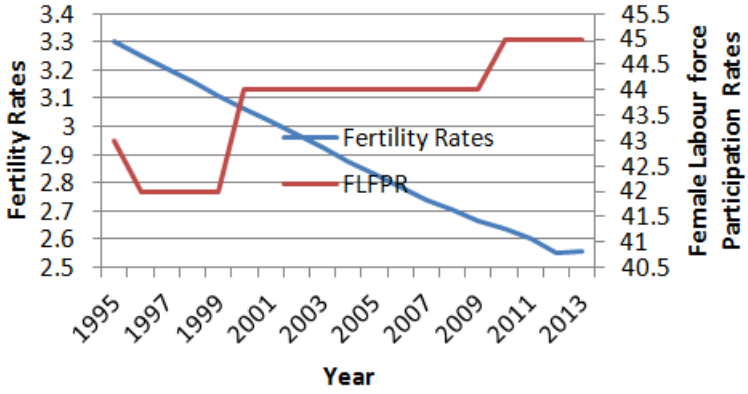


Figure 2. FTFR and FLFP comparison of MALAYSIA

1.1.3 Philippines

The total Philippines population 94,013,200, is recorded and ranked 12th as most populated countries in the world. Among ASEAN-6 countries Philippines was ranked 1st with the average FTFR 3.511 per woman and ranked 5th with the average FLFP 49 percent during the period of 1995-2013. In 1995 FTFR start from 4.006 and continued decreasing up till 2.881 in 2013, While, FLFP begin with 49 percent and shows some fluctuations with increasing and decreasing trend but overall its show decreasing trend.

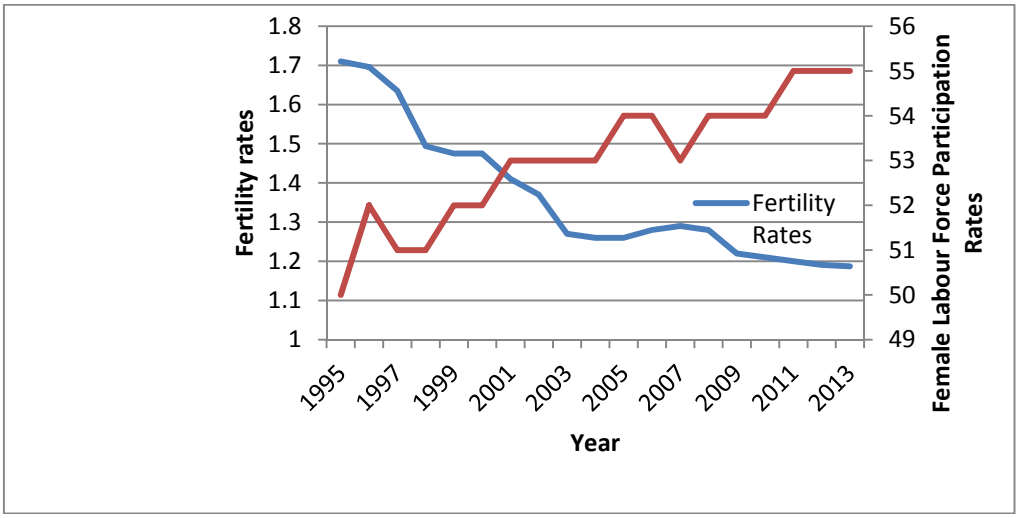


Figure 3. FTFR and FLFP comparison of PHILIPPINES

1.1.4 Singapore

The total Singapore population 4,987,600, is recorded and ranked 114th as most populated countries in the world. Among 6-ASEAN countries, Singapore was ranked 6th with the average FTFR 1.363 per woman and ranked 3rd with the average FLFP 53 percent during the period of 1995-2013. In 1995 FTFR start of 1.71 and continued decreasing up till 1.187 in 2013, While, FLFP begin with 50 percent and increased up to 55 percent.

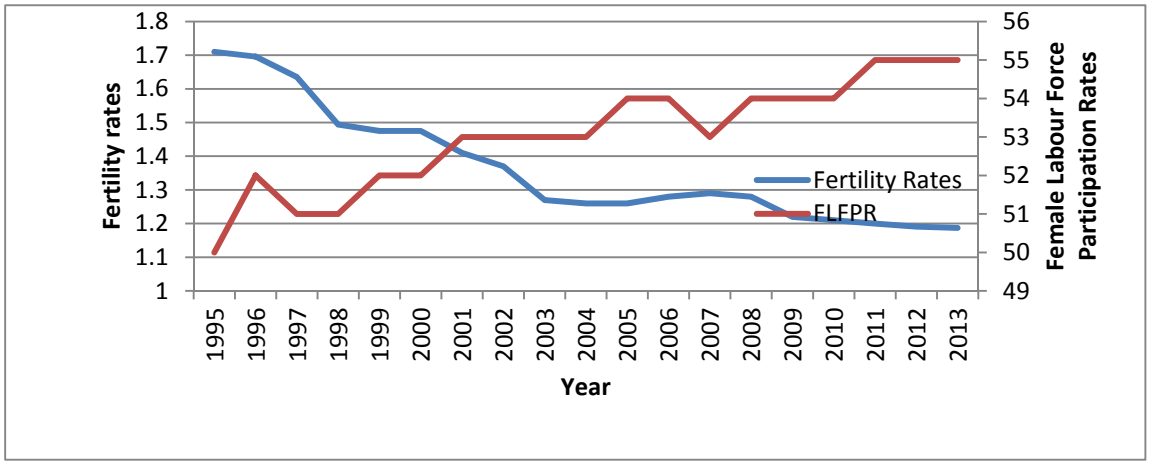


Figure 3. FTFR and FLFP comparison of SINGAPORE

1.1.5 Thailand

The total Thailand population 63,447,374, is recorded and ranked 21st as most populated country in the world. Among ASEAN-6 countries, Thailand was ranked 5th with the average FTFR 1.671 per woman and ranked 2nd with the average FLFP 66 percent during the period of 1995-2013. In 1995 FTFR start of 1.862 and continued decreasing up till 1.534 in 2013, While, FLFP begin with 66 percent and shows variation up till 2013 to 65 percent.

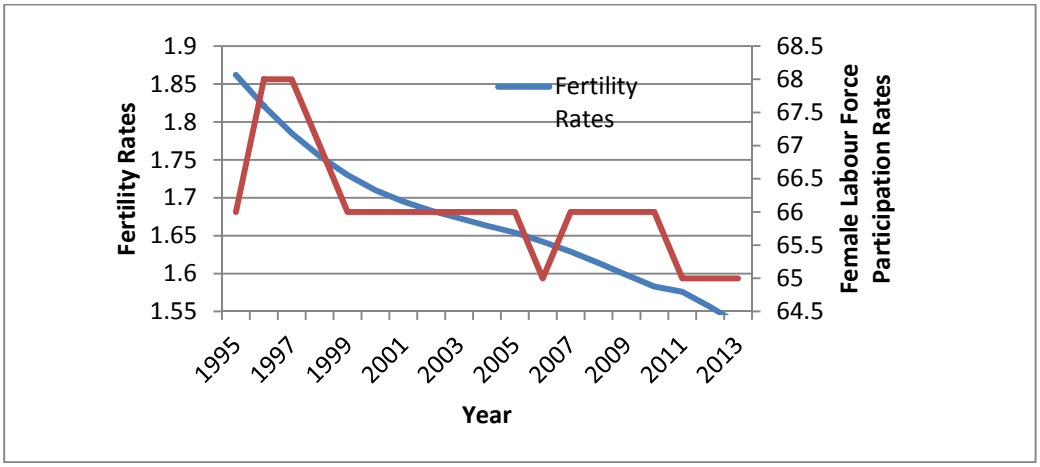


Figure 4. FTFR and FLFP comparison of THAILAND 1.1.6 Vietnam

The total Vietnam population 85,789,573, is recorded and ranked 13th as most populated countries in the world. Among ASEAN-6 countries Vietnam was ranked 4th with the average FTFR 1.991 per woman and ranked 1st with the average FLFP 69 percent during the period of 1995-2013. In 1995 FTFR start of 2.666 and continued decreasing up till 1.791 in 2013, While, FLFP begin with 73 percent and decrease up to 67 percent.

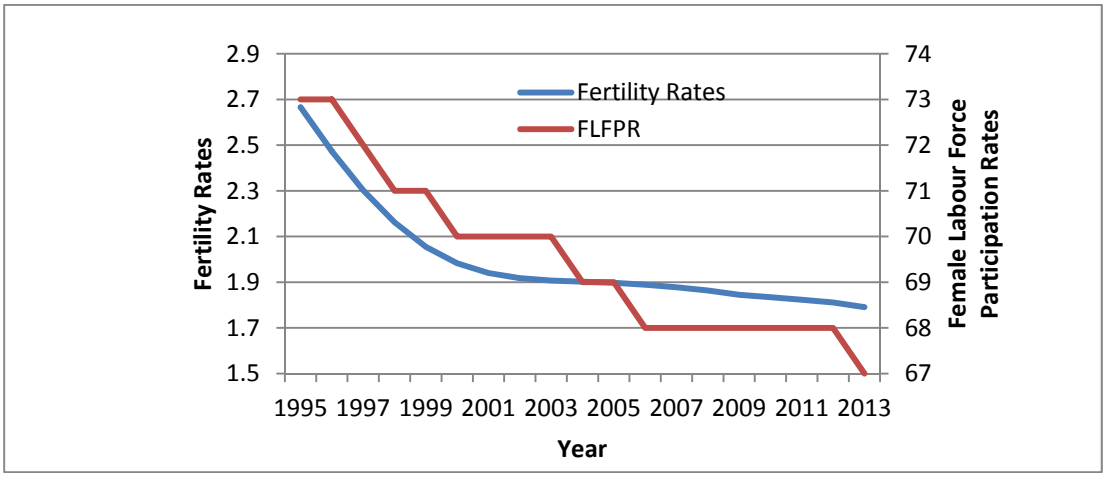


Figure 5. FTFR and FLFP comparison of VIETNAM

Among ASEAN-6 countries three countries (Indonesia, Malaysia and Singapore) showing negative relationship between FTFR and FLFP while three countries (Thailand, Vietnam and Philippines) showing positive relationship between two variables. This complex, controversial and inconsistent results open the research gate for the researchers for further investigation.

1.2 Theoretical Perceptions on Fertility and Female Labor Force Participation

The different researchers such as, Amador, Bernal, and Peña (2013) and Tam (2011) investigates increase in FTFR in two ways, like negative and positive effects on each other. On the negative side, increase in number of children can increase the amount of home work, which causes to reduce the chance of mother to seek work in the market and support the hypothesis of incompatibility role. However, on the positive side, if the number of children increase there is need to increase in household's income too and it may cause for mother to seek outside employment.

On the other hand, if FLFP is increase and effect negative on FTFR it means its support the hypothesis of incompatibility role. Further, if female employees having children and they admission in child care centre than it is opportunity cast so, FLFP negatively affect FTFR. Additionally increase in FLFP cause in the disruption towards female's career success, apparent in the hammering of a higher prospective future income tributary and

non-pecuniary settlement, including appreciation and status related with a more senior post in her chosen profession. An additional opportunity cost of increase in numbers of children may perhaps the loss of camaraderie and social associations in the place of work that serve as a point of liberate outside the home. However, researcher investigates that the relationship between FLFP and FTFR revolutionize from a negative to a positive value in the 1980s (Ahn & Mira, 2002).

The studies such as, Engelhardt, Kögel, and Prskawetz (2004) and Narayan and Smyth (2006) explore that the countries with lowest FTFR also have lowest level of FLFP and vice-versa. It is concluded that causality run from FTFR to FLFP and FLFP to FTFR it may has negative effect of FTFR on FLFP and FLFP on FTFR which support the hypothesis of incompatibility role. However, if causality runs from FTFR to FLFP and FLFP to FTFR it may have a positive effect FTFR on FLFP and FLFP on FTFR which support the public response hypothesis.

2. Review of Existing Literature

The nexus between female labor force participation and female total fertility rate is a recent issue among researchers. Similarly studies of Borderías (2013), de Laat and Sevilla-Sanz (2011), Humphries and Sarasúa (2012) and Mahmoudian (2006) found correlation between FLFP and FTFR. While only few studies such as, S.-H. Lee, Ogawa, and Matsukura (2009) and Mishra and Smyth (2010) explore causation between both variables. The study of de Laat and Sevilla-Sanz (2011) explore nexus of female fertility and labor force participation in OECD countries. The results are demonstrates that one effect of Southern Europe's rapid fertility decline is the emergence of a positive cross-country correlation between women's labor force participation and fertility across developed countries, despite the continuing negative correlation between these factors within countries.

The study of Tsani, Paroussos, Fragiadakis, Charalambidis, and Capros (2013), utilized annual time series data of south Mediterranean from the year 1960 – 2008; Ridao-Cano and McNown (2005), employed annual time series data of USA from 1948 – 1997 and Sugawara and Nakamura (2014), used annual time series data of Japan from 1950 – 1993 found that there is evidence of causality between both variables.

Furthermore, Amador et al. (2013) investigated the causation between FTFR and FLFP American females. The results confirm the causality from FLFP to FTFR. Engelhardt et al. (2004)used annual time series data from Sweden, Italy, UK, Germany and France from the year 1960 to 1994 and utilized cointegration and Granger causality. The results showed that there long run causality for all countries. Further, Narayan and Smyth (2006) explored the relationship between FTFR, FLFP and infant mortality rates in Australia from 1960–2000 and investigate that FTFR and infant mortality rate both jointly Granger cause FLFP. As mentioned earlier there not various studies utilized panel cointegration and panel causality between these two purposed variables. So, this the contributions of present study to apply panel cointegration and panel causality to get better results.

3. Econometric Methodology and Data Source

3.1 Data Source

The annual time series data have been used from 1995 to 2013 for the ASEAN-6 countries, namely, Malaysia, Indonesia, Thailand, Singapore, Philippines and Vietnam. The data on total fertility rates and female labor force participation collected from the World Bank data base and converted into natural logarithm before analysis. The FTFR is the weighted mean of age specific fertility rates and FLFP is defined as the adult female population in the age group 15 – 65 years in the labor force.

The different econometric approaches have been applied to test the causality as well as relationship between FTFR and FLFP.

4. Results

4.1 Unit Root Test

Before apply cointegration and causality, it is preliminary to test stationary and level of integration. The augmented Dickey-Fuller (ADF) and Phillips Perron test for unit root have been applied. The results in Table 1 suggest that all the variables of each country are non-stationary at level and become stationary at first difference it means the level of integration of all variables are I(1).

Table 1. ADF and PP Unit Root Test

Country	ADF Unit root test			
	TRF		FLFP	
	Level	First Difference	Level	First Difference
Indonesia	0.761	0.032*	1.981	0.000*
Malaysia	0.321	0.045*	0.987	0.045*
Philippines	0.061	0.001*	0.891	0.032*
Singapore	0.098	0.030*	0.762	0.010*
Thailand	0.754	0.012*	1.987	0.023*
Vietnam	1.896	0.000*	0.056	0.049*
Country	PP Unit root test			
			TRF	FLFP
	Level	First Difference	Level	First Difference
Indonesia	0.761	0.032*	1.981	0.000*
Malaysia	0.321	0.045*	0.987	0.045*
Philippines	0.061	0.001*	0.891	0.032*
Singapore	0.098	0.030*	0.762	0.010*
Thailand	0.754	0.012*	1.987	0.023*
Vietnam	1.896	0.000*	0.056	0.049*

Note: * statistically significant at 5 percent.

4.2 Cross Sectional Dependence (CD) Test

It is observed that the Panel unit root tests are more powerful because of joint information from cross section data and time series data. But it is observed that panel unit test are facing enormous problem of cross-sectional dependence. The test applied for check panel cross sectional dependence (CD) is proposed by Pesaran (2004). The statistics based on univariate AR (p) specification with the level of variables $p \leq 4$. The null hypothesis (H_0) stated that output innovation is independent by cross sectional. The critical values for CD test are 1% = 2.57, 5% = 1.96 and 10% = 1.64. The results of Pesaran (2004) test are reported in Table 2.

Table 2. Cross Section Dependence Test

ADP regression across Six-Asian countries				
	1995-2013 (T = 19, N = 6)			
	p = 1	p = 2	p = 3	p = 4
FTFR				
Î	0.134	0.118	0.321	0.104
CD	4.532***	3.312***	3.041***	3.041***
FLFP				
Î	0.091	0.043	0.051	0.055
CD	2.413**	1.214	1.314	1.312

Note: **, *** denoted for statistically significant at 5 %and 1% level, respectively.

The CD test is significant at 1%, and FTFR correlation coefficient observed around 0.1.

4.3 Panel Unit Root Test

The panel unit root is based on different tests. The prominent tests are Im *et al*. (2003) Levin *et al*. (2002) and also include IPS test statistic (CIPS) proposed by Pesaran (2007). According to LLC tests it is presumes that countries include in the test are unite towards the equilibrium value with the same velocity under the alternative hypothesis (H_1). The panel unit root test results are reported in Table3. The results of utilized tests confirm that FLFP and FTFR are I (1).

Table 3. Panel Unit Root Test Results

	FTFR		FLFP	
	Level	Difference	Level	Difference
IPS	2.62	-7.31***	3.21	-4.82***
LLC	-0.18	-6.89***	0.54	-5.12***
CIPS	-2.312	-5.012***	-1.675	-3.123***

The IPS test is less restrictive as compare to the LLC because the IPS test does not make assumptions like LLC test. The IPS test looks towards the solution of CD problem. For transformed the data IPS test subtracts the cross sectional means and apply the *T-bar* statistic. However, Strauss and Yigit (2003) suggested that humiliate across the panel does not habitually eradicate CD. Strauss and Yigit argue that CIPS test is more powerful as compare to IPS test and LLC test because CIPS unambiguously permit for CD by suitably truncating the IPS *t*-bar statistic.

4.4 Cointegration Test

To investigate the cointegration between FTFR and FLFP this study applied Johansen maximum likelihood (JML) approach developed by Johansen (1988) instead of Kao's cointegration approach. The results for maximum likelihood reported in Table4. The results suggest that the null hypothesis (H_0) of no cointegration rejected for Indonesia, Vietnam and Singapore and null hypothesis of one cointegration cannot be rejected in these countries. Furthermore, the null hypothesis of no cointegration and null hypothesis of one cointegration cannot be rejected for Malaysia, Philippines and Thailand.

Table 4. Johansen cointegration test results

Maximum Eigen value statistics for H_0 : *rank* = *r*			
Country	$r = 0$ (17.3, 19.4, 23.8)	$r \leq 1$ (10.7, 12.6, 16.6)	*p*
Indonesia	21.03**	11.44	2
Malaysia	17.78	7.15	1
Philippines	9.87	7.74	2
Vietnam	31.35***	8.93	2
Thailand	10.12	5.44	1
Singapore	25.35***	9.12	2

Note: **, *** rejected null at 5% and 1% level, respectively.

It is concluded that there is a single cointegration vector for three (Indonesia, Vietnam, Singapore) out of ASEAN-6 countries. The Pedroni (2000) panel cointegration test has been applied to confirm the long run relationship between both variables. Pedroni (2000) presents seven different statistics to test the null of no cointegration in heterogeneous panels. Pedroni divided these seven tests into two groups. Group one called within the dimension panel test and second group called between dimension group tests.

$$TFR_{i,t} = \alpha_i + \beta_i FLFP_{i,t} + \varepsilon_{i,t} \quad (1)$$

Here FLFP and FTFR are defined as above and $\varepsilon i\ t = \eta i\ \varepsilon i\ (t-1) + \mu i\ t$ is the estimated residuals from the panel regression. The values of seven statistics and two groups are tabulated in Pesaran (2004). If the calculated values are higher than tabulated value the null hypothesis of no cointegration can be rejected and it is suggested that there is exist a long run relationship between FTFR and FLFP. The results of panel cointegration presented in Table 5.

Table 5. Panel Cointegration Tests

Pedroni Test			
	Statistics	$r = 0$	$r \leq 1$
FTFR has cointegration			
Panel of v-statistics	0.32431	*35.12***	13.19
Panel of rho test-statistics	0.33400		
Panel of PP test-statistics	0.42071		
Panel of ADF test-statistics	-0.14563		
Group rho test-statistics	1.43251		
Group PP test-statistics	1.34210		
Group ADF test-statistics	0.56213		
FLFP has cointegration			
Panel of v-statistics	1.34512		
Panel of rho test-statistics	-1.67123*		
Panel of PP test-statistics	-2.41231**		
Panel of ADF test-statistics	-2.13210**		
Group rho test-statistics	-2.42131**		
Group PP test-statistics	-3.43231***		
Group ADF test-statistics	-3.34412***		

The results suggested that the majority of the variables confirms the panel cointegration in case of FLFP as dependent variable, whereas, none of the variable is cointegrated in case of FTFR as the dependent variable. These seven statistics are residual-based; to aggregate the probability-values of the individual JML cointegration test statistics Fisher $\chi 2$ cointegration test utilized.

4.5 Causality Test

After the confirmation of cointegration study examine the direction of causality between the FTFR and FLFP.As it is confirmed that there is a long run relationship between both variables, Granger causality for a long run relationship run with the dynamic error correction model (DECM) specification. The DECM is estimated by following a two-step procedure. First, the study estimates the cointegration between FTFR and FLFP follow the JML procedure. Second, the study utilizes the results of this cointegrating relation to estimate the EC term $\hat{ECT}_{it} = FTFR_{it} - \hat{\alpha}_i - \hat{b}_t - \hat{\beta}_i FLFP_{it}$.

Now study estimate ECM:

$$\Delta TFR_{it} = \alpha_{1i} \sum_{l=1}^{P} \alpha_{11ip} \Delta TFR_{it-p} + \sum_{l=1}^{p} \alpha_{12ip} \Delta FLFP_{it-p} + \phi_{1i} ECT_{it-1} + \varepsilon_{1it} \quad (2)$$

$$\Delta FLFP_{it} = \alpha_{2i} \sum_{l=1}^{P} \alpha_{21ip} \Delta FLFP_{it-p} + \sum_{l=1}^{p} \alpha_{22ip} \Delta TFR_{it-p} + \phi_{2i} ECT_{it-1} + \varepsilon_{2it} \quad (3)$$

where FTFR and FLFP are already defined, denoted first difference,

Δ	=	First Difference
ECT	=	Error correction term
p	=	Lag length

The results of long run panel causality are presented in Table6. The Wald test shows that variables are significant in the Equation 2 and 3. According to results reported in Table6 the null hypothesis (H_0) FTFR does not cause FLFP is rejected, while the null hypothesis (H_0) FLFP does not cause FTFR cannot be rejected at the 5% significance level. Finally, findings of these results suggested that there is a unidirectional long-run Granger causality relationship running from FTFR to FLFP. In case of large N and small T the dynamic panel data usually face the problem of Nickell (1981) bias. In this study, we used large $T = 19$ and small $N = 6$ so the Nickell (1981) bias is negligible and can be ignored.

Table 6. Long-Run Panel causality Tests

H_0:No Causality	X^2	Probability- value
FLFP does not cause FTFR	0.66	0.44
FTFR does not cause FLFP	3.81	0.01**

** Rejection of the H0 at the 5% critical value.

4.6 Fully Modify Least Square (FMOLS) Test

After the confirmation of cointegration and the direction of causality, the study able to test the structural coefficients by using FMOLS test. The recently numerous studies such as Abidin, Bakar and Haseeb (2014) and Haseeb et al. (2014) has been applied the same methodology in the case of Malaysia. The FMOLS test proposed by Pedroni (2000) and tackles the dilemma of regressors with non-stationary specification, as well as the dilemma of simultaneity biases. For example, I = 1, 2, . . . , N countries over time t = 1, 2, . . . , M: than

$$Y_{it} = \alpha_{it} + \beta X_{it} + \varepsilon_{it} \quad \text{and} \quad X_{it} = X_{it-1} + \mu_{it} \tag{4}$$

The panel FMOLS estimator for coefficient β is

$$\hat{\beta}_{FMOLS} = [\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x_{it}^{'}]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\left(x_{it}y_{it}^{+} - \lambda_{i,\varepsilon u}\right)\right] \tag{5}$$

The results of the panel long run relationship estimator by using FMOLS are presented in Table7. The results explain that FTFR of all ASEAN-6 countries and as a whole have a statistically significant negative effect on FLFP.

Table 7. FMOLS estimators Results

FLFP is dependent variables		
Country	Coefficients	*t*-statistics
Indonesia	-0.66***	-9.45
Malaysia	-0.43***	-6.79
Thailand	-0.12***	-3.34
Singapore	-0.45***	-16.12
Vietnam	-0.13***	-6.13
Philippines	-0.41***	-6.13
Panel	-0.44***	-19.13

***significant at 1% critical value.

The largest effect of FTFR on FLFP is observed in Indonesia, where 1% increase in FTFR cause decreases FLFP by 0.66%. While in three countries, among ASEAN-6 countries effect of FTFR on FLFP is almost similar, which is a 1 % increase in FTFR cause decrease FLFP by 0.45%, 0.43% and 0.41% in Singapore, Malaysia and Philippines respectively. Furthermore, the smallest effect of FTFR on FLFP observed in Thailand, where increases 1% in FTFR cause reduces FLFP by 0.12%. The result of the panel long run elasticity suggested that 1% increase in FTFR causes FLFP decrease by 0.44%.

The negative equilibrium relationship between FLFP and FTFR are the witness of the lack of effective access to childcare. Chevalier and Viitanen (2002) investigate the causality between FLFP with young children and the supply of childcare in the UK. The results are suggested that lack of childcare services confines FLFP and currently increase in demand for childcare only serve to increase costs or queues rather than have an effect on the supply. The results are leads that if the Government of UK have aim to increase FLFP than change policy in the child care market is required. Similarly, Bratti (2003) also suggested childcare improvement and increased access to affordable childcare in Italy. Furthermore, B. S. Lee, Jang, and Sarkar (2008) investigates FLFP in Korea and Chen, Shao, Murtaza, and Zhao (2014) in Germany. The results are suggested that FLFP is affected by childcare supply.

5. Conclusion

The main aim of this study is to investigate the direction of causality between FLFP and FTFR for the ASEAN-6 countries. For the purpose to check the stationarity and level of integration panel unit root test have been applied. After confirmed the stationarity of each variable, panel cointegration and Granger causality test used to examine

the cointegration and direction of causality between FLFP and FTFR. The panel based cointegration has advantage against individual cointegration that it is more powerful, especially in case of small size with less than 50 observations. The results of cointegration confirm the cointegration between proposed variables. In addition, results of panel causality suggested that there is causality run from FTFR to FLFP in the ASEAN-6 countries. The results also suggested that there is an opposite relationship between FLFP and FTFR. The FMOLS techniques are used to examine the long run relationship between FLFP and FTFR and results confirm the long run relationship between FLFP and FTFR. This study is multivariate setting further it can be extended investigating causality between FLFP and FTFR from the bivariate context. Further FTFR and FLFP also can consider as factor influence the opportunity cost of having more children like household total income, female education and male unemployment.

References

Abidin, I. S. Z., Bakar, N. A. A., & Haseeb, M. (2014). An Empirical Analysis of Exports between Malaysia and TPP Member Countries: Evidence from a Panel Cointegration (FMOLS) Model. *Modern Applied Science*, *8*(6), 238.

Ahn, N., & Mira, P. (2002). A note on the changing relationship between fertility and female employment rates in developed countries. *Journal of Population economics, 15*(4), 667-682. http://dx.doi.org/10.1007/s001480100078

Amador, D., Bernal, R., & Peña, X. (2013). The rise in female participation in Colombia: Fertility, marital status or education? *Ensayos sobre Política Económica, 31*(71), 54-63.

Borderías, C. (2013). Revisiting Women's Labor Force Participation in Catalonia (1920–36). *Feminist Economics, 19*(4), 224-242. http://dx.doi.org/10.1080/13545701.2013.831181

Bratti, M. (2003). Labor force participation and marital fertility of Italian women: The role of education. *Journal of Population economics, 16*(3), 525-554. http://dx.doi.org/10.1007/s00148-003-0142-5

Chen, J., Shao, X., Murtaza, G., & Zhao, Z. (2014). Factors that influence female labor force supply in China. *Economic Modelling, 37*(0), 485-491. http://dx.doi.org/10.1016/j.econmod.2013.11.043

Chevalier, A., & Viitanen, T. K. (2002). The causality between female labor force participation and the availability of childcare. *Applied Economics Letters, 9*(14), 915-918. http://dx.doi.org/10.1080/13504850210138469

de Laat, J., & Sevilla-Sanz, A. (2011). The Fertility and Women's Labor Force Participation puzzle in OECD Countries: The Role of Men's Home Production. *Feminist Economics, 17*(2), 87-119. http://dx.doi.org/10.1080/13545701.2011.573484

Engelhardt, H., Kögel, T., & Prskawetz, A. (2004). Fertility and women's employment reconsidered: A macro-level time-series analysis for developed countries, 1960–2000. *Population Studies, 58*(1), 109-120. http://dx.doi.org/10.1080/0032472032000167715

Haseeb, M., & Azam, M. (2015). Energy Consumption, Economic Growth and CO2 Emission Nexus in Pakistan. *Asian Journal of Applied Sciences*, *8*, 27-36. http://dx.doi.org/10.3923/ajaps.2015.27.36

Haseeb, M., Bakar, N. A. A., Azam, M., Hassan, S., & Hartani, N. H. (2014). The Macroeconomic Impact of Defense Expenditure on Economic Growth of Pakistan: An Econometric Approach. *Asian Social Science*, *10*(4), 203. http://dx.doi.org/10.5539/ass.v10n4p203

Haseeb, M., Hartani., N. H., Abu-Bakar, N. A., Azam, M., & Hassan, S. (2014). Exports, foreign direct investment and economic growth: Empirical evidence from Malaysia (1971-2013). *American Journal of Applied Sciences, 11*(6), 1010-1015. http://dx.doi.org/ 10.3844/ajassap.2014.1010.1015

Humphries, J., & Sarasúa, C. (2012). Off the Record: Reconstructing Women's Labor Force Participation in the European Past. *Feminist Economics, 18*(4), 39-67. http://dx.doi.org/10.1080/13545701.2012.746465

Kenjoh, E. (2005). New mothers' employment and public policy in the UK, Germany, the Netherlands, Sweden, and Japan. *Labor, 19*(s1), 5-49. http://dx.doi.org/10.1111/j.1467-9914.2005.00322.x

Lee, B. S., Jang, S., & Sarkar, J. (2008). Women's labor force participation and marriage: The case of Korea. *Journal of Asian Economics, 19*(2), 138-154. http://dx.doi.org/10.1016/j.asieco.2007.12.012

Lee, S. H., Ogawa, N., & Matsukura, R. (2009). Is Childcare Leave Effective in Raising Fertility in Japan? *Asian Population Studies, 5*(3), 349-369. http://dx.doi.org/10.1080/17441730903351750

Mahmoudian, H. (2006). Socio-demographic Factors Affecting Women's Labor Force Participation in Iran, 1976–96. *Critique: Critical Middle Eastern Studies, 15*(3), 233-248. http://dx.doi.org/10.1080/10669920600997043

Mishra, V., & Smyth, R. (2010). Female labor force participation and female total fertility rates in the OECD: New evidence from panel cointegration and Granger causality testing. *Journal of Economics and Business, 62*(1), 48-64. http://dx.doi.org/10.1016/j.jeconbus.2009.07.006

Narayan, P. K., & Smyth, R. (2006). Female labor force participation, fertility and infant mortality in Australia: some empirical evidence from Granger causality tests. *Applied Economics, 38*(5), 563-572. http://dx.doi.org/10.1080/00036840500118838

Ridao-Cano, C., & McNown, R. (2005). The effect of tax-benefit policies on fertility and female labor force participation in the United States. *Journal of Policy Modeling, 27*(9), 1083-1096. http://dx.doi.org/10.1016/j.jpolmod.2005.07.001

Sugawara, S., & Nakamura, J. (2014). Can formal elderly care stimulate female labor supply? The Japanese experience. *Journal of the Japanese and International Economies, 34*(0), 98-115. http://dx.doi.org/10.1016/j.jjie.2014.05.006

Tam, H. (2011). U-shaped female labor participation with economic development: Some panel data evidence. *Economics Letters, 110*(2), 140-142. http://dx.doi.org/10.1016/j.econlet.2010.11.003

Tsani, S., Paroussos, L., Fragiadakis, C., Charalambidis, I., & Capros, P. (2013). Female labor force participation and economic growth in the South Mediterranean countries. *Economics Letters, 120*(2), 323-328. http://dx.doi.org/10.1016/j.econlet.2013.04.043

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Economic Mathematical Modeling of Attributed Costs of Production in Industrial Enterprises

Mikhail Nikolaevich Dudin[1], Nikolaj Vasil'evich Lyasnikov[1], Dzhurabaeva Gulnora Kahramanovna[2] & Dzhurabaev Kahraman Tursunovich[2]

[1] Russian Academy of Entrepreneurship, Moscow, Russian Federation

[2] Novosibirsk State Technical University, Novosibirsk, Russian Federation

Correspondence: Mikhail Nikolaevich Dudin, Russian Academy of Entrepreneurship, Moscow, 105005, Russian Federation. E-mail: dudinmn@mail.ru

Received: December 18, 2014 Accepted: December 29, 2014 Online Published: July 6, 2015

doi:10.5539/mas.v9n8p40 URL: http://dx.doi.org/10.5539/mas.v9n8p40

Abstract

Statement of a problem: the present article is theoretical and methodological research focused on analysis and provisioning the basics for using different approaches to economic mathematical modeling of attributed costs of production in industrial enterprises. Approach: in its methodological part the article is based on a set of economic mathematical modeling methods. In particular, method of graphical simulation was used as well as regression modeling method.

Research **results** allow making conclusion that using graphical approaches to modeling of attributed costs of production does not always provides relevant and objective results. Authors prove that the best and the most practical way is to use regression and correlation modeling methods in managing cost prices.

Conclusion/recommendations: materials provide in this article are an addition to general management theory as well as extension of theoretical and methodological basis of production management. Main theoretical and methodological results obtained in the work are recommended as development platform for making high quality and effective management decisions in modeling cost price, volume of production and revenue of an enterprise.

Keywords: cost price, manufacturing facilities, economic mathematical modeling, econometric approach, graphical analysis, regression analysis

1. Introduction

Production activity management of today industrial enterprises is getting more and more complicated (Limitivski, 2010). Both internal and external factors contribute complicating specifics of production activity management of today industrial enterprises. Among external factors there is at first shift of global social and economic trend and transition to knowledge economy. Not the less important factor is change of management paradygm for the one with dominating role of innovative activity and sustainable development of industrial enterprises (Mochalova, 2010). Internal factors that affect specifics of industrial enterprise's functioning in external environment may be viewed in functional, process and system dimensions. But the most important is income, costs and revenue management because availability and sufficiency of financial resources is the basis of innovative-driven sustainable development of today industrial enterprises (McConnell & Bru, 2009).

The ongoing geopolitic processes reduce capability of today enterprises to generate high earnings growth rate of operational (core) activity. Besides, many enterprises have limited capability to raise external funding (Limitivski, Lobanova, Palamarchuk & Minasian, 2012). So changes of external environment point on the necessity not only of the search for new markets but also of preserving and increase of profitability. Preserving of relevant profitability level in the situation of limited capabilities to gain revenue depends on attributed costs management.

From economic point of view cost price is a money term of all costs and expenses that enterprise has bear for production and sales (Drury, 2010). Enterprise's costs management (or attributed costs management) is the process focused on streamlining of operational budget related to core activity. Streamlining of operational budget is one of the main conditions of revenue level growth while preserving or reducing of operational budget level. It is worth noting that in scientific sources it is often stressed the fact that costs and expenses are not identical

concepts (Limitivski, Lobanova, Palamarchuk & Minasian, 2012). On one hand it is true because costs may be both tangible (including money) and intangible. Expenses (including attributed costs of industrial enterprises) are represented only in money terms.

But on the other hand considering spending bear in operational activity enterprise put all of them in one form and in this case this form is price. So one may speak that expenses of industrial enterprise related to forming of cost price of products are evaluated in money terms all the resources spent.

2. Methodology

This article is methodologically represented by a set of methods of economic mathematical modeling. As related to practical calculations, an approach based on the correlation of two or more factors describing the specific nature of a separately taken enterprise's activity was used.

Bifactorial economic mathematical model describing dependence between cost price and production volume may be the following (Popov & Sotnikov, 2012):

$$Y = \alpha + \beta_x \tag{1}$$

Relation between variables $X\ (x_1, x_2 \ldots x_n)$ and $Y(y_1, y_2 \ldots y_n)$ is set by means of correlation analysis and factors' impact level is determined ad also it is determined the reliability of the proposition that factors are interrelated. Such indicators as co variation factor, pair correlation coefficient, dispersion coefficient, coefficient of determination, coefficient and others may be calculated in the course of analysis.

Coefficient β is calculated according to the formula (Popov & Sotnikov, 2012):

$$\beta = \frac{\sum_{i=1}^{n}(x_i - \acute{x}) * (y_i - \acute{y})}{\sum_{i=1}^{n}(x_i - \acute{x})^2} \tag{2}$$

Coefficient αis calculated according to the formula, respectively, [15]:

$$\alpha = \acute{y} - \beta_x \tag{3}$$

3. Results

Activity of today industrial enterprises is highly determined. Industrial production in Russia in the last couple of years is not characterized by sufficient positive dynamics.

At the same time there was no significant falls of industrial production index in Russia during 2014. Moreover, for some months industrial production indexwas higher comparing the same periods of 2013 (see Figure 1).

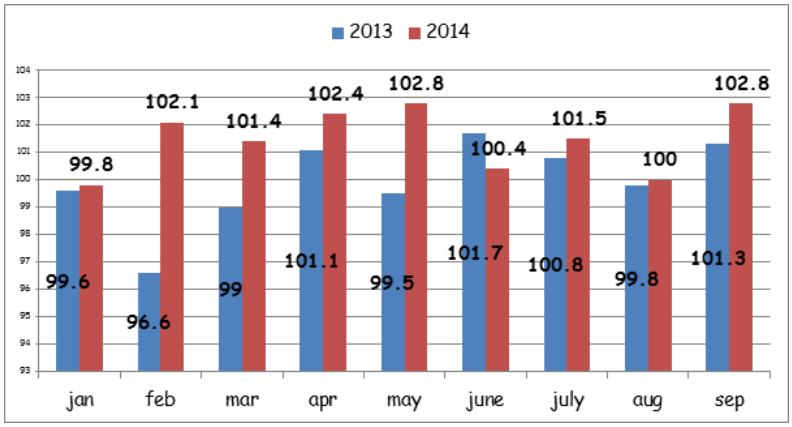


Figure 1. Industrial production index dynamics in Russia in 9 months of 2013–2014 (Index of industrial production. Federal Statistics Agency. www.gks.ru/bgd/free/b00_24/IssWWW.exe/Stg/d000/I000010R.HTM)

It is obvious that there is no forecasted stagnation of Russian industry against ongoing geopolitic changes in 2014, although the value of Russian index is somehow lower that American and European indicators (see Figure 2). Rate of industrial production volumes decrease in Germany in last two months of 2014 may be considered

significant. It is obvious that geopolitic factors affecting activity and development of industrial enterprises have bilateral character.

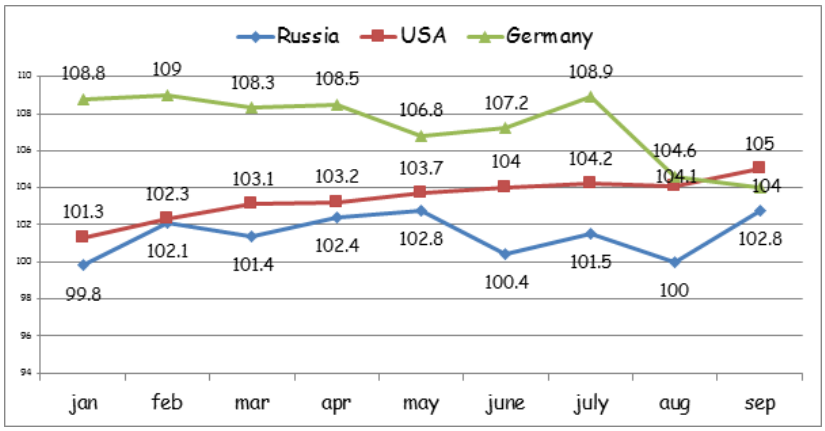


Figure 2. Industrial production indices in the USA, Germany and Russia in 9 months of 2014 (USA Industrial Output. Quote RosBusiness Consulting. www.quote.rbc.ru/macro/indicator/24/98.shtml; German Industrial Output. Quote Ros Business Consulting. www.quote.rbc.ru/macro/indicator/26/228.shtml)

So, today industrial enterprises in the world now face the same problems related to limitation of capabilities to generate revenue and at the same time necessity to preserve or increase revenue level that may be invested in further development.

So effective costs management (or expenses management) to provide appropriate and necessary to support the growth level of profitability of industrial enterprise operation becomes the most important managerial function. It is also necessary to make sure that all the costs (expenses) of industrial enterprise attributed to cost price are manageable. The answer is obvious — industrial enterprise (more precisely its management) is not always capable to manage all operational costs and expenses in planned and purposeful manner.

The reason is that in practice management of costs attributed to cost price is viewed as the necessity to reduce it. But managers often lose of sight that cut of expenses in operational activity of industrial enterprise may cause partial decrease in quality of products and loss of competitive advantage. Correlation between the greatest possible cut off attributed costs and such indicators as the quality of products, competitive advantage of an enterprise is objective and obvious.

Purchases of resources (both tangible and intangible) that support operational requirements of production are necessary to produce required volume of products of relevant quality (Total Cost Management Framework, 2012). Reduction of attributed costs usually causes choosing more cheap resources that affects the quality of products. Decrease of quality in turn leads to the loss of competitive advantage of industrial enterprise in external environment because the products of lower quality (even when these products are cheaper) means loss of customers' loyalty and draught of customers to competitors (Ross, Westerfield & Jordan, 2000).

So, critical cut of costs of industrial enterprise for resource provisioning in operational activity is absolutely wrong approach. To keep the balance of economic growth and sustainable development industrial enterprise should be capable to model effects of optimization of costs for resources and of sales products in managing operational costs.

Modeling as a special instrument of management is widely used in managing activity of today industrial enterprises (Limitivski, 2010).

In theory model is analogous to real physical process or object with keeping features and patterns of its functioning and development. So economic modeling (including modeling operational attributed costs) may be viewed as process focused on reproduction of analogue of operational activity of industrial enterprise with keeping the relation of resources utilization and production of necessary volume of products with required quality.

It is accepted that attributed costs in the scope of production activity of industrial enterprise are classified by items of calculation and items of expenses (Cooper, 1990). Theoretical classification of operational budget (attributed costs and costs of sales) of industrial enterprise is shown in Figure 3.

It is obvious that in some cases calculation items and expense items attributed to price cost of operational activity of industrial enterprise are the same. So the most relevant approach is dividing costs by permanent and variable (i.e. dependent and independent of production volumes). Classification of costs attributed to cost price of production and sales of industrial enterprise is the base of selection of the method of costing.

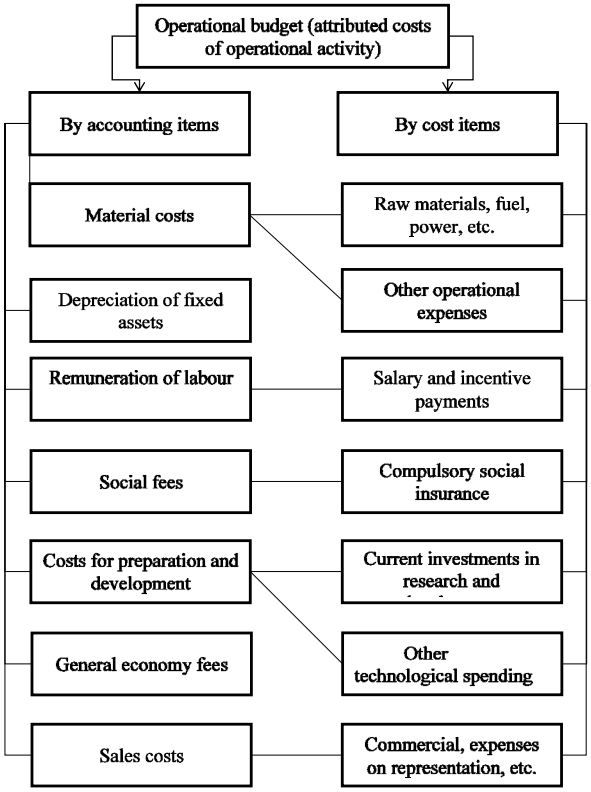


Figure 3. Classification of operational budget (attributed costs) of industrial enterprise

In particular, it is accepted to distinguish between several methodological approaches to costing (Cooper & Kapla, 1991):

- standard costing is costing method that may be defines as normative. In other words standard costing is normative resource usage in money terms necessary to product an item of production (and planned volume of products, respectively). Key idea of standard costing method is that industries in which resources usage cost for production and sales is not subjected significant variations may develop standards of costs and expenses. These standards are used as the basis for planning activity of industrial enterprises and evaluate deviations;
- direct costing is costing method based on separation of the latter on direct and indirect costs. Direct costs are costs that are directly related to production volume (i.e. are direct variable costs). Indirect costs, respectively, are not related to cost price of production and sales and do not depend on production volume (i.e. are indirect permanent costs). Direct variable costs are usually included in cost price in direct costing although there is a concept of advanced direct costing when both direct variable and indirect variable costs are included in cost price;
- target costing is costing method based on calculation of target cost price. Target costing is usually used in development and marketing new products, including innovative products. Unlike normative approaches to costing target costing method presupposes setting target reference point for production and sales to offer competitive product not only by quality but also by price. In other words target costing method allows making optimal price/quality ratio basing on target expenses related to production and sales.

This list of costing methods is far from being comprehensive. In particular, such methods as «Justin Time», «ABC-cost» and many others methods and models may be named. To our mind each of listed costing methods for evaluation of price cost of operational activity has its advantages and drawbacks. In particular, normative

costs of production per one item or the whole volume of products in condition of turbulent changes of environment and resources' prices does not allow adaptively approaching planning of operational activity of industrial enterprise. Setting target volume of attributed costs of operational activity reduces the number of variants of customer stimulation (in this case price stimulation) that may have negative impact on customers' perception of products.

At the same time separation of costs and expenses related to cost price of operational activity of industrial enterprise in permanent and variable allows not only adapting planning procedures to changing external conditions but provide required balance of resources considering possible linear and non-linear trends of prices variations and prices of these resources.

4. Discussion

It is commonly thought in macro- and microeconomy that increase of production volumes means reduction of cost unit (price cost of production of production unit) due to economy of scale (Deakins, Logan & Steele, 2001). But at the same time standard CVP model (model costs-volume-profit) is based on the principle of variation of one parameter with invariability of two other parameters (or with directly proportional variation of these parameters depending on changes of variable parameter) that makes this model linear (see Figure 4).

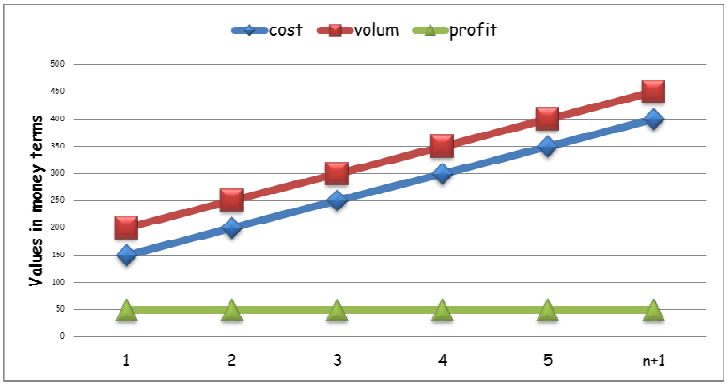


Figure 4. Linear representation of CVP model (model costs-volume-profit)

In ideal view of CVP model in each further period with increase (or cut) of costs production volume proportionally grows and revenue level remains the same.

4.1 Graphic Simulation of Cost Price

But taking into consideration non-linear character of changes going on in external environment of industrial enterprise cost unit per unit of production may be characterized by converse effects (see Figure 5). It is obvious that growth of production v (f per production unit may reduce but after passing a certain turning point further se rom V_2 to V_3) may cause reverse effect.

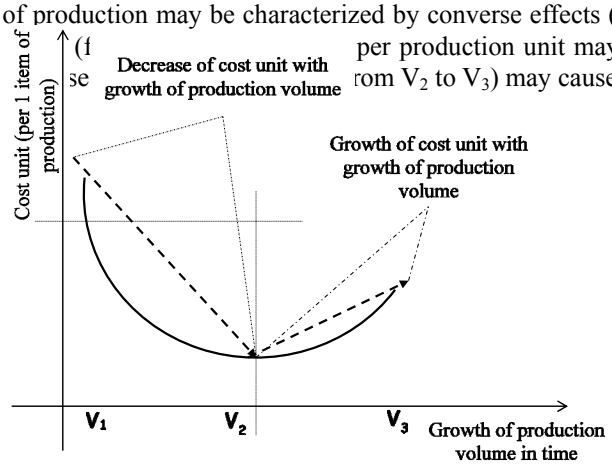


Figure 5. Non-linear change of cost unit of production with production volume growth

The essence of this converse effect is that despite any increase of production volume cost unit and, respectively, aggregated attributed costs always demonstrate higher comparative growth. And moreover, it is obvious that always the level of aggregated costs and expenses attributed to price cost of production and sales grows with the

growth of production volume.

The following dependence may be derived as the ratio of cost price and production volume (sales) changing. Figure 6 shows buoyant dynamics that characterizes the growth of production volume and proportional growth of cost price.

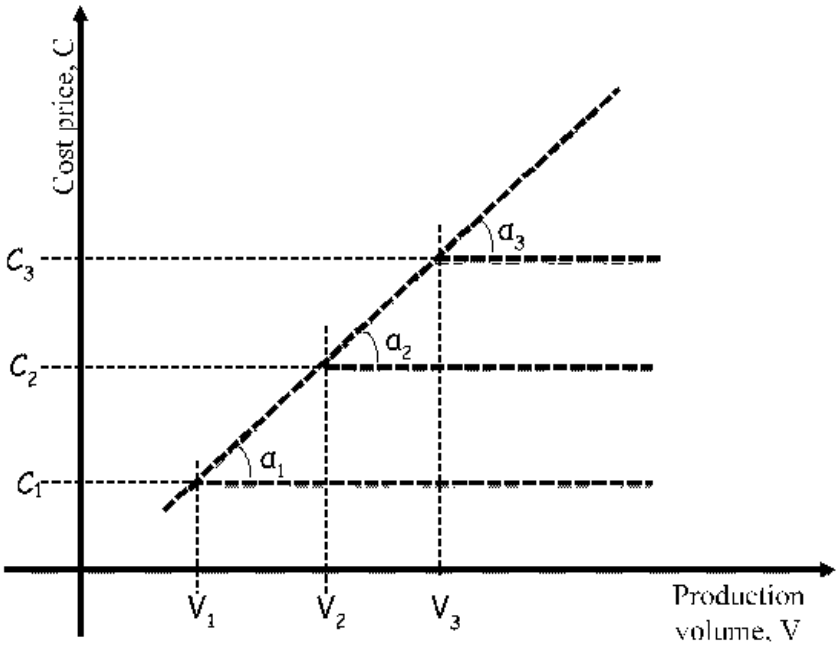


Figure 6. Graphic of dependence of cost price change in respect to change of production volume on industial enterprise

Reverse situation (decrease of production volume and proportional decrease of cost price) requires independent analysis, but bearish trend is described in the same way. So each next crossing of lines and trend of sales volume change will form an angle (in our case this angle is denoted as $\alpha_1 - \alpha_3$). It is commonly considered that angle formed by lines C_n;V_n and their change trend is 45º and it has conditionally constant value defined as α (tg_α) angle tangent, i.e.:

$$\angle\alpha(45^\circ) = \frac{sin_\alpha}{cos_\alpha} = tg_\alpha = 1 \tag{4}$$

But this is not constant because usage of resources included in cost price and change of production volume are not characterized by a priori proportional or linear character. So angle formed by lines C_n;V_n and their change trend (angle α) is not always 45˚ and is 1. From geometry it is known that tangent of any angle varies from 0 (for angles 0˚; 180˚; 360˚) to infinity (for angles 90˚; 270˚). So change of cost price depends on changes of production volumes, adjusted by the factor of business activity of an enterprise. Factor of business activity in this case equals the value of tangent of angle α ($tg_{\alpha n}$). So, calculation of costs and expenses attributed to cost price of production and sales may be represented as:

$$C_t = C_p + \Delta V_t * k(ortg_a) \tag{5}$$

Where:

C_t – attributed cost of production and sales in time point (t);

C_p – planned cost price of production and sales;

ΔV_t – production volume change in time point (t);

k – business activity factor, defines as angle α tangent in time point (t).

In theory, graphic of the sum of attributed price of cost and sales that depends on business activity level of industrial enterprise may vary in the range 0˚ - 180˚, i.e. be a parabola (see Figure 7). But in practice neither the sum of attributed costs of production nor production volumes and sales of that production can be negative figures.

So variations of the sum of attributed costs from geometrical point of view are possible in the range of variations of angle α tangent — from 0˚ to 89,9˚ (the task has no solution for tangent of α = 90˚). One should agree that graphical solution of the task of modeling of optimal attributed costs of production and sales is not always possible in practice, besides solutions are not always correct because strict unification of costs data with their graphical representation is required.

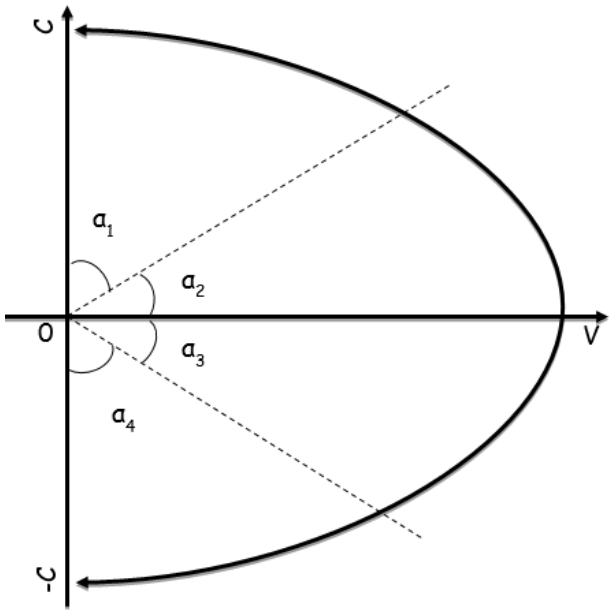


Figure 7. Theoretical interpretation of attributed costs variation depending on production volume and sales on industrial enterprise

It is also necessary to understand that constant and variable costs in attributed costs has different growth and decrease rate and it may depend and may not depend on production volumes and sales (Geroski & Gugler, 2004).. So in practice it is optimal to use regression models that describe dependence of two and more indicators.

4.2 Correlation Modeling of Cost Price

Let us assume, that there is industrial enterprise with production volumes and attributed costs of this production are characterized by the following parameters (see Table 1).

Table 1. Volumes and attributed cost of production on industrial enterprise

Period	Production volume (Y)	Attributed costs of production (X)
1	37221	33456
2	27201	26465
3	32594	31539

Then economic mathematical model of regression dependence of attributed costs and production volumes on industrial enterprise is derived (Table 2, Figure 8).

Table 2. Economic mathematical model of regression dependence between attributed costs and production volume on industrial enterprise

#	Y	X	$y_i - \bar{y}$	$x_i - \dot{x}$	$(y_i - \bar{y})^* (x_i - \dot{x})$	$(x_i - \dot{x})^2$	$(y_i - \bar{y})^2$
1	37221	33456	4882.3	2969.3	14497275.11	8816940.4	23837178.8
2	27201	26465	-5137.7	-4021.7	20661982.78	16173802.8	26395618.8
3	32594	31539	255.3	1052.3	268695.7778	1107405.4	65195.1
Total	**97016**	**91460**	**7E-12**	**7.3E-12**	**35427953.67**	**26098148.7**	**50297992.7**
Average	32338.67	30486.67					
Pair correlation coefficient					0.978	---	

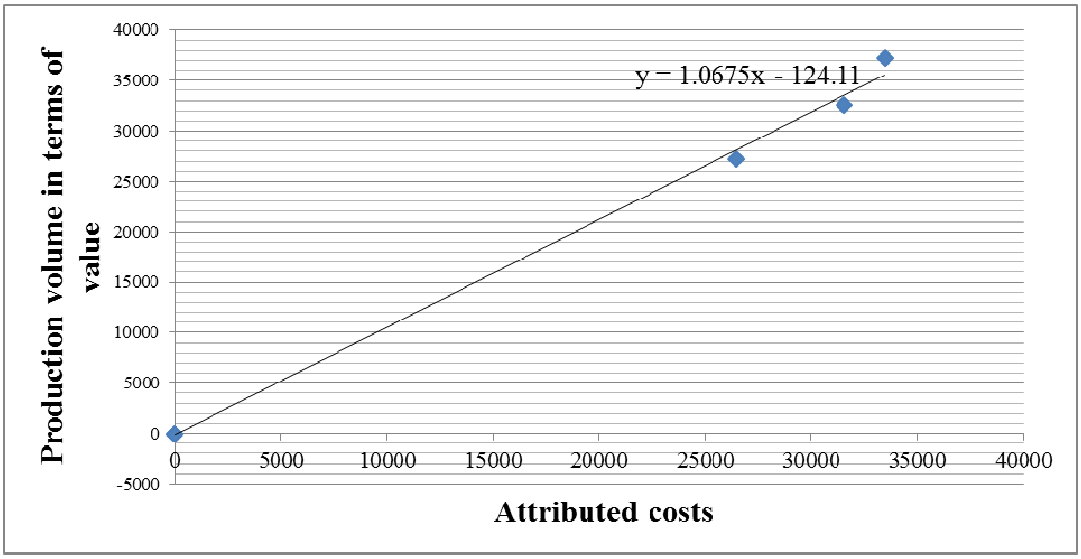


Figure 8. Model of regression dependence between attributed costs and production volumes on industrial enterprise

Calculation in presented economic and mathematical model allows making the following conclusions:

– firstly, attributed costs of production dependence of variations of production volume is relatively high, because in reverse model there will be equivalent results. Correlation between factors is high, according to Chaddock criterium;

– secondly, the model shows that in each iteration step growth of production volume on 1 conditional unit will potentiate attributed costs growth by 1.0675 conditional price units, i.e. growth rate of expenses are always in average 6% higher than sales growth rate. Revenue of an enterprise will be decreasing, respectively.

Calculations shows that in this example attributed costs grow disproportionately with the lower rate of production volume growth. Besides, disproportionate and consequently uncontrolled growth of spending causes reliable bearish trend of the revenue. To reveal and clarify the problem of surpassing attributed costs growth rate comparing production volume production and sales volume should be correlated with constant and(or) variable costs' items. In this step it may be empirically determined that the higher the value of some cost item in money terms the better is correlation between this item and production volume and consequently revenue and income.

So optimization solutions related not with setting normative values of spending but with setting limited thresholds are best solution in managing attributed costs of production and sales on industrial enterprise to achieve required economic benefits (revenue and income). In these thresholds attributed costs of production may vary from preset minimum to preset maximum depending on analogous variations of production volumes and sales. Enterprise should define critical production volume or control point. Beyond this point economy of scale is lost and costs per unit for production and sales of a unit of production do not decrease but on the contrary grow. Defining control point of critical production volume allows enterprise to develop and realize plans of sustainable functioning and development in the most adaptive manner (Stewart, 1991; Dudin, Ljasnikov, Pankov & Sepiashvili, 2013).

5. Conclusion

Theoretical and methodological conclusions made in the work allow stating that management of attributed costs of production and sales on industrial enterprise is complicated task that requires different approaches to respective decisions. According to authors, attributed costs and sales management toolkit used now on industrial enterprise now requires adaptive and natural basis. Graphical modeling method and dependence deriving method (both proportional and non-proportional) between variation of attributed costs and production volumes may be used as such methodological basis. Besides, method of correlation regression modeling may be used that allows revealing patterns of changing of attributed costs depending on variability of production volume changes.

Theoretical and methodological thesis of the article allows stating that the greatest effectiveness from the point of view of relevance of result obtained have methods of economics and mathematical modeling based on regression deriving and determining correlation in variations of two interrelated economic indicators (attributed

costs and production volume).

Graphical methods of modeling attributed costs of production require universalization and unification of used criteria, indicators and calculation methods. So use of graphic methods of modeling of attributed costs of operational activity of industrial enterprise in practice is limited.

References

Cooper, R. (1990). Cost classification in unit-based and activity-based manufacturing cost. *Journal of Cost Management*. Fall, 4–14.

Cooper, R., & Kapla, R. S. (1991). *The design of cost management systems.* Cases and Readings. Prentice-Hall.

Deakins, D., Logan, D., & Steele, L. (2001). *The financial management of the Enterprise.* London: Certified Accountants Educational Trust.

Drury, C. (2010). *Managerial and production accounting*. Moscow: UNITI-Dana.

Dudin, M. N., Ljasnikov, N. V., Pankov, S. V., & Sepiashvili, E. N. (2013). Innovative foresight as the method for management of strategic sustainable development of the business structures. *World Applied Sciences Journal, 8,* 1086-1089. http://dx.doi.org/10.5829/idosi.wasj.2013.26.08.13550

German Industrial Output. Quote Ros Business Consulting. Retrieved October, 26, 2014, from http://quote.rbc.ru/macro/indicator/26/228.shtml

Geroski, P., & Gugler, K. (2004). Corporate growth convergence in Europe. *Oxford Economic Papers, 56,* 597–620. http://dx.doi.org/10.1093/oep/gpf055

Index of industrial production. *Federal Statistics Agency.* Retrieved October, 26, 2014, from http://www.gks.ru/bgd/free/b00_24/IssWWW.exe/Stg/d000/I000010R.HTM

Limitivski, M. A. (2010).Sustainable growth of a company and leverage effects. *Russian Management Magazine*, *2*, 35-46.

Limitivski, M. A., Lobanova, N. E., Palamarchuk, V. P., & Minasian, V. B. (2012). Corporate financial management. Moscow: Urait.

McConnell, K. R., & Bru, S. L. (2009). *Economics*. Moscow: Infra-M.

Mochalova, L. A. (2010). Model of risk-oriented financial strategy of a corporation. *M.F. Reshetnev Siberian State Aerospace Uniersity Gerald.*

Popov, A. M., & Sotnikov, V. N. (2012). *Economic mathematical methods and models.* Moscow: Urait.

Ross, S., Westerfield, R., & Jordan, B. (2000). Basics of corporate finances. Moscow: Laboratory of Basic Knowledge.

Stewart, G. B. (1991). *The Quest for Value*. NY: Harper Business.

Total Cost Management Framework. (2012). An Integrated Approach to Portfolio, Program, and Project Management. *PE CCE CEP. AACE International.*

USA Industrial Output. *Quote RosBusiness Consulting*. Retrieved October, 26, 2014, from http://quote.rbc.ru/macro/indicator/24/98.shtml

Copyrights

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

A Study of the Relationship between Internet Dependence and Social Skills of Students of Medical Sciences

Hossein Jenaabadi[1] & Ghazal Fatehrad[2]

[1] Psychology Department of Education, University of Sistan and Baluchestan, Iran

[2] MA. Degree on educational Administration. Tehran Islamic Azad University, Science and Research Branch, Iran

Correspondence: Hossein Jenaabadi, Associate professor of Psychology Department of Education, University of Sistan and Baluchestan, Iran. E-mail: hjenaabadi@ped.usb.ac.ir

Received: November 29, 2014 Accepted: January 8, 2015 Online Published: July 6, 2015

doi:10.5539/mas.v9n8p49 URL: http://dx.doi.org/10.5539/mas.v9n8p49

Abstract

Introduction: Internet dependence is a topic of interest that has been discussed as a behavior-based addiction in recent years and has become a growing issue in the information technology era. This addiction has caused many problems for college students. In this regard, the current study aimed to investigate the relationship between Internet addiction and social skills of students of Medical Sciences.

Methods: This is a descriptive-correlational study. The sample included 354 medical students who were selected through applying stratified random sampling method and were tested using two questionnaires of Internet Addiction and Social Skills. Data were analyzed applying the Pearson correlation coefficient and stepwise regression analysis.

Results: The findings indicated that there were significant positive relationships between Internet dependence and social skills. Internet dependence has a reversed relation with initiation and termination, assertiveness, social reinforcement, empathy, and cooperation. Increasing Internet dependence, these skills weakened. However, no significant correlation was found between Internet dependence and orientation skills. Moreover, the results of the regression analysis showed that these five variables predicted about 66% of the criterion variable (internet dependence).

Conclusion: Since Internet addiction can falter students' social skills and has strong negative effects on interpersonal communication and social interaction, it is essential to make efforts to give students' use of the Internet a specific direction to avoid its probable adverse effects.

Keywords: internet, internet dependence, social skills, university students

1. Introduction and Objective

The essence of society is sustainable communications between human beings and communication as the first need of any system is a special form of social interaction (Doran, 2006). Nowadays, the Internet has become an essential part of everyday life around the world (Bayraktar & Gun, 2007) and its use, especially among the young and educated, is steadily increasing (Ozcan & Buzlu, 2007) and brings with itself various promises and threats (Mc Quail, 2005). Creating new social cohesion, increasing social isolation, reduction of social participation, Internet dependence (Mc Kenna and Bargh, 2004) and replacing social activities with the Internet (Kestenbaum et al., 2002) are among the numerous effects of the Internet. These effects are so extensive and widespread that some researchers consider the role of the Internet in today's life as the role of electricity in the industrial revolution (Liesler et al., 2002).

Internet dependence is an interesting issue being recently considered as a behavior-based addiction (Mohammad-beigi et al., 2010) and has become the growing issue of the information technology era (Davis and Besser, 2002; Hur, 2006). Internet dependence refers to one's inability to control himself/herself in using the Internet which gradually leads to create uncomfortable feelings, debilitation of daily functional activities (Shapira et al., 2003; Douglas et al., 2008) and eventually mental disturbance or stress (Weinstiein and Lejoyeux, 2010). Douglas et al. introduced Internet dependence as an obsessive-compulsive use of the Internet, that if the individual is excluded from such use, he/she becomes so irritated and shows behaviors with tantrum (Douglas et

al., 2008). Internet dependence is associated with symptoms such as anxiety, depression, irritability, obsessive thoughts, isolation and disruption of social relations (Kooraki et al., 2012).

On the other hand, it seems that Internet dependence influences individuals' social skills (Preece, 2000; Wellman and Gulia, 1999). Social skills are a series of purposeful, interrelated behaviors based on the status, which are acquired and are under the control of individuals (Momayezi et al., 2011). Social skills are important factors in shaping relationships, enhancing the quality of social interaction and even mental health of individuals. These skills can be defined as a perfect model of behaviors exhibited by an individual in interpersonal relations (Teodoro, 2005). Social skills encompass a variety of capabilities and capacities including emotional expression, self-regulation, social flexibility, social sensitivity, assertiveness (Segrin & Taylor, 2007). Emotional expression is defined as the ability to understand and express emotions in interpersonal relationships. Self-regulation is an integrated behavioral process consisted of useful behaviors for learning. Social sensitivity is defined as the ability to decode non-verbal signs and social flexibility is the ability to inhibition emotions of self in stressful conditions. Assertiveness is another social skill that enables an individual to act in favor of him/herself and express his/her feelings honestly without abusing others' rights (Kooraki et al., 2012).

Since Internet connections, unlike face to face communication, are more flexible and individuals can easily remove or edit their negative information, such relationships continue and lead to shortcomings in social skills. A study on the relationship of social self-efficacy, locus of control and Internet addiction among students revealed that there was a negative relationship between social self-efficacy and Internet addiction (Iskender and Akin, 2010). During the past decade, the number of studies conducted on Internet addiction has increased rapidly (Bayraktar & Gun, 2007; Yen et al., 2007; Kubey, Lavin and Barrows, 2001). For example, researchers found significant relationships between Internet dependence, anxiety, restlessness, aggression and hostility (Yen et al., 2007), drugs experience (Yen et al., 2007; Ko et al, 2006), losing control and endurance, retreat, performance disorder, reduction of decision-making ability (Ko et al, 2005), loneliness (Kraut et al., 1998; Nalwa and Anand, 2003), sensation seeking (Lin & Tsai, 2002), poor mental health (Yang, 2001), reduction of family communications (Armstrong, Phillips and Saling, 2000), reduction of students' GPA and their educational failure (Mohammad-beigi et al., 2010), low self-concept (Shoa Kazemi, 2008), reduction of social communications circle (Young and Rogers, 1998), loss of social skills and social isolation (Moody, 2004; Whitty and McLaughlin, 2007), low social quality of learners (Nadi & Sajjadian, 2010) and low social skills (Kooraki et al., 2012), etc.

In many countries, medical university students form a substantial group of Internet users (Lazinger, Bar-Ilan and Peritz, 1997; Salkovic-Petrisic et al., 2001) and this group of the users see the Internet as a reliable source (Eitel, Yankowitz and Ely, 1998) and use it for personal as well as professional goals (Scheleyer, Spallek and Torres-Urquidy, 1998; Mohagheghzadeh and Abdollahi, 2002). However, in spite of multitudes of benefits and capacities for its users, it seems this technological and information tool has attracted youth's attention and lead them to spend long hours at the computer that results in hindering from other aspects of life especially social activities. Given the young age structure of the country and considering that the youth are the most vulnerable stratum of the society against risky behaviors such as improper use of the computer and the Internet, obviously each damage to their physical and mental health and also reduction of the abilities of this active group, inevitably leads to slowing down the progress of the society. On the other hand, in the long-term, their physical and mental health may have negative effects on the health of the whole people of the society. Statistics in Iran (2004) indicated that the number of the Internet users was more than 6 million and this number is increasing day by day (Sadeghian, 2008). Considering the increasing use of the Internet in Iran especially by the youth and the adolescent, if preventive measures are not foreseen, soon we will witness an increase in problems of this area especially Internet dependence. However, very few studies are conducted on the subject. Therefore, there is an emergent need to carry out more extensive, detailed and rigorous research on Internet dependence and social skills.

In many countries, medical university students form a substantial group of Internet users (2001) and this group of the users see the Internet as a reliable source (1998) and use it for personal as well as professional goals (2006). However, in spite of multitudes of benefits and capacities for its users, it seems this technological and information tool has attracted youth's attention and lead them to spend long hours at the computer that results in hindering from other aspects of life especially social activities. Given the young age structure of the country and considering that the youth are the most vulnerable stratum of the society against risky behaviors such as improper use of the computer and the Internet, obviously each damage to their physical and mental health and also reduction of the abilities of this active group, inevitably leads to slowing down the progress of the society. On the other hand, in the long-term, their physical and mental health may have negative effects on the health of the whole people of the society. Statistics in Iran (2004) indicated that the number of the Internet users was more

than 6 million and this number is increasing day by day (2013). Considering the increasing use of the Internet in Iran especially by the youth and the adolescent, if preventive measures are not foreseen, soon we will witness an increase in problems of this area especially Internet dependence(2014). However, very few studies are conducted on the subject(2013). Therefore, there is an emergent need to carry out more extensive, detailed and rigorous research on Internet dependence and social skills(2014). Thus, in this regard, the present study sought to answer the question that whether there is any relationship between dimensions of social skills and Internet dependence among students and what is the contribution of each dimension of social skills in prediction of Internet dependence?

2. Methods and Materials

Since it examines the relationships between the variables, this is a descriptive-correlational study. The population consisted of all medical science students of Islamic Azad university of Tehran, Sciences and Research Branch (N=4277), among which based on the Cochran's formula, 354 students were selected using stratified random sampling method (to place on the sample students from all fields and branches of medical sciences). Sample participants included 212 female students and 142 male students. Two tools were used in this study.

A) Social Skills Inventory: The inventory was designed and developed by Abbasi for higher education in 2003. The inventory has 143 items. It was standardized on 142 female students and 212 male students. Subjects should specify their response among a 5-point Likert type scale, (always, quite often, sometimes, rarely and never). Six factors including initiation and termination, self-assertiveness, social reinforcement, empathy, orientation and cooperation are analyzed. The internal consistency of the inventory was calculated using the Cronbach's alpha that was 0.96 and standardized scores of t and z were calculated on two groups of male and female subjects of Ahvaz Azad University that revealed good construct validity and predictive power (Harjy et al., 2005).

B) Internet Addiction Test: The instrument is developed and standardized by Young (Kimberly, 2007). The test consists of 20 items and subjects should specify their response among a 5-point Likert-type scale, (always, often, sometimes, rarely, does not apply to me). The internal consistency of the test was calculated using the Cronbach's alpha that was 0.92. It also has a good construct validity and predictive power. In the present study, before the final run, a pilot study was implemented on a sample of 50 students which revealed that the test is reliable at the alpha level of 0.01 with the Cronbach's alpha of 0.92 which indicates that the reliability coefficient is high.

3. Results

Q1: Is there any significant relationship between dimensions of social skills and Internet dependence among students?

Before proceeding to the test results, first descriptive statistics are presented in Table 1.

Table 1. Descriptive characteristics of the variables

Variable	Mean	SD	Variance	Skewness	Kurtosis
Internet dependence	77/1	682/0	466/0	50/0	09/0-
Initiation and termination skills	71/2	601/0	361/0	44/0-	14/0
Assertiveness skills	54/2	549/0	302/0	64/0-	54/0
social reinforcement skills	92/2	645/0	416/0	87/0-	47/1
Orientation Skills	54/2	656/0	430/0	59/0-	62/0
Empathy skills	76/2	647/0	419/0	15/0-	84/1
cooperation skills	76/2	647/0	419/0	48/1-	17/1

Mean and standard deviation values in the above table indicate the appropriate dispersion and Skewness and Kurtosis indices indicate a normal distribution of the variables. The above table also indicates that among the dimensions of social skills, the highest scores goes for the component of social reinforcement skills (2.92) and the lowest scores goes for the components of assertiveness skills (2.54) and orientation skills (2.54). The Pearson correlation test was used to examine the hypothesis. The results are presented in Table 2.

Table 2. Correlation coefficients of Internet dependence and social skills

Variable	Internet addiction	Sig
Initiation and termination skills	152/0-	01/0
Assertiveness skills	143/0-	01/0
Social reinforcement skills	185/0-	001/0
Orientation Skills	091/0-	08/0
Empathy skills	183/0-	001/0
Cooperation skills	133/0-	01/0

The table above indicates that the results of all dimensions of social skills except for orientation skills have significant negative correlation with Internet dependence. Negative correlation indicates that the increase of Internet dependence among students reduces their social skills and vice versa.

Q2: what is the contribution of each dimension of social skills in prediction of Internet dependence?

To answer this question a stepwise multiple regression analysis was used. In the first step, the variable of empathy skills, in the second step, initiation and termination skills, in the third, assertiveness skills, in the fourth social, reinforcement skills and in the fifth step, cooperation skills entered into the equation and maintained their significance in five steps. Only the variable of orientation skills was not significant, and removed from the equation. The results of the regression analysis are reported in Table 3.

Table 3. Results of stepwise regression analysis of Internet dependence from predictor variables

Predictor variable	R	R^2	Δr^2	SE	df	F	Δf	B	Beta	T
Empathy skills	66/.	44/.	44/.	32/5	352/1	***61/280	61/280	31/.-	***87/.-	86/17-
Initiation and termination skills	73/.	54/.	10/.	82/4	351/1	***32/210	33/78	50/.-	***49/.-	85/13-
Assertiveness skills	76/.	58/.	04/.	62/4	350/1	***86/162	45/31	29/.-	***35/.-	97/8-
social reinforcement skills	78/.	62/.	04/.	44/4	349/1	***83/139	11/30	30/.-	***31/.-	83/6-
Orientation Skills	81/.	66/.	04/.	18/4	348/1	***31/135	66/45	45/.-	***26/.-	76/6-

*** p <0.001

Stepwise regression analysis results in Table 3 indicate that empathy skills predicted 44% of the variance of Internet dependence. Initiation and termination skills, assertiveness skills, social reinforcement skills and cooperation skills predicted 10%, 4%, 4%, and 4% of the variance of Internet dependence, respectively. These five variables predict a total of 66% of criterion variable variance. ANOVA results presented in the above table and the significance f value indicates that these five variables have a significant effect on Internet dependence. Due to the beta coefficient which is negative, it can be argued that students, who have high scores on measures of social skills, are more dependent on the Internet and vice versa.

4. Discussion and Conclusion

The present study aimed to investigate the relationship between Internet dependence and social skills of students. To examine the research hypotheses, the Pearson correlation and stepwise multiple regression were used. In the following, the results of these two tests used to explain each of the dimensions of social skills, are discussed separately. Results of the present study about the reduction of social skills among students in the light of Internet dependence are consistent with the results of Kooraki et al. (2012) Young and Rogers (1998), Moody (34), Whitty et al. (2007) and Nadi and the Sajjadian (2010).

As the findings indicated, there was a significant negative relationship between Internet dependence and initiation and termination skills. Therefore, if students' Internet dependence increases, their initiation and termination skills decreases. Results of regression analysis revealed that this component predicted 10% of the variance of Internet dependence. According to the results of the statistical data analysis, it seems that the relationship of Internet dependence and initiation and termination skills resulted from the interventions in learned behaviors. Verbal and non-verbal behaviors cause individuals' interactions with others. Excessive use of the Internet makes a person use less verbal behavior and progressively his speech is impaired and his/her contact with the outside fades. Increased Internet use increases loneliness, depression, and social isolation. It also affects

one of the verbal skills, i.e., initiating and terminating a social conversation (Kraut et al., 1998; Nalwa and Anand, 2003).

In the current study, the relationship between Internet dependence and assertiveness skills was also negative. Multiple regression analysis also suggested that this component was able to predict 4% of the variance of Internet dependence. Assertiveness is another social skill that enables an individual to act in favor of him/herself and express his/her feelings honestly without abusing others' rights. Such an individual is able to stand on his/her own feet, show his/her true feelings honestly and simply reject unreasonable requests. These individuals have more control over their own lives and have firm and steadfast personalities. The statistical data suggested that there was a significant negative relationship between Internet dependence and assertiveness that means individuals who are addicted to the Internet lose such characteristics and cannot firmly reject unreasonable requests. The study found that increased Internet usage reduces the ability of assertiveness. However, the relationship was not robust and other variables were not controlled.

The findings also suggested that increased dependence on the Internet among students decreased their social reinforcement skills (another dimension of social skills). The predictive power of this component was 4%. It should be argued that Internet dependence and social reinforcement skills arise from positive reinforcements that the Internet suggests. Strengthening social relationships on the side of family and friends increases and maintains interaction and marks interest in the ideas, thoughts and feelings of other people. When individuals do not receive reward or reinforcement from their family or the society and their interests and emotions are reacted through the Internet, they prefer to answer to such reinforcements and thus their addiction to the Internet rises and they withdraw from friends and the community.

However, the important point in investigating social skills is the fact that unlike other dimensions, no significant relationship was observed between orientation skills and Internet dependence. Regression results indicated the inability of this component to predict Internet dependence. The lack of correlation between the above variables can be explained in this way that leading individuals have more ability and talent compared to their peers, they have higher confidence and morale and can lead a group. Acquiring such an ability arises either from their family in which some responsibilities are taken by the individual or their academic discipline (Kooraki et al., 2012). Individuals with orientation skills do their own tasks and duties without other's help, since the ability of doing things independently has made them responsible and aware of their own ability and talent and this leads to high confidence. Extroverts receive more social support than introverts. Individuals with strong social ties use the Internet in order to complement and strengthen their communication. Considering the statistical results, leading individuals are not Internet dependence and use the Internet positively to complement their communication.

In addition, the present study indicated a significant negative relationship between Internet dependence and empathy skills. According to statistical results, it is found that dependence to the Internet reduced individuals' empathy skills. Individuals without empathy skills can hardly understand others' emotions and discomforts, they cannot feel others' sorrow, and they try to be seen as less as possible in the community which leads to their withdrawal and isolation. Therefore, Internet dependence is associated with separation from community and spending more time on the Internet (Moody, 2004).

Finally, based on the results, there was a significant negative relationship between cooperation skills and Internet dependence. It can be argued that cooperation skills are among acquired skills, such as the ability to ask others' help and effective communication so that this connection is acceptable, valuable and lucrative for both parties. The time spent on an activity inevitably leads to lower time on similar activities. Thus, there is still concern that the use of the Internet, especially non-social uses, such as web surfing, spending much time on the Internet, is associated with a decrease in social activities including cooperation skills which eventually leads to social isolation. Such individuals refuse to participate in group activities and avoid collaboration with peers. An increase in Internet dependence decreases cooperation skills.

Today, social skills are highly regarded; so that in social and educational systems if individuals lack social skills, they are often labeled inappropriately and are considered as people with social problems. Inability in social skills even creates problems in terms of academic achievement, because lack of such skills may destroy self-worth and self-esteem and make the individual seem different from others or be considered as abnormal. In such a situation, the individual is easily excluded and loses desirable positions for development along with his/her peers. However, one of the factors that could cause impairment or loss of social skills may be Internet dependence.

Although the development of electronic devices and the Internet is remarkable in all scenes, today's modern civilized man needs to appropriately use these electronic means, including the Internet in order to avoid flaws in his social relations. Excessive use of the Internet and severe interest in it causes Internet dependence. It seems

that one of the factors that led to significant relationship between Internet dependence and social skills is the lack of the mentioned skills on the side of students. Internet dependence causes disorders in life, loss of social skills or social isolation. Overall, the results of this study suggested that there was a significant negative relationship between Internet dependence and students' social skills. Therefore, it can be concluded that Internet dependence decreases students' social skills.

References

Autor, D. H., & David, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review, 103*(5), 1553-97. http://dx.doi.org/10.1257/aer.103.5.1553

Azher, M., Behram K. R., Salim, M., Bilal, M., Hussain, A., & Haseeb, M. (2014). The Relationship between Internet Addiction and Anxiety among students of University of Sargodha. *International Journal of Humanities and Social Science, 4*(1).

Cho , Sung , Shin , Lim and Shin (2013). Does psychopathology in childhood predict internet addiction in male adolescents*? Child Psychiatry Hum Dev., 44*(4), 549-555. http://dx.doi.org/10.1007/s10578-012-0348-4

DeMatteo, F. J., Arter, P. S., Sworen-Parise, C., Faseiana, M., & Panihamus, M. A. (2012). Social Skills Training for Young Adults with Autism Spectrum Disorder: Overview and Implications for Practice. *National Teacher Education Journal, 5*(4), 57-65

Heckman, James, J., & Kautz, T. (2012). Hard Evidence on Soft Skills. *Labour Economics, 19*, 4451-464. http://dx.doi.org/10.1016/j.labeco.2012.05.014

Hillary, A. (2013). Problematic Internet Use among Persons with Social Anxiety Disorder. *PloS One, 8*(2), e57831. http://dx.doi.org/10.1371/journal.pone.0057831

Hong , Zalesky, Cocchi, Fornito, Choi , Kim , Suh, Kim , Kim and Yi (2013). Decreased functional brain connectivity in adolescents with internet addiction. *PloS One, 8*(2), e57831. http://dx.doi.org/*10.1371/journal.pone.0057831.*

Tankersley, M., & Cook, B. (Eds.) (2013). *Effective practices in special education*. Boston, MA: Pearson.

Armstrong, L., Phillips, J., & Saling, L. (2000). Potential determinants of heavier internet usage. *International Journal of Human–Computer Studies, 53*, 537–550. http://dx.doi.org/10.1006/ijhc.2000.0400

Bayraktar, F., & Gun, Z. (2007). Incidence and correlates of internet usage among adolescents in North Cyprus. *Cyber Psychology & Behavior, 10*(2), 191–197. http://dx.doi.org/10.1089/cpb.2006.9969

Davis, R., Flett, G., & Besser, A. (2002). Validation of a new scale for measuring problematic Internet use: Implications for pre-employment screening. *Cyber Psychology and Behavior, 5*(4), 331–345. http://dx.doi.org/10.1089/109493102760275581

Doran, B. (2006). Effect of cybernetic space on identity. PHD Thesis's. Field of Sociology. University of Teacher Training Tehran.

Douglas, A., Mills, J., Niang, M., Stepchenkova, S., Byund, S., & Ruffini, C. (2008). Internet addiction: Meta-synthesis of qualitative research for the decade 1996-2006. *Computers in Human Behavior, 24*(6), 3027-44. http://dx.doi.org/10.1016/j.chb.2008.05.009

Douglas, A. C., Mills, J. E., Niang, M., Stepchenkova, S., Byun, S., Ruffini, C., … Blanton, M. (2008). Internet addiction: Meta-synthesis of qualitative research for the decade 1996–2006. *Computers in Human Behavior, 24*, 3027–3044. http://dx.doi.org/10.1016/j.chb.2008.05.009

Eitel, D. R., Yankowitz, J., & Ely, J. W. (1998). Use of internet technology by obstetricians and family physicians. *JAMA, 280*(15), 1306-7. http://dx.doi.org/10.1001/jama.280.15.1306

Harjy, U. et al. (2005). Social skills in interpersonal communication. (Translation Khashayar Beigi and Merdad Firuz bakht). Tehran: growth.

Hur, M. (2006). Demographic, habitual, and socioeconomic determinants of Internet addiction disorder: An empirical study of Korean teenagers. *Cyber Psychology and Behavior, 9*(5), 514–525. http://dx.doi.org/10.1089/cpb.2006.9.514

Iskender, M., Akin, A. (2010). Social self-efficacy, academic locus of control, and internet addiction. Computers & Education, 54, 1101–1106. http://dx.doi.org/10.1016/j.compedu.2009.10.014

Kestenbaum, M., Robinson, G. P., Neustadtl, A., & Alvarez, A. (2002). Information technology and social time displacement. *IT& society, 1*(1), 21-37 Retrieved from http://www.Itand society.org

Kiesler, S., Kraut, R., Cummings, J., Boneva, B., Helgeson, V., & Crawford, A. (2002). Internet evolution and social impact. *IT& society, 1*(1), 120 – 134.

Kimberly, Y. (2007). *Internet Addiction Test (IAT).* Retrieved from http://www. netaddiction. com/ index.php

Ko, C. H., Yen, J. Y., Chen, C. C., Chen, S. H., & Yen, C. F. (2005). Gender differences and related factors affecting online gaming addiction among Taiwanese adolescents. *Journal of Nervous and Mental Disease, 193*(4), 273–277. http://dx.doi.org/10.1097/01.nmd.0000158373.85150.57

Ko, C., Yen, J., Chen, C., Chen, S., Wu, K., & Yen, C. (2006). Tri-dimensional personality of adolescents with internet addiction and substance use experience. *Canadian Journal of Psychiatry, 5*1(14), 887–894.

Kooraki, M., Yazdkhasti F., Ebrahimi A., Oreizi H. R. (2012). Effectiveness of Psychodrama in Improving Social Skills and Reducing Internet Addiction in Female Students. *Iranian Journal of Psychiatry and Clinical Psychology, 17*(4), 279-288.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being. *American Psychologist, 53*(9), 1017–1031. http://dx.doi.org/10.1037/0003-066X.53.9.1017

Kubey, R., Lavin, M., & Barrows, J. (2001). Internet use and collegiate academic performance decrements: Early findings. *Journal of Communication, 51*, 366–382. http://dx.doi.org/10.1111/j.1460-2466.2001.tb02885.x

Lazinger, S.S., Bar-Ilan, J., Peritz, B.C. (1997). Internet use by faculty members in various disciplines: a comparative case study. *JASIS, 48*(6), 508-18. http://dx.doi.org/10.1002/(SICI)1097-4571(199706)48:6<508::AID-ASI4>3.0.CO;2-Y

Lin, S. S. J., & Tsai, C. C. (2002). Sensation seeking and internet dependence of Taiwanese high school adolescents. *Computers in Human Behavior, 18,* 411–426. http://dx.doi.org/10.1016/S0747-5632(01)00056-5

Mc Kenna, K. Y. A., & Bargh, J. A. (2004). Plan9: from cyberspace. The implication of the internet for personality and social psychology. *Personality and social psychology review, 4*(1), 57-75. http://dx.doi.org/10.1207/S15327957PSPR0401_6

Mc Quail, D. (2005). Mc Quail`s mass communication theory, 5th edition. Sage publications. London.

Mohagheghzadeh, M. S., & Abdollahi, M. (2002). Scopes and using type of internet center members of Shiraz University of Medical Sciences about center facilities and its impact on research works of them. *Communication sciences, 18*(1, 2), 1-10.

Mohammad-beigi, A., Ghazavi, A., Mohammad-salehi, N., Ghamari, F., & Saeedi, A. (2010). Effect of internet addiction on educational status of Arak University of medical sciences students, spring 2009. *Arak Medical University Journal (AMUJ), 12*(4, Supp 1), 95-102.

Momayezi, F., Abdi Zarrin S., Eglima M., Raheb Gh. (2011). Social Skills: The Role of this Skills in Prevention of Adolescents Abuse. *Journal of Tolooe Behdasht, 10*(1), 96-108.

Moody, E. J. (2004). Internet use and its relationship to lineless. *Cyber Psychology & Behavior, 4*(3), 393-401. http://dx.doi.org/10.1089/109493101300210303

Nadi MA, & Sajjadian I. (2010). Path analysis of relationship between personality traits and internet addiction with quality of life of internet users in Isfahan city. *Journal of Behavioral Sciences, 8*(1), 34-45.

Nalwa, K., & Anand, A. (2003). Internet addiction in students: A cause of concern. *CyberPsychology & Behavior, 6*(6), 653–656. http://dx.doi.org/10.1089/109493103322725441

Ozcan, N. K., & Buzlu, S. (2007). Internet use and its relation with the psychosocial situation for a sample of university students. *Cyber Psychology & Behavior, 10*, 767–772. http://dx.doi.org/10.1089/cpb.2007.9953

Preece, J. (2000). Online communities: Designing usability, supporting sociability. Chichester: Wiley.

Sadeghian, E. (2008). Effect of computer and internet on children and teenagers. *Iran doc Scientific Communication Monthly Journal, 4*(4), 78-86.

Salkovic-Petrisic M, Mrzljak A, Lackovic Z. (2001). Usage of the internet pharmacology resources among European Pharmacologists: a preliminary investigation. *Fundam Clin Pharmacol, 15*(1), 55-60. http://dx.doi.org/10.1046/j.1472-8206.2001.00002.x

Schleyer, T.K., Spallek, H., Torres-Urquidy, M.H. (1998). A profile of current internet users in dentistry. *J Am*

Dent Assoc, 129(12), 1748-53. http://dx.doi.org/10.14219/jada.archive.1998.0146

Segrin, C., & Taylor, M. (2007). Positive interpersonal relationships mediate the association between social skills and psychological well-being. *Personality and Individual Differences, 43,* 637-646. http://dx.doi.org/10.1016/j.paid.2007.01.017

Shapira, N., Lessig, M., Goldsmith, T., Szabo, S., Lazoritz, M., & Gold, M. (2003). Problematic Internet use: Proposed classification and diagnostic criteria. Depression and Anxiety. *Computers in Human Behavior, 17*(4), 207-16. http://dx.doi.org/10.1002/da.10094

Shoa Kazemi, M. (2008). A comparison of chat behavior and its relation with self- concept university student (s) academic year from 2006 - 2007. *Journal of Psychological Studies, 4*(2), 27 - 36.

Teodoro M. L. & Kappler Ch, Rodrigues J. L, et al. (2005). The Matson Evaluation of Social Skill with Youngsters, (MESSY) & its Adaptation for Brazilian Children & Adolescents. *Interamerican Journal of Psychology, 39*(22), 239-246.

Weinstein, A., & Lejoyeux, M. (2010). Internet addiction or excessive internet use. *American Journal of Drug & Alcohol Abuse, 36*, 277-283. http://dx.doi.org/10.3109/00952990.2010.491880

Wellman, B., & Gulia, M. (1999). Net surfers don't ride alone: Virtual communities as communities. In P. Kollock & M. Smith (Eds.), *Communities and Cyberspace*. New York: Routledge.

Whitty, M. T., & McLaughlin, D. (2007). Online recreation: The relationship between loneliness, Internet self-efficacy and the use of the internet for entertainment purpose. *Computers in Human Behavior, 23,* 1435-1446. http://dx.doi.org/10.1016/j.chb.2005.05.003

Yang, C. K. (2001). Sociopsychiatric characteristics of adolescents who use computers to excess. *Acta Psychiatric a Scandinavica, 104,* 217–222. http://dx.doi.org/10.1034/j.1600-0447.2001.00197.x

Yen, J., Ko, C., Yen, C., Wu, H., & Yang, M. (2007). The co morbid psychiatric symptoms of internet addiction: Attention deficit and hyperactivity disorder (ADHD), depression, social phobia, and hostility. *Journal of Adolescent Health, 41*, 93–98. http://dx.doi.org/10.1016/j.jadohealth.2007.02.002

Young, K. S., & Rogers, R. C. (1998). The relationship between depression and internet addiction. *Cyber Psychology & Behavior, 1*, 25–28. http://dx.doi.org/10.1089/cpb.1998.1.25

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Is Employability Orientation More Enhanced by Career Self-Efficacy or Leadership Attribute?

Neda Tiraieyari[1] & Jamaliah Abdul Hamid[2]

[1] Institute for Social Science Studies, Universiti Putra Malaysia, Malaysia

[2] Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia

Correspondence: Jamaliah Abdul Hamid, Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia. E-mail: aliah@upm.edu.my

Received: January 3, 2015 Accepted: January 19, 2015 Online Published: July 15, 2015

doi:10.5539/mas.v9n8p57 URL: http://dx.doi.org/10.5539/mas.v9n8p57

Abstract

This paper aim to determine the association between employability orientations, leadership attribute and career self- efficacy. This study further investigates the influences of employability orientation among Malaysian university students. Data were collected randomly from 711undergraduate students from five Malaysian public universities. The dependent variable of this study was employability orientation and the independent variables were career self- efficacy and leadership attribute. Supporting the hypotheses, the results of study showed career self- efficacy and leadership attribute was positively related to employability orientation. MLR results showed that variables selected for this study explained 43.1% of the variance in employability orientation. Leadership attribute proved a stronger predictor for employability orientation.

Keywords: employability orientation, career efficacy, leadership attribute, Malaysia, university students

1. Introduction

Recently focus is given on employability orientation in organizations. Organizations need to adopt flexibility to keep up with fast changes in technology and globalization. This requires new abilities of the employees (Legge, 1995). Highlighting on organizational flexibility, employability orientation is very important (Van Dam, 2004). Employees need to maintain their employability for the organization when changes in the work condition occur and these involve developing new knowledge and skills and changing task. On the other hand, Career Self –Efficacy has been shown to be a very beneficial for evaluating the career development process (Gainor, 2006). Self-efficacy has been applied well to career choice, job decision-making, and career indecision (Hackett, 1995). Employability reflects the perceived opportunity of getting new employment whereas self-efficacy carefully related to the self and an overall feeling of how to perform jobs. Leadership is another work related skill that is the ability to influence others to act in order to achieve a common goal. It helps team members to work productively together. Despite difference among the three concepts, a relationship may exist since individuals who have more leadership attribute and self- efficacy skills may have positive attitudes towards employability interventions.

Work relevant skills are important for students, especially where competition for employment is very high. To compete successfully, it is imperative for graduates to have more than subject-specific skills as employers seek a wide range of skills when recruiting graduate students (Deeley, 2014; Hinchliffe & Jolly, 2011). According to Deeley (2014), there is an emphasis on employability skills and a belief that it should be a main part of student's university experience. Overall, it is clear from the literature that employability is a complex concept covering more than a set of practical skills because it also involves students' development of intellectual skills, competencies and personal attributes (Gravells, 2010; Knight & Yorke, 2002; Knight, 2006). Employability, therefore, refers to more than practical skills. It also refers to being able to learn from new experiences, to apply new knowledge, and thus includes competencies, capabilities and attributes (Knight & Yorke, 2002).

Considering university students as the potential employees in future so there is essential for them to have employability orientation skill prior become an employee in any organization. To our knowledge, no study carried out to seek the relationship between employability orientations, career decision making self –efficacy and leadership attribute among university students in Malaysia. Thus, this study aims to seek the influence of career

decision making self –efficacy and leadership attribute on employability orientation of university students. Hence, the central question that this article seeks to address is whether employability orientation more enhanced by career self –efficacy or leadership attribute. To address this, we begin with a literature on employability orientation, leadership attribute and career efficacy.

2. Literature Review

2.1 Employability Orientation

Employability refers to establish the personal and professional capacity to increase one's employment potential (Smith, 2010). According to Kelan (2008), people may feel stressed to improve their skills and knowledge when they get new positions. However, within an unreliable economy it is necessary that employees constantly improve their skills to enhance their marketability. Therefore, people who work for incomes need consistently update their skills and search for jobs more frequently. Literature recommends that employability is more than getting a job. Employability orientation refers to the attitudes toward changes that aim to boost the organizational flexibility (Van Dam, 2004). Therefore, employability is a critical requirement for both organizations and individuals. It is a psycho-social construct that nurture adaptive cognition, behaviour, and increase the individual-work interface (Fugate, Kinicki, & Ashforth, 2004). According to Fugate, Kinicki, & Ashforth (2004) the concept of employability refers to work adaptability that help employees to identify career opportunities. As such, employability helps the movement among jobs and within organization. It actually increases an individual's chance of getting job. In summary, employability reveals the perceived probability of getting new employment (Berntson, Näswall, & Sverke, 2008; Forrier & Sels, 2003).

2.2 Career Decision Making Self –Efficacy

Self-efficacy is a self- perceived ability to perform a task (Rollins & Valdez, 2006). It related to the overall feeling of how to perform tasks (Bandura, 1997). It determines how much effort he or she will dedicate for the activity that chose to undertake, and how long will continue when confronted with challenges (Bandura, 1982). People with a high level of self-efficacy are not give up easily when face obstacles. Nevertheless, those with low level of self-efficacy tend to skip out on uncertain activities and easily discourage when challenges occur (Bandura, 1982). Scholars consider self-efficacy as significant parts of career choice, tenacity, career performance, and career development (Betz & Hackett, 1997). According to Choi et al. (2012) career self –efficacy is one of the most factors studied in career literature. Several studies using the social cognitive career theory framework have proved that career self –efficacy plays an important part in an individual's career development (Guay, Ratelle, Senécal, Larose, & Deschênes, 2006; Lease, 2006; Lent, 2001). Choi et al. (2012) suggested two areas of career self –efficacy: the content and the process. The content part refers to self-efficacy in some career area such as math or science; whereas the process refers to self-efficacy in using the required strategies for successfully navigating a decision-making process. The concept of career decision-making self-efficacy is considered a way of measuring self-efficacy with respect to career decision-making tasks and behaviour (Taylor & Betz, 1983; Taylor & Popma, 1990). McAuliffe (1992) suggested that career decision-making self-efficacy as an important instrument in recognizing personal emotional barriers and helping employees who are having trouble with the career decision-making process.

2.3 Leadership Attribute

The concept of leadership attributes is indeed an old one exist before the scientific study of leadership (Zaccaro, Kemp, & Bader, 2004). According to Machiavelli (1997), the main leadership attribute is the ability to understand social situations and to manipulate them in the practice of leadership. Several authors have written about the importance of character as a leadership attributes (Barlow, Jordan & Hendrix, 2003). Kirkpatick & Locke (1991) suggested that leadership character involves six elements such as; desire to lead, drive, honesty and integrity, self-confidence, cognitive ability, and knowledge of the business. Tait (1996) found that leadership character consisted of honesty, fairness, compassion, humility, and being one's own person. Similarly, Barker & Coy (2003) identified seven qualities such as humility, courage, integrity, compassion, humour, passion, and wisdom. (Cycyota et al. (2011) suggest that students' leadership skills are involved in the academic evaluations of classroom assignments such as participation in team, service learning processes, and other very relevant areas. Moreover, leadership is an essential part of the physical education evaluation with a certain emphasis on the capability to work in teams while maintaining personal performance.

2.4 Association among the Three Concepts

The literature review suggested that individual characteristics are important contributors of employees' attitudes toward career development activities. Two personality traits were considered important antecedents of

employability orientation namely openness and initiative. Openness defined as being open to new ideas and changes. Initiative defined as being self-started and proactive instead of being reactive. Researchers have revealed relationships between initiative and some career related variables, such as career planning, career initiative, and employability (Bateman & Crant, 1993; Frese, Fay, Hilburger, Leng & Tag, 1997; Seibert, Kraimer, & Crant, 2001). According to Berntson, Näswall, & Sverke (2008) employability refers to individuals' beliefs about their opportunities of getting new employment, it is accepted to associate it to self-efficacy. Similarly, Knight & Yorke (2002) mentioned that efficacy beliefs are an important aspect of employability. Several scholars suggested that employability is mostly depending on self-efficacy (Bandura, 1997; Velde & Berg, 2003). Additionally, self-efficacy has been shown to be related positively to job search behavior among unemployed (Kanfer, Wanberg, & Kantrowitz, 2001; Moynihan, Roehling, LePine & Boswell, 2003). They are more likely to change jobs, which changing the job would affect their level of self-efficacy positively (Berntson, Näswall & Sverke, 2008). Similarly, result of another study showed that self-efficacy positively related to employability. It appeared that employees who have higher level of self-efficacy perceived themselves as highly employable (Dacre Pool & Qualter, 2013). Finding of study by Berntson et al. (2008) reveal that although employability and self-efficacy are two separate concepts but separate qualities. Furthermore their data supported the opinion that employability causes self-efficacy, rather than the reverse relationship that reported by scholars in past studies (Bandura, 1997; Velde & Berg, 2003). Finding of study conducted by (Nauta, Kahn, Angell, & Cantarelli, 2002) revealed that relationship between employability and self- efficacy is mutual. In relation to leadership attribute according to Camps & Rodríguez (2011) there is a connection between leadership and the employability.

In line with these findings, self- efficacy and leadership attribute expected to positively related to employability orientation (Hypothesis 1 & 2).

The following hypotheses are formulated:

Hypothesis 1: There will be a positive correlation between employability and career self-efficacy.

Hypothesis 2: There will be a positive correlation between employability and leadership attribute.

3. Objectives of Study

1) To determine respondents' employability orientation, career self -efficacy and leadership attribute

2) To determine the relationship between variables

3) To determine those variables that help to explain variation of employability ordination

4. Method

4.1 Participants

Public universities namely; USM Penang, UPM Selangor, UTM, UMT, and UPNM were randomly selected from four zones of peninsular Malaysia. Data were collected using a questionnaire that was distributed among 800 respondents. A pool of 764 questionnaires were returned (response rate 95.5%) and 53 questionnaires were excluded from the analysis due to incomplete data (N=711).

4.2 Validity and Reliability of the Instrument

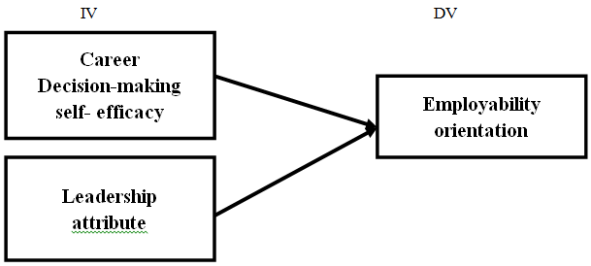
A panel of experts at Universiti Putra Malaysia validated the instrument. Reliability analysis was conducted to determine the reliability of the questionnaire. The pilot-tested conducted at one public university prior to conduct the actual study. Following pilot-test, some items were modified. The results of reliability statistics for variables were all above 0.70.

4.3 Measuring Variables

Employability: the dependent variable was measured with a 6-item scale. Respondents were asked to mark their degree of agreement with each statement on a Likert scale ranging from 1 (‘‘strongly disagree’’) to 5 (‘‘strongly agree’’).

Career decision-making self-efficacy: the independent variable of study was assessed using three dimensions namely; seeking occupational information, goal setting, and self- appraisal. Again, the response scale ranged from 1 (‘‘strongly disagree’’) to 5 (‘‘strongly agree’’).

Leadership attribute: was measured by two sub domain; acceptance of leadership and visionary and change leadership with eight statements on a Likert scale ranging from 1 (‘‘strongly disagree’’) to 5 (‘‘strongly agree’’). Leadership attribute is the second independent variables of the study.



5. Results

5.1 Descriptive Analysis of Variables (Objective1)

Means and standard deviations for employability, career self- efficacy, and leadership attribute are listed in Table 1. For the employability orientation mean score was (M=4.37, SD=0.53), for leadership attribute, the mean score of (M=4.01, SD=.48), and for career decision-making self-efficacy (M=4.08, SD=0.511). The overall mean of response on employability orientation, leadership attribute and career self- efficacy divided into three levels for reporting purpose. Majority of the respondents, 48.9% (n=632) had high level of employability, 5.6% (n=76) had moderate of employability and only 2 respondents (0.2%) reported low level of employability. Regarding leadership attribute, 55 % (n=711) of students had high level of Leadership attribute while 12.2% (n=158) had moderate level and only one (% 0.1) reported low level of leadership attribute. For the career self -efficacy (n=42.7) % 552 reported high level of self- efficacy, %12.2 (n=158) had medium level and only 1 student (% 0.1) reported low level of career self- efficacy.

Table 1. Means and standard deviations for variables of study (N=711)

Variables	N	M	SD
Employability Orientation	711	4.37	.539
Leadership attribute	711	4.01	.483
Career Decision-making self-Efficacy	711	4.08	.511

5.2 Correlation Coefficient Analysis (Objective2)

To measure the relationship among the variables of study, the Pearson correlation coefficient was used. The statistical analyses were conducted using IBM® SPSS® Statistics 20. The results show that employability orientation is positively related to leadership attribute and career decision-making self- efficacy. There is a significant positive relationship between employability orientation and leadership attribute (r=0.646, p=0.001). There is a significant positive relationship between employability orientation and career decision-making self-efficacy (r=0.512, p=0.001). The highest correlation coefficient was .646 between leadership attribute and employability orientation, followed by career decision-making self-efficacy and employability orientation (r=.512, p=.000). (See table 2).

Table 2. Pearson correlations coefficient of independent and dependent variables

Variables	**Y**	**X1**	**X2**
Employability Orientation	1.000		
Leadership attribute(X1)	.646**	1.000	
Career-Decision-making self-Efficacy(X2)	.512**	.655**	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

5.3 Regression Analysis (Objective3)

A regression analysis was conducted to determine the predictor variables of employability orientation. We applied MLR to assess two-factor regression model explaining the variation of employability orientation.

The equation proposed by the MLR is as follows:

Y= 1.271+ .606 (X1) + .165 (X2) + e

According to this method, two variables found to be significant in explaining employability orientation. The predictor variable is leadership attribute (t=14.47, p=0.000) and career decision-making self- efficacy (t=4.155, p=0.000). As depicted in table 2, the largest beta coefficient is (.543), which is for leadership attribute followed by career decision-making self- efficacy (.156). The R2 value of 0.431 implies that two predictors explain 43.1% of the variances in the employability orientation. In other words, 43.1% of the variability of employability orientation is explained by leadership attribute and career decision -making self-efficacy. (See table 3).

Table 3. Estimates of coefficients for the regression model

Coefficients[a]					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	1.271	.138		9.213	.000
Leadership attribute	.606	.042	.543	14.473	.000
CareerDecision-making self-Efficacy	.165	.040	.156	4.155	.000

R=0.656; R2=0.431; Adjusted R^2 = 0.429

6. Discussions and Conclusion

Career decision -making self- efficacy involves planning and seeking information about one's career, and making important decisions regarding career options. Positive levels of ability to plan and make decisions about career will likely also encourage a developed level of awareness of the need to acquire and enhance more or higher skills and knowledge in order to gain employment or to advance further in career in one's choice of employment. The more one has skills in seeking information and planning one's career, the more likely one would demonstrate the ability and commitment to invest in up scaling of skills and knowledge to also advance in one's career or to remain employed. Hence, it makes sense that career decision-making skills have quite a strong relationship with employability orientation. Leadership attributes on the other hand demonstrates self-confidence in one's ability to take on responsibilities, and to work with others to create and achieve visions and goals. Formulating visions of what may be achieved, and strategically planning how to achieve those goals indicate high levels of awareness of the importance of team cohesion which effectively deploys all the different types of knowledge and skills to achieve the goals. Leadership attributes, namely visionary leadership, supports the principle that people are constantly trying to improve their lives through better work, and better careers. Hence, the strong relationship between leadership attributes and employability orientation.

In terms of career preparation programs in universities, it is recommended that educators develop an integrated program to enhance the employability orientation of their undergraduates that assimilate leadership development within career decision activity. By helping undergraduates to focus on what they hope to achieve for themselves in terms of transforming their individual identity, their personal capacity, their social responsibility, and the social status and they might aspire towards in the future through the work or career they choose, universities will then become more successful in enhancing the meaning of work itself as a transformative agent or medium to achieve fulfilment in the development of the individual's identity, vision and capacity. Only then will the purpose of career decision making and developing positive employability orientation go beyond just securing financial stability. Instead the integrated employment orientation program will offer a blended medium through which undergraduates seek to find, and in turn bring forth value, to the meaning of careers, and of being employed in their future lives.

Acknowledgment

We would like to thank Universiti Putra Malaysia to grant this study.

References

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122. Retrieved from http://psycnet.apa.org/journals/amp/37/2/122. http://dx.doi.org/10.1037/0003-066X.37.2.122

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Barker, C., & Coy, R. (2003). The 7 heavenly virtues of leadership. McGraw-Hill.

Barlow, C. B., Jordan, M., & Hendrix, W. H. (2003). Character Assessment: An Examination of Leadership

Levels. *Journal of Business and Psychology, 17*(4), 563–584. http://dx.doi.org/10.1023/A:1023408403204

Bateman, T. S., & Crant, J. M. (1993). The proactive component of organizational behavior: A measure and correlates. *Journal of Organizational Behavior, 14*(2), 103–118. Retrieved from http://onlinelibrary.wiley.com http://dx.doi.org/10.1002/job.4030140202

Berntson, E., Näswall, K., & Sverke, M. (2008). Investigating the relationship between employability and self-efficacy: A cross-lagged analysis. *European Journal of Work and Organizational Psychology, 17*(4), 413–425. http://dx.doi.org/10.1080/13594320801969699

Betz, N. E., & Hackett, G. (1997). Applications of self-efficacy theory to the career assessment of women. *Journal of Career Assessment, 5*(4), 383–402. Retrieved from http://jca.sagepub.com/content/5/4/383.short. http://dx.doi.org/10.1177/106907279700500402.

Camps, J., & Rodríguez, H. (2011). Transformational leadership, learning, and employability. *Personnel Review, 40*(4), 423–442. http://dx.doi.org/10.1108/00483481111133327

Choi, B. Y., Park, H., Yang, E., Lee, S. K., Lee, Y., & Lee, S. M. (2012). Understanding Career Decision Self-Efficacy A Meta-Analytic Approach. *Journal of Career Development, 39*(5), 443–460. http://dx.doi.org/10.1177/0894845311398042

Cycyota, C. S., Ferrante, C. J., Green, S. G., Heppard, K. A., & Karolick, D. M. (2011). Leaders of Character: The USAFA Approach to Ethics Education and Leadership Development. *Journal of Academic Ethics, 9*(3), 177–192. http://dx.doi.org/10.1007/s10805-011-9138-z

Dacre Pool, L., & Qualter, P. (2013). Emotional self-efficacy, graduate employability, and career satisfaction: Testing the associations. *Australian Journal of Psychology, 65*(4), 214–223. http://dx.doi.org/10.1111/ajpy.12023

Deeley, S. J. (2014). Summative co-assessment: A deep learning approach to enhancing employability skills and attributes. *Active Learning in Higher Education, 15*(1), 39–51. http://dx.doi.org/10.1177/1469787413514649

Forrier, A., & Sels, L. (2003). The concept employability: a complex mosaic. International Journal of Human *Resources Development and Management, 3*(2), 102–124. http://dx.doi.org/10.1504/IJHRDM.2003.002414.

Frese, M., Fay, D., Hilburger, T., Leng, K., & Tag, A. (1997). The concept of personal initiative: Operationalization, reliability and validity in two German samples. *Journal of Occupational and Organizational Psychology, 70*(2), 139–161. http://dx.doi.org/10.1111/j.2044-8325.1997.tb00639.x

Fugate, M., Kinicki, A. J., & Ashforth, B. E. (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior, 65*(1), 14–38. http://dx.doi.org/10.1080/09585192.2015.1004100.

Gainor, K. A. (2006). Twenty-Five Years of Self-Efficacy in Career Assessment and Practice. *Journal of Career Assessment, 14*(1), 161–178. http://dx.doi.org/10.1177/1069072705282435

Gravells, A. (2010). Delivering employability skills in the lifelong learning sector. SAGE. Retrieved from http://books.google.com.my/books?hl=en&lr=&id=CXEm5d6PfKUC&oi=fnd&pg=PT6&dq=Delivering+Employability+Skills+in+the+Lifelong+Learning+Sector&ots=rW8e7FvmfW&sig=vhklyv1CQ6HZeOVoaurV4kIlEmI

Guay, F., Ratelle, C. F., Senécal, C., Larose, S., & Deschênes, A. (2006). Distinguishing Developmental From Chronic Career Indecision: Self-Efficacy, Autonomy, and Social Support. *Journal of Career Assessment, 14*(2), 235–251. http://dx.doi.org/10.1177/1069072705283975

Hackett, G. (1995). Self-efficacy in career choice and development. In Self-efficacy in Changing Societies. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511527692.010

Hinchliffe, G. W., & Jolly, A. (2011). Graduate identity and employability. *British Educational Research Journal, 37*(4), 563–584. http://dx.doi.org/10.1080/01411926.2010.482200

Kanfer, R., Wanberg, C. R., & Kantrowitz, T. M. (2001). Job search and employment: A personality–motivational analysis and meta-analytic review. *Journal of Applied Psychology, 86*(5), 837. http://dx.doi.org/10.1177/1069072714535031.

Kelan, E. (2008). Gender, risk and employment insecurity: The masculine breadwinner subtext. *Human Relations, 61*(9), 1171–1202. http://dx.doi.org/10.1177/0018726708094909

Kirkpatick, S. A., & Locke, E. A. (1991). Leadership: do .traits matter? *The Executive, 5*(2), 48–60. Retrieved from http://amp.aom.org/content/5/2/48

Knight, P. (2006). Assessing complex achievements. Beyond Mass Higher Education: Building on Experience, 96–104. Retrieved from http://books.google.com.my/books?hl=en&lr=&id=Oy_ZlqdIAlgC&oi=fnd&pg=PA96&dq=Assessing+complex+achievement&ots=FOasAsH6yN&sig=nGG_AzpTwyv3OK5alFZmNNa9ZTo

Knight, P. T., & Yorke, M. (2002). Employability through the curriculum. *Tertiary Education and Management, 8*(4), 261–276. http://dx.doi.org/10.1080/13583883.2002.9967084

Lease, S. H. (2006). Factors Predictive of the Range of Occupations Considered by African American Juniors and Seniors in High School. *Journal of Career Development, 32*(4), 333–350. http://dx.doi.org/10.1177/0894845305283003

Legge, K. (1995). Human resource management: rhetorics and realities (Management, work and organisations). London.

Machiavelli, N. (1997). The prince. Yale University Press.

McAuliffe, G. J. (1992). Assessing and changing career decision-making self-efficacy expectations. *Journal of Career Development, 19*(1), 25–36. Retrieved from http://link.springer.com/article/10.1007/BF01323002.http://dx.doi.org/10.1007/BF01323002

Moynihan, L. M., Roehling, M. V., LePine, M. A., & Boswell, W. R. (2003). A longitudinal study of the relationships among job search self-efficacy, job interviews, and employment outcomes. *Journal of Business and Psychology, 18*(2), 207–233. http://dx.doi.org/10.1023/A:1027349115277

Nauta, M. M., Kahn, J. H., Angell, J. W., & Cantarelli, E. A. (2002). Identifying the antecedent in the relation between career interests and self-efficacy: Is it one, the other, or both? *Journal of Counseling Psychology, 49*(3), 290. http://dx.doi.org/10.1037/0022-0167.49.3.290.

Rollins, V. B., & Valdez, J. N. (2006). Perceived Racism and Career Self-Efficacy in African American Adolescents. *Journal of Black Psychology, 32*(2), 176–198. http://dx.doi.org/10.1177/0095798406287109

Seibert, S. E., Kraimer, M. L., & Crant, J. M. (2001). What do proactive people do? A longitudinal model linking proactive personality and career success. *Personnel Psychology, 54*(4), 845–874. http://dx.doi.org/10.1111/j.1744-6570.2001.tb00234.x

Smith, V. (2010). Review article: Enhancing employability: Human, cultural, and social capital in an era of turbulent unpredictability. *Human Relations, 63*(2), 279–300. http://dx.doi.org/10.1177/0018726709353639

Tait, R. (1996). The attributes of leadership. *Leadership & Organization Development Journal, 17*(1), 27–31. Retrieved from http://www.emeraldinsight.com/journals.htm?articleid=1410408&show=abstract

Taylor, K. M., & Betz, N. E. (1983). Applications of self-efficacy theory to the understanding and treatment of career indecision. *Journal of Vocational Behavior, 22*(1), 63–81. Retrieved from http://www.sciencedirect.com/science/article/pii/0001879183900064

Taylor, K. M., & Popma, J. (1990). An examination of the relationships among career decision-making self-efficacy, career salience, locus of control, and vocational indecision. *Journal of Vocational Behavior, 37*(1), 17–31. Retrieved from http://www.sciencedirect.com/science/article/pii/000187919090004L

Van Dam, K. (2004). Antecedents and consequences of employability orientation. *European Journal of Work and Organizational Psychology, 13*(1), 29–51. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13594320344000237

Velde, M., & Berg, P. (2003). Managing functional flexibility in a passenger transport firm. *Human Resource Management Journal, 13*(4), 45–55. http://dx.doi.org/10.1111/j.1748-8583.2003.tb00104.x

Zaccaro, S. J., Kemp, C., & Bader, P. (2004). Leader traits and attributes. The Nature of Leadership, 101, 124.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Regional Factors in Boosting the Efficiency of Inviting Investments in Entrepreneurial Activity

Mansoor Maitah[1], Almatova D.S.[2], Kholnazar Amonov[3] & Luboš Smutka[1]

[1] Czech University of Life Sciences Prague, Prague 6, Czech Republic

[2] Tashkent Financial Institute, Uzbekistan

[3] Central Bohemia University, o.p.s., Czech Republic

Correspondence: Mansoor Maitah, Faculty of Economics and Management, Czech University of Life Sciences Prague, Prague, ON., Kamýcká 129, 165 21 Praha 6 - Suchdol, Czech Republic. Tel: 420-23438-2139. E-mail: maitah@ pef.czu.cz

Received: January 22, 2015 Accepted: March 13, 2015 Online Published: July 20, 2015

doi:10.5539/mas.v9n8p64 URL: http://dx.doi.org/10.5539/mas.v9n8p64

Abstract

This article is devoted to analyze the current condition of elaboration of regional investment projects at the present-day stage of economic reforms being led in Uzbekistan. Article sets forward proposals on working out prospective and present-day programs on inviting international financial institutions, foreign financial and insurance organizations and donor countries in order to implement with them investment projects of high importance for Uzbekistan. Practical recommendations on working out the mechanisms for implementation and efficiency assessment of investment projects, investment projects portfolio revolving to choose high priority investments were also suggested.

Keywords: small business, investment processes, entrepreneurial activity, regional economy, industry-specific and region-specific features of investments, investment attractiveness

1. Introduction

Overall review of investment processes development in Uzbekistan and economy of the regions, and inviting investments in entrepreneurial activity. Today enactment of a variety of regulations aimed at boosting investment climate attractiveness of the regions and intensifying investment processes is ensuring the growth of the country's investment potential and increase of foreign investments year after year, and thus, the growth of county's macroeconomic indicators, including, its gross domestic product volume. During the period of 2005-2013 in Uzbekistan, gross domestic product has increased at least 8.5 percent every year. Compared to the year of 2000, GDP volume grew by 2.9-fold, industrial production increased by 2.6-fold, volume of investments in the economy increased by 3.4-fold, particularly, volume of foreign direct investments increased by 20-fold. Small business sector, being a significant part of country's economy, is serving as a main source in supplying domestic market with necessary goods and services. Small business not only bridges some gaps in the economy, but also plays a vital role in diversification and development of the economy at stable paces. Uzbekistan is attaching great importance to building an institutional and market infrastructure which supports the development of small businesses and private enterprises. Since 2012, for the purpose of radical reduction of government intervention in the activity of business entities, 80 procedures related to various kind of permissions, as well as, licensing requirement of 15 types of entrepreneurship activities have been cancelled; also, number of financial statement submitting forms and periods of submitting them have been reduced by 1.5-fold. In addition, new regulations introduced as of January 1, 2013, which canceled 65 statistical report forms and 6 tax report forms, and decreased the number of periods for submitting these reports by more than 2-fold.

2. Literature Review

Uzbekistan established an economic and legal base aimed at organizing investment processes and investment activity as per the market economy principles. But an investment climate, which promotes investment attracting in the economy of the regions, including, in small businesses, has not been fully created yet. This requires working out effective strategies aimed at boosting the volume of investments to be invited in the economy of the regions.

In this view, development of investment activity of small businesses plays an important role in attaining key objectives directed to stable development of production potential of the regions' economy based on the priority directions of innovative renovation. Primary purpose of research activities in this topic is to study and determine the conditions and factors which opening the doors to development of investment activity of small enterprises in the present-day context, as well as, to determine the real prospects of investment cooperation between small and large enterprises.

International experience and theoretical and research conclusions confirm that entrepreneurial activity is the power that eliminate problems in the economy, and investment policy serves as its main source and factor. In turn, flexibility of production for market conditions, import substitution and export-orienteers of production is ensured by coordination of investment policy directions, fields, and sources. Rational investment policy increases labor and capital productivity, opens doors to satisfying the needs of consumers with high quality products.

3. Method

Our research studies indicate that after setting regional policy on small business development, often, it will be possible to get the results of implementation of this policy only after a long time, when it has been late to change something.

Therefore, we think that it is necessary to model various variants by means of economic and mathematical models in order to avoid negative consequences of made decisions.

Many models are built on a basis of reflecting small enterprises in the form of R. Stone's production function (Leontyev function):

$$P(t) = min \{F(t)/a, C(t)/b, L(t)/\theta\}$$

where P(t) – production of goods and services by small enterprises (in physical volume);

$F(t)$ – fixed assets, $C(t)$ – current capital, $L(t)$ – labor resources;

a, b, θ – expenditure rate of related production factors per production unit $P(t)$ (soums);

t – time (simulation step).

Reflecting a small enterprise in this way is convenient; moreover, production function also includes the employment as its structural component. But, in practice, it is impossible to build models for every established enterprise. In modeling the regional level decision making, taking regional small business sector as a single object makes good results. For this, it is necessary to use simulation modeling methods.

As a typical problem to be solved in process of supporting small businesses in a region, we can take the issue of financing some small enterprises chosen in a particular sector in a context of limited financial resources. In general, this model looks like as follows:

$$INV = \sum_{i=1}^{n} inv_i \quad (1)$$

$$PR = \sum_{i=1}^{n} rev_i \; - \; INR \; - \; NR \quad (2)$$

$$n = q + k + l + \ldots + m, \text{ here} \quad (3)$$

All investments are directed to financing small enterprises (equation 1). Profits from investments (equation 2) equals to the amount of money resources returned from each enterprise, minus invested investments, minus non-return of investments. Finally, regional investment policy (equation 3) is determined by allocating the number of enterprises for per investment direction. If an average number of employees of the enterprises in each investment direction is known, we can assess an additional labor market capacity emerging as a result of implementation of programs on supporting small businesses.

$$\text{Tr} = \sum_{j=1}^{q} tr_j + \sum_{j=1}^{k} tr_j + \ldots + \sum_{j=1}^{m} tr_j \quad (4)$$

The modeling principles that reflected by the 1–4 equations indicated above were used in modeling the results of various region-wide measures taken to develop small business sector. The first model is a simulation model for the development of tourism sector in the regions of Uzbekistan.

Main elements constituting the tourism sector are the following systems:

Accommodation system - major part of this system is, usually, consisted of hotels, and this complex also includes some small family passionate and rented private houses;

Catering system– it includes big restaurants, small cafeterias, fast food points (tents), retail stores selling ice-cream, bakery and other products;

T*ransportation system* – presently, private sector is dominating in this system, and they are providing passenger transportation services by automobiles, minibuses, and even by large buses;

Entertainment services system – small artistic troupes, attractions, private art galleries, museums play a big role in this system;

Retail sales system – in this system, retails sales are run mostly by small enterprises and private entrepreneurs in majority of the regions;

Souvenirs complex – it includes folk crafts and souvenir goods producing enterprises. Services, such as laundry, dry cleaning etc. which don't directly relate to tourism, but make profit from it, are modeled separately.

4. Results and Discussion

During this research studies, we determined the correlation between the indicators reflecting the investment and economic activity results of an enterprise.

Regression equation is statistically important in our studies. Because determination coefficient has 45.6 % value (table 1), and it reflects a close relationship between dependent variables (investments) and independent variables (profit, average number of workers, arable land areas, production expenditures, fixed assets):

$$У=2169.7+1.8X3+55.1X2+0.04X6 \quad (5)$$

Parameters of regression equation show a direct proportional relationship of production factors which lead to growth of investments.

As volume of investments per farm business increases, amount of workforce wages and main indicators of production in an enterprise also will increase.

One of the important indicators to be used in the assessment of investments efficiency is a recovery of production costs. In addition, as we study investment attractiveness of agricultural enterprises, we determined that several production indicators have a great influence on the activity of enterprises.

We determined that area of agricultural lands, value of material current assets per hectare, number of workers per 100 hectares of agricultural land, amount of average yearly salary of workers have an influence on productivity of a small business's activity. Increase of agricultural land areas at all types agricultural enterprises has a positive effect on productivity of agricultural business, but this growth rate is not the same at all enterprises. As land areas increase by 100 ha, recovery of production costs in farms increases by 0.01 thousand uzbek soums. We can see that in a regression equation and coefficient of correlation. Regression equation for livestock farming units looks like as follows: У=0.563+0.0001x, linear coefficient value (R=0.492), correlation ratio is moderate. Regression equation for arable farming units is У=0.920+0.0001x, the linear coefficient value (R=0.552), and correlation ratio is moderate.

When we studied the dependency of recovery amount of production costs on the number of workers per ha land area, it is determined that if the number of workers per 100 ha land area increases by 1 worker, recovery of production costs increases by 0.208 thousand uzbek soums. And the regression equation for this looks like as follows:

У=0.180+0.208x and R=0.844 (correlation ratio is high).

Table 1. Multifactorial correlation and regression analyses of the dependency of investments on several selected indicators

Parameters	Value of parameters	Standard error	Statistical error	Level of reliance
Free coefficient	2169.57	1022.13	2.122	0.0348
X1-X2	1. 7863	0.3683	4.85	0
X3-X1	55.067	6.192	8.894	0
X4-X5	0.037	0.120	0.311	0.05

Source: own calculation.

Determination coefficient (D) = 45.65%. Explanations of reference characters:

У1 – investments, mln uzbek soums; X1 – fixed assets, mln uzbek soums; X2 – average yearly number of workers, quantity; X3 – planted acres, ha; X4 – production costs, mln uzbek soums; X5– profits, mln uzbek soums.

Correlation ratio between recovery of costs and investments is high R=0.855 and the regression equation will be as follows:

$$У=0.231 + 0.217x \quad (6)$$

If volume of investments per 100 ha land areas increase by thousand uzbek soums, recovery of production costs increases by 0.237 uzbek soums. In the process of this research studies, we studied the investment inviting processes in entrepreneurial activity in a region, and also discovered several aspects of labor market by analyzing the distribution of employment level. Thus, small business is a complicated social and economic institution, and it should be studied at a full scale. For this purpose, we will build an integral indicator that reflects a development level of entrepreneurship in the regions.

The indicators reflecting the development level of small business in the regions serve as a data base to be used in building this integral indicator.

- Number of small enterprises registered in a region (SE);
- Employment of population in small business sector (EP). This index, at the same time, includes three indicators, accounting of which is maintained at the state statistics: number of workers in SE, number of workers having a second job in SE, and number of workers employed in SE on a contract basis;
- Volume of production (goods and services) produced by SE (Pr);
- Volume of investments in small business (In).

We assume that all indicators have equal influence on the development level of small business in a region, and thus, these indicators will be included into integral indicator with the assumption that all of these indicators carry an equal weight. The formula for calculation of the integral indicator that reflects the development level of small business in the region will look like as follows:

$$I_j^{SBdev} = \left(\frac{SE_j}{\max SE} + \frac{EP_j}{\max EP} + \frac{\Pr_j}{\max \Pr} + \frac{In_j}{\max In}\right)/4 =$$

$$= (SE_j + EP_j + \Pr_j + In_j)/4, \quad (7)$$

here I_j^{SBdev} – an integral indicator reflecting a development level of entrepreneurship in the j region, $1 < j \le 85$;

$\overline{SE}_j$, EP_j, $\Pr_j$, In_j - normed values of the indicators.

Table 6 shows the calculation results about regions and leaders in regional development of small business. Tashkent city has an absolute leadership in small business development level, and its integral indicator value is 0.863. But while Tashkent city had a maximal value in all four indicators in 2012 ($I_{Tash\_2012}^{SE} = 1$), now Tashkent region has been ahead of Tashkent city in production turnover ($I_{Tash.reg}^{SE} = 0{,}368$).

In the meantime, Tashkent region also secured the second place in the overall ranking, outstripping Andijan region ($I^{And} = 0.363$).

As was previous periods, Tashkent city is still ahead of the other regions in the ranking by the value of integral indicator that reflect a development level of small business, but the value of this difference has been decreased a bit, up to 0,495, since 2012 (that was 0.602 in 2012).

The top ten of the ranking was secured by the regions in which big cities are located. This situation is explained by the following facts: firstly, there is usually huge demand for goods and services produced in big cities, and

this creates favorable conditions for the development of small business sector; secondly, an infrastructure for supporting small business functions more efficiently in city areas. Now we discuss how much small business sector contributes to the development of the regions and we will compare the regions by the volume of this contribution. For this purpose, we will use the same method which we already used to analyze development level of small businesses in the regions when we built the integral indicator. We will use the following indicators to build an integral indicator that indicates the contribution of small enterprises to development of the regions:

- number of small enterprises per 1000 population of a region;
- share of employees employed in small business sector within the total number of employed population in a region;
- share of small business sector turnover within the total turnover of enterprises in a region;
- share of investments in fixed assets of small enterprises within the total volume of investments invested in a region.
- The formula for calculation of the integral indicator that reflects the contribution of small enterprises to development of a region will look like the one we saw above:

$$I_j^{SBcont} = \left(\frac{SBnum_j}{\max SBnum} + \frac{\Delta ShEmp_j}{\max \Delta ShEmp} + \frac{\Delta SBturn_j}{\max \Delta SBturn} + \frac{\Delta SBin_j}{\max \Delta SBin}\right) / 4 =$$

$$= \left(SBnum_j + \bar{\Delta} ShEmp_j + \bar{\Delta}\overline{SBturn}_j + \bar{\Delta} SBin_j\right) / 4, \quad (8)$$

where I^{SBcont}_{j} – an integral indicator reflecting the contribution of small businesses for development of the j region, $1 < j \leq 85$;

$SBnum_j + \bar{\Delta} ShEmp_j + \bar{\Delta}\overline{SBturn}_j + \bar{\Delta} SBin_j$ - changed (normed) values of the indicators. Contribution of small businesses in Uzbekistan to the economic growth of the country in 2007 was as follows: at average, the number of small enterprises per 1000 population was 12.9, and 51.3% of able-bodied population employed in small business sector, turnover of small business sector made up 31% of the total volume of the country's turnover, but investments in small business sector made up only 3,9%.

Table 2. Integral indicators showing the development level of small business sector in the regions and the contribution of small businesses to development of the regions

Territories and regions	Year of 2012					
	Development level of entrepreneurial activity in a region			Contribution of small businesses to development of a region		
	I^{SMBdev}	place	Place in the territory	I^{SMBdev}	place	Place in the territory
Tashkent territory	0.956	1	-	0.742	1	-
Fergana territory	0.598	2	-	0.628	2	-
Zarafshan territory	0.371	5	-	0.603	3	
Southern territory	0.390	3	-	0.602	4	-
Mirzachul territory	0.389	4	-	0.596	5	-
Lower Amudarya territory	0.234	6	-	0.427	6	-
Tashkent city	0.863	1	1	0.588	2	1
Andijan region	0.364	3	1	0.622	1	1
Navoiy region	0.338	4	2	0.329	7	2
Tashkent region	0.284	2	2	0.321	8	2

Fergana region	0.207	6	3	0.305	9	2
Namangan region	0.369	5	2	0.252	11	3
Samarkand region	0.095	7	3	0.317	10	3
Bukhara region	0.144	8	1	0.359	3	1
Kashkadarya region	0.211	9	1	0.358	4	1
Khorezm region	0.168	10	1	0.251	12	1
Surkhandarya region	0.095	11	2	0.250	13	2
Djizzakh region	0.144	12	2	0.357	5	1
Sirdarya region	0.211	13	1	0.356	6	2
Karakalpak Republic	0.168	14	2	0.249	14	2

Source: own calculation

Table 2 shows the ranking of the regions made based on calculations of integral indicator that describes contribution of small enterprises to economic development of the regions in 2007 and 2012: The secured places of the regions in the ranking in 2012 coincide with the ones secured by the regions in 2007. Tashkent and Fergana territories led the ranking in contribution of small businesses to development of the regions. $I_{Tash}^{SBcont} = 0.742$ (0.659 in 2007), this value came from as follows: the number of small enterprises per 1000 population of Tashkent territory was 17.7 (14.4 in 2007), and this indicator for Tashkent city was 19.5; percentage of able-bodied population employed in these small enterprises was 74.7% (70,3% in Tashkent city); turnover of small enterprises made up 52.2% of the total volume of territory's turnover (48,1% in Tashkent region); only 12% of the total investments invested in fixed assets was directed to small enterprises (14% in Tashkent region). Zarafshan territory secured one of the leading places in the ranking in contribution (share within the total volume of production turnover) of small enterprises to the production in the territory ($I^{\Delta SBturn}_{Zaraf} = 0.603$).

4. Discussion

It is necessary to define an investment policy of small enterprises in the regions based on the following principles:

- Directing investments primarily to economic sectors that have a competitive advantage (competitive advantage in production costs, competitive advantage coming from production of rare and not readily available goods which have a high demand in domestic and world markets).
- Implementing structural reforms in the economic complexes of the regions and economic sectors.
- Directing investments goal-oriented. Today investments should be directed to restructuration of the regional economies, production in real sector of the economy, building a market infrastructure.
- Priority of investing in innovative technologies. This enables to boost profits, improve the conditions of extended reproduction, rebuild the structure of regional enterprises on the basis of innovative industry, as well as, to enhance competitiveness of produced goods.
- Investing in human capital. Because it is impossible to create and use modern innovative technologies without highly skilled workforce.
- Market relations require development of market infrastructure (banks, sales and hotel complexes, exchange houses etc.).

Studies of international experience indicate that the most important factors that open doors to cooperation between large and small enterprises in the context of modern innovative economy are as follows:

- High level of specialization of small business enterprises, and their active innovative activity (because competition in each market segment forces them to enhance their activeness of innovative activity);

- endeavor of business entities to minimize transactional costs and extend market prospects by building a flexible technological and sales chain;
- endeavor of big firms to introduce innovative technologies in their activity by means of small innovator firms;
- existence of special institutions cooperating and intermediating in innovations, as well as, existence of efficient financial credit technologies that support venture entrepreneurship;
- existence of goal-oriented state measures that stimulate stable relations of big firms with small enterprises in production and cooperation;
- efficient distribution of authority and resources among central, regional and local government bodies for supporting small business sector.

These principles should be reflected on the new strategy of the state (above all, on the related goal-oriented programs, regulatory measures in budgetary and tax matters) for supporting small business sector. In our opinion, it is necessary to take the following countrywide measures to boost the efficiency of using domestic resources for financing regional investment projects:

- implementation of investment projects where return of capital investments is fast, reduction of energy and material consumption in industrial production;
- stimulation of private capital participation in financing the construction of industrial facilities, supporting the emergence and development of small enterprises;
- intensification of establishment of industrial enterprises, radical reduction of the number of facilities in the economy, construction of which is incomplete;
- assessment of industrial enterprises' performance, calculation of profitability level, dynamically analyzing economic indicators;
- provision of more free access to data and information related to investment activity, monitoring development of national production;
- promoting economic entities to carry out their activities without restrictions for etc.

As a final conclusion, it is important to point out that we can solve many issues associated with building a developed market economy in Uzbekistan by boosting investment potential of the regions and efficiently implementing investment projects.

References

Abrams, S. (2002). A Practical Approach to the International Valuation and Capital Allocation Puzzle. Salomon Smith Barney.

Bruner, R., et al. (2002). Introduction to 'Valuation in Emerging Markets. *Emerging Markets Review, 3*.

Colin, J. (1998). State Keeps a Tight Rein: Market Reform in Uzbekistan. *The Banker, 14*(8), 47-50.

Estrada J. (2007). Discount rates in emerging markets: four models and an application. *Journal of Applied Corporate Finance, 19*, 72-78.

Estrada, J. (2002). Systematic risk in emerging markets: the D-CAPM. *Emerging Markets Review, 3*, 365–379.

James, M., & Koller T. (2000). Valuation in emerging markets. *Corporate Finance, 4*, 78-85.

Koller T., & Murrin, J. (2000). Valuation: Measuring and Managing the Value of Companies (3rd ed.). New York: John Wiley & Sons.

Luehrman, T. (1997). General manager's guide to valuation. Harvard Business Review.

Luehrman, T. (2009). Business valuation and the cost of capital. Harvard business school publishing, Boston, 9-210-037.

Peimani, H. (2009). Conflict and Security in Central Asia and Caucasus. Santa Barbara, Calif.

PRICEWATERHOUSECOOPERS. (2013). Guide to doing business and investing in Uzbekistan. Retrieved from www.pwc.com/uz

Tsamenyi, M., & Tauringana, V. (2004). Capital budgeting and budgeting practices of foreign operations in Uzbekistan: an exploratory study. *International Journal of Strategic Cost Management. Spring,* 20-36.

UNs Economic and Social Commission for Asia and Pacific. (2003). Managing globalization in selected

countries with economies in transition.ST-ESCAP/2274. United Nations. New York.

Yescombe, E. (2007). Public-Private partnerships. Principles of Policy and Finance. Published by Elsevier ltd. UK.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Study of the Relationship between Dependent and Independent Variable Groups by Using Canonical Correlation Analysis with Application

Thanoon Y. Thanoon[1,3], Robiah Adnan[1] & Seyed Ehsan Saffari[2]

[1] Department of Mathematical Science, Faculty of Science, Universiti Teknologi Malaysia, Malaysia

[2] Education Development Centre, Sabzevar University of Medical Sciences, Sabzevar, Iran

[3] Department of Operation Management Technique, Technical College of Management, Foundation of Technical Education, Mosul, Iraq

Correspondence: Thanoon Y. Thanoon, Department of Mathematical Science, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia. E-mail: Thanoon.younis80@gmail.com/ robiaha@utm.my/ehsanreiki@yahoo.com/

Received: January 22, 2015 Accepted: February 2, 2015 Online Published: July 20, 2015

doi:10.5539/mas.v9n8p72 URL: http://dx.doi.org/10.5539/mas.v9n8p72

Abstract

Canonical correlation analysis is used to study the relationship between two groups of variables (dependent and independent). Since each group represents the linear combination to a number of variables, canonical correlation analysis measures the relationship between these variables that maximally correlate with linear combinations of another subset of variables. Statistical analysis involves canonical correlation between two groups of variables, canonical variates, standard canonical variates, canonical factor loadings, canonical cross factor loadings for both groups. Test of significance of canonical correlation using Wilk's Lambda showed that the first and second canonical correlation was significant and the third and fourth canonical correlation were insignificant. This method is illustrated by using a real data set. Results obtained by using SPSS program.

Keywords: canonical correlation analysis, dependent variables, independent variables, canonical variates

1. Introduction

Canonical correlation analysis (see Weenink, 2003; Fan & Konold, 2010; Vía et al., 2007) is one of the multivariate statistical analysis methods which measures the strength of the overall relationship between the linear structures (Canonical variables) of the dependent and the independent variables. It is a bivariate correlation between two Canonical variables, for example: a group of personal variables and a group of potential measures, group of price indices and a group of production indices, a group of psychological characteristics and a group of physiological characteristics, a group of academic achievement variables and a group of measures for business success (Rencher, 2002).

Canonical correlation analysis is a generalization of the concept of regression analysis, but rather than being a relationship between one variable Y and a group of variables $X_1, X_2, \ldots, X_q$, the Canonical correlation measures (with respect to) the relationship between a group of independent variables $X_1, X_2, \ldots, X_q$, and another group of dependent variables $Y_1, Y_2, \ldots, Y_p$ (Hair, 2009).

More statistically sound methods in the field are based on canonical correlation analysis and involve linear and nonlinear relationships between the groups of variables proposed by (Böckenholt and Böcknholt, 1990; Cook et al., 1996; Thorndike, 2000; Hardoon et al., 2004; Thompson, 2005).

Canonical correlation depends on finding a linear function (linear fitting) in the $X_1, X_2, \ldots, X_q$ (variable U) and a linear function (linear fitting) in the $Y_1, Y_2, \ldots, Y_p$ (variable V). The selected function which represents the correlation between the two largest (correlation between U and V is the greatest), in that case, there will be r correlation relations, which is equal to the smallest value between p and q (Gittins, 1985).

The aim of the canonical correlation analysis is to get a simple description of the structure of the relationship between subgroups of variables.

The paper is organized as follows. Canonical correlation analysis is described in section 2. Some important definitions are explained in Section 3. Test of Significance for Canonical Correlation is discussed in section 4. Statistical analysis which involves data collection and empirical results are presented in Section 5. Some concluding remarks are given in section 6.

2. Canonical Correlation Analysis

Suppose that there are two groups of variables $X' = [X_1, X_2, ..., X_q]$, $Y' = [Y_1, Y_2, ..., Y_P]$

Each of them has a variance matrix $\sum xx$, $\sum yy$ respectively, where $s = \min(p, q)$. The basic objective of the Canonical correlation is to find the canonical variables $U = \alpha' Y$ and $V = \beta' X$, where it should be a correlation between the U and V whichever is greater.

Suppose that the joint variance-covariance matrix of the vector

$$(Y_1, Y_2, ..., Y_P, X_1, X_2, ..., X_q) = (Y', X')$$

known as:

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \tag{1}$$

Also, the joint variance-covariance matrix of the sample

$$S = \begin{bmatrix} S_{yy} & S_{yx} \\ S_{xy} & S_{xx} \end{bmatrix} \tag{2}$$

Therefore, the correlation between U and V are as follows: -

$$\rho_{UV} = \frac{\alpha' \Sigma_{yx} \beta}{\left[(\alpha' \Sigma_{yy} \alpha)(\beta' \Sigma_{xx} \beta) \right]^{1/2}} \tag{3}$$

because ρ_{UV} includes canonical variables U and V so-called canonical correlation.

$$R = [diag(S_{yy}, S_{xx})]^{-1/2} [diag(S_{yy}, S_{xx})]^{-1/2} \tag{4}$$

From the previous two equations the correlation matrix can be expressed as follows:

$$R = \begin{bmatrix} R_{yy} & R_{yx} \\ R_{xy} & R_{xx} \end{bmatrix} \tag{5}$$

and we can prove it by second method as follows: -

$$\left| S_{yx} S_{xx}^{-1} S_{xy} - \rho^2 S_{yy} \right| = 0 \tag{6}$$

$$\left| R_{yx} R_{xx}^{-1} R_{xy} - \rho^2 R_{yy} \right| = 0 \tag{7}$$

where (a_i, b_i) is the Eigen vector of matrix S, (c_i, d_i) is the Eigen vector of matrix R, and can be expressed in canonical variables as a vector of variables

$$U_i = a_i' y \tag{8}$$

$$V_i = b_i' x \tag{9}$$

Standard parameters can also calculated according to the following formula:

$$U_i = c_i' z_y \tag{10}$$

$$V_i = d_i' z_x \tag{11}$$

where they can express on (Z_x, Z_y) as standardized variables (Timm, 2002).

3. Some Important Definitions (Black *et al.*, 1998)

3.1 Canonical Function

Represents the relationship (correlation) between two structures (Canonical variables). Each Canonical function has two variables, one of the Canonical group of independent variables and the other to the group of dependent variables. The strength of this relationship is given by the canonical correlation.

3.2 Canonical Loadings

A measure of simple linear correlation between the independent variables and canonical variables. Interpretation of the canonical loadings are similar to the interpretation of factor loadings in factor analysis.

3.3 Canonical Cross-Loadings

Represents the correlation between the independent or dependent corresponding canonical variables, for example: the independent variables associated with the canonical dependent variables, the dependent variables associated with the canonical independent variables.

3.4 Canonical Variates

Represents the linear structure of the total weighted sum of two variables or more and can be defined as either independent or dependent variables.

3.5 Canonical Roots

Represent the square of canonical correlation which is used to estimate the amount of the variance between the weighted optimal canonical variables for independent variables and dependent and can be named Eigen values.

4. Test of Significance for Canonical Correlation (Härdle and Simar, 2007)

We can test the following alternative hypothesis

$$\begin{aligned} H_o &: \rho_1 = \rho_2 = ... = \rho_m = 0 \\ H_1 &: \rho_1 \neq \rho_2 \neq .... \neq \rho_m \neq 0 \end{aligned} \tag{12}$$

for canonical correlation coefficient and by using the most common measures used a Wilk's lambda described as follows:

$$\lambda = \prod_{i=1}^{s} (1 - \rho^2) \tag{13}$$

where ρ^2 represent the eigenvalues of the sample:

$$\left| S_{yx} S_{xx}^{-1} S_{xy} - \rho^2 S_{yy} \right| = 0 \tag{14}$$

which is a square canonical correlation of the sample.

Equation (13) proves that the two variables X, Y are uncorrelated linearity and mathematical application (Wilk's λ) will be described almost as a variable distribution following a Chi-square with degrees of freedom v = (p-k) (q-k), as well statistically significant application (Wilk's λ) needs to account for the following statistics:

$$\chi^2 = -\left[(n-1) - (p+q+1)/2 \right] \log \lambda \tag{15}$$

where:

n: the number of cases.

Log: the natural logarithm function.

q: the number of variables in the first group.

p: the number of variables in the second group.

5. Statistical Analysis

Statistical program SPSS was used to find a canonical correlation analysis through finding the correlation matrix between all independent variables and the correlation matrix between all dependent matrix; the correlation

matrix between independent and dependent variables in both groups; finding Wilks' Lambda test to see significance of the canonical correlation; finding a canonical correlation between the two groups; finding standard canonical coefficients in the first and second groups; creating factor loadings matrix in the first and second groups; finding cross loadings matrix in first and second groups; and finding the canonical scores of the first and second groups.

5.1 Data Collection

The data has been taken from Ibn Sina Hospital (surgery and fractures ward) for (80) patients with infection of the urinary tract were selected as a group of variables which could be influential on the disease. These variables were divided into two groups: the first group was a group of personal variables and the second group was a group of pathological variables as follows:

5.2 Results and Discussion

By Equation (5) will be displayed a correlation matrix for personal variables in addition to the correlation matrix of pathological variables, the joint correlation matrix between personal and pathological variables as shown in tables (1, 2, 3) respectively.

Table 1. Correlation Matrix between the Independent Variables

Var.	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1	-0.0832	-0.4823	-0.0185	-0.0885	-0.3640
X_2	-0.0832	1	0.0448	0.0618	-0.6364	0.0051
X_3	-0.4823	0.0448	1	-0.2834	-0.0348	0.4984
X_4	-0.0185	0.0618	-0.2834	1	-0.0745	-0.1673
X_5	-0.0885	-0.6364	-0.0348	-0.0745	1	0.0951
X_6	-0.3640	0.0051	0.4984	-0.1673	0.0951	1

Table 2. Correlation Matrix between the Dependent Variables

Var.	Y_1	Y_2	Y_3	Y_4
Y_1	1.0000	0.1701	-0.0771	0.1283
Y_2	0.1701	1.0000	0.2090	0.1541
Y_3	-0.0771	0.2090	1.0000	0.1014
Y_4	0.1283	0.1541	0.1014	1.0000

Table 3. Correlation Matrix between the Dependent and the Independent Variables

Var.	Y_1	Y_2	Y_3	Y_4
X_1	0.4121	0.1711	0.2281	0.1273
X_2	-0.1183	0.5063	0.1838	0.0064
X_3	-0.1820	-0.1325	-0.3085	-0.0391
X_4	0.1230	0.1691	0.0685	-0.1768
X_5	-0.0305	-0.3617	-0.1165	-0.0797
X_6	-0.1205	0.0490	-0.1743	-0.0623

Table 4. Wilks' Lambda Test

Wilk's Lambda	Chi-SQ.	DF	Sig.
0.415	64.571	24	0.000
0.683	28.033	15	0.021
0.893	8.280	8	0.407
0.961	2.893	3	0.408

Table 4 represents the first and second canonical correlations were significant but the rest of the canonical correlations were not significant, based on the test results (Wilk's lambda).

Table 5. Canonical Correlations

F_1	F_2	F_3	F_4
0.626	0.485	0.266	0.196

Table 5, which contains the first, second, third and fourth canonical correlation, the canonical correlations are strong between the first canonical variable which is extracted from the canonical correlation function. The rest of the canonical correlation was weak and it refers to the weakness of the relationship between the canonical variables extracted from the functions and canonical correlation.

Table 6. Standardized Canonical Coefficients for Group-1

	F_1	F_2	F_3	F_4
X_1	-0.460	0.762	-0.359	0.457
X_2	-0.796	-0.535	0.166	-0.222
X_3	0.260	-0.020	-0.710	0.536
X_4	-0.282	0.131	-0.881	-0.434
X_5	0.021	-0.101	-0.086	-0.621
X_6	-0.315	-0.114	-0.360	0.225

Table 7. Canonical Coefficients for Group-1

	F_1	F_2	F_3	F_4
X_1	-0.922	1.525	-0.719	0.916
X_2	-0.045	-0.030	0.009	-0.013
X_3	0.191	-0.015	-0.523	0.395
X_4	-0.658	0.307	-2.057	-1.013
X_5	0.050	-0.245	-0.210	-1.510
X_6	-0.351	-0.127	-0.402	0.251

Table 8. Standardized Canonical Coefficients for Group-2

	F_1	F_2	F_3	F_4
Y_1	-0.122	0.947	-0.372	0.088
Y_2	-0.877	-0.434	-0.240	0.292
Y_3	-0.320	0.407	0.633	-0.631
Y_4	0.173	0.082	0.653	0.762

Table 9. Canonical Coefficients for Group-2

	F_1	F_2	F_3	F_4
Y_1	-0.089	0.690	-0.271	0.064
Y_2	-0.563	-0.279	-0.154	0.188
Y_3	-0.255	0.324	0.504	-0.502
Y_4	0.044	0.021	0.167	0.195

Table 10. Canonical Loadings for Group-1

	F_1	F_2	F_3	F_4
X_1	-0.401	0.864	0.125	0.199
X_2	-0.778	-0.528	0.163	0.134
X_3	0.368	-0.502	-0.457	0.562
X_4	-0.345	0.117	-0.596	-0.599
X_5	0.550	0.153	-0.104	-0.486
X_6	0.027	-0.435	-0.444	0.338

Table 10 represents the factor loadings for the first group (personal variables). Through the first factor we show that the independent variables (X_2, X_5) have a simple linear correlation with the corresponding canonical independent variables, while variables appeared (X_1, X_2, X_3) with a linear relationship with the corresponding independent variables canonical in Group II. However the rest of the factors (third and fourth) cannot be relied upon to describe the data, because the canonical correlation coefficients were not significant, according to the test (Wilk's lambda).

Table 11. Cross Loadings for Group-1

	F_1	F_2	F_3	F_4
X_1	-0.251	0.419	0.033	0.039
X_2	-0.51	-0.256	0.043	0.026
X_3	0.230	-0.244	-0.121	0.110
X_4	-0.216	0.057	-0.158	-0.118
X_5	0.344	0.074	-0.028	-0.095
X_6	0.017	-0.211	-0.118	0.066

Table 11 represents cross loadings of the first group (personal variables). Through the first factor it can be seen that the independent variable (X_2) has a linear relationship with the canonical dependent variables. However, the second, third and the fourth factors could not be relied upon in the description of the data, because the canonical correlation coefficients were not significant, according to the test (Wilk's lambda).

Table 12. Canonical Loadings for Group-2

	F_1	F_2	F_3	F_4
Y_1	-0.224	0.852	-0.378	0.284
Y_2	-0.938	-0.175	-0.070	0.292
Y_3	-0.476	0.251	0.678	-0.500
Y_4	-0.010	0.177	0.632	0.754

Table 12 above represents canonical loadings for the second group (pathological variables) and through the first factor it is shown that the variable (Y_2) has a simple linear correlation with the corresponding canonical dependent variables, while the variable (Y_1) has a linear relationship with the corresponding canonical dependent variables in the second group. The rest of the factors (third and fourth) cannot be relied upon to describe the data, because the canonical correlation coefficients were not significant, according to the test (Wilk's lambda).

Table 13. Cross Factor Loadings for Group-2

	F_1	F_2	F_3	F_4
Y_1	-0.140	0.414	-0.100	0.056
Y_2	-0.587	-0.085	-0.019	0.057
Y_3	-0.298	0.122	0.180	-0.098
Y_4	-0.006	0.086	0.168	0.148

Table 13 represents cross factor loadings for the second group (pathological variables). Through the first factor it is shown that the dependent variable (Y_2) has a linear relationship with the canonical independent variables. The second cross factor loading does not appear to have any significant effect variable with any of the supported canonical changes. The third and fourth factor could not be relied upon in the description of the data because the canonical correlation coefficients were not significant, according to the Wilk's lambda test. The canonical values of the first and second groups are shown in Table 14 in Appendix A**.**

6. Conclusions

Canonical correlation analysis method (CCA) is very useful in interpretation of data by discovering the structures and similar relationships between two sets of multi-dimensional variables and categories of those variables are often used in medical data. The significant values (sig.) are selected when the value is less than or

equal to 0.05. Wilk's lambda test showed significant first and the second canonical correlation and the rest of the canonical correlations were not significant.

There is a strong relationship between the first group (personal variables) and the second group (pathological variables), because the correlation function has worked to maximize the correlation between two groups, and through factor loading matrix have been identified canonical variables that have a relationship with the original values. There are significant and non signifigant relationships in the results of factor loadings and cross factor loadings. Theses results are are very important to interpretation the correlation relationships between the depandent and the independent variable groups.

Reference

Black, W., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). Multivariate data analysis. *New Jersey*.

Böckenholt, U. and Böcknholt, I. (1990). Canonical analysis of contingency tables with linear constraints. *Psychometrika, 55*(4), 633-639.

Cook, J. A., Razzano, L., & Cappelleri, J. C. (1996). Canonical correlation analysis of residential and vocational outcomes following psychiatric rehabilitation. *Evaluation and Program Planning, 19*(4), 351-363.

Fan, X., & Konold, T. R. (2010). Canonical correlation analysis. *The reviewer's guide to quantitative methods in the social sciences*, 29-40.

Gittins, R. (1985). *Canonical analysis: a review with applications in ecology*: Springer Berlin.

Hair, J. F. (2009). Multivariate data analysis.

Härdle, W., & Simar, L. (2007). *Applied multivariate statistical analysis* (Vol. 22007): Springer.

Hardoon, D. R., Szedmak, S., & Shawe-Taylor, J. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural computation, 16*(12), 2639-2664.

Rencher, A. C. (2002). *Methods of multivariate analysis*. New York: J. Wiley.

Thompson, B. (2005). Canonical correlation analysis. *Encyclopedia of statistics in behavioral science*.

Thorndike, R. M. (2000). Canonical correlation analysis. *Applied multivariate statistics and mathematical modeling*, 237-263.

Timm, N. H. (2002). *Applied multivariate analysis*: Springer.

Vía, J., Santamaría, I., & Pérez, J. (2007). A learning algorithm for adaptive canonical correlation analysis of several data groups. *Neural Networks, 20*(1), 139-152.

Weenink, D. (2003). Canonical correlation analysis. Proceedings of the 2003 *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 81-99.

Appendix A

Table 14. Canonical scores for first and second Groups

Canonical scores for Group-1					Canonical scores for Group-2			
	F_1	**F_2**	**F_3**	**F_4**	**F_1**	**F_2**	**F_3**	**F_4**
Obs1	-3.81	-1.46	0.41	2.09	-8.55	0.98	-0.32	-1
Obs2	-6.41	-3.07	1.16	2.31	-6.03	-0.53	-1.93	0.2
Obs3	-2.57	-0.94	0.43	1.89	-4.41	1.48	-2.39	0.68
Obs4	-2.16	-0.94	0.57	1.89	-4.98	1.48	-1.93	0.68
Obs5	-4.61	-3.7	-0.48	1.43	-5.69	1.87	-0.05	0.42
Obs6	-4.44	-3.93	1.6	1.88	-4.73	0.17	-0.21	1.74
Obs7	-3.69	-1.38	-0.18	3.15	-4.46	-0.06	-0.84	-0.37
Obs8	-4.22	-1.41	1.75	0.93	-4.78	1.6	-0.15	0.8
Obs9	-3.11	-1.38	0.96	3.15	-6.01	-0.06	-3.71	-0.37
Obs10	-4.22	-1.95	1.75	2.87	-4.78	-0.22	-0.15	-0.18
Obs11	-4.43	-1.38	1.19	3.15	-5.42	-0.06	-1.39	-0.37
Obs12	-3.52	-1.64	1.93	3.47	-5.89	0.44	0.73	-0.87
Obs13	-3.85	-2.14	2.61	4.57	-6.77	2.95	-2.74	-1.21
Obs14	-3.77	-1.07	2.05	1.1	-4.87	0.92	-0.03	-0.36
Obs15	-4.22	-1.59	1.75	3.04	-4.78	1.12	-0.15	1.38
Obs16	-5.55	-1.68	1.6	3	-6.69	0.78	-1.35	0.99

Obs17	-3.51	-1.12	2.37	3.27	-4.44	0.94	-0.3	0.8
Obs18	-4.23	-2.24	-0.5	2.72	-6.74	0.63	0.74	1.18
Obs19	-4	-0.86	-0.4	2.95	-4.4	0.43	-0.92	1.31
Obs20	-4.61	-2.2	-0.52	0.54	-3.3	0.61	-1.78	0.01
Obs21	-5.29	-3.37	1.92	2.16	-6.26	0.32	-1.62	1.55
Obs22	-3.5	-1.03	0.83	3.32	-6.21	1.27	-1.96	1.19
Obs23	-2.94	-0.94	0.9	1.89	-7.49	1.48	-2.63	0.68
Obs24	-2.65	-0.73	0.51	3.01	-4.68	0.94	-0.55	1.89
Obs25	-4.1	-3.19	-0.32	2.24	-3.89	0.99	-1.29	2.33
Obs26	-3.74	-1.21	2.22	3.96	-4.4	0.67	-0.36	0.87
Obs27	-5.97	-3.64	1.47	2.03	-6.12	-0.68	-1.81	0.38
Obs28	-3.95	-1.64	0.53	3.47	-6.12	0.44	-2.08	-0.87
Obs29	-3.27	-1.37	2.11	3.6	-5.18	1.44	-1.41	0.3
Obs30	-2.46	-0.81	0.5	1.96	-5.2	1.98	-0.15	1.26
Obs31	-5.97	-3.64	1.47	2.03	-6.12	-0.68	-1.81	0.38
Obs32	-3.32	-1.46	-0.22	2.09	-4.74	0.98	-2.26	-1
Obs33	-5.31	-1.94	1.16	3.32	-4.07	1.29	-0.8	0.49
Obs34	-3.77	-1.37	2.05	3.6	-4.87	1.44	-0.03	0.3
Obs35	-2.2	-1.11	0.79	2.26	-4.96	2.32	-1.94	0.56
Obs36	-4.44	-1.89	1.6	3.8	-4.73	0.94	-0.21	-1.37
Obs37	-5.04	-1.94	-0.06	3.32	-5.41	1.29	-2.73	0.49
Obs38	-3.19	-2.93	2.01	3.56	-6.44	1.54	1.16	1.9
Obs39	-6.19	-4.46	1.31	2.08	-6.08	-0.33	-1.87	0.07
Obs40	-3.27	-1.37	2.11	3.6	-5.18	1.44	-1.41	0.3
Obs41	-3.32	-1.12	2.35	3.27	-4.97	0.94	0.1	0.8
Obs42	-4.22	-1.68	1.75	4.18	-4.78	0.34	-0.15	-0.61
Obs43	-4.9	-3.06	-0.96	2.76	-6.6	0.98	0.56	0.86
Obs44	-2.65	-1.37	0.51	3.6	-4.68	1.44	-0.55	0.3
Obs45	-3.35	-1.37	2.19	3.6	-5.44	1.44	0.43	0.3
Obs46	-3.14	-3.37	0.04	2.16	-5.06	0.32	-0.34	1.55
Obs47	-4.44	-0.42	1.6	1.7	-4.73	1.98	-0.21	2.35
Obs48	-4.63	-1.94	-0.45	1.12	-4.73	1.1	-0.74	-0.87
Obs49	-2.66	-0.82	0.16	0.78	-6.79	0.41	-0.69	0.14
Obs50	-3.55	-1.99	-0.09	0.2	-4.49	-0.06	-0.8	0.32
Obs51	-3.81	-1.33	-0.41	0.97	-4.92	1.93	-0.53	1.19
Obs52	-4.36	-1.68	-0.6	0.8	-6.71	0.6	0.71	-0.37
Obs53	-4.98	-2.09	1.23	2.39	-4.62	2.93	-0.36	-2.37
Obs54	-5.28	-0.8	-0.97	1.67	-3.16	3.42	-1.96	1.47
Obs55	-5.29	-1.55	1.92	3.06	-6.26	1.29	-1.62	1.58
Obs56	-3.66	-1.12	1.84	1.08	-5.86	0.75	0.69	-0.56
Obs57	-5.12	-4.21	1.14	3.54	-4.59	3.01	-0.4	0.32
Obs58	-3.7	-1.16	0.85	1.06	-5.69	0.58	-2.35	-0.75
Obs59	-3.62	-1.33	-0.43	1.7	-5.44	1.99	-0.13	1.64
Obs60	-3.79	-1.94	-0.11	1.12	-3.96	1.1	-1.21	-0.87
Obs61	-5.31	-1.95	1.16	2.87	-4.07	-0.22	-0.8	-0.18
Obs62	-2.18	-0.86	0.52	0.76	-5.93	0.25	-1.25	-0.06
Obs63	-3.62	-0.81	-0.43	1.23	-5.44	1.92	-0.13	0.81
Obs64	-5.12	-3.09	1.14	0.54	-4.59	2.64	-0.4	2.03
Obs65	-3.54	-1.37	0.67	1.4	-6.69	1.25	-1.62	-1.06
Obs66	-3.27	-1.37	-0.05	1.4	-5.03	1.25	-0.38	-1.06
Obs67	-4.81	-3.12	-0.9	1.83	-6.62	-0.18	0.58	2.06
Obs68	-2.29	-1.12	0.73	1.08	-4.94	0.75	-1.96	-0.56
Obs69	-4.29	-1.62	0.02	2	-7.01	2.33	-1.49	0.95
Obs70	-4.71	-1.33	-1.02	1.42	-4.73	1.42	-0.78	-0.87
Obs71	-3.58	-0.61	2.04	2.63	-5.4	-0.07	0.37	1.81
Obs72	-3.16	-0.61	2.15	2.63	-6.16	-0.07	-0.69	1.81

Obs73	-4.28	-2	-0.12	1.94	-7.5	-1.38	-1.15	1.01
Obs74	-2.72	-0.87	0.18	2.5	-5.63	-1.07	0.12	0.64
Obs75	-3.54	-0.7	0.67	1.12	-6.69	-0.53	-1.62	0.51
Obs76	-2.14	-0.31	0.55	2.78	-5.94	-0.92	-1.24	0.45
Obs77	-2.19	-0.31	0.65	2.78	-5.45	-0.92	-1.59	0.45
Obs78	-4.23	0.13	-0.5	1.53	-6.74	0.63	0.74	1.49
Obs79	-3.87	-0.27	-0.38	2.8	-6.34	-0.75	0.49	0.65
Obs80	-3.08	0.09	2.34	2.97	-5.69	0.59	-1.01	2.2

Appendix B

X: Group of Personal Variables

X_1: Gender (0: Male, 1: Female)

X_2: Age

X_3: Cultural level (1: Not able to read or write, 2: Primary, 3: Secondary1, 4: Secendary2, 5: University)

X_4: Residence (1: City, 2: Rural)

X_5: Marital Status (1: Married, 2: Single)

X_6: Occupation (1: earner or a housewife, 2: Student, 3: Employee, 4: Retired)

Y: Group of Pathological Variables

Y_1: current disease (1: skeletal system, 2: muscular system, 3: nervous system, 4: digestive system, 5: reproductive system 6: lymphatic system.

Y_2: Chronic disease (1: None, 2: D.M, 3: H.T, 4: I.H.D, 5: 2 + 3, 6: 3 + 4.

Y_3: Day incidence (1: Unknown, 2: the second day, 3: the third day, 4: the fourth day, 5: fifth day, 6: The Sixth Day, 7: The Seventh Day.

Y_4: drugs taken (1: cannot take an antibiotic, 2: cephalosporines, 3: penicillins, 4: aminoglycosides, 5:antiprotozoal, 6: quinolones + cephalosporines + antiprotozoal, 7: antituberculosis, 8: 5 +2, 9: 3 + 2, 10: 5 +4 +3, 11: 5 +3, 12: 5 +4 +2, 13: 4 +3, 14: 5 +3 +2,15: 5 +4.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The Development Of Mercury Ion Selective Electrode With Ionophore 7,16-Di-(2-Methylquinolyl)-1,4,10,13-Tetraoxa-7,16-Diazacyclooctadecane (DQDC)

Eidi Sihombing[1], Manihar Situmorang[2], Timbangen Sembiring[1] & Nasruddin[1]

[1] Department of Physics, Faculty of Mathematics and Natural Science, Universitas Sumatera Utara, Medan, North Sumatra, Indonesia

[2] Department of Chemistry, Faculty of Mathematics and Natural Science, Universitas Negeri Medan, Medan, North Sumatra, Indonesia

Correspondence: Manihar Situmorang, Department of Chemistry, Faculty of Mathematics and Natural Science, Universitas Negeri Medan, Medan, North Sumatra, Indonesia. Tel: 62-61-663-6757. E-mail: msitumorang@lycos.com

Received: February4, 2015 Accepted: February17, 2015 Online Published: July 20, 2015

doi:10.5539/mas.v9n8p81 URL: http://dx.doi.org/10.5539/mas.v9n8p81

The research is financed by Research Project HibahBersaingUniversitasNegeri Medan (UNIMED), Directorate General of Higher Education (DGHE), Ministry of Research and Higher Education2015

Abstract

The development of mercury ion selective electrode (ISE-Hg) by using ionophore 7,16-Di-(2-methylquinolyl)-1,4,10,13-tetraoxa-7,16-diazacyclooctadecane (DQDC) in an electrode membrane sensing is explained. The study is aimed to construct an ion selective electrode for mercury by immobilization of ionophore DQDC in a polymeric PVC membrane electrode. Membrane electrode was constructed by casting and spin coated methods from the mixture of DQDC and plastisizer in a polymeric PVC matrix dissolved in a tetrahydofuran solvent. The membrane has been characterized and used as a sensing material in an ISE-Hg. The developed ISE-Hggive sensitive and selective response to mercury, working range linearity lies between 0.005 mM – 1.0 mM Hg^{2+}, slope 25.82 mV per decade concentration of Hg^{2+} ($r^2 = 0.998$).

Keywords: potentiometric; membrane electrode, ionophore, ion selective electrode, ISE-Hg, mercury

1. Introduction

The development of potentiometric method with using ion selective electrode (ISE) for mercury is a great challenge to obtain an alternative analytical tool for low cost, simple instrumentation, fast response, high selectivity, and wide linear detection system. Potentiometric method has been considered as an attractive detection method because of the cheap and simple instrumentation required, commonly a pH-meter (Situmorang, et al., 2008). Tungsten metal electrode due to its responsive to potential changes has been used as sensing device, and its compatible for sensing component has been demonstrated in the construction of potentiometric sensor (Situmorang, 2001). Tungsten electrode has also been used as transducer for potentiometric biosensors with immobilization of enzyme in polytyramine for malic acid (Situmorang, et al., 2001). It was demonstrated that electrodeposited polytyramine (Situmorang, et al., 1999) could act as a substrate electrode and retain its ability to determine potential change (Situmorang, et al., 2002).

Various analytical methods have been developed for determination of mercury, they are using spectrophotometric methods with adsorbing agent of 1,5-diphenylthiocarbazone (Khan, et al., 2005) or with o-carboxy phenyl diazoamino p-azobenzene (Chatterjeeetal, 2002). Spectrofluorometry with 2-hydroxy-1-naphthaldehydene-8-aminoquinolinefluorescent reagent has also introduced (Lietal, 2006). Chromatographic method has also been reported for the determination of mercury (Mondal& Das, 2003; Hu, 2002). Amperometricmethod for the determination of mercury have been reported (Majid, et al., 2002; Lu, et al, 2003). Inductively coupled plasma mass spectrometry (ICP-MS) has also been introduced for mercury determination (Zhang, et al, 2004). Another method by using capillary electrophoresis (Páger&Gáspár, 2002),

and capillary electrophoresis inductively coupled plasma mass spectrometry with microconcentric nebulization has also been reported (Lee & Jiang, 2000). The determination of mercury are still dominated by atomic absorption spectroscopy (AAS), those are withgraphite furnace atomic absorption spectrometry (Da-Silva, et al., 2002; Izgi, et al., 2000), electrothermal atomic absorption spectrometry (Moreno, et al., 2002), and cold vapor generation-electrothermal atomic absorption spectrometry (Moreda-Piñeiro, et al., 2002). In practice, cold vapor atomic absorption spectrophotometry (CV-AAS) is a good choice because it has superiority in accuracy and sensitivity (Silva, et al., 2006, Rizea, et al., 2007). However, those are with relatively high cost instruments.

Potentiometric determination of mercury by using diazacrownionophore membrane electrode has been reported (Yang, et al., 1997; Yang, et al., 1998). Potentiometric device with ISE by incorporation of ionophore in membrane electrode is a robust analytical method for the determination of mercury. It was believed the key aspects on the success for the construction of ISE is based on its design with the integration of ionophore onto the membrane electrode. The development of ISE-Hg reported in this study is based upon successfully in the construction of ISE by using ionophore sensing device for the determination of lead (Situmorang, 2005) and mercury (Purba, et al., 2013). The construction of PVC membrane electrodes with incorporating ionophore is very important in the development of ion selective electrode for mercury (ISE-Hg). The purpose of this paper is to develop a potentiometric sensor for mercury with incorporation of ionophore 7,16-Di-(2-methylquinolyl)-1,4,10,13-tetraoxa-7,16-diazacyclooctadecane (DQDC) that is immobilized in a polymeric PVC membrane electrode. Synthesis of sensing agents of ionophore DQDC, the strategy to incorporate ionophore on to the PVC membrane electrode, the construction of ISE-Hg, and the characterization of the electrode for its performance to mercury are demonstrated in this study. The application of the electrode to ion-sensing components for mercury ion in batch analysis is also described.

2. Experimental

2.1 Reagents and Materials

All chemicals are laboratory reagent grade and used as received except where indicated. Starting material of 1,4,10,13-tetraoxa-7,16-diazacyclooctadecane (DC), Poly(vinyl chloride) (PVC: low molecular weight type) as membrane matrix material, 2-nitrophenyl octyl ether (NPOE) as membrane plasticizer, potassium tetrakis (p-chlorophenyl) borate (KTpClPB) as anionic additive, 2-chloromethylquinoline, mercury nitrate, solvent of acetonitrile, tetrahydrofuran (THF), dichloromethane, and tungsten wire (99.9 %, Ø 0.5 mm) were all purchased from Aldrich Chemical Company.

2.2 Instrumentation

Potentiometric measurements were made with a Keithley 177-Microvolt Digital Multi Meter (Keithley Instrument, USA) connecting with MacLab/8 Analog Digital Instrument, Sydney, Australia, for data acquisition. The ion selective electrodes assembled with a tungsten wire electrode was constructed followed the procedures explained in previous report (Purba, et al., 2014). Commercial Ag/AgCl electrode containing 3 M KCl internal solution from Bioanalytical System (BAS) USA was used as reference electrode in all electrochemical experiments. High-resolutionscanning electron microscope, SEM (Zeiss Neon) is used for membrane morphology studies.

2.3 Synthesis of Ionophore DQDC

Synthesis of ionophore as a sensing agent for mercury was very important part of the work because diazacrown ethers bearing a cation ligating pendant group on the nitrogen atom have successfully been used in cation membrane transport. Ionophore 7,16-Di-(2-methylquinolyl)-1,4,10,13-tetraoxa-7,16-diazacyclooctadecane is synthesized by the alkylation of 1,4,10,13-tetraoxa-7,16-diazacyclooctadecane with 2-chloromethylquinoline in acetonitrile containing sodium carbonate followed the modification of the procedures explained by Yang, et al., (1998). Another reaction was carried out in the dried analytical grade THF and dichloromethane solvents and done under an atmosphere of nitrogenas explained previously (Situmorang, et al., 2014).

2.4 Preparation of Membrane Electrodes

Preparation of membrane electrode is carried out by casting method and spin coated procedures to obtain sheet of PVC membrane from the mixture composition of matrix polymer PVC, plasticizer NPOE, ionophoreDQDC, with added the anionic additive for polymeric membraneKTpClPB in THF solution. The casting method procedure to prepare membrane electrode containing of ionophore was conducted followed the procedures explained previously (Purba, et al., 2014; Purba, et al., 2013), it was the modification from the procedure explained by Yang, et al., (1998). The membrane composition (in weight %) was 3% ionophore DQDC, 30% PVC, 67% NPOE and 50 mol% KTpClPB were dissolved in 5 mL of dry tetrahydrofuran (THF) at constant

stirring, and poured onto casting apparatus (petridis) followed by evaporation of the solvent to form a thin, flat, and clear membrane sheet. Spin coating procedure in the preparation of the electrode membrane was also carried out. The same quantity of the mixture of ionophore DQDC, PVC and plasticizer are dissolved in 5 mL THF as explained above, and the liquid was put on sputtering apparatus, followed by two steps spinning, at 600 rpm for 10 second and at 800 rpm for 30 second, then let the solvent to evaporate to form a thin, flat, and clear membrane sheet. The use of spin-coated techniques in the preparation of membrane electrode has advantage in the good reproducibility of membrane thickness, and the membrane sheet is compatibel for the construction of ion selective electrode.

2.5 Preparation of ISE-Hg

The strategy by using membrane type system is a considered choice to construct ISE-Hg due to the simple fabrication of the electrode. Preparation of ISE-Hg is carried out followed the procedures explained in previous study (Situmorang, et al., 2005) with modification to obtain a better sensing for mercury. The ISE-Hg with membrane electrode was fabricated by attaching a disc of Ø 5 mm diameter PVC membrane, which is cut from the PVC membrane sheet prepared above, to PVC body (approximate Ø 5.0 mm internal diameter and 60 mm length) at one edge is glued by using a light paste of PVC in THF solution. The electrode is then filled with internal solution of 10 mM$Hg(NO_3)_2$ and 10 mMKCl solution. Pure tungsten (W) wire was firstly cleaned with 1 M nitric acid for 1 minute, followed by sonication, and finally rinsed with distilled water. The W-wire is then inserted into the electrode body that is immersed in an internal solution, and sealed the top with rubberseptum to construct an ISE-Hg. The Ag/AgCl (BAS) was used for external reference electrode. The construction of PVC membrane liquid-contact electrodes is selected in the construction of electrode body design because the electrode assembly is simple in the incorporation of ionophore in to the electrode. The design of ISE-Hg is illustrated in Figure 1.

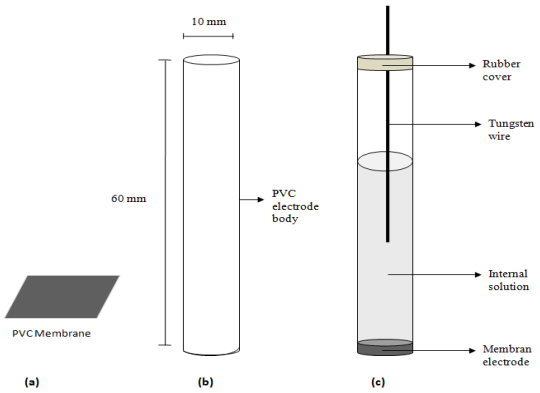


Figure 1. The design of ISE-Hg electrode containing of ionophore with PVC membrane liquid-contact electrode: (a) membrane electrode containing of ionophore DQDC, (b) a PVC electrode body, and (c) the ISE-Hg

2.6 Potentiometric Measurements

An ISE-Hg and Ag/AgCl reference electrode were immersed in a stirred bulk background solution of 10 mL nitric acid solution, pH 4.0, containing 10 mM KNO_3. Standard solution of $Hg(NO_3)$ at different concentration (accurately measured) was introduced to the electrochemical cell from a syringe followed by stirringafter each addition, successively starting from low concentration to higher concentrations of mercury standard solutions. Potential change was monitored on a digital mV meter interfaced with PowerLab, and the data were collected in personal computer. All measurements were carried out at 25± 0.5 ˚C. The calibration curve was done in the absence of interfering cations.

3. Result and Discussion

3.1 Optimization of Membrane Composition

Polymeric membrane of poly(vinyl chloride) (PVC) is popular to be used in the construction of ion selective electrode. The PVC is chosen as membrane matrix in this study with respect to the mechanical stability, chemical stability, chemical inertness (there is no competition with solution), the resulting membrane has no pores, and has low electrical resistance. The composition of ion-selective membranes have been optimized for mechanical properties to be suited for the construction of ISE-Hg. Membrane electrode at different ingredients consisted of

matrix polymer PVC, plasticizer NPOE and KTpClPB and ionophore DQDC has been fabricated by using casting and spin coated procedures. The mechanical properties of the membrane and the performance of the ISE-Hg electrodes prepared with these compositions are listed in Table 1.

Table 1. The composition of membrane ingredients, consisted of ionophore DQDC, matrix polymer PVC, and plasticizer NPOE, and anionic additive KTpClPB toward the membrane properties and electrode performances. The values given are to nearest weight percentage

No	Composition in 5 mL THF				Mechanical Properties of Membrane		ISE-Hg Performance	
	PVC (wt %)	NPOE (wt %)	DQDC (wt %)	KTp-ClPB (mol %)	Casting	Spin coating	Response	Slope (mV)
1	20	77	3	50	Clear, very elastic, thin	Clear, very elastic, very thin	very fast	25.42
2	30	67	3	50	Clear, very elastic	Clear, very elastic, thin	very fast	25.26
3	40	57	3	50	Clear, elastic	Clear, elastic	Fast	25.81
4	50	47	3	50	Clear, elastic, thick	Clear, elastic, thick	Slow	22.11
5	60	37	3	50	White, rigid, thick	White, rigid, thick	Slow	18.26
6	70	27	3	50	White, rigid, very thick	White, rigid, very thick	Very slow	13.38

The mechanical properties of membrane electrode are influenced by the composition of polymeric PVC and the plasticizer. It is observed that the thickness and hardness of the electrode membrane depended upon the amount of PVC used. When using higher amount of PVC the membrane became rigid, too dense, and resistive that results in longer response time needed in the potential measurement. However, a lower PVC content made the membrane too thin, results in poor mechanical strength, it is swelled up in the solution very fast, and broken easily. The membrane electrode prepared from low concentration of PVC (20 wt%) with high composition of plasticizer (77 wt%) produce very thin membrane sheet with clear in color, where membrane texture is very elastic. This condition is adequat for the preparation of membrane electrode for ISE-Hg, but with short life time. Increasing the amount of PVC will increase the thickness of the membrane electrode but sacrifice in the elasticity of the membrane sheet. The membrane electrode prepared from high amount of PVC (70 wt%) with low concentration of plasticizer (27 wt%) produce very thick membrane sheet with white in color, accompanied with rigid texture.The compositions of the membranes listed in Table 1 are compatible with the construction of ISE-Hg. The membrane prepared from very low amount of PVC (less than 20 wt%) could not be used for electrode materials because it was swelling very fast in the solution and broken easily. Therefore, compromise in the composition of polymer and plasticizer has to be considered in the preparation of the electrode membrane to obtain the membrane with good mechanical properties for ISE-Hg. The membrane electrode made from 30% PVC and 67% plasticizer NPOE give good quality membrane.The membrane is then to be used to produce an ISE with good electrode performance based on its response sensitivity to mercury, and enhance the life time of the electrode for continuous uses.

The performance of ISE-Hg for mercury is influenced by the membrane composition in an electrode. Variation in the composition of polymer and plasticizer (PVC and NPOEa fixed amount of 3% DQDC and 50 mol% KTpClPB in the ingredient) on the electrode performanceis observed from the speed of sensing response to target analyte and the signal recovery to get constant baseline. The sensitivity of the electrode, it is based on the slope from the calibration curve of standard solution for the electrodes containing different ratios of plasticizer to PVC were summarized in Table 1. It is observed that the thickness of the membrane influenced the sensitivity of the electrode. The sensitivity of the electrode increased slightly with a decrease in the PVC content. The ISE-Hg made from the membrane electrode consisted of 30% PVC and 67% plasticizer NPOE is the most sensitive electrode for mercury. The higher the amount of plasticizer in the membrane improved the speed response of the electrode, but the mechanical properties of the electrode became much worse. The membrane was too soft, more permeable to water, easily swelling,internal solution is leaching, and results in a steady degradation of electrode performance.

The effect of the amount of KTpClPB (anionic additive) on the ion-selectivity of the membrane was examined in the PVC membrane electrode. The presence of the anionic additive aided the efficiency of the electrodes towards

the target cation by facilitating of the cation through the membrane. The addition of the anionic additive caused an increase in the response, a reduction of membrane resistance, an improvement of the speed of response, and a decrease in the effects of anions at high sample ionic strength. The best electrode response accompany with high selectivity was obtained when using 50 mol% KTpClPB of the ionophore content of DQDC in the electrode membrane. Increase the amount of KTpClPB in the membrane formula caused an increase in the sensitivity, but the selectivity towards primary ion reduced, because the selectivity was strongly governed by the ion selective feature of the $KTpClPB^-$ anions. The membrane electrodes prepared with 30% PVC, 67% plasticizer NPOE, 3% DQDC, and 50 mol% KTpClPB relative to the ionophore content provided good response to mercury. The membrane prepared at optimum condition has been observed by using SEM to obtain its surface at different amplification as shown in Figure 2. The SEM microscopy results showed that the contour of the membrane electrode containing ionophore**DQDC** is not very smooth but it has no pore.

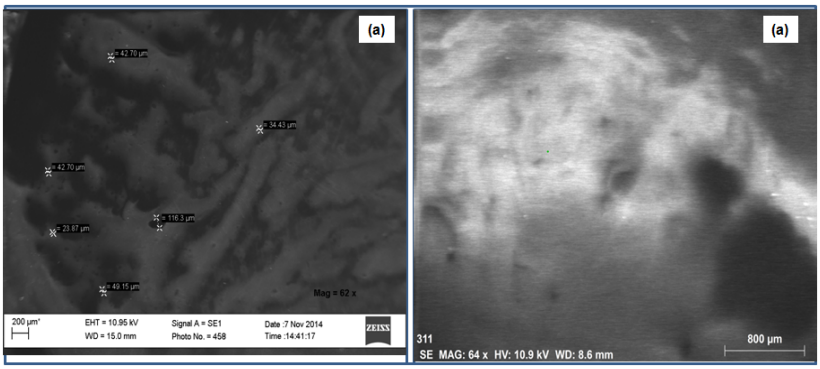


Figure 2. SEM images of PVC membrane electrode containing ionophore DQDCat different amplifications: (a) 200 µm, and (b) 800 µm

3.2 Determination of Hg^{2+} by Using ISE-Hg Electrodes

The electrode has been used for the determination of mercury in static system, and the potential was measured *versus* Ag/AgCl in solution containing of 10 mM potassium nitrate in nitric acid (pH 4.0). Typical signal output for the determination of series concentrations of mercury standard solutions and their calibration curvesare presented in Figure 3. Potential responses given by the ISE-Hg electrode to mercury ion increaseswhen increasing the concentration of Hg^{2+} in the solution (Figure 3A). The ISE-Hg has good performance in nitric acid solution (pH 4.0) where the response signal gives positive values, and constant base line was achieved at 30 seconds after the addition of mercury. There was no drifting signal in this condition, and the electrode sensitivity shows nearly Nernstian response. The calibration profile of ISE-Hg for mercury standard solution is presented in Figure 3C. The electrode gave linear response for 0.005 mM – 1.0 mM Hg^{2+}. A typical 95% confidence limit on the slope of ISE-Hg electrode is 25.82 mV per decade concentration of Hg^{2+}. The detection limit of an ISE-Hg electrode is 0.001 mM Hg^{2+}, where signal to noise ratio is two (S/N = 2). For comparison purposes, an ion selective electrode in the absence of ionophore in the membrane electrode (ISE-without DQDC) is also prepared. The electrode is used to determine mercury the same procedures as for ISE-Hg, and the signal profile for an ISE is shown in Figure 3B. There was no signal observed for mercury when the concentration of Hg^{2+} increased in the solution. The results have proved that the electrode responses to mercury are mainly generated by the presence of ionophore DQDC in the membrane electrode.

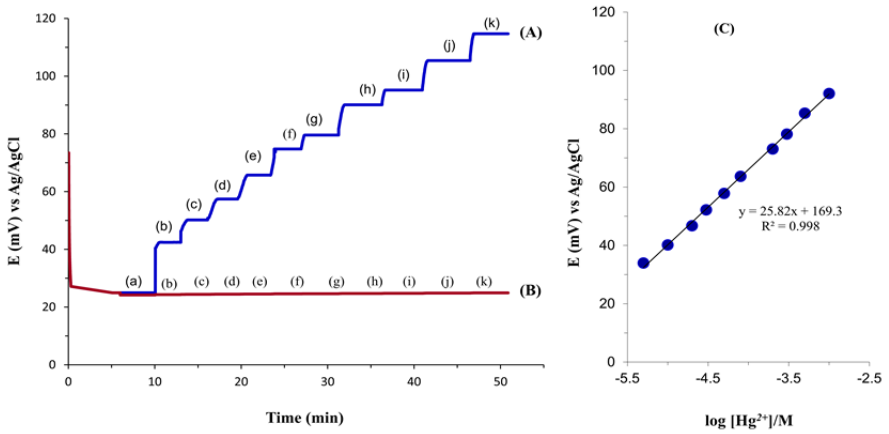


Figure 3. Typical response of ISE-Hg corresponding to successive addition of mercury in to nitric acid solution (pH 4.0) containing 10 mM KNO_3: (a) 0 µM, (b) 5 µM, (c) 10 µM, (d) 20 µM, (e) 30 µM, (f) 50 µM, (g) 0.1 mM, (h) 0.2 mM, (i) 0.3 mM, (j) 0.5 mM, and (k) 1.0 mM Hg^{2+}. (A) ISE-Hg containingionophore DQDC, (B) ISEwithout ionophore DQDC, (C) The calibration curve for mercury standard solution

3.3 Effect of DQDC on Electrode Performance

The effect of the quantity of ionophore in the membrane electrode was also investigated for electrode's sensitivity for mercury ion. The electrodes containing 1.5%, 2%, 3% and 4% DQDC respectively exhibited almost similar profileswith different in detection sensitivity (Figure 4). The electrode prepared with 3% ionophore DQDC in the electrode membrane gave good performance to mercury ion and good sensitivity was achieved, the slope is 25.82 mV per decade concentration of Hg^{2+}. Response speed of the ISE-Hg at this condition is very fast, is only need 30 second to reach baseline for determination of mercury. An electrode constructed with 1.5% DQDC in the membrane electrode give low sensitivity (slope of 24.13 mV). The lifetime of the electrode with low concentration of ionophore was reduced due to slow leaching of entrapped DQDC from membrane when the electrode was washed during repeated uses.The electrode give good response to mercury when the amount of ionophore in the membrane electrode is 2% DQDC (slope of 25.26 mV). However, the response sensitivity is obtained lower than the sensitivity given by the ISE-Hg electrode containing 3% DQDC in the membrane (slope is 25.82 mV). Furthermore, the crystals were observed in the electrode membrane when the content of ionophoreis increased to 4% DQDC. The electrode sensitivity at this condition was greatly reduced (slope of 22.74 mV). Therefore, the best condition on the preparation of ISE-Hg is chosen by immobilisation of ionophore 3% DQDC in the membrane electrode.

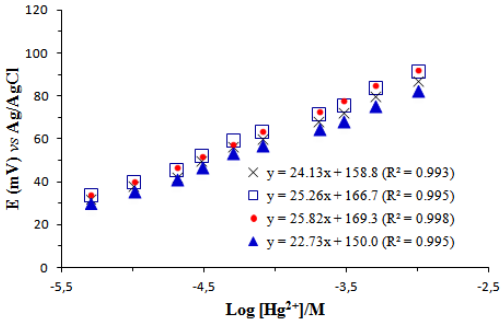


Figure 4. Calibration graphs for ISE-Hg for mercury when using different concentration of ionophore DQDC in electrode membrane: (×) ionophore 1.5%, (□) ionophore 2.0 mg, (●) ionophore 3.0 mg, and (▲) ionophore 4.0%

3.4 Selectivity of ISE-Hg

Selectivity of ISE-Hg has been tested for various types of cations, which are predicted to be interfering agents for mercury determination in real samples of polluted water samples. The response of ISE-Hg for 0.5 mM Hg^{2+} and the mixture of 0.5 mM $Hg^{2+}$with 0.1 mMeach of these cationsare listed Table 2. Those cations are

successively Mg^{2+}; Ca^{2+}; Zn^{2+}; Cu^{2+}; Cd^{2+}; Ni^{2+}; Al^{3+}; Fe^{3+}; and Sn^{3+}. It was found that all cations added into mercury standard solution gave small change in the potential compare to a standard 0.5 mM Hg^{2+}. Potential changes are vary between 1 - 5%, where copper ion was the highest(5%) response. However, at moderate concentration (0.1 mM) of all those cations added into mercury standard solution were not interfere the measurement of the ISE-Hg. The results reveal that the developed ISE-Hg electrode is selective for mercury. The presence of ionophore DQDCin the membrane electrode has proven to give sensitive and selective response to mercury ion.

Table 2. The potential value of an ISE-Hg to 0.5 mM Hg^{2+} standard solution and the mixture of 0.5 mM interfering cations and 0.5 mM Hg^{2+}

Mixture of Mercury and Interfering Cations	Electrode responses (E, mV)		
	Mean	Change in E*	% Change*
Hg^{2+}	85.83	-	-
Mg^{2+}	87.93	2.10	2
Ca^{2+}	88.75	2.92	3
Zn^{2+}	87.60	1.77	2
Cu^{2+}	90.40	4.57	5
Cd^{2+}	88.44	2.61	3
Ni^{2+}	88.44	1.90	2
Al^{3+}	88.50	2.67	3
Fe^{3+}	87.97	2.14	2
Sn^{3+}	86.51	0.68	1

*The electrode response for the mixture of 0.5 mM Hg^{2+} + 0.5 mM interfering agents is compared with 0.5 mM Hg^{2+}.

3.5 Stability of ISE-Hg

An ISE-Hg has been tested for its stability and the lifetime of the membrane electrode when the electrode is used to determine mercury at long period of time. Repeat determination of 0.5 mM Hg^{2+} (one measurement per day) was carried out for 30 days, and the relative resposess (%) of the electrode is presented in Figure 5. The electrode exhibited good stability at normal uses for one month. The electrode signal is steady reduce to become 94% after 30 days. This results is promising where the PVC is compatible to be used as polymeric matrix in the immobilisastion of ionophore in the membrane electrode. The study is still now still under investigation to see if any response decreased due to the ionophore leaching out from the membrane for longer time of uses.

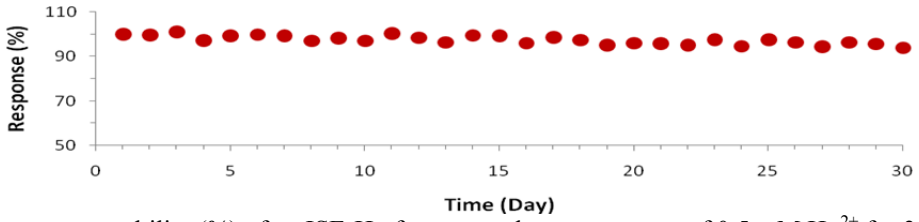


Figure 5. The response stability (%) of an ISE-Hg for repeated measurements of 0.5 mM Hg^{2+} for 30 days

4. Conclusion

The preliminary work on the construction of ISE-Hg by using ionophore DQDC in a membrane electrode has successfully done. This work demonstrated the compatibility of polymeric PVC membrane electrodes with ionophore DQDC as sensing agent in an ion selective electrode. The ISE-Hg give sensitive response to mercury ion, the working linearity range lies between 0.005 mM-1.0 mM Hg^{2+}, slope 25.82 mV per decade concentration of Hg^{2+}, and the detection limit 0.001 mM Hg^{2+}. This work is now still under investigation to obtain optimum conditions of the sensor for mercury.

Acknowledgement

The funding from Research Project Hibah Bersaing 2015 Universitas Negeri Medan (UNIMED), Directorate

General of Higher Education (DGHE), Ministry of Research, Technology and Higher Education is acknowledged. Special thank is also expressed to Jamalum Purba, Melinda Lena Lamria, Henni Cintiya, and Kiki Agnesia Putri Br Sinulingga, staff of The Department of ChemistryUniversitas Negeri Medan, for their help to synthesis DQDC compound used in this study.

References

Chatterjee, S., Pillai, A., & Gupta, V. K. (2002). Spectrophotometric determination of mercury in environmental sample and fungicides based on its complex with *o*-carboxy phenyl diazoamino*p*-azobenzene, *Talanta, 57*(3), 461 - 465. http://dx.doi.org/10.1016/S0039-9140(02)00037-1

da Silva, A. F., Welz, B., & Curtius, A. J. (2002). Noble metals as permanent chemical modifiers for the determination of mercury in environmental reference materials using solid sampling graphite furnace atomic absorption spectrometry and calibration against aqueous standards, *SpectrochimicaActa Part B: Atomic Spectroscopy, 57*(12), 2031-2045. http://dx.doi.org/10.1016/S0584-8547(02)00175-1

de la Riva, B. S. V., Costa-Fernández, J. M., Jin, W. J., Pereiro, R., & Sanz-Medel, A. (2002). Determination of trace levels of mercury in water samples based on room temperature phosphorescence energy transfer, *AnalyticaChimicaActa, 455*(2), 179-186. http://dx.doi.org/10.1016/S0003-2670(01)01597-5

de la Riva, B. S. V., Costa-Fernández, J. M., Pereiro, R., & Sanz-Medel, A. (2000). Fluorimetric method for the determination of trace levels of mercury in sea water using 6-mercaptopurine. *AnalyticaChimicaActa, 419*(1), 33-40. http://dx.doi.org/10.1016/S0003-2670(00)00979-X

Hassan, S. S., Mahmoud, W. H., Mohamed, A. H., &Kelany, A. E. (2006). Mercury(II) ion-selective polymeric membrane sensors for analysis of mercury in hazardous wastes. *Anal Sci., 22*(6), 877-881. http://dx.doi.org/10.2116/analsci.22.877

Hu, O., Yang, G., Yin, J., & Yao, Y. (2002). Determination of trace lead, cadmium and mercury by on-line column enrichment followed by RP-HPLC as metal-tetra-(4-bromophenyl)-porphyrin chelates. *Talanta, 57*(4), 751-756.

Izgi, B., Demir, C., & Güçer, S. (2000). Application of factorial design for mercury determination by trapping and graphite furnace atomic absorption spectrometry, *SpectrochimicaActa Part B: Atomic Spectroscopy, 55*(7), 969-975. http://dx.doi.org/10.1016/S0584-8547(00)00234-2

Khan, H., Ahmed, M. J., & Bhanger, M. I. (2005). A simple spectrophotometric determination of trace level mercury using 1,5-diphenylthiocarbazone solubilized in micelle. *Anal Sci.*, *21*(5), 507-512. http://dx.doi.org/10.2116/analsci.21.507

Lee, T. H., & Jiang, S. J. (2000). Determination of mercury compounds by capillary electrophoresis inductively coupled plasma mass spectrometry with microconcentric nebulization, *AnalyticaChimicaActa, 413*(1-2), 197-205. http://dx.doi.org/10.1016/S0003-2670(00)00807-2

Li, H., Zhang, Y., Zheng, C., Wu, L., Lv, Y., & Hou, X. (2006). UV irradiation controlled cold vapor generation using $SnCl_2$ as reductant for mercury speciation., *Anal Sci.*, *22*(10), 1361-1365. http://dx.doi.org/10.2116/analsci.22.1361

Li, J., He, F., & Jiang, C. Q. (2006). Highly sensitive spectrofluorometric determination of trace amounts of mercury with a new fluorescent reagent, 2-hydroxy-1-naphthaldehydene-8-aminoquinoline. *Anal Sci.*, *22*(4), 607-611. http://dx.doi.org/10.2116/analsci.22.607

Lu, J., He, X., Zeng, X., Wan, O., & Zhang, Z. (2003). Voltammetric determination of mercury (II) in aqueous media using glassy carbon electrodes modified with novel calix[4]arene. *Talanta, 59*(3), 553-560. http://dx.doi.org/10.1016/S0039-9140(02)00569-6

Majid, S., Rhazi, M. E., Amine, A., & Christopher, M. A. (2002). An amperometric method for the determination of trace mercury(II) by formation of complexes with L-tyrosine. *Analytica Chimica Acta, 464*(1), 123-133. http://dx.doi.org/10.1016/S0003-2670(02)00422-1

Mondal, B. C., & Das, A. K. (2003). Determination of mercury species with a resin functionalized with a 1,2-bis(*o*-aminophenylthio)ethane moiety. *Analytica Chimica Acta, 477*(1), 73-80. http://dx.doi.org/10.1016/S0003-2670(02)01372-7

Moreda-Piñeiro, J., López-Mahía, P., Muniategui-Lorenzo, S., Fernández-Fernández, E., & Prada-Rodríguez, D. (2002). Direct mercury determination in aqueous slurries of environmental and biological samples by cold vapour generation–electrothermal atomic absorption spectrometry. *Analytica Chimica Acta*, *460*(1), 111-122.

http://dx.doi.org/10.1016/S0003-2670(02)00137-X

Moreno, R. G. M., de Oliveira, E., Pedrotti, J. J., & Oliveira, P. V. (2002). An electrochemical flow-cell for permanent modification of graphite tube with palladium for mercury determination by electrothermal atomic absorption spectrometry. *Spectrochimica Acta Part B*: *Atomic Spectroscopy*, *57*(4), 769-778. http://dx.doi.org/10.1016/S0584-8547(02)00009-5

Páger, Cs., &Gáspár, A. (2002). Possibilities of determination of mercury compounds using capillary zone electrophoresis. *Microchemical Journal*, *73*(1-2), 53-58. http://dx.doi.org/10.1016/S0026-265X(02)00050-4

Purba, J., Sibuea, G. V., Tarigan, M. L., Fonica, A., & Situmorang, M. (2013). Sintesis Ionofor Sebagai Bahan Aktif Ion Selektif Elektroda (ISE) UntukAnalisis Penentuan Ion Logam Berat Di Dalam Sampel Lingkungan. *Jurnal Penelitian Saintika*, *13*(2), 94-104.

Purba, J., Sinaga, M., & Situmorang, M. (2012). Sintesis Ionofor Sebagai Bahan Aktif Ion Selektif Elektroda (ISE) Untuk Analisis Penentuan Ion Logam Berat, *Prosiding Seminar dan Rapat Tahunan BKS PTN-B Bidang MIPA di Hotel Madani Medan,*Tgl 11-12 Mei 2012, p. 181-185.

Purba, J., Tarigan, M. L., & Situmorang, M. (2014). Pembuatan Elektroda Ion Selektif Merkuri Menggunakan Bahan Aktif Ionofor DTODC (Construction Of Ion Selective Electrode For Mercury By Using Ionofor DTODC, *Prosiding Seminar dan Rapat Tahunan BKS PTN-B Bidang MIPA di Bogor,* Tgl 9-11 Mei 2014.

Purba, J., Zainiati, & Situmorang, M. (2013). Sintesis Ionofor Sebagai Bahan Aktif Ion Selektif Elektroda (ISE) Untuk Analisis Penentuan Logam Merkuri (Hg), *Prosiding Seminar Hasil Penelitian Lembaga Penelitian Unimed Tahun 2013 Bidang Sain, Teknologi, Sosial, Bahasa dan Humaniora,* Tgl 14-16 November 2013, 28-35.

Purba, J., Zainiati, Samosir, E. A., &Situmorang, M. (2013). Pembuatan Ion Selektif Elektroda Menggunakan Ionofor DTODC Untuk Penentuan Merkuri (ISE-Hg), *Prosiding Seminar dan Rapat Tahunan BKS PTN-B Bidang MIPA di Bandar Lampung,* Tgl 10-12 Mei 2013, 207-211.

Rizea, M. C., Bratu, M. C., Danet, A. F., & Bratu, A. (2007). Determination of mercury in fish tissue using a minianalyzer based on cold vapor atomic absorption spectrometry at the 184.9 nm line. *Anal Sci., 23*(9), 1121-1215. http://dx.doi.org/10.2116/analsci.23.1121

Silva, M. F., Tóth, I. V., & Rangel, A. O. (2006). Determination of mercury in fish by cold vapor atomic absorption spectrophotometry using a multicommuted flow injection analysis system. *Anal Sci., 22*(6), 861-864. http://dx.doi.org/10.2116/analsci.22.861

Situmorang, M. (1992). *Enzymatic Assay For Glucose And Cholesterol Using Flow Injection Potentiometry With A Tungsten Wire Electrode,* M.Sc Thesis, The University of New South Wales, Australia.

Situmorang, M., (2005). Pembuatan Sensor Potensiometri Dalam Sistem Flow Injeksi Analisis Untuk Penentuan Timbal Menggunakan Ionofor Diazacrown. *JurnalSain Indonesia, 29*(2), 55-61.

Situmorang, M., Alexander, P. W., & Hibbert, D.B., (1998). Flow injection potentiometry for enzymatic assay of cholesterol with a tungsten electrode sensor, *Talanta, 49*,639-649.http://dx.doi.org/10.1016/S0039-9140(99)00057-0

Situmorang, M., Gooding, J. J., & Hibbert, D. B. (1999). Immobilisation of Enzyme Throughout a Polytyramine Matrix: A Versatile Procedure for Fabricating Biosensors. *Analytica Chimica Acta, 394*(2-3), 211-223. http://dx.doi.org/10.1016/S0003-2670(99)00291-3

Situmorang, M., Gooding, J. J., Hibbert, D. B., & Barnett, D. (2001). Development of Potentiometric Biosensors Using Electrodeposited Polytyramine as the Enzyme Immobilisation Matrix, *Electroanalysis, 13*(18), 1469-1474. http://dx.doi.org/10.1002/1521-4109(200112)13:18<1469::AID-ELAN1469>3.0.CO;2-U

Situmorang, M., Gooding, J. J., Hibbert, D. B., & Barnett, D. (2002). The Development of a Pyruvate Biosensor Using Electrodeposited Polytyramine. *Electroanalysis, 14*(1), 17-21.http://dx.doi.org/10.1002/1521-4109(200201)14:1<17::AID-ELAN17>3.0.CO;2-O

Situmorang, M., Purba, J., Lamria, M. L., Cintiya, H., Sinulingga, K. A. P., & Sihombing, E. (2014). Sintesis Ionofor DTODC Sebagai Bahan Aktif Dalam Elektroda Ion Selektif Penentuan Merkury(ISE-Hg), *JurnalPenelitianSaintika*.

Situmorang, M., Simarmata, R., Napitupulu, S. K., Sitanggang, P., & Sibarani, O. M. (2005). Pembuatan Elektroda Ion Selektif Untuk Penentuan Merkuri (ISE-Hg). *Jurnal Sain Indonesia, 29*(4), 126-134.

Yang, X. H., Hibbert, D. B., & Alexander, P. W. (1997). Continuous flow analysis of lead (II) and mercury (II) with substituted diazacrown ionophore membrane electrodes. *Talanta, 45*, 155-165. http://dx.doi.org/10.1016/S0039-9140(97)00122-7

Yang, X. H., Hibbert, D. B., & Alexander, P. W. (1998). Flow Injection Potensiometry by PVC – Membrane Electrodes with Substituted AzacrownIonophore for the Determination of Lead (II) and Mercury (II) Ion. *Analitica Chemica Acta, 372*, 387-398. http://dx.doi.org/10.1016/S0003-2670(98)00382-1

Yantasee, W., Lin, Y, Hongsirikarn, K., Fryxell, G. E., Addleman, R., & Timchalk, C. (2007). Electrochemical sensors for the detection of lead and other toxic heavy metals: The next generation of personal exposure biomonitors, *Environ Health Perspect, 115*(12), 1683-1690. http://dx.doi.org/10.1289/ehp.10190

Ye, G., Chai, Y., Yuan, R., & Dai, J. (2006). A mercury(II) ion-selective electrode based on N,N-dimethylformamide-salicylacylhydrazone as a neutral carrier. *Anal Sci., 22*(4), 579-582. http://dx.doi.org/10.2116/analsci.22.579

Yoon, S., Albers, A. E., Wong, A. P., & Chang, C. J. (2005). Screening mercury levels in fish with a selective fluorescent chemosensor. *J Am Chem Soc., 127*(46), 16030-16031. http://dx.doi.org/10.1021/ja0557987

Zhang, Z., Chen, S., Yu, H., Sun, M., & Liu, W. (2004). Simultaneous determination of arsenic, selenium, and mercury by Ion exchange-vapor generation-inductively coupled plasma-mass spectrometry. *Analytica Chimica Acta, 513*(2), 417-423. http://dx.doi.org/10.1016/j.aca.2004.03.006

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Developing a Mobile Based Automated Testing Tool for Windows Phone 8

Albert Mayan. J[1], Julian Menezes. R[2] & John Bruce. E[3]

[1] Associate Professor, Department of Computer Science and Engineering, Sathyabama University, India

[2] P.G Scholar, Department of Information Technology, Sathyabama University, India

[3] Associate Professor, Department of Computer Science and Engineering, Sathyabama University, India

Correspondence: Albert Mayan. J, Associate Professor, Department of Computer Science and Engineering, Sathyabama University, Chennai-119, India. E-mail: albertmayan@gmail.com/mtechit2k13@gmail.com/johnbruce@sathyabamauniversity.ac.in

Received: February 6, 2015 Accepted: February 18, 2015 Online Published: July 25, 2015

doi:10.5539/mas.v9n8p91 URL: http://dx.doi.org/10.5539/mas.v9n8p91

Abstract

Smart phones, or Mobile phones are quickly fetching the essential computer and communication tool in people's life. Every phone has lots of applications and every application has its own different characteristics. Before producing these applications to the end user's use, the developer should confirm that the applications are working smoothly, sans any technical glitches, and user friendly in every feature, and for that we use testing tools to check the compatibility of the software by using test applications like Eggplant, silk test, etc. But nowadays, every tool is designed for the desktop environment. In this project an application is being proposed for the windows phones by which an end-user can install the application in the mobiles directly. After this work the end user will be able to know if the apps are working properly or not, this will help us to catch all the information by executing the data, detecting the type and the kernel. This application will help us to install any application on our windows phone by the other devices and will help to grab the elements in the present applications. It is much faster and user friendly, because it will work without the need of a desktop, and we can run some more test cases as well.

Keywords: testing, mobile testing, windows phone, phone testing, windows 8

1. Introduction

The Software development technologies, usually makes an apparent difference amidst functional testing and security testing. An automatic software testing is a function which verifies, that an executing component of source code is working well or not. Mobile application, also known as Mobile apps is a word to describe the application that was developed for low-powered devices. For, getting executed on the mobile, tab, computer, iPad or other electronic computer devices, a sort of application software is designed; it is termed as Mobile Apps or just a mobile app. An Application that creates a specific logic is preferred, if the aim is to interact with users, or to deliver an app that needs more work similar to a program than a web URL. A usual mobile structure differs in several conditions, the functionality of the device over the user's requirements, the accessing time, the operation in several locations and how it gets connected to the server in different ways.

The cell phone structure basically contains, a cell phone device, wired or wireless connectivity, backend server and an underlying network infrastructure. The application might be a stand-alone, or self-contained. The app depends on the capabilities of the hardware, or the application will run at the server end by remote and the mobile acts as a platform for interacting and viewing.

The mobile app will run on the server or the mobile device, the work of the mobile is dependent on the characteristics of the structure, the device potential and the behavior of the mobile applications. From the core business apps to the Enterprise environment there arises a requirement for an extra complicated situation to run these applications, which are providing mobile app interfaces in mobiles to communicate with the features, of some classic apps by the mobile clients.

Nowadays, Mobile devices have been built with the powerful Processors, Software execution abilities, storage, connectivity features, multi-media delivery with powerful Audio / Video, Accelerometers, Sensors and GPS

facilities.The inbuilt camera in the phone, is not only helping the users to take a photo or record the video, but it also helps the user in face identification. The variety among the mobile apps which catches the imagination of the human brain is increased by the software development team. These applications are widely spread through several platforms such as Nokia, OVI, Lumia stores and the country service providers. The software which has been made for any specific platform may or may not run on all the devices that have been vended by the developers. Even if the platform is same, but the device compatibilities are different, then the apps will be unsuitable for other devices.

The mobile apps can be developed easily with a little bit of testing or without testing and dispersed without brief testing, because the mobile app developer focuses on the software functionalities instead of the device configuration, and approximately all the software testing is carried on emulators instead of the devices. The errors that exist in the app may harm the device, along with the application in which it is present. The requirement of the testing procedure assures the developer for the successfulness; venture to make the footprints in the mobile field. There are some determining studies that describe how to make a successful application.

The existing applications are all desktop based and not a mobile based application and it can test generalized windows 8 mobile applications through desktops alone. The main drawback lies in the complexity of the applications. Most of them are difficult to comprehend. As a result, this research work introduces a new mobile based automated testing tool for windows phone to resolve the above mentioned issues. In this proposed system, the automated testing tool will run on the windows phone itself. By utilizing the windows' simulation software (emulator) and Microsoft visual studio for Phone app development, the end-user can launch an application using another Windows application. The product thus created will be a mobile to mobile testing with both the testing applications and the application to be tested to be present on the same windows phone itself.

2. Related Work

Eggplant is a black-box GUI testing tool, which was given inception by Redstone Software in 2002. Redstone was acquired in 2008 by a UK-based company called Test Plant. Its approach uses image matching technology as opposed to looking at the object-level of the application being tested. This permits for a full system automation, of a system-under-test as compared to that of a solution related to an application. It also states that the technology, which is used to build the application to be tested is not a concern, nor does the system that the application runs on.

Currently, some of the apps which are available for downloads, can run on various mobile platforms such as NOKIA, and windows phones. However, the developers aren't making the public become aware, about the information, how the application will function on various platforms. Agile software is a group of development methods which is based on the growth of the apps, where the needs and the solution has been involved from the collaboration method amid cross-functional team. It is a theoretical framework which promotes interaction between the development cycles.The known author Dr. Jeff Sutherland and Ken Schwaber formed agile software development which is used by approximately, all the companies.

Kent Beck introduced a technique called Extreme programming, which is a well known process in agile process, that is used for lightweight, low-risk, predictable, flexible and scientific way for software development. In this process the client or user works with the development team and gets a high priority for software testing on iteration basis.

The most lightweight methods for developing the software are crystal methodology. It has some types such as crystal clear, crystal orange, crystal yellow, by which characteristics are obtained by project priorities and team size. The crystal method needed policy standard, certain roles and products. This method is applied to the team size of 6-8 members. This method mainly focuses on the requirement of the client or user instead of the artifacts. The author in has proposed a new technique that is known as FDD (Feature Driven Development). This technique combines some practice in one methodology. These are all the techniques which are functioning through the client viewpoint.

The methodology for documenting and modeling the software is based on practice. It is based on the collection of principles and the values of Agile modeling and it is more flexible. This technique is the supplement of the XP (Extreme processing), Scrum and RUP (Rational Unified process).

3. Proposed Work

3.1 Architecture

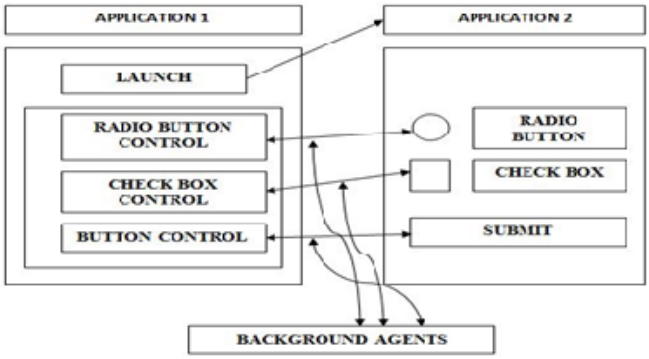


Figure 1. Proposed System Architecture

3.2 Proposed Architecture Overview

The proposed system developed a mobile based automated testing tool for handling various installed applications in the windows mobile. The proposed testing tool uses the following components such as Microsoft visual studio 2013, Windows Phone 8, Software Development Kit 8. Initially, a source code is developed to launch one application using another application in windows operating system version 8.1. After that, a foreground as well as a background application is created. The foreground windows application is installed in one Windows PC or Windows phone, as mentioned in Fig.1. Then, the background application is installed on the same windows phone. After installing foreground and background applications, the user launches foreground application on the Windows phone. Using a URI (Uniform Resource Identifier) an end-user can launch the background application. Similarly, with the help of user Interface (UI) the user can also launch various background applications. The proposed application launches any other application present in our windows8 phone and it grabs the objects present in the particular application.

3.3 Implementation

This stage describes the details about the implementation of our proposed mobile based automated testing tool. Our proposed automated tool utilizes Emulator for connecting Windows Phone with PC. Detail description of Emulator is given below:

3.3.1 Emulator

The proposed tool is developed with the great help of an EMULATOR. The use of these emulators are extremely useful in the initial stages of the development of the application, because they allow quick and efficient checking of the progress, of the application. An emulator is a system that runs software from one environment to another without changing the integrity of the software itself. It duplicates the features and works of a real system.

Some advantages of using an emulator are:

- All applications can be launched and subjected to test on an emulator, without the need for investing in a new mobile handset for various versions of windows phone.
- Emulators are mostly available for free, and we can also perform stress, UI and performance testing using it.
- Emulators are easy to use without the need to have knowledge of any complex features in it.
- The version of the Emulator involved in the project is "**Emulator WVGA 512MB**".

3.4 Working Procedure

3.4.1 Launching One Windows form Application

Initially two simple windows form applications are created, one application is Installed in windows 8 phone and another application is installed on windows PC using Microsoft visual studio 2013 and windows emulator. After that, one windows form application is launched from another windows window form application. It is made quite easier for us to understand, how one application launches another application, whether it may be desktop to

desktop, or mobile based application to another mobile based application.

3.4.2 Connecting One Application to another Application

Generally, the function call helps the user to access two functions simultaneously. The Function call works based on the function command which starts the execution, when the users calls the respected function. After launching an application, we can successfully call the functions of the elements of another application from the windows forms application.

3.4.3 Automating the Tests Performed on Windows Forms.

This module describes the details of the automated testing process using windows forms. The next approach adopted was to automate all the tests which was performed by calling the functions of the elements, of the other foreground application from the background application. The automated process was done by looping method. The Looping method helps the testing process by making a loop in the background application/form for all the functions in the foreground application/form. It was done by launching a windows 8 phone application on emulator 8/8.1.

3.5 Pseudo Code for Proposed Work

```
namespace PhoneApp7
{
public partial class MainPage : PhoneApplicationPage
{
  Public MainPage() // Constructor
{
  InitializeComponent();}
Public async void Button_Click_1(object sender, RoutedEventArgsrea)
{ // Launch URI.
  awaitWindows.System.Launcher.LaunchUriAsync(newSystem.Uri("PhoneApp7"));
} } }
```

4. Results and Discussions

To evaluate the performance of the proposed automated Testing tool, a series of experiments on the installed applications in windows 8 phone were conducted. In these experiments, the proposed automated tool was implemented and evaluated, and it requires the following configurations. The Table 1 below lists out all the Hardware and the Software required for the development of the Mobile Testing Tool.

Table 1. Hardware & Software Requirements

S.No	REQUIREMENTS
1.	Windows 8 Mobile Phone.
2.	Processor: Intel(R) COre(TM) i7-3770K CPU @ 3.90 GHz. RAM: 8.00 GB Hard Disk Drive: 500GB. Operating System: Windows 8.1 x64 (Professional).
3.	Microsoft Visual Studio 2013. Software Development Kit 8.
4.	Emulator WVGA 512MB

4.1 Experiments

By using the above mentioned components the proposed testing tool can efficiently test the applications.

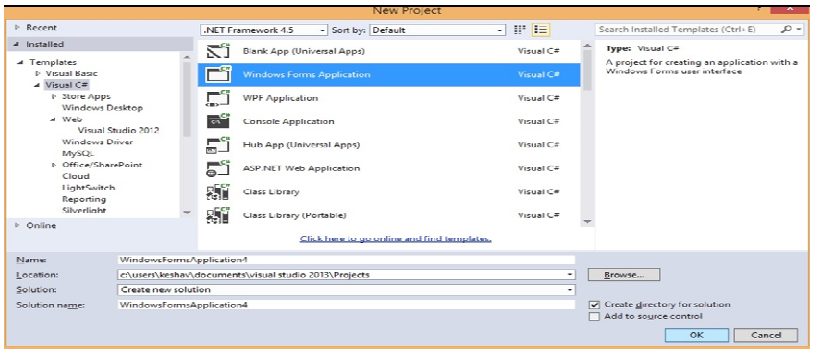


Figure 2. Windows Forms Application (Foreground)

In the first step, the Windows forms application in .Net framework 4.5 with the Visual C# is chosen. This framework is used for launching one application from another application.

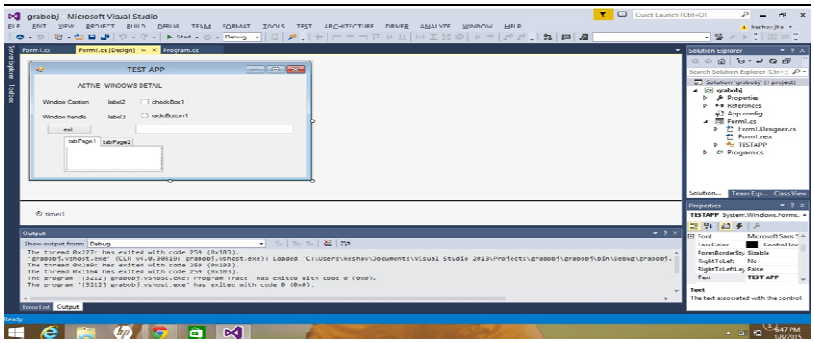


Figure 3. Test Application

After choosing the Windows form application the developer is in need to draw a set up for the Test application for both the devices, (Mobile and Desktop) as well. In this step the developers generate, the internal Modeling for the software which is communicating with other designed software.

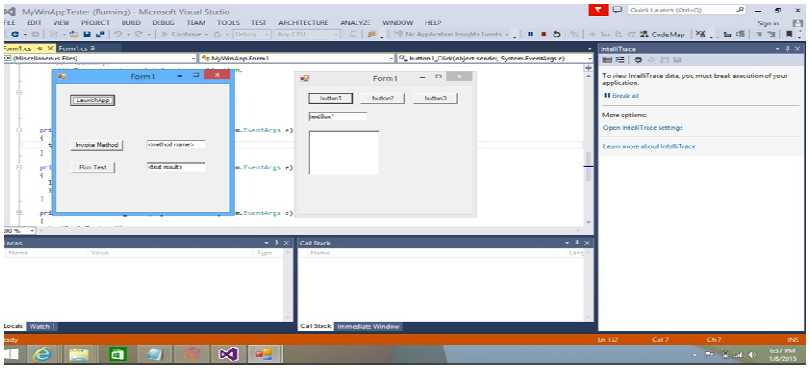


Figure 4. Launching Foreground application

Both the software Designs are finalized and the software is processed for installing in the Windows phone and Desktop. These softwares will communicate with each other for reading the contents of the other device.

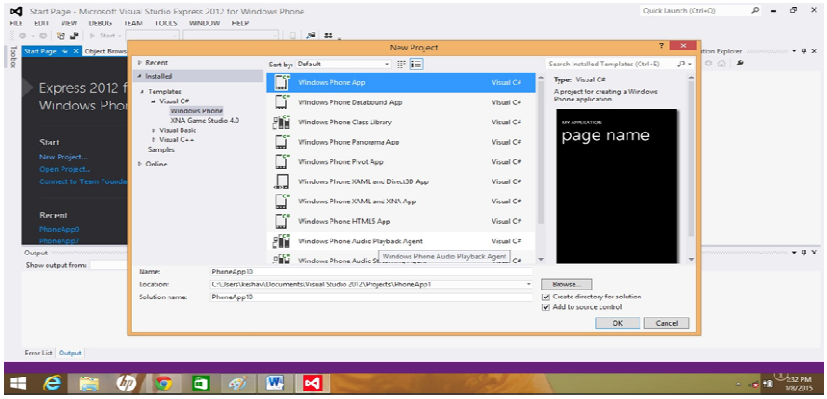


Figure 5. Connecting Background Application using URI

The Apps created by .Net framework 4.5 is accessed on the Mobile side (background) which will communicate with the (Desktop side) foreground application. The installed application shows the status for the other installed application as well. The (PC side) foreground installed application has to start the new page title or the page name.

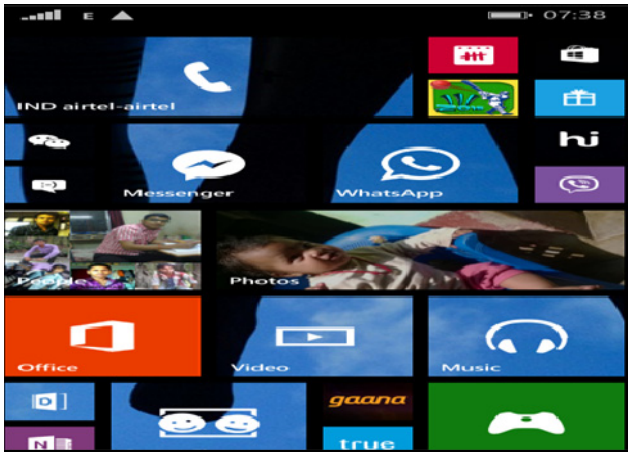


Figure 6. Launching Background Application

The installed application is getting connected with the background application in Windows phone. After launching an application, we can successfully call the functions of the elements of another application from the tester windows forms application.

Figure 7. Expected Output

Finally the communication has been established over both ends. By using this application the end-user can launch any other application installed in windows mobile and after launching that app, it will grab all the objects and elements present in the launched application.

4.2 Performance Evaluation

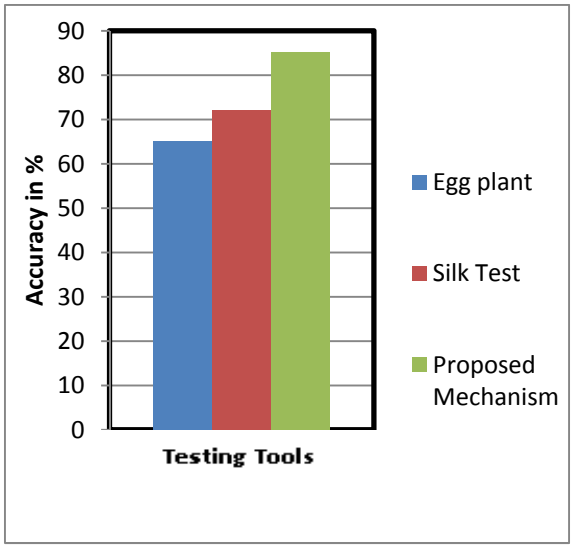


Figure 8. Accuracy of Testing Tools

The Figure 8 shows the testing accuracy of proposed mobile based Automated Testing Tool. It does not depend on the use of a Hardware apart from the Mobile itself. It tailors the need of an organization for one unique purpose, unlike other testing tools. It can be launched from within a Mobile and takes less or no time to trigger the testing process. As a result, our proposed testing mechanism achieves maximum testing accuracy than other traditional testing tools like eggplant, silk test, etc.

5. Conclusion

In this research, the proposed System implements a mobile based automated testing to efficiently test the applications in Windows phone 8. This system utilized the mentioned tools for testing the foreground application by using the background application on windows phone. In this proposed automated testing tool, the user can clutch the techniques and substances present in an active window on a desktop screen, the same technique is used to perform the action similar for an active window in a Windows application, but the way is unclear on how to install this application from desktop to windows phone. Utilizing the proposed work it is possible to do a direct installation directly in the windows phone through the desktop environment.

References

A Guide to Emulators . Retrieved from http://mobiforge.com

Agile Modelling. Retrieved from http://en.wikipedia.org/wiki/Agile_Modeling

Albert, M. J., & Ravi, T. (2015). Structural Software Testing: Hybrid Algorithm for Optimal Test Sequence Selection during Regression Testing. *International Journal of Engineering and Technology (IJET), 7*(1), 2015.

Crystal Clear (2013). Retrieved from http://en.wikipedia.org/wiki/Crystal_Clear_(software_development)

Extreme Programming. (1996). Retrieved from http://www.extremeprogramming.org

Feature Driven Development. (2009). Retrieved from http://en.wikipedia.org/wiki/Feature_Driven_Development

Hyo-Eun, E., & Lee, S. W. (2013). Human-centered software development methodology in mobile computing environment: agent supported agile approach. Eom and Lee EURASIP. *Journal on Wireless Communications and Networking, 111*. Retrieved from http://jwcn.eurasipjournals.com/content/2013/1/111

IEEE. (2000). IEEE Standard for Software Test Documentation. IEEE Std 829.

Ken, S., & Sutherland, J. (2011). The Definitive Guide to Scrum: The Rules of the Game.

Kent, B., (2004). Extreme Programming explained (2nd ed.). Addison- Wesley, Boston.

Malloy, A. D., Varshney, U., & Snow, A. P. (2002). Supporting mobile commerce applications using dependable wireless networks. *Mobile Networks and Applications, 7*(3), 225-234.

Mobile Stats. Retrieved from http://www.slideshare.net/vaibhavkubadia75/mobile-web-vs-mobile apps-27540693? from_search=1

Ondrej, K., Jakub, J., & Dalibor, J. (2011). Use of Mobile Phones as Intelligent Sensors for Sound Input Analysis and Sleep State Detection. ISSN 1424-8220. http://dx.doi.org/10.3390/s110606037

Selvam, R., & Karthikeyani, V. (2011). Mobile Software Testing-Automated Test Case Design Strategies. *International Journal on Computer Science and Engineering (IJCSE), 3*(4). ISSN: 0975-3397

Testlabs Blog. (2010). Top 10 Tips for Testing iPhone Applications. Retrieved September, 2010, from http://blog.testlabs.com/search/label/iPhone

Wooldridge, D., & Michael, S. (2010). The business of iPhone app development: Making and marketing apps that succeed. Apress.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Introduction to Bridge Construction in a Mountainous City

Xiaoxia Zhai[1]

[1] Chongqing Jiaotong University, China

Correspondence: Room 105, Unit 2, Building 12, Knowledge Garden Community, Chongqing Jiaotong University, Chongqing, China. E-mail: 286026317@qq.com

Received: March 28, 2014 Accepted: May 7, 2015 Online Published: July 27, 2015

doi:10.5539/mas.v9n8p99 URL: http://dx.doi.org/10.5539/mas.v9n8p99

Abstract

Bridges in a mountainous city are different in style from other bridges. They are grand in opening and closing and are often recognized as important symbols of a city. Together with other elements, they form beautiful scenery of a mountainous city. This paper focuses on Chongqing, a representative mountainous city to study and analyze characteristics and features of bridges in a mountainous city, learn from experience in designing bridges in a mountainous city and provide theoretic reference for future bridge construction in a mountainous city.

Keywords: Mountainous city, features of bridge in a city, bridge design

1. Introduction

Bridge Construction in a Domestic Mountainous City. In Western China where mountains are gathered together, most mountainous cities are built along rivers, for which typical examples are given to Chongqing City and its other cities like Wanzhou City and Peiling City. In addition, cities that are mountainous include Yibin City and Panzhihua City in Sichuan Province, Guiyang City in Guizhou Province, Nanning City in Guangxi Province, and Hong Kong, a mountainous city built along mountains and seas in the coast of South China.

1.1 Brides of Chongqing

Chongqing City locates between Zhongliang Mountain and Tongluo Mountain and in river valleys and hilly areas where the Jialing River and the Yangtze River flow over. It is built by relying upon mountains, with roads being uneven and buildings well-arranged. Therefore, it is recognized as the largest and the most famous mountainous city in China. Major rivers flowing through Chongqing City are the Yangtze River, the Jialing River, the Wujiang River, the Fujiang River, Qijiang River, Daninghe River, etc. Main stream of the Yangtze River flows through the whole city from west to east, and its flow path is 665 km long; it passes three anticlines of Wushan Mountain, creating famous Qutang Gorge, Wuxi Gorge, and Xiling Gorge (located in Hubei Province), i.e., the Three Gorges on the Yangtze River. The Yangtze River and the Jialing River flows through the main urban zone of Chongqing City.

(1) Jiayue Bridge

Jiayue Bridge is a Y type extradosed cable-stayed bridge. Its main bridge is 778 meters long and designed with dual 6-lane carriage and speed per hour of 80 km. Jiayue Bridge is structured with two stories, the upper storey for vehicles and the lower storey for pedestrians. At each side of the road, there is a 3.5m wide “sightseeing corridor”, which is first seen in China. At the two bridge towers, there will be more “sightseeing balconies”, 4 inn total, each being about over 20 m^2. Citizens can take a walk on the sightseeing corridor and stand for sightseeing.

Figure 1. Jiayue Bridge

(2) Chaotianmen Yangtze River

Figure 2. Chongqing Chaotianmen Yangtze River

Chongqing Chaotianmen Yangtze River locates in Wangjiatuo River Section of the Yangtze River, about 2.4 km on the downstream of the junction between the Jialing River and the Yangtze River, in the central business district of the main urban zone of Chongqing and on the east to west rapid transit radiated from the main urban area of Chongqing City.

As an extra-large arch bridge for both road and pathway, the bridge is of double-deck type, with the upper deck designed with dual 6-lane carriage and the lower deck designed with 2 reserving lanes and 2 dual traffic rails. The bridge is 1.74 km long in total. Its main bridge adopts three-span continuous steel truss tied arch bridge (190m+552m+790m); the main bridge is 36.5m wide, the truss is 29m wide; the two side spans are structured with trussed beams of variable height; the mid-span is structured with steel truss tied arch. The height between arch top and middle support point is 142m, arch rib's lower bottom chord adopts second-degree parabola, vector height is 128m and rise-span ratio is 1/4.3125; arch rib's upper chord also adopts second-degree parabola and it smoothly transits to side span's upper chord line through a circular curve (R=700m). The main truss adopts N

type truss of variable height, arch rib's truss mid-span is 14m high, and truss height at the middle support point is 73.13m high (of which the arch rib's reinforced chord is 40.65m high) and the truss height at side-span's support point is 11.83m.

Technically, Chaotianmen Yangtze River is the first flying swallow multi-rib steel truss mid-level arch bridge with the main span of 552m. It is the largest one among similar bridges in span and has successfully developed 145000KN anti-shock support base of the largest tonnage in the world. The bridge combines tied rod for steel structure and pre-stressed tied rod; structurally, steel structure tie rod is part of steel truss girder and its surface coincide with with the main truss surface. The connecting structure between tie rod and main truss arch is simple, with clear load bearing. Besides, this bridge adopts deck unit erection crane, and cable stayed knotting technology, as well as the theory of raising beam elevation to enable the main bridge to rotate, so as to realize arch first, beam second and zero stress closure mode, which is the first example in the world and the bridge line is easily guaranteed.

(3) Caiyuanba Yangtze River Bridge

Figure 3. Caiyuanba Yangtze River Bridge

Chongqing Caiyuanba Yangtze River Bridge is a large-span arch bridge used for both public transport and urban light rail. Its main span is 240m, enabling it to be the arch bridge of the second largest span. Its steel structure is 18,000 tons in total. It is a center-supported zero thrust steel pipe concrete tie rod arch bridge in structure and a modern bridge that has integrated various new bridge structure types (e.g. steel pipe arch, steel box girder and steel truss girder) and scientific and technological achievements.

2. Features of Bridges in Mountainous Cities at Home and Abroad

We can tell that main features of bridges in mountainous cities are as follows, through analyzing bridges in mountainous cities at home and abroad:

(1) Bridges in a mountainous city must be planned according to urban zone road network of the mountainous city. Layout of a mountainous city will have a great impact upon traffic of a mountainous city. In a mountainous city, roads are built by using mountains; complicated topography, landform and geology would affect setting-out and layout of roads of a mountainous city.

(2) In the urban central zone of a large mountainous city, bridges would undertake complicated loads; affected by high-rise buildings in the city, wind direction is uncertain and wind flows at a relatively high speed; there are too many decorative objects, producing relatively high decoration load; water velocity is relatively high; anti-collision requirements of bridge piers are high. Bridge body is usually light and thin; most of it adopt steel girder as main girder, which brings high flexibility.

(3) Bridges in a mountainous city vary in structural type; their structures are very complicated, posing a real challenge to structural design. Those bridge types applicable to mountainous cities include beam bridge, arch bridge, cable-stayed bridge, suspension bridge and composite bridge; this is mainly procured by mountainous cities, and mainly closed to bridge area's geology, topographical condition and sea route navigation level. Geological conditions of different cities vary, and bridge area's river width and navigation requirement vary as

well; therefore, there are various bridges built.

3. Difficult Points for Designing Bridges of a Mountainous City

3.1 Design of Yudong Yangtze River Bridge

The bridge is 1,541.6m long in total. Its main bridge is a four-span (145.32+2×260+145.32) continuous prestressed concrete rigid frame bridge. Laterally, it is a paralleled twin-deck bridge, bearing 6 lanes and 2 light rail lanes. The main bridge's main girder, structured with single-box dual-chamber section, falls into upper and lower decks; the bridge deck is 41.6m wide in total, each half being 20.3m in width; its box is 12.9m in width, maximum cantilever is 4.8m, girder at the root is 15.1m high, girder at mid-span is 4.6m high; all box girders are measured for height as per external edge of external web plate; girder height from closure middle point to cantilever root changes in the form of 1.8 parabola, while the thickness of bottom plate of cast-in-situ section of side span from the closure to support end changes linearly. There are three separated interchanges, two level crossings and one interchange. It started construction formally on December 29, 2004. On December 26, 2008, its upper half bridge opened to traffic, while the lower half bridge opened to traffic on August 13, 2011.

Figure 4. Yudong Yangtze River Bridge

(1) As an ultra-large-span continuous rigid frame bridge for both road and rail traffic, its two main spans are both 260m, making it at the top list of similar bridges. Its comprehensive technology can be ranked no. 1 in continuous rigid frame bridges.

(2) For the first time, this rigid frame bridge bears automobile and light rail, but with asymmetrical structural section and load. The bridge is 41.6m wide, which enables it to be no. 1 among Yangtze River Bridges.

(3) At the main piers, bed rock is complete, but cracks develop. Therefore, if extended foundations are adopted completely, then their large dimension would cause heavy excavation, which is not economic; besides, erosion of cracked bed rock by river water can have adverse impact to the bridge; but if it is designed as per normal pile group foundation, its work for pile foundation will increase; to make full use of bedrock bearing stratum under pile cap, it's better to coordinate external force distribution as per pile and bed rock, that is to say, to consider the combined action of pile cap and pile foundation.

(4) Thickness of bottom plate of 0-9 section changes from 2.2m to 1.5m, gradually. The box girder is constructed with C60 concrete. In order to prevent temperature crack caused by heat of hydration, cooling pipes are added in bottom plate, which is the first time.

(5) Design requirements are as follows to prevent crack of block no.: ①The connecting bracket between double-limb piers has certain capacity of adapting to concrete shrinkage to prevent cracking of bottom plate due to constraint; ② When block no. 0 is poured in multiple times, exert temporary prestress and at the same time, set strengthened circular steel bars at connecting parts of different layers to prevent cracking due to interlayer load and desynchronized shrinkage of new and old concrete surfaces; ③ In design structure, D8 crack control mesh reinforcement is added and polypropylene fiber is mixed into concrete to effectively prevent cracking of block 0.

(6) Add H-shaped steel into double-limb thin wall pier and hollow interface pier: ① Replace H-shaped steel

with main bars of partial pier, which reduces steel bar's set density and is easy for concrete pouring; ② H-shaped steel already installed in place can be used as skeleton to ensure the stability during construction of main bar of pier; ③ Set H-shaped steel surrounding the pier externally to strengthen the anti-ship collision action of pier and possibly enhance anti-shock capacity of pier.

(7) Only the main bridge is set in the whole riverbed scope, where there are only three T-type pier and 2 junction piers. For this reason, there are good unobstructed sights and multiple bridge holes under the bridge; two main spans are helpful for ships at Foeryan Wharf at the 1.2km of the upper stream to navigate and change direction; only main piers are set in the river for the cause of bridge collision resistance.

(8) Web plate and bottom plate vent hole are set longitudinally for each box girder segment; the junction pier is set into pillar pier cap shape, with good ventilation effect; these are good for reducing the temperature between inside and outside box temperature, improving structural load bearing, and promoting the working environment for post maintenance inside the box.

(9) For No. 0 beam section, the concrete can reach up to 1,638 m3, making it the largest no. 0 beam section in the world; the cantilever casting section is up to 510t in weight, becoming the cantilever casting section of the highest weight.

(10) Precamber is set as per the equation (constant load (prestressing force) + live load + span length/1000); concrete shrinkage and creep are calculated as per 20 years, and then properly adjusted as per relevant standards and in combination with down-warping phenomena of other similar bridges. Meanwhile, to prevent insufficient comprehensive stress reserve due to excessive actual comprehensive stress and similar causes, try to maintain the section at the comprehensive stress above 1 MPa under any adverse load in design; besides, use backed beam hole to leave adequate leeway for post adjustment of prestressing tendon.

3.2 Key Technologies in Designing Caiyuanba Yangtze River Bridge

3.2.1 Y-type Rigid Frame System of Space

The bridge scheme dexterously conceives a combined structural system of rigid frame, steel truss girder and tied arch by making full use of the features of strong main beam required for "road-rail use" and basing upon field condition, structural load bearing and material feature; one pair of side span prestressed concrete rigid frames and one mid-span steel box handle-basket arch's three relatively separated substructure are connected into 420 m tie rod arch through mid-span tie rod and side-span tie rod. As to side-span's Y-type concrete rigid frame, because of cancellation of supporting column below the main beam and linearization of Y-type cantilever, its load bearing features has changed from the features of arch to beam features and the member bar's load bearing status has changed from "mainly under pressure" to "mainly under bending"; its structural design has changed fundamentally, with prestressed structure adopted; especially, to realize the space's handle-basket arch structure, Y-type rigid frame is also designed to a space structure with front and back cantilever ends adducted on the plane surface and torsion of cantilever surface. Y-type rigid frame primarily comprises front and back cantilevers, main crossbeam, front and back secondary crossbeams, front and back main crossbeams, and tie rod cable anchoring part; set heteromorphic transition block on the op of main pier and rigid frame bottom; transition block cross-bridge connects to rigid frame laterally, while internal and external lateral faces are curved. The rigid frame's front and back cantilevers are of thin-walled hollow structure with variable cross section; on the front end of the front cantilever, tie rod cable anchorage element and one front secondary crossbeam; tie rod anchorage parts are separately set at the anchorage points of side-span and mid-span tie rod; the front secondary crossbeam is an important component to keep Y structure space's stability, share Y structure's lateral component force, and set pivot point sling for the cross point between trussed girder and Y structure; as to back cantilever's end, tie rod anchorage structure is set at the solid part of side pier top, at anchorage end of side-span tie rod, and meanwhile, vertical prestressed tie rod is also set to connection the back cantilever end to side pier; this tie rod makes it possible to exert active control over rigid construction and internal force after bridge formation. On the top of the two piers upstream and downstream, main crossbeams of hexagonal gemstone type are set to link together the main pier and both sides' rigid frame.

The Y-type steel structure system of this space , get good shape, stressing definite, complete functions, the precision and technical requirements of the design and construction are very high.Its implementation will be important grarantee of the accurate location of steel box arch rib, the effective of steel truss main girder , the effective force application of The whole bridge tie bar.

3.2.2 Steel and Mixed Contact Design of Y-Type and Arch Rib Joint Part

In the bridge construction, with the improvement of using function, the increase of spanning ability, the

optimization of the structural system, give full play to the material properties, realize optimal combination of the structure and reduce the cost, the segment combination of the material

is the inevitable development trend of bridge construction in the future. The arch rib of Caiyuanba Yangtze River Bridge using steel box structure of equal section, so there must be a transition junction between steel box and Y-type rigid frame system, that is the Steel and mixed contact, is another key to bridge structure. The contact should not only bearing great pressure and repeated moment after the completed bridge, and resistance greater construction load internal force, steel mixed materials have different material characteristics, the effective combination of different materials and the reliability of joint force are very important, Rationality design of the contact directly affects the use of state and structure safety of the main arch rib and bridge.

3.2.3 The Design of Large Segment Whole Steel Truss Girder

At present, the construction of large span steel truss bridges in our country basically adopt the member assembly system erection, while the Chongqing Caiyuanba Yangtze River bridge first adopted the design idea of integral girder block in domestic, Where a standard section for 16 m, width of about 40 m long steel truss girder integral segment as the basic unit of site assembly, requested the segment rods, beams, contact, orthotropic bridge deck are assemble in our factory, then shipped to the construction site hoisting splicing. This whole section design, will make a lot of site bar splicing can achieve inside the factory, which can guarantee the quality of structure within the segment, reduce a lot of bar splicing process, speed up the construction site of frame beam, the increase in the construction site's ability to fight bad weather.

3.3 Technical Difficulties of Zhongzhou Yangtze River Bridge

(1) Deep water foundation (foundation water depth 25m), and failure of side span counter-balance weight of balanced anchor box, auxiliary pier and temporary pier, are main technical difficulties of the main bridge. During construction, the mid-span of double cantilever reaches 216m and the side span reaches 194m and side span reaches 194m; such a large double cantilever status is rarely seen in China, which enables its design and construction difficulty to be higher than general stayed cable.

(2) The main beam’s side span weight is exerted by gradual addition of wide ribbed plate.

(3) The main beam’s temporary consolidation bracket separates the main beam from the main tower with steel plate, but combines them into an entity through external prestressing tendon and sandglass bracket; at the closing of main span, unlease external prestressing tendon and sandglass bracket to remove longitudinal constraint of the main beam.

(4) After closing of the main span, adjust the force of the last several pairs of stayed-cable to reduce temporary pre-press weight at mid span closing and lower the closing difficulty.

(5) The main tower pier is 247.5m high, next only to Sutong Yangtze River Bridge. The diameter of double wall steel cofferdam adopted for the main tower’s foundation construction is 36, enabling this bridge to be ranked as No.1 of bridges of similar scale.

(6) The first segment of steel cofferdam and steel box girder is assembled on floating platform and hoisted wholly after consignment into place; afterwards, back from the floating platform and the first segment immerse into water; this reduces construction processes of shore assembly, launch, consignment into place and construction, thus saving time for construction period.

(7) Both steel cofferdam and suspended steel box are sealed twice. Steel cofferdam sealed twice can effectively prevent the leakage of bottom concrete; suspended steel box back-sealed for the first time can completely isolate water in suspended water box and river water, while the concrete for secondary bottom sealing can make full use of water buoyancy; water drawing can be used to substitute for concrete weight to reduce rod’s load and lower the difficulty for bottom sealing construction. Sealing for twice lowers technical difficulty of the critical step of cofferdam.

(8) When the water storage of the Three Gorges’ Reservoir reaches 175m, 10,000-ton class vessels can sail to Chongqing directly, so there can be higher risk of collision between vessel and bridge. The Ministry of Communications and relevant department of Chongqing City have performed “Ship Collision Resistant Measure Design and Early Warning System Study of Piers of Cross-River Bridge in the Three Gorges Reservoir Region”, and set floating casing box anti-collision design for piers with higher collision risks. The bridge can bear the impact force up to 2,765 tons, so it is the bridge with the highest anti-collision capacity in Chongqing City.

4. Conclusion

It’s not simple to build a bridge in a mountainous region. It’s more difficult to build a bridge in a mountainous

city, because of impacts and requirements of a mountainous city in, among other factors, landform, hydrometeorology, building features, comprehensive pipe network, historical humanism, overall scene and overall planning. Why are many existing large-span bridges in a mountainous city recognized as technologically leading ones in China, even in the world? This is not that constructors only pay attention to the breakthrough of technical rank, but because they overcome difficulties above.

References

Airong, C., Yong, S., & Feng, Q. (2005). Bridge Modeling . Beijing: China Communications Press, 26-31.

Fan, L. C. (2001). Bridge Construction . First Version Beijing: China Communications Press.

Hua, Y. S. (2011). Chongqing Caiyuanba Yangtze River Bridge. Chongqing. *Publishing House.*

Li, B., & Yu, W. (2008). Briefly Discussion of Bridge Construction Technology of Mountain Region. *Shanxi Building, 34*(31).

Sun, J. S. (2011). Bridge Annals of Chongqing. *Chongqing: Chongqing University Press,* 94-175.

Wang, F. M., & Luo, Q. (2002). Discussion on designs and considratios of city bridges according to construction of Chongqing Yuao Bridge.Technology of Highway and Transport, (3), 33-34.

Wang, H., & Bao,Y. L. (2011). Urban Bridge Planning and Design. *Science & Technology Association Forum.*

Wei, Y. (2012). Bridge Position Selection in Bridge Design. *Technology Innovation and Application.*

Xiang, Z. F. (2008). Study on Chongqing Bridge Construction Level and Feature. Chongqing Jiaotong University.

Zhai, C. X. (2012). Study on Traffic Characteristics and Development Strategy of Group City Taking as an example . *Journal of Chongqing Jiaotong University: Social Sciences Edition, 12*(12), 9-11.

Zhang, J. (2009). Introduction to Chongqing Jiahua Bridge. *Technology of Highway and Transport,* (5), 153-155.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Mathematical Model of System of Protection of Computer Networks against Attacks DOS/DDOS

Shangytbayeva G. A.[1], Karpinski M. P.[2], Akhmetov B. S.[3], Yerekesheva M. M.[1] & Zhekambayeva M. N.[3]

[1] K. Zhubanov Aktobe Regional State University, Aktobe, Kazakhstan

[2] Academy of Technologies and the Humanities in Bielsko-Biala, Bielsko-Biala, Poland

[3] Kazakh National Technical University named after K.I.Satpayev, Almaty, Kazakhstan

Correspondence: Shangytbayeva Gulmira, K. Zhubanov Aktobe Regional State University, Aktobe, Kazakhstan. E-mail: shangytbaeva@mail.ru

Received: January 6, 2015 Accepted: February 18, 2015 Online Published: July 10, 2015

doi:10.5539/mas.v9n8p106 URL: http://dx.doi.org/10.5539/mas.v9n8p106

Abstract

In practice the most part of connections of subjects of a wide area network uses the virtual connections as this method is the dynamic protection of network connection and won't use static key information. Therefore, the interaction without establishing a virtual channel is one of the possible reasons for the success of remote attacks such as DoS / DDoS. An abstract considers the mathematical model of system of protection of computer networks against attacks such as DoS/DDoS, allowing in practice to detect attacks such as DoS / DDoS. Grounded way to prevent attacks through the use of network reconfiguration procedures, it is difficult for practical implementation attacks such as DoS / DDoS. The algorithm to create new virtual data channels to ensure a minimum amount of traffic regardless of reconfiguring computer network. The article presents an approach to detection of the distributed network attacks to refusal in service, the offered method increases efficiency of use of the calculated resource of a computer network at the big distributed network attacks to "Denial of Service".

Keywords: computer network, routing, attacks, network attacks, DoS – attack, DDoS – attack, "Denial of Service", detection of network attacks

1. Introduction

Computer network providing every opportunity for exchanging data between the client and server, but now widely distributed attack denial of service clients, the determination of distributed attacks in the network is particularly acute. The most common types of such attacks are DoS / DDoS attacks, which deny certain users of computer network services. With the constant development of computer networks and the increasing number of users grows and the number of new types of attacks to denial of service. DoS / DDoS attacks are characterized by a straightforward implementation complexity and resistance, which poses new problems of researchers, who are still not yet resolved (Yang Z. X. et al., 2014).

Wide use of computer networks creates conditions for implementation of the attacks using standard algorithms of routing. It is known that the routing protocol in data transfer is rule set and arrangements on an exchange of information network between routers to determine the data path transfer which satisfies to the given parameters of quality of service and provides the balancing load of all computer network in general therefore the research problem of network traffic acquires special relevance (Wang J., & Chien, A, 2003).

To questions of the organization and creation of computer networks including to questions of routing, are devoted to the work of scientists M.Yu.Ilchenko, S.G. Bunina, A.S. Petrova and D. Davis, D. Barber, V. Price, V. Vilinger, D. Wilson, D. Rakhson, etc.

In most respects it is similar to a DoS attack but the results are much, much different. Instead of one computer and one internet connection the DDoS attack utilises many computers and many connections. The computers behind such an attack are often distributed around the whole world and will be part of what is known as a botnet. The main difference between a DDoS attack vs a DoS attack, therefore, is that the target server will be overload by hundreds or even thousands of requests in the case of the former as opposed to just one attacker in the case of the latter (Hussain, A., et al., 2003).

Therefore it much harder for a server to withstand a DDoS attack as opposed to the simpler DoS incursion. Do not process a large number of requests, the server first starts just slowly and then stops completely. Inquiries most often have smart and senseless character that even more complicates operation of the server (Denial-of-service attack, 2015).

DDoS-attack – the distributed attack like refusal in service which is one of the most widespread and dangerous network attacks. DDOS is a type of DOS attack where multiple compromised systems which are usually infected with a Trojan are used to target a single system causing a Denial of Service (DoS) attack. Victims of a DDoS attack consist of both the end targeted system and all systems maliciously used and controlled by the hacker in the distributed attack (Kihong Park & Heejo Lee, 2001).

DDoS attack, the incoming traffic flooding the victim originates from many different sources – potentially hundreds of thousands or more. This effectively makes it impossible to stop the attack simply by blocking a single IP address; plus, it is very difficult to distinguish legitimate user traffic from attack traffic when spread across so many points of origin (Goodrich, 2002).

The majority of DDoS-attacks use vulnerabilities in the main Internet Protocol (TCP/IP), namely, a method of processing a query systems SYN (Wang, H., et al., 2002).

Allocate two main types of attacks which cause refusal in service .

As a result of carrying out attack of the first type, work of all system or a network stops. The hacker sends to system data or packages which she doesn't expect, and it leads to a stop of system or to its reset.

The second type of DDoS-attacks leads to overflow system or a local network using the vast amount of information that can’t be processed.

DDoS-attack consists in the continuous appeal to the site from many computers which are located in different parts of the world. In most cases these computers are infected with viruses which are operated by swindlers is centralized and eaten in the botnet. Computers which enter in botnet, send to spams, participating, thus, in DDoS-attacks (Li Muh *et al.,* 2008).

The most common type of Denial of Service attack involves flooding the target resource with external communication requests. This overload prevents the resource from responding to legitimate traffic, or slows its response so significantly that it is rendered effectively unavailable. Resources targeted in a DoS attack can be a specific computer, a port or service on the targeted system, an entire network, a component of a given network any system component. For implement of attacks like DoS / DDoS in the modern computer networks is characteristic multilevel routing in of which the computer network certain way breaks into the subnets and they working on standard protocols (Savage, S., et al., 2000). Most implementations of attacks like DoS / DDoS are calculated on the net with a homogeneous structure or on network with the fixed structure of domains. Frequent change of components of a computer network leads to change of its topology, composition and quantity of routing domains, influences efficiency of procedure of routing, and promotes operation of algorithms like DoS / DdoS (Park, K., & Lee, H., 2000). Therefore there is a need for development of new methods of protection of a computer network that will provide information transfer with the given parameters of quality of service at the minimum volume of traffic by development of a method of protection of traffic against the redundant information on the basis of determination of criterion of network transmission capacity and computing resources. The analysis of practical implementation of the attack in the wide area network allows to determine information security mechanisms in the computer networks on the basis of use of algorithms like DoS / DDoS. For elimination of the reasons of attacks to infrastructure and basic protocols of a network is advisable to change a configuration of computer system. At the first investigation phase it is necessary to analyze network traffic for preventing of unauthorized reading from physical transmission channel of data, it will allow to avoid interception of information leakage. This task successfully treated by creating of the virtual networks, of the the algorithms the tunneling, identification, authentication (Bhuyan, M. H., et al., 2015).

Danger of the majority of DDoS-attacks is that at the first stages don't violate the exchange protocol of data. They manifest themselves when computable network resource becomes insufficiently. For preventing of such attacks quite properly configure the router and firewall.

For simplification of implementation of DoS / DDoS of algorithms the user should observe of rated speeds of data transfer, it will allow in practice to avoid algorithms of optimization of network traffic on which are realized of many attacks (Wu, Y., et al., 2015).

2. Methodology

The Method section describes in detail how the study was conducted, including dependence of the volume of traffic on the types of attacks DoS / DDoS, efficiency of methods of routing, etc (Klimenko, I. A., 2005).

From attacks such as flood can effectively dissociate themselves by division of the main communication channel into multiple virtual. This will allow to create other network interfaces in case of defeat of the channel DoS / DDoS of algorithms. Firewalls advisable to continuously strengthen and adjust so that internal network services were unavailable to the external user. It is expedient to set the analyzer of network traffic, value of its parameters will allow in time to identify the beginning of the attacks. Before the direct beginning of attack bots gradually increase a flow of packets on system. Therefore is necessary the continuous observation of over the router connected to the external network (Dean, D., et al., 2001).

The effectiveness of routing methods is directly dependent on the topology of the network and its size. Multilevel routing substantially depends on optimum partition of a computer network on domains routing. Thus, one of the main objectives of protecting the functioning of a computer network has network traffic, which is based on the principle of the minimum amount of network data. The task of protecting network data reduced to the problem of minimization of parameters of information transfer (Law, T., et al., 2005).

It is known that when using known routing protocols in the domain network topology change leads to growth of traffic the non-linear law, therefore reconfiguration of domains routing in the course of topology change of a network promotes reduction of volume of traffic and time of formation of ways of data transfer, and also complicates implementation of attack like DoS / DDoS. Based on an advanced mathematical model is defined the choice of quantity and the size of routing domains. For the purpose of support of maximum efficiency of functioning of a computer network, procedure should take into account changes the network topology. However the majority of routing protocols don't provide procedure of change of structure of routing domains. In this regard there is a need of development of new model of routing which at the expense of the accounting of attacks of a failure in service of resources of routing will allow to increase efficiency of information transfer in the computer networks. Thus it is necessary to consider conditions of feasibility of parameter values of network transmission capacity and computing resources. It is expedient to determine the parameters regulating the amount of the packets transferred on each communication link separately and total amount of the packets transferred during the update of routing tables (Stone, R., 2000).

The total volume of traffic is determined by this model (1):

$$V = \frac{T_{sys}}{\Delta t_{sys}} \sum_{i,j=1}^{N} P_i Q_j \tag{1}$$

Where:

Δt_{sys} – time of one clock period of system;

Q_j– amount of information, transferred for one clock period on each certain canal;

N – quantity of nodes on the computer networks;

P_j – the degree of a node is compromised;

T_{sys} – time during which in case of topology change of a network nodes distribute messages of message on restoration of routes (Bu, T., et al., 2004).

3. Results

In the Results section, summarize the collected data and the analysis performed on those data relevant to the discourse that is to follow. Report the data in sufficient detail to justify your conclusions. Mention all relevant results, including those that run counter to expectation; be sure to include small effect sizes (or statistically nonsignificant findings) when theory predicts large (or statistically significant) ones. Do not hide uncomfortable results by omission. Do not include individual scores or raw data with the exception, for example, of single-case designs or illustrative examples. In the spirit of data sharing, raw data, including study characteristics and indivldual effect sizes used in a meta -analysis, can be made available on supplemental online archives (Karpinski, M. P., 2011).

Researches showed that during implementation of attack of a type of DoS / DDoS increases the quantity of the compromised nodes and grows the total amount of traffic of V.

The results of numerical experiment presented in Figure 1.

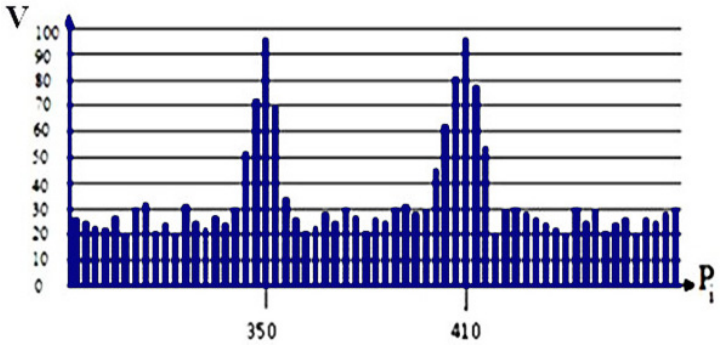


Figure 1. The dependence of the volume of traffic on the type of attacks DoS / DDoS

The results of the model experiment are presented in Figure 2.

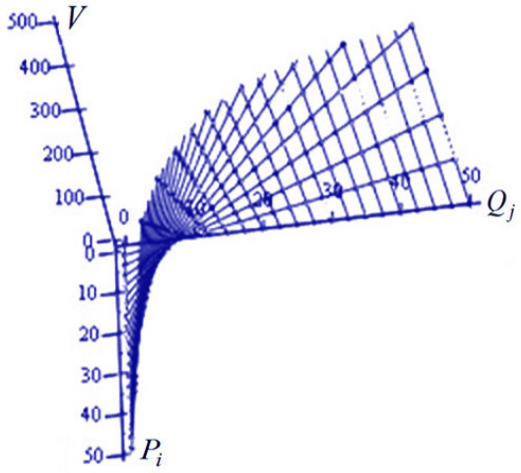


Figure 2. The dependence of the volume of traffic from the degree of a compromise node

The analysis of a figure shows that in attack promptly increases traffic volume in channels of a network, the most part of traffic uses an algorithm like DoS / DDoS. It's pretty much slows down the work network. To prevent such situations, it is advisable to use the traffic analyzers that control the amount of packets in the network (John, D. H., 1998). Also it is necessary to execute procedure of routing by means of distributed system of agents of routing provided that routing in the domain is carried out by the agent who is a part of this domain, and routing between domains is executed at the interoperability layer of agents of routing. This is due to the fact that the exchange of official information on the network is carried out by the same channels as the transmission of useful information (Aleksander, M. A., et al., 2012).

4. Discussion

In this paper for the detection of DDoS-attacks provided a method for estimating the probability of loss of any requests during its passage through of networks (Karpinski, N., & Shangytbayeva, G., 2005).

After presenting the results, you are in a position to evaluate and analyze, interpret their implications. Here you will examine, analyze, interpret, and qualify the results and draw inferences and conclusions from them.

The analysis of the data provided research shows that in attack time traffic volume in channels of a network promptly increases, and the most part of traffic is used by attack like DoS / DDoS. it's pretty much slows down the network (Baba, T., & Matsuda, S., 2002).

For increase of efficiency of procedure of routing in operation is offered to separate data on the virtual links. For this purpose, a plurality of workstations and virtual channels between them arranged in the form of a local area

computer network. A method of forming and dynamic reconfiguration of routing domains can increase the efficiency procedure routing of computer network. Formation and dynamic reconfiguration of domains is carried out by means of specialized system of routing which basic functions is determination of quantity and layout of workstations, updates of the routing information and a choice of a way, meets the requirements of stability and the minimum temporal time delay.

To prevent attacks appropriate to introduce in the system of additional detectors of monitoring of traffic of a network. These detectors instruct the executing modules in different network segments. As a result, before the attacked flow it is formed the screen, the separating attacks from an internal network. Routes are elected dynamically or statically so as to use only the physical security of the subnet, nodes of switching and channels. Transmission of data having security tags through certain subnet, switching nodes and channels advisable to prohibit the security policy.

This research is directed on studying of the distributed network attacks such as DoS / DDoS and methods, models and architecture of network attacks to refusal in service.

On the basis of the presented methodology the developed architecture and is constructed program realization of system of detection of DDoS-attacks. (Shangytbayeva, G., et al., 2015) According to the results is published scientific articles. The methodology developed in this study got considerable support. (Shangytbayeva, G., & Beysembekova, R., 2015).

Conclusion

In article improved mathematical model of total traffic, allows in practice to reveal attacks like DoS / DDoS. Use of the received results allows to raise the network security level at the expense of the organization of multi-level of protocols routing. For preventing specified attacks appropriate to introduce a system of additional detectors traffic monitoring network. Reasonably a method of preventing of attacks on the basis of use of procedure of reconfiguration of a network that allowed to complicate practical implementation of attack like DoS / DDoS, by creation of new virtual links of data transfer. This method provides the minimum volume of traffic irrespective of reconfiguration of a computer network.

References

Aleksander, M. A., Karpinski, M. P., & Yatsykovska, U. O. (2012). Features of Denial of Service attacks in information systems. *Informatics and Mathematical Methods in Simulation, 2*(2), 129-130.

Baba, T., & Matsuda, S. (2002). Tracing network attacks to their sources. *IEEE Internet Computing, 6*(2), 20-26.

Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2015). An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection. *Pattern Recognition Letters,* (51), 1-7.

Bu, T., Norden, S., & Woo, T. (2004). Trading resiliency for security: Model and algorithms. *In Proc. 12th IEEE International Conference on Network Protocols,* pp. 218-227),

Dean, D., Franklin, M., & Stubblefield, A. (2001). *An algebraic approach to IP traceback,* pp. 3-12. In Network and Distributed System Security Symposium (NDSS).

Denial-of-service attack. (2015, February). In Wikipedia, the free encyclopedia. Retrieved February 2, 2015, from http://en.wikipedia.org/wiki/Denial-of-service_attack

Goodrich, M. T. (2002). *Efficient packet marking for large-scale IP traceback.* In 9th ACM Conf. on Computer and Communications Security (CCS). pp. 117–126.

Hussain, A., Heidemann, J., & Papadopoulos, C. (2003). *A framework for classifying denial of service attacks.* (pp. 99-110), Proc. ACM SIGCOMM. Karlsruhe, Germany.

John, D. H. (1998, August). *An Analysis of Security Incidents on the Internet*. Ph.D. thesis, Carnegie Mellon Univerisity.

Karpinski, M. P. (2011). *Modeling network traffic computer network in implementation attacks such as DOS / DDOS.* (5), 143-146. Information Security, American Psychological Association. Ethical standards of psychologists. Washington, DC: American Psychological Association.

Karpinski, N., & Shangytbayeva, G. (2015, 05-06 January). *Architecture and Program Realization of System of Detection of Network Attacks to Denial of Service.* (pp. 55), International Conference on "Global Issues in Multidisciplinary Academic Research" GIMAR-2015, Dubai, UAE.

Kihong, P., & Lee, H. (2001, August). *On the Effectiveness of Route-Based Packet Filtering for Distributed DoS*

Attack Prevention in Power-Law Internets, (pp. 15-26), Proc. of ACM SIGCOMM '01.

Klimenko, I. A. (2005). *A method of dynamic routing with support for required level of quality of service in mobile networks without a fixed infrastructure.* (Vol. 15. pp. 102-112). Problems of Information and Control: Sat. Sciences. pr. M.: NAU.

Law, T., Yau, D., & Lui, J. (2005). *You can run, but you can not hide: An effective statistical methodology to trace back ddos attackers,* 16(9), 799-813. IEEE Transactions on Parallel and Distributed Systems.

Li, M., Li, M., & Jiang, X. (2008). DDoS attacks detection model and its application (Vol. 7, No. 8, pp. 1159-1168). WSEAS Trans. Computers.

Park, K., & Lee, H. (2000, June). *On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack,* Tech. Rep. CSD-00-013, Department of Computer Sciences, Purdue University.

Savage, S., Wetherall, D., Karlin, A. R., & Anderson, T. (2000). *Practical network support for IP traceback* (pp. 295-306). In Conference on SIGCOMM.

Shangytbayeva, G., & Beysembekova, R. (2015, 21-22 March). *Integrated approach to the detection of distributed network attacks. International Conference on Mechanical*, Electronic and Information Technology Engineering, Chongqing, China.

Shangytbayeva, G., Akhmetov, B., & Beysembekova, R. (2015, 25-26 February). *Analysis of Methods Organization of the Modelling of Protection of Systems Client-Server.* International Conference on "Multidisciplinary Innovation in Business Engineering Science & Technology". Manila, Philippine.

Stone, R. (2000, August). Centertrack: An IP overlay network for tracking DoS floods. In Proc. of 9th USENIX Security Symposium.

Wang, H., Zhang, D., & Shin, K. G. (2002). *Detecting SYN flooding attacks.* (pp. 1530-1539), Proc. IEEE INFOCOM'2002. New York.

Wang, J., & Chien, A. (2003, October). *Using overlay networks to resist denial of service attacks.* Submitted to ACM Conference on Computer and Communication Security.

Wu, Y., Zhao, Z., Bao, F., & Deng, R. H. (2015). *Software puzzle: A countermeasure to resource-inflated denial-of-service attacks.* IEEE Transactions on Information Forensics and Security.

Yang, Z. X., Qin, X. L., Li, W. R., & Yang, Y. J. (2014). *A DDoS detection approach based on CNN in cloud computing.* Applied Mechanics and Materials.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Power Factor Control of Matrix Converter Based Induction Motor Drive

Settar S Keream[1], Ahmed N Abdalla[1], Mohd Razali Bin Daud[1], Ruzlaini[1] &youssif Al Mashhadany[1]

[1] University of Anbar- Univrsity, Pahang, Malaysia

Correspondence: Settar S Keream, University of Anbar- Univrsity, Pahang, Malaysia. E-mail: settar_msc@yahoo.com

Received: November 29, 2014 Accepted: January 12, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p112 URL: http://dx.doi.org/10.5539/mas.v9n8p112

Abstract

The Induction Motors utilizing Matrix Converters with Direct Torque Control (DTC) is a great responsibility for motor control scheme with expeditious torque and flux replications. This paper shwed a new power factor control along with the existing DTC matrix converter induction motor drive. The core benefits of the DTC matrix converter are improved with those of the power factor technique, the required voltage vectors is producing under 0.86 input power factor operations. The implementation of this kind of controller is done by using the TMS320F28335. The results validate the good quality and robustness in the proposed system dynamic response and reduction in the transient motor ripple torque.

Keywords: power factor, DTC, Matrix converter, induction motor

1. Introduction

In last years the direct torque control (DTC) strategy develops the control approaches and one of the high-performances for induction motor due to high response for torque and flux control (Lee, H. H., Nguyen, H. M., Chun, T. W., & Choi, W. H. 2007, pp. 51-55). The torque and flux of the induction motor can be made by the cull through a look-up table in dirct torque control of the puissance converter voltage space vectors. The highest amelioration of DTC is its structural facileness, then no current controllers, coordinate conversions with PWM are required. Furthermore, the controller does not be influenced by motor factors. Direct torque control can make to be a simple and robust system which reaches a rapid and exact torque control reaction. For such progressive reasons, the advantages of combination matrix converter with DTC method is effectually promising (Chen, D. F., Liao, C. W., & Yao, K. C. 2007, pp. 101-101). However, some research is quiet being done to decrease the electromagnetic torque ripple, which is its chief problem that leads to the growing stator current deformation noise (Alesina, A., & Venturini, M. A. R. C. O. G. B. 1989, p101-112). The next approaches are applied to improve the effects of the ripple on the output torque: fuzzy logic controller, the modulation procedures of the SVM, multilevel inverter (Casadei, D., Serra, G., & Tani, A. 2000, p769-777, Lascu, C., Boldea, I., & Blaabjerg, F. 2000, p122-130, Buja, G. S., & Kazmierkowski, M. P. 2004, p744-757, Ghoni, r., abdalla, a. N., & sujod, z. 2010) and so on. DTC technique was implemented using digital signal processor. Then this algorithm was implemented using ASIC (Application Specific Integrated Circuit) design (Buja, G. S., & Kazmierkowski, M. P. 2004, p744-757). It is hard to implement DTC utilizing prevalent IC hardware.. The DTC method is normally implemented by sequential calculations on a DSP panel. Nevertheless, as a predictive control scheme, the DTC has a steady-state error formed by the time delay of the long calculations, which depends largely on the control method and hardware performance. Unusual DSP (TMS32010) implementation time of the DTC method for a VSI-fed induction motor is more than 250μs (Habetler, T. G., Profumo, F., Pastorelli, M., & Tolbert, L. M. 1992, p1045-1053.The induction motors are used in seventy to eighty percent of all industrial drive applications due to their simple mechanical construction, reliability, ruggedness, low cost and low maintenance requirement compared to other types of motors. Also it operates at essentially constant speed. These advantages are however suppressed from control point of view. When using an induction motor in industrial drives with high performance demands, the induction motors are non linear high order systems of considerable complexity (Zaghloul, M. E., Meador, J. L., & Newcomb, R. W. 1994). So, DTC of VSI-fed induction motor based on ANN had been introduced out (Shi, K. L., Chan, T. F., Wong, Y. K., & Ho, S. L. 2001, p1290-1298, Dung, P. Q., & Thuong, H. T. N. 2004, pp. 60-63). Also, the originators must possess plentiful experiences on related theories.

The MC is a captivating converter topology for applications where elements such as elimination of electrolytic capacitors, potential to achieve high power density, reduced size and weight, sinusoidal input and output currents and unity power factor operation are desired. This would include applications ranging from industrial application up to megawatt level, renewable energy applications and additionally conveyance applications. MC applications documented in literature include field oriented control and DTC of motor drives, wind power generation topologies in both squirrel cage induction machine (SCIG) and doubly-alimented induction topologies. The MC is withal being investigated for use in auxiliary drive system for diesel locomotives (Mei, Y., Sun, K., Zhou, D., Huang, L., & Matsuse, K. 2005, pp. 2476-2483). Although matrix converters (MCs) share the same basic functionality with cycloconverters: to perform single stage ac power conversion with variable magnitude and variable frequency, the MCs differ from the cycloconverters in many aspects. Equipped with four-quadrant bidirectional switches made up of force-commutated devices such as IGBT, a typical MC is able to generate output voltages with either higher or lower frequencies than that of the input. A conventional MC features inherent four quadrant operation capability, sinusoidal input/output waveforms, and a controllable input power factor (Kolar, J. W., Friedli, T., Rodriguez, J., & Wheeler, P. W. 2011, pp 4988-5006).

In this paper, the new power factor control is introduced along with the existing DTC control for matrix converter and leading to the reduction of the electromagnetic ripple torque. The experimental results demonstrate the effectiveness of the proposed control scheme was presented.

2. Modelling of Induction Motor

Induction motor is nonlinear system and strong coupling multivariable system. In order to enable analysis of the mathematical model of the induction motor, some assumptions must be made, which are; ignore the space harmonics, assuming symmetrical three phase winding, the sine distribution is genrateing the air gap magnatic field, the phenomenon of magnetic saturation is also ignored, eliminating the core loss, disregard the frequency and the influence of temperature change on the winding. Voltage equation of induction motor in stationary dq frame is given in (1):

$$\begin{bmatrix} V_{sd} \\ V_{sq} \\ V_{rd} \\ V_{rq} \end{bmatrix} = \begin{bmatrix} R_s + L_s\delta & 0 & L_m p & 0 \\ 0 & R_s + L_s\delta & 0 & L_m\delta \\ L_m\delta & \omega L_m & R_r + L_r\delta & \omega L_r \\ -\omega L_m & L_m\delta & -\omega L_r & R_r + L_r\delta \end{bmatrix} \begin{bmatrix} i_{sd} \\ i_{sq} \\ i_{rd} \\ i_{rq} \end{bmatrix} \tag{1}$$

Where, L_s, L_r and $L_m$are the stator, rotor and the stator magnetizing inductance, ω is the angular velocity, δ is the different symbols for δ=d/dt , subscript s and r respectively for the stator and rotor; d and q are for dq frame. The flux equation is given in (2).

$$\begin{bmatrix} \psi_{sd} \\ \psi_{sq} \\ \psi_{rd} \\ \psi_{rq} \end{bmatrix} = \begin{bmatrix} L_s & 0 & L_m & 0 \\ 0 & L_s & 0 & L_m \\ L_m & 0 & L_r & 0 \\ 0 & L_m & 0 & L_r \end{bmatrix} \begin{bmatrix} i_{sd} \\ i_{sq} \\ i_{rd} \\ i_{rq} \end{bmatrix} \tag{2}$$

Where, ψ_{sd} and ψ_{sq} is the dq stator flux, ψ_{rd} and ψ_{rq} are the rotor flux. The electromagnetic torque equation is as (3);

$$T_e = \frac{3}{2} p_n L_m \left(i_{sq} i_{rd} - i_{sd} i_{rq} \right) \tag{3}$$

According to the flux equation in (2), (3) can be rewritten as (4):

$$T_e = \frac{3}{2} p_n L_m \left(\psi_{sd} i_{sq} - i_{sd\psi_{sq}} \right) \tag{4}$$

The mathematical model of the electric drive system equations of motion is done by disregard the electric drive transmission mechanism in the viscous friction and torsional flexibility as,

$$T_e = T_L + \frac{J}{p_n} d\omega / dt \tag{5}$$

Where, T_L is the load torque, T_e is the electromagnetic torque for the motor, J is the moment of inertia and p_n is the number of pole pairs for motor.

3. Induction Motor (IM) Efficiency and Power Factor Control

From (1), the IM efficiency and power factor will be calculated. The calculation of the motor active power, P and reactive power Qare given by,

$$P = \left(V_{ds}i_{ds} + V_{qs}i_{qs}\right) \tag{6}$$

$$Q = \left(V_{qs}i_{ds} - V_{ds}i_{qs}\right) \tag{7}$$

Using (1), (6) and (7), the IM efficiency, η and power factor, $\cos\varphi$ are derived, without considering the mechanical losses. Therefore, the efficiency can be expressed as,

$$\begin{aligned} \eta &= \frac{T_e\Omega_r}{P} = \frac{T_e\omega_r}{p_nP} \\ &= f_1(\Delta\omega, \omega_r) \end{aligned} \tag{8}$$

Where, f_1 is the stator frequency. The active power and power factor is equal to the ratio of the apparent power as in (9),

$$\begin{aligned} \cos\varphi &= \frac{P}{\sqrt{P^2 + Q^2}} \\ &= f_2(\Delta\omega, \omega_r) \end{aligned} \tag{9}$$

Where, $L_\sigma = \frac{L_sL_r - L_m^2}{L_r} a_2 = \Gamma_r(R_s\Gamma_r + L_s - L_\sigma)$, $a_1 = \Gamma_r(L_s - L_\sigma)$, $a_0 = R_s$, $b_3 = b_2 = L_\sigma\Gamma_r^2$, $b_1 = b_0 = L_s$, f_2 is the rotor frequency and $\Gamma_r = \frac{L_r}{R_r}$, is the rotor time constant.

Equation (8) and (9) show that the IM efficiency and power factor is the relationship of the rotor and a slip function angular frequency. In other words, when the rotor and slip angular frequency is constant, the value of the efficiency and power factor is also constant. Thus, the maximum efficiency of the IM power factor values can be calculated.

The relationship between the IM motor efficiency and power factor analysis was done using Matlab 10.0 software. The selected motor parameters are as follows;

$$P = 0.75kw, V = 215V, n = 1000rpm, T = 5.52\ Nm, R_1 = 5.89\Omega, L_{11} = .061H, R_2' = 5.375\Omega, L_{12}' = 0.013H, L_m = 0.149H, J = 0.02kg.m^2, p = 4.$$

Figure 1 is the graph of efficiency versus the rotor and slip frequency. When $\Delta\omega$ is fixed, the motor efficiency increases with the increasing of the rotor angular frequency. However, when ω_r is fixed, the motor efficiency is increases at first, and starting to show the decreasing trend with the maximum changing process. The value is the maximum efficiency with the increasing of the rotor angular frequency.

Figure 2 shows that the power factor has small changes when the speed is large, but when the speed is lowered to its maximum of about 10%, the power factor increases rapidly. The curve bends of Figure 2 shows the clear trend. The power factor increases with the increases of the slip frequency when ω_r is fixed.

From the previous analysis, it was concluded that the power factor is not at its highest when the efficiency is maximized, but the relationship is one to one. Thus, the power factor can be used as the control volume. The implementation of closed-loop power factor control by comparison of the given power factor with the actual can be used to adjust the motor terminal voltage in real time and allowing the system to achieve the optimum efficiency. From (8), the value of the efficiency derivation can be drawn as (10):

$$\frac{\delta\eta}{\delta\Delta\omega} = \frac{\delta f(\Delta\omega)}{\delta\Delta\omega} = 0 \tag{10}$$

And

$$\Delta\omega^B = \sqrt{\frac{R_s R_r^2}{R_r L_m^2 + R_s L_r^2}} \tag{11}$$

$\Delta\omega^B$ is the efficiency of a rotor angular frequency corresponding to the maximum point of the slip frequency. From (9), the efficiency of a rotor angular frequency corresponding to the maximum point of the power factor is given in (12):

$$\cos\varphi^B = f_2(\Delta\omega^B, \omega_r) \tag{12}$$

From (12), the $\Delta\omega^B$ and$\cos\varphi^B$ are the motor parameters to achieve the maximum efficiency. From (12), the best power factor angle of the rotor frequency curve can be calculated and shown in Figure 3.

It can be seen in the normal operating speed range, the best power factor changes from 0.55 to 0.75. From the control perspective, the power factor should be controlled along with the changes of a given speed so that it is always running at a maximum efficiency state. As can be seen from the above analysis, the power factor control, is the motor slip frequency control for different operating conditions by adjusting the input voltage to keep the slip frequency $\Delta\omega^B$around$\Delta\omega$.

The use of the power factor closed-loop control system will optimize the efficiency of the establishment in constant pressure based on the frequency ratio control to control the amount of the power factor, compromising in the stability of the actual power factor. The system block diagram is shown in Figure 4.

Efficiency optimization control system includes the following main modules: the best power factor calculation, function generator, power factor measurement, the reference wave generation, SVPWM, the drive circuit, MC, efficiency calculations, IM and other parts. The system model is shown in Figure 5.

The best angle for a given power factor according to the method described previously calculates the optimum power factor for a given value as the changes of the rotor speed.

4. Results and Discussion

The proposed algorithms were implemented in a TMS320F28335 DSP and the system setup is shown in Figure 6. Figure 7 shows the simulation and experimental results for the flux at a given amplitude of1.5Wb. The time taken for the flux motor to reach the steady-state is less than 0.05s with very small fluctuations for the proposed method. Figure 8 shows the response of torque when the torque is changed from 6Nm to 12Nm. Figure 8 (a) is the result without using the proposed method with large among of torque ripple as compared to the proposed method. Figure 9 shows the response of phase a stator current when the torque is changed from 6Nm to 12Nm with the stator current response to the changes of torque command. Some noise was observed for the stator current without using the proposed method as in Figure 9 (a).

Figure 10 (b) shows the torque response with the speed increased from 500rpm to 1000rpm with. The greater torque impact is achieved with the speed increases at less than 0.05s. When the curve of the speed changes in Figure 11 (a), the torque response changes with the speed changes. The electromagnetic torque is reversed following the speed command as in Figure 11 (c).

5. Conclusions

The proposed power factor control in this paper will optimize the efficiency of the establishment in constant pressure based on the frequency ratio control to control the amount of the power factor, compromising in the stability of the actual power factor. The power factor control reduced the torque ripple hence improving the matrix converter performance.

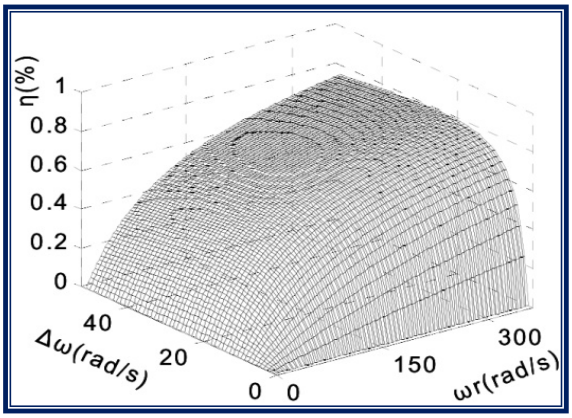


Figure 1. Efficiency versus rotor and slip frequency

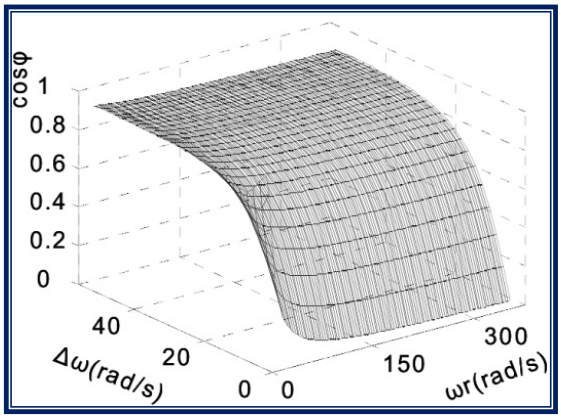


Figure 2. Power factor versus rotor and slip frequency

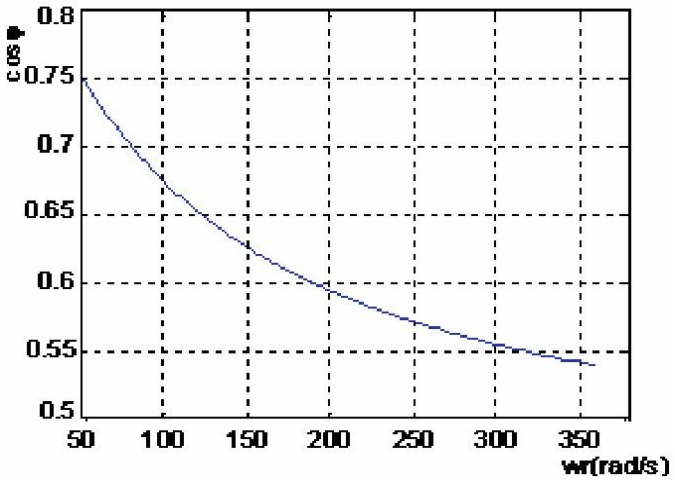


Figure 3. Relation between the rotor frequency and best power factor

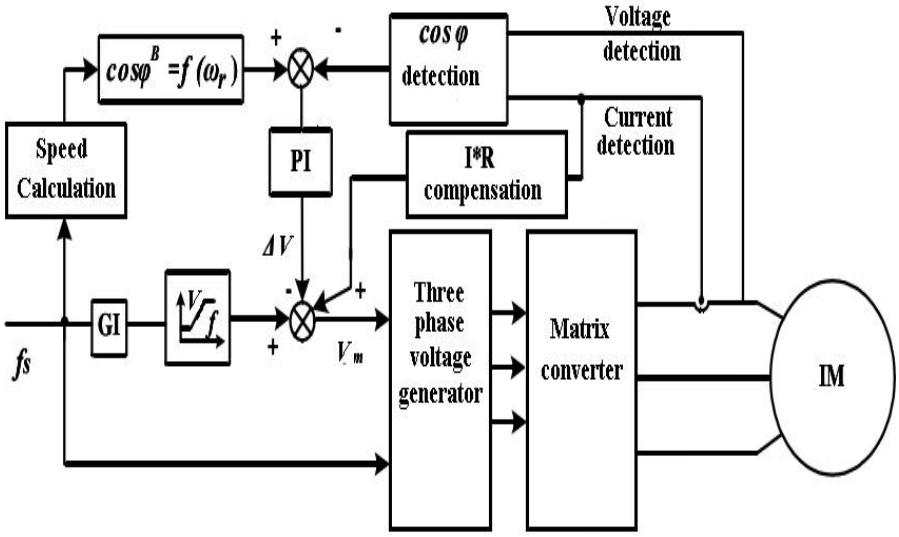


Figure 4. Efficiency optimization control system

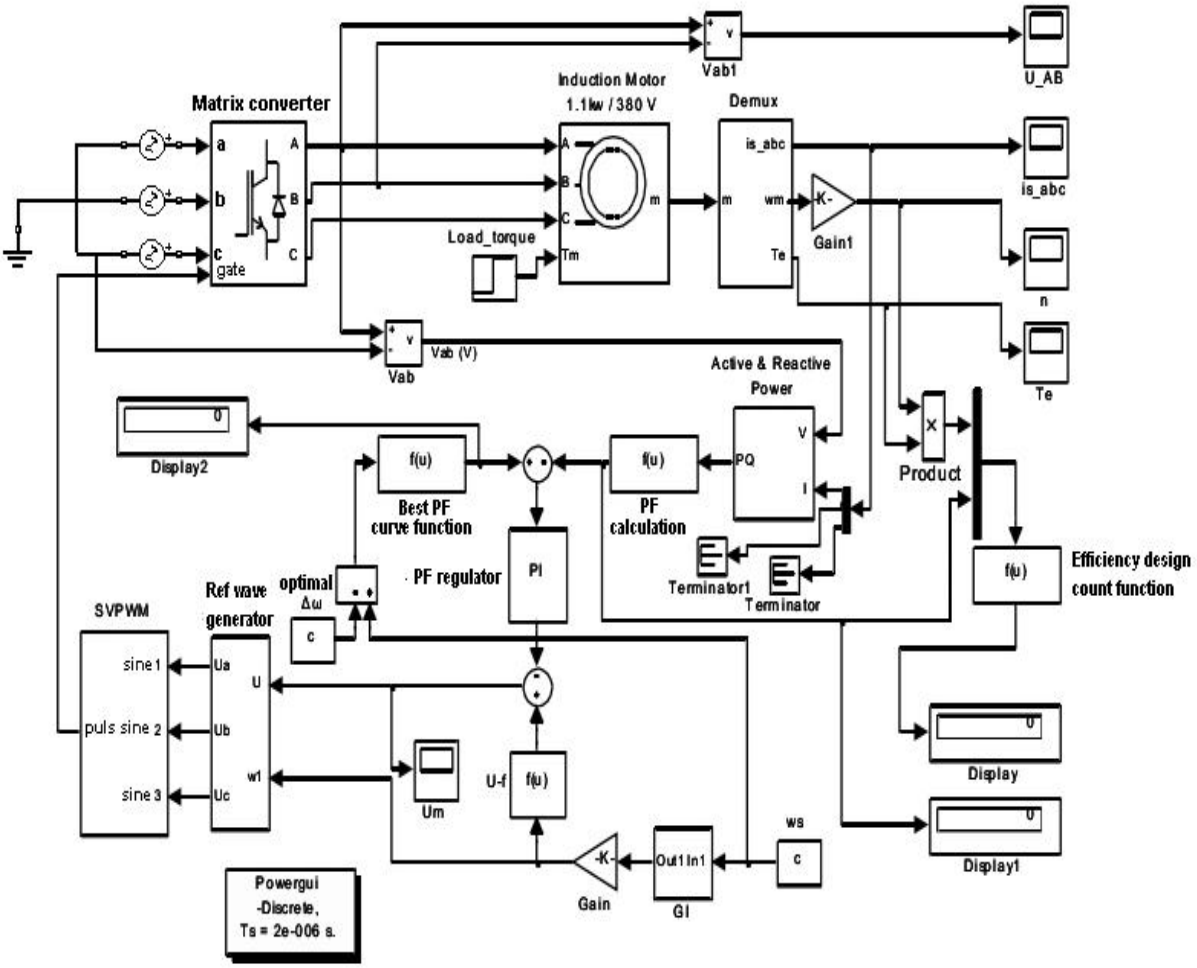


Figure 5. Modelling of efficiency optimization system

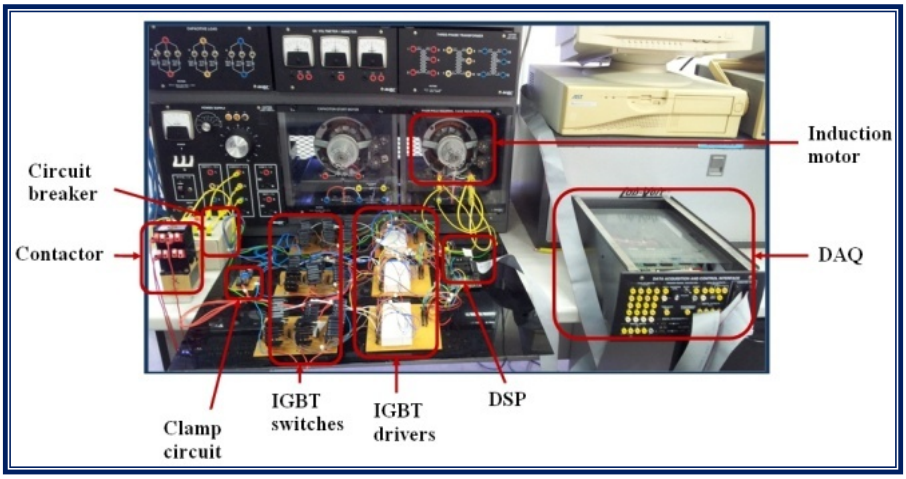


Figure 6. Complete MC hardware

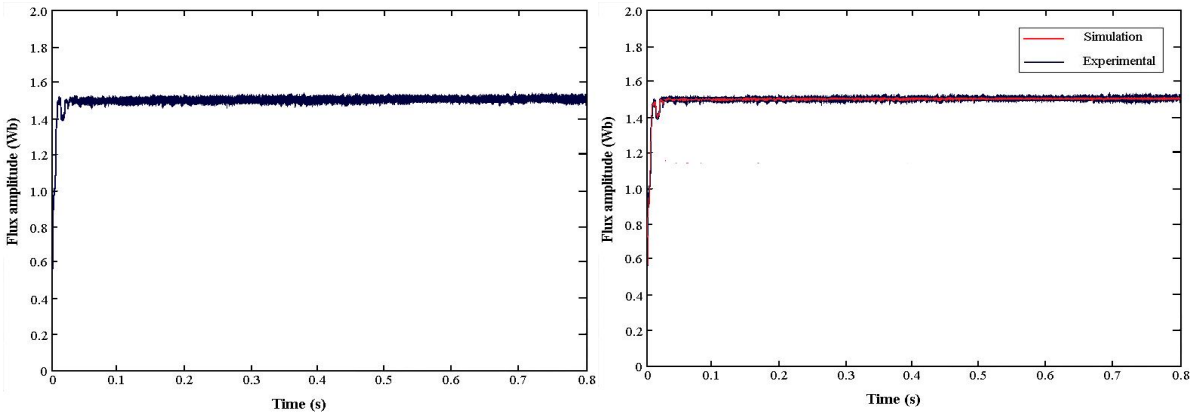


Figure 7 (a) Figure 7 (b)

Figure 7. Flux amplitude; (a) without proposed method; (b) with the proposed method

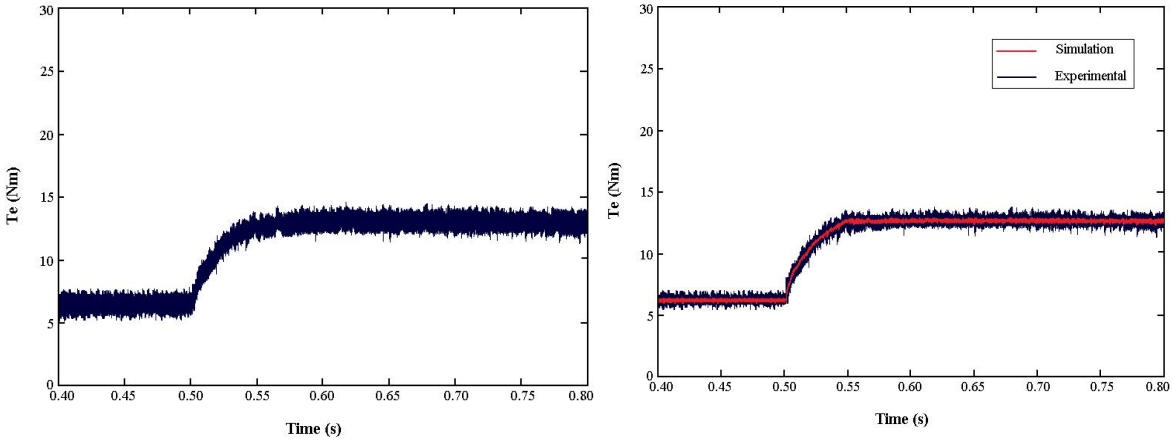


Figure 8 (a) Figure 8 (b)

Figure 8. Motor torque; (a) Torque changes from 6Nm to 12Nm without proposed system; (b) Torque changes from 6Nm to 12Nm with the proposed system

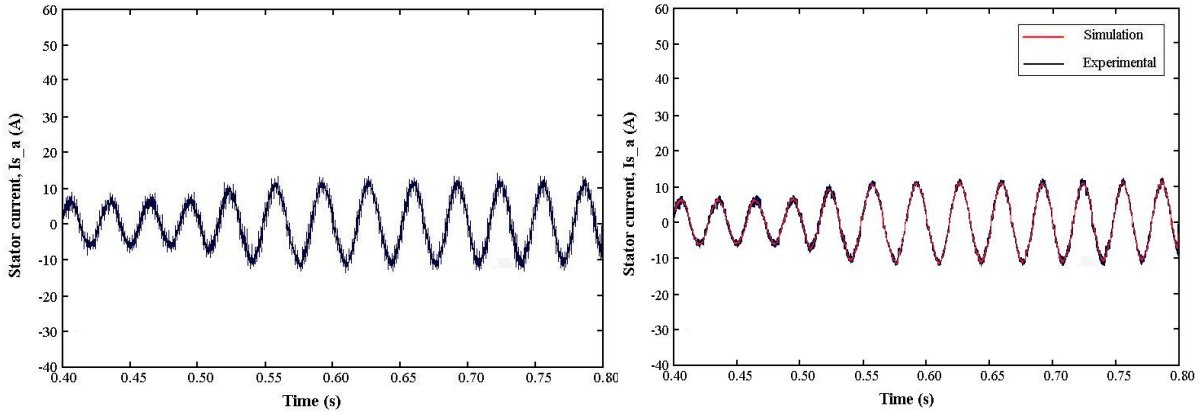
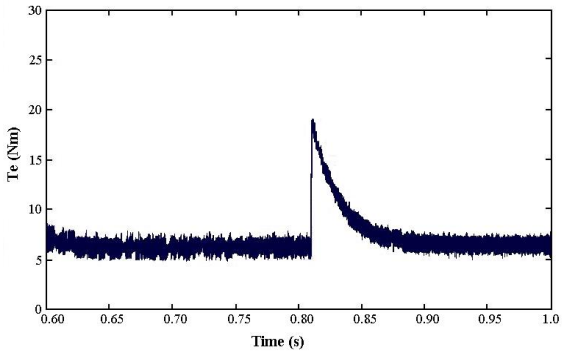


Figure 9 (a) Figure 9 (b)

Figure 9. Stator current (a) without proposed method; (b) with the proposed method



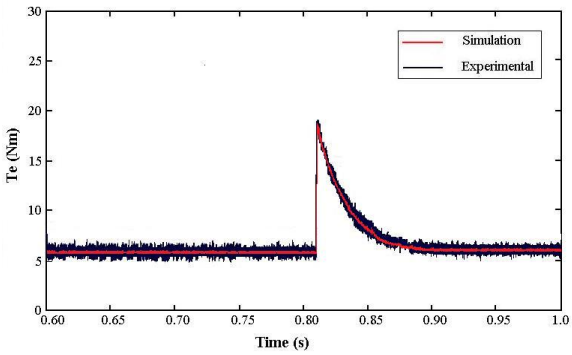
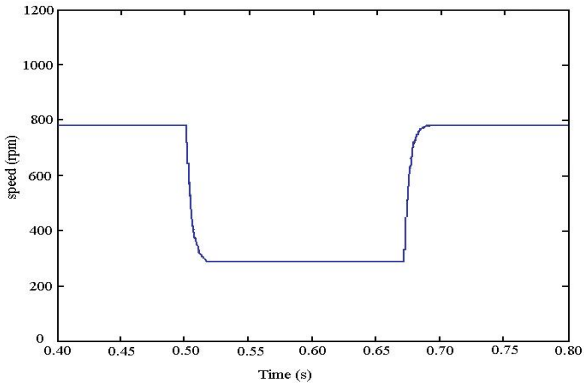


Figure 10 (a) Figure 10 (b)

Figure 10. Effect of changes the motor speed of the torque response (a) speed from 500rpm to 1000rpm (b) Torque response without proposed method



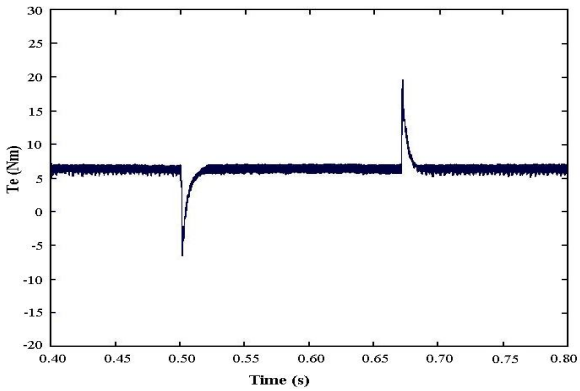


Figure 11 (a) Figure 11 (b)

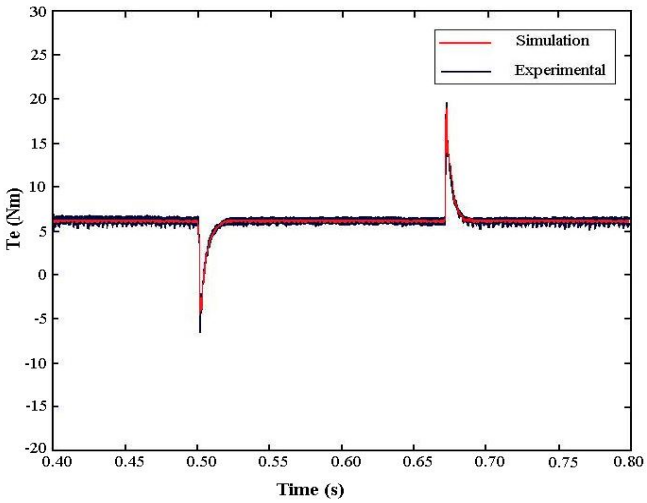


Figure 11 (c)

Figure 11. Dynamic changes of motor speed (a) speed (b) torque responses without proposed system (c) torque responses with the proposed system

References

Alesina, A., & Venturini, M. A. R. C. O. G. B. (1989). Analysis and design of optimum-amplitude nine-switch direct AC-AC converters. *IEEE Transactions on Power Electronics, 4*(1), 101-112.

http://dx.doi.org/10.1109/63.21879

Buja, G. S., & Kazmierkowski, M. P. (2004). Direct torque control of PWM inverter-fed AC motors-a survey. *IEEE Transactions on Industrial Electronics, 51*(4), 744-757. http://dx.doi.org/10.1109/TIE.2004.831717

Casadei, D., Serra, G., & Tani, A. (2000). Implementation of a direct control algorithm for induction motors based on discrete space vector modulation. *IEEE Transactions on Power Electronics, 15*(4), 769-777. http://dx.doi.org/10.1109/63.849048

Chen, D. F., Liao, C. W., & Yao, K. C. (2007, September). Direct torque control for a matrix converter based on induction motor drive systems (pp. 101-101). In Innovative Computing, IEEE.Second International Conference on Information and Control, 2007. ICICIC'07. http://dx.doi.org/10.1109/ICICIC.2007.268

Dung, P. Q., & Thuong, H. T. N. (2004). Direct torque control for induction motor using ANN. In The 2004 International Symposium on Advanced Science and Engineering, HCM City, Vietnam (pp. 60-63).

Ghoni, R., Abdalla, A. N., & Sujod, Z. (2010). Direct torque control for matrix converter-fed three phase induction motor with hybrid pso. *Journal of Theoretical & Applied Information Technology, 13*. Retrieved from http://www.jatit.org/

Habetler, T. G., Profumo, F., Pastorelli, M., & Tolbert, L. M. (1992). Direct torque control of induction machines using space vector modulation. *IEEE Transactions on Industry Applications, 28*(5), 1045-1053. http://dx.doi.org/10.1109/28.158828

Kolar, J. W., Friedli, T., Rodriguez, J., & Wheeler, P. W. (2011). Review of three-phase PWM AC–AC converter topologies. *IEEE Transactions on Industrial Electronics, 58*(11), 4988-5006. http://dx.doi.org/10.1109/TIE.2011.2159353

Lascu, C., Boldea, I., & Blaabjerg, F. (2000). A modified direct torque control for induction motor sensorless drive. *IEEE Transactions on Industry Applications, 36*(1), 122-130. http://dx.doi.org/10.1109/28.821806

Lee, H. H., Nguyen, H. M., Chun, T. W., & Choi, W. H. (2007, October). Implementation of direct torque control method using matrix converter fed induction motor (pp. 51-55). IEEE.International Forum on Strategic Technology, 2007. IFOST 2007. http://dx.doi.org/10.1109/IFOST.2007.4798517

Mei, Y., Sun, K., Zhou, D., Huang, L., & Matsuse, K. (2005, October). Application of matrix converter in auxiliary drive system for Diesel locomotives. IEEE Conference Record of the 2005 Fourtieth IAS Annual Meeting In Industry Applications Conference, 2005 (Vol. 4, pp. 2476-2483). http://dx.doi.org/10.1109/IAS.2005.1518808

Shi, K. L., Chan, T. F., Wong, Y. K., & Ho, S. L. (2001). Direct self control of induction motor based on neural network. *IEEE Transactions on Industry Applications, 37*(5), 1290-1298. http://dx.doi.org/10.1109/28.952504

Zaghloul, M. E., Meador, J. L., & Newcomb, R. W. (1994). Silicon implementation of pulse coded neural networks. Kluwer Academic Publishers. ISBN:0792394496.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Determination of the Content Heavy Metals in of Introduced Tree Stand of Astana City

Erzhan Zhunusovich Kentbayev[1], Botagoz Aidarbekovna Kentbayeva[1] & Talgat Sagidollaevich Abzhanov[1]

[1] Department of forest resources and hunting, The Kazakh National Agricultural University, Almaty, Kazakhstan

Correspondence: Talgat Sagidollaevich Abzhanov, Department of forest resources and hunting,The Kazakh National AgriculturalUniversity, 050010, Abai Avenue 8, Almaty, Kazakhstan. Tel: 871-7272-4350. E-mail: taka_777@mail.ru

Received: January 9, 2015 Accepted: March 20, 2015 Online Published: June 25, 2015

doi:10.5539/mas.v9n8p121 URL: http://dx.doi.org/10.5539/mas.v9n3p121

Abstract

Nowadays, the importance of determining the numericalvaluation of environmental and genetic structure of populations of woody tree species in dealing with the introduction, without any doubt, is recognized by many researchers. The growing numbers of works associated with the introduction, examination of geographic cultures focused on the study of interspecific, inter- and intrapopulation, subpopulation genetic structure of quantitative and qualitative attributes which are important for forestry production and industrial wood processing. Heavy metals that were included in the scientific literature in the middle of the last century under such a negative name, at current stage already occupy the second place in the hazard level. In the future they might become the most hazardous or even more hazardous than APS and solid wastes. Contamination with heavy metals occurs due to their wide use in industry, causing them to fall into the environment bringing enormous damage. Human activity always brings to formation of numerous volumes of waste, including toxic. Industrial enterprises and large number of vehicles - obligatory accompaniments of urban environment are the major and permanent “suppliers” of toxic waste into the atmosphere. Under normal conditions, heavy metals are contained in soil in small quantities and are not hazardous. However, their concentration may increase due to vehicle exhaust, waste, residues and emissions during operation of industrial plants, during fertilize, etc.

Keywords: heavy metals, caragana arborescens, introduced

1. Introduction

An introduction is a transfer of plants from one region to the region where they were missing with the use of impactmethods to the nature. The possibility of introduced plants to acclimate was pointed out by many researchers (Nicholaevsky V.S., 1979, Sergeichik S.A., 1997). However, the naturalization (the transfer of plants to the similar habitats) was more recognized for a long time, which also denies the ability of the plant to adapt to the new conditions.

In 1868, at the initiative of the nobility, the local intelligentsia, amateur gardenersfrom Voronezh, Kharkov, Penza province, as well as Nicholas Botanical Garden, plants, which are several varieties of apple, grape, lilac, oak, marpleplatanoides, mountain ash, horse chestnut, quince, ailanthus, Japanese acacia, were brought. It was the spontaneous introduction. In 1868, Vernenskaya Grove was established by order of the General Governor Kolpakovsky and the scientist and forester E.O. Baum. By 1879 it had an area of 152 acres, including a forest nursery. In 1871, the seeds of pine and larch were produced, and, in 1874, seeds of white and yellow acacia, hawthorn, ash, mulberry were experimentally planted. In the same year, Moscow received seeds of Pinus pinaster, P.peuce, P. strobus, P. cembra, Thuja occidentalis, Juniperus virginiana (Semakin V.P., 1968, Vazhenin I.G., 1987, GOST (State Standards) 1983, SanPin 2002, Clausen, K.E., 1984).

Climatic conditions: Factors and ecological conditions of soil formation and landscape environment. The geographical position of Astana is led by the severity of extreme continental climate and its instability. Winter is cold and long with a steady snowing. Summer is relatively short, but hot with low precipitation and strong evaporation. Sharp continental climate is due to the remoteness of the area from major water ponds and the closeness to the desert and semi-desert areas of Central Asia and the polar regions of Siberia (Bridgwater F.E., et.al., 1983, Birot Y., et.al., 1983, Pollard D.F. et.al., 1974, Stastny T. 1971).

2. Materials and Methods

2.1 Basic Methods of Introduction and Acclimatization of Plants

After studying the flora of Central America (1799-1804) A. Humboldt drew attention to the relationship between the distribution of plants and climatic conditions. He was the first to point out that the acclimatization of plants should take into account not just the average temperature but the amount of temperature above 0 ° C during the growing season (Shimadzu 2008, Wilcox M.D., 1991, Burdon R. D., et.al., 1973, Maronek D.M., et.al., 1974, Campbell R. K., 1974,). For successful acclimatization it should not be less than it was at home conditions. Humboldt paid attention to other climatic factors of habitats of plants, such as average temperature, humidity, pressure and air transparency. He believed that these factors would affect the distribution of plants. In this regard, he suggested the vertical and horizontal zonation of vegetation. Humboldt proposed the method of gradual acclimatization of plants, later called the step acclimatization, which is moving plants from one climate to another, by growing them at intermediate stations (Vincet G., 1974, Krutzsch P., 1975, Eriksson G., et al., 1975, Giertych M., 1976, Szonyi L., et al.,1975, Guidelines., All-Union Scientific Research Institute of Mineral Raw Materials. (AIMRM), 1991).

3. Results and Discussion

3.1 Evaluation of introduced species on Bio Morphological Parameters of Leaf Blades

Plants growing in a large city are influenced by many different factors. It is especially painful for plants to tolerate exposure to man-made factors, an essential condition of civilization. Conditions of the natural habitat of plants are radically different from those of the city.

The plants, which were grown at the best environments, adapt and often survive in extreme conditions by changing not only the biological and physiological processes, but biological and morphological parameters, the anatomical structure that is reflected in the change of habitus, shape and size of the leaf blades, fruit, etc.

3.2 The Length of the Leaf Blades

The length of the leaf blades is the main parameter which determines their size and shape. We studied the length of the leaf blades of 13 species of plants which were growing at three test sites. Research and statistics of materials are given in Tables 1, 2, 3. The value of the error in mean showsrelatively narrow confidence limits, which indicatesa highly accurate obtained value. The arithmetic means was calculated as a criterion for evaluation of the entire group of plants.

On the first test site, all plants formed a complete leaf blade except Acer ginnala, on the second site - all the plants, and the third section –Juglans mandshurica and Saragana arborescens did not form leaf blades. Length of the leaf blades of Fraxinus excelsior L., which was planted on the test sites, varies within the following limits: 1test site - 66.90 mm, 2 test site - 71.27 mm, 3 site - 86.00 mm. Populus simonii had the longest leaves on the first test site - 55.83 mm. Prunus divaricata, Salix babilonica, Saragana arborescens, Populus simonii, Radus aviumare the leaders on the length of the leaf blades of the first test site (Figure 1).

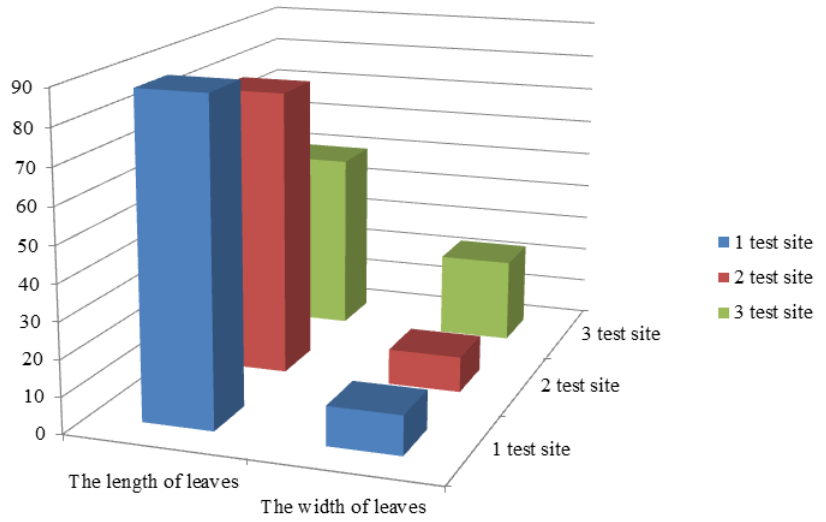


Figure 1. Linear parameters of leave blades of Salix babilonica from tree test sites

Ranking the objects of study according to the length of leaves revealed the plants, which had maximum value: Amorpha fruticosa, Phellodendron amurense, Gled itsiatriacanthos, Acer ginnala, Mahonia aquifolia, Juglansmands hurica, Forsythia xintermedia Zabel. According to the arithmetic mean data, they formed the longest leaves on the second site (Figure 2).

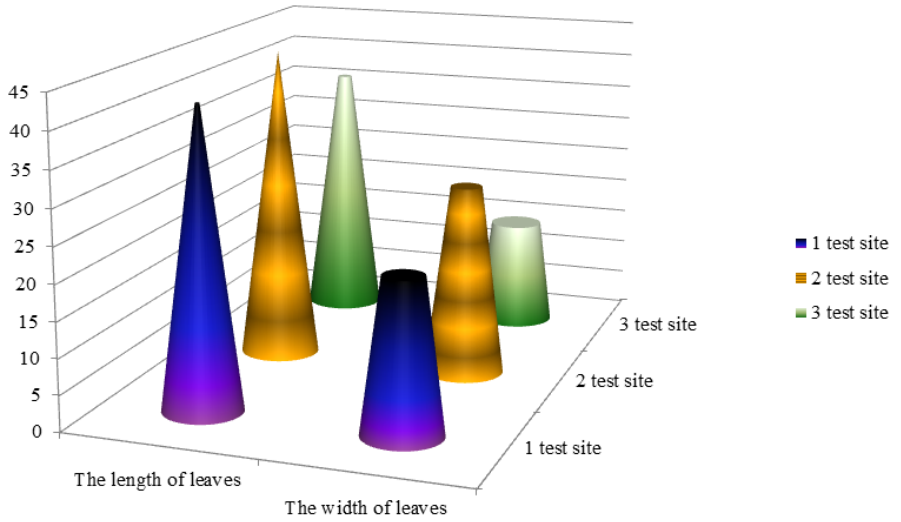


Figure 2. Linear parameters of leave blades of Mahonia aquifolia from tree test sites

Leading positions according to the length of the leaf blades of the first test siteare occupied by Saragana arborescens, Prunus divaricata, Salix babilonica, Populus simonii, Radus avium.

The highest values of the arithmetic means of the width of the second site of leaves belong to Amorpha fruticosa, Phellodendron amurense, Gled itsiatriacanthos, Acer ginnala, Mahonia aquifolia, Juglans mandshurica.

In the third site, only two species, a Fraxinus excelsior L. and Forsythia xintermedia Zabel, have a predominance of the width of the leaf blade.

Percentage ratio of the number of plants with a maximum length of leaf blades on the first site is 38.47%, in the second site - 46.16%, and the third site - 18.19%.

The average level of variation coefficients dominates on the first and third sites in 8-cases. The width of the leaf blades of the studied plants have a low, medium and upper level of variability, measured by the values of the variation coefficients according to the scale of variability. On the first site the medium level of variation is 5 cases out of 12, a low level - 4 cases out of 12, an upper level is 3 out of 12 cases (Acer ginnala has not formed leaves). According to the second site results, 5 plants have the low level of variability, 8 plants – the average level, and no plants have an upper level. In the third site, 2 types of experimental plants did not form leaves (Saragana arborescens, Juglans mandshurica), 1 case with low level of variation, 8 cases of medium variation and 3 cases of upper variation. Limits vary widely.

Accuracy of the experiment in all cases is less than 5% level, which indicates the high accuracy of the experiments.

From the tables and visual observations we can say that the subjects are characterized by linear parameters of the leaf blades. The wide range of studied plants is characterized by a high degree of environmental heterogeneity. Differences were found in all areas which are mostly characterized as significant within the same type of planting. However, significant differences were found among species growing on different sites. Relatively favorable conditions for planting in the park area, which is protected from the negative effects of transport and other human impact, had a positive effect on the growth and development of plants, in general, and on the change in individual organs, in particular.

Table 1. Length and width of the leaf blades of inducted woody plants on the 1 test site

N	Name of woody plant species	Arithmetic mean	Variability	Test accuracy,	Limits	
		M ± m, mm	index, Cv, %	P, %	min	max
1	Caragana arborescens	20,80±0,65	17,08	3,12	14	25
	(f. pendula)	10,03 ± 0,26	14,44	2,64	7	13
2	Prunus divaricata	34,13±1,59	25,58	4,67	20	50
		14,87 ± 0,70	25,86	4,72	10	22
3	Amorpha fruticosa	22,80±0,91	21,85	3,99	11	30
		9,83 ± 0,45	25,20	4,60	7	13
4	Phellodendron amurense	75,57±2,20	15,96	2,91	60	102
		32,73±0,70	11,72	2,14	27	40
5	Gled itsiatriacanthos	28,33±1,17	22,66	4,14	15	36
		11,57±0,34	15,89	2,90	8	14
6	Salix babilonica	88,17±4,18	25,95	4,75	55	145
		10,77±0,50	25,34	4,63	8	18
7	Acer ginnala	-	-	-	-	-
		-	-	-	-	-
8	Mahonia aquifolia	42,10±0,87	11,31	2,07	35	52,0
		21,70±0,77	19,52	3,56	15	30,0
9	Juglans mandshurica	49,20±1,09	12,17	2,22	41	63,0
		25,30±065	14,15	2,58	19	32
10	Populus simonii	55,83±2,08	20,43	3,7	29	75
		34,00±1,38	22,23	4,08	21	47
11	Forsythia xintermedia Zabel	39,37±1,57	21,05	3,84	20	49
		20,50±0,80	21,29	3,89	10	25
12	Padus avium	59,43±1,63	14,93	2,74	46	74
		32,13 ± 0,48	8,26	1,51	29	37
13	Fraxinus excelsior L.	66,90±1,69	13,81	2,52	50	80
		18,43 ± 0,69	20,38	3,72	14	25

Only one type of test plant has the highest value of the length of the leaves in all three areas - a Fraxinus excelsior L. The weak development of the parameters of the leaf blades on the third experimental site is most likely connected with the relative rigidity of the environment in this area of research. All three sites are located in the parks, recreation areas; eco factor complex and mainly watering had a positive effect on the development and status of the leaf blades, which is reflected on the length in this case.

Percentage ratio of the number of plants with a maximum length of leaf blades on test site: 1 test site (5 of 12 plants with formed leaves) - 41.67%, 2 test site (7 of 13 plants with formed leaves) - 53.85%, 3 test site (1 of 11 plants with formed leaves) - 9.09%. Plants which were planted on the 1 test site (located in the park near the Presidential Park near Palace of Peace and Reconciliation) and the second test site (located in the park near the shopping and entertainment center "Khan Shatyr") even by visual inspection showed a better rate.

Table 2. Length and width of the leaf blades of inducted woody plants on the 2 test site

№	Name of woody plant species	Arithmetic mean M±m, mm	Variability index, Cv, %	Test accuracy, P, %	Limits min	max
1	Caragana arborescens	16,10 ± 0,64	21,66	3,95	10	21
		8,37 ± 0,13	8,44	1,54	7	9
2	Prunus divaricata	33,07 ± 1,20	19,92	3,64	21	43
		13,40 ± 0,63	25,75	4,70	8	20
3	Amorpha fruticosa	28,63 ± 0,57	10,91	1,99	23	33
		10,93 ± 0,23	13,31	2,06	9	13
4	Phellodendron amurense	84,23±1,53	9,95	1,82	73	100
		35,17±0,83	12,97	2,37	31	45
5	Gled itsiatriacanthos	36,97±1,59	23,53	4,30	24	57
		13,43±0,49	19,95	3,64	10	18,0

6	Salix babilonica	79,27±3,95	27,32	4,99	50	108,0
		9,87±0,42	23,22	4,24	6	13
7	Acer ginnala	44,70±0,98	11,98	2,19	37	54
		39,93±1,07	14,69	2,68	27	47
8	Mahonia aquifolia	44,53±0,92	11,28	2,06	37	52
		27,10±0,75	15,20	2,77	20	34
9	Juglans mandshurica	51,17±1,38	14,74	2,69	39	65
		28,40±0,85	16,43	3,00	15	36,0
10	Populus simonii	51,37±1,61	17,18	3,14	35	65
		33,00±1,14	18,91	3,45	20	42
11	Forsythia xintermedia Zabel	62,07±2,62	23,13	4,22	33	84
		31,57±1,28	22,26	4,06	9	42,0
12	Padus avium	52,10±0,96	10,12	1,85	45	60
		31,93±0,90	15,45	2,82	25	40
13	Caragana arborescens	71,27 ± 1,44	11,10	2,03	58	82
		21,87±0,87	21,89	4,00	14	36

It is important to mention the differences of maximum and minimum average of leaf length in Phellodendron amurense between sites. The first site's leaf length is 75.57 mm, the second's - 84.23 mm, the length of the leaves of the third portion is 48.63 mm. The range of variation is observed at the level of 35.60 mm. In Salix babilonica, the limit of variability of arithmetic mean is equal to 38.34 mm, in Forsythia xintermedia Zabel - 22, 20 mm.

The trait variability scale according to S.A. Mamaeva levels is estimated by the values of the coefficients of variation and refers to the low, medium and upper level. On the first test site the medium level of variation is dominated - 6 of 12 cases (one kind of plant has not formed leaves - Acer ginnala). According to the second site results 7 plants have the low level of variation, 5 plants – the medium level, the upper level - 1 plant. In the third site there were 4 cases of low level variation, 5 cases of medium variation and 2 cases of upper variation of the 11 plants which have formed leaves.

Table 3. Length and width of the leaf blades of inducted woody plants on the 3 test site

№	Name of woody plant species	Arithmetic mean M±m, mm	Variability index, Cv, %	Test accuracy, P, %	Limits	
					min	max
1	Caragana arborescens	-	-	-	-	-
		-	-	-	-	-
2	Prunus divaricata	30,53 ± 1,10	19,79	3,61	20	39
		12,93 ± 0,55	23,19	4,23	8	18
3	Amorpha fruticosa	24,03±1,10	24,98	4,56	15	30
		9,53 ± 0,45	25,80	4,71	5	16
4	Phellodendron amurense	48,63±2,16	24,30	4,44	29	65
		25,70±1,18	25,10	4,58	16	38
5	Gled itsiatriacanthos	22,50 ± 1,06	25,83	4,72	13	31
		11,27±0,53	25,72	4,70	5	15
6	Salix babilonica	49,83 ± 2,37	26,03	4,75	30	66
		6,47±0,22	18,61	3,40	5	8
7	Acer ginnala	31,77 ± 0,82	14,09	2,57	25	41
		25,10±0,73	15,85	2,89	20	33
8	Mahonia aquifolia	36,47 ± 0,88	13,17	2,40	29	44
		18,97±0,55	15,96	2,91	15	24
9	Juglans mandshurica	-	-	-	-	-
		-	-	-	-	-
10	Populus simonii	44,90 ± 1,64	19,98	3,65	30	60
		24,37±0,69	15,55	2,84	17	30
11	Forsythia xintermedia Zabel	61,57 ± 0,96	8,57	1,56	53	70
		33,17±1,31	21,60	3,94	27	59

12	Padus avium	53,90 ± 2,15	21,82	3,98	32	70
		30,23±1,43	25,90	4,73	15	41
13	Caragana arborescens	86,00 ± 1,60	10,17	1,86	73	104
		30,73±0,46	8,23	1,50	25	35

Accuracy of the experiment, which shows the correct formulation and implementation of research in all cases, is within an acceptable range and does not exceed 5%. Maximum and minimum limits, which are the limits can vary within wide range, determines the data indicated in the tables of arithmetic means of 13 kinds of plants.

The width of the leaves. As well as the length, width of the leaves is the main parameter that creates the appearance of plants, reflecting in its shape and area (Table 1, 2, 3). Leaf width determines its elasticity and promotes opposition to external influence of the nature. During the dry season the leaves of some plants fold their edges, thereby reducing the exposure to the sun and minimizing the rate of evaporation while increasing water-holding capacity. The width of the leaves is a regulatory mechanism and has great practical significance.

Depending on the site the width leaves of Prunus divaricata varies within the following limits: 1 site - 14.87 mm, 2 site - 13.40 mm, 3 site - 12.93 mm. Fraxinus excelsior L. has the widest leaves formed on the third section - 30.73 mm.

4. Conclusion

The following conclusions can be made based on performed studies:

1. Analysis of the current range of Astana woody plants and forest nursery JSC “Astana Zelenstroy” indicates lack of species diversity. The identified lack admits introduction events in Astana. Based on the study of biological characteristics of woody and shrubby plants, an assortment of 50 woody species can be used in landscaping Astana was formed.

2. Test landings of 13 of 50 species of woody plant introductions were conducted in Astana. Plants imported from the JSC “Forest Nursery” of Almaty region located 53 km east of Almaty. Pilot group of plants was planted on three experimental sites: *Caragana arborescens (f. pendula); Prunus divaricate; Amorpha fruticosa; Phellodendron amurense;Gleditsia triacanthos; Salix babilonica; Acer ginnala; Mahonia aquifolia; Juglans mandshurica; Populus simonii; Forsythia xintermedia Zabel; Padus avium; Fraxinus excelsior L.*

The separation of plants for three sites will allow for further full comparative observations of introduced plants growth and development.

The greatest impact of the complex environmental factors were felt by plants of the third test site, which are reflected in the linear parameters of the studied plants.

Thus, we can conclude that the plants which were planted in the environmental sections produce the leaf blades differentiated by the linear parameters. Considering the uniformity of the studied plants we can say that the identified differences in the experiments are strongly influenced by environmental conditions.

References

Birot, Y. (1983). Christophe C Genetic structures and expected genrtic gains multitrait selection in wild populations of Douglas fir and Sitka spruce. I. Genetic variaton between and within populations. *Silvae genet., Bd.32*(5-6), 141-151.

Bridgwater, F. E., Talbert, J. T., & Jahromi, S. (1983). Index selection for increased dry weight in a young loblolly pine population. *Silvae Genet., Bd.32*(N5-6), 157-161.

Burdon, R. D., & Bannister, M. H. (1973). Provenances of Pinus radiata: Their earlyperformance and silvicultural potential. *N.Z.J. Forestry, 18*(N2), 217-232.

Campbell, R. K. (1974). A provenance - transfer model for boreal regions. *Medd.Norsk Inst. Skogiorsk., Bd.31*(10), 543-566, ill.

Clausen, K. E. (1984). Nonlinear regressions and contour plots: techniques for selection and transfer of white ash provenances. *Forest. Sci., 30*(N2), 441-453.

Eriksson, G., Andersson, S., & Schelander, B. (1975). Lovande tillvaxt hos introduceradegranprovenienser i en kombinerad klonoch froplantsplantage i norra Uppland. *Sver. skogsvardsford. tidskr., Bd.73*(N3), 277-286.

Giertych, M. (1976). Summary results of the IUFRO 1938 Norway spruce (Piceaabies (L.) Karst) provenance experiment. Height growth. *Silvae Genet., Bd. 25*(N5-6), 154-164.

GOST (State Standards)(1983). Environment protection. Soils. Classification of chemicals for pollution control.

Krutzsch, P. (1975). Die Pflanzschulenergebniss eines inventierenden Fichtenherkunftsversuches. (Picea abies Karst. und Picea obovata Ledeb.). *Rapp. ochUppsats. Inst. skogsgenet. Skogshosk., N14*, 141.

Maronek, D. M., & Flint, H. L. (1974). Cold hardiness of needles of Pinus strobus L. as afunction of geographic source. *Forest Sci., 20*(N2), 135-141.

Moscow (1991). Method of metals determination in plants. Fedorovsky All-Union Scientific Research Institute of Mineral Raw Materials. (AIMRM).

Nicholaevsky, V. S. (1979). Biological basis of plants gas resistance. - Novosibirsk: Nauka, 278.

Pollard, D. F. W., & Londan, K. T. (1974). The role of free growth in the differentiation ofprovenances of black spruce Picea mariana (Mill.). *Can J. Forest Res., 4*(N3), 308-311.

SanPin (2002). The sanitary-epidemiological rules and regulations Hygienic requirements to the safety and mcg/g nutritional value of foods.

Semakin, V. P. (1968). Clonal selection in gardening. - Moscow: Kolos, 136.

Sergeichik, S. A. (1997). Plants and Ecology. - Minsk: Uradzhay, 224.

Shimadzu. (2008). Atomic absorption spectrophotometer A.A.,– 7000, “Shimadzu” corporation, 1-5.

Stastny, T. (1971). Modifikovanie prejavu genetickej podstaty rastu Larix deciduasMill, vplyvom rozdielnych podmienok prostredia. *Lesn. Stud., 10*, 101.

Szonyi, L., & Ujvari, F. (1975). First results of the international (IUFRO) Norway spruceprovenance experiment. *Erdesz. Kut., Kot. 71*(N2), 139-147.

Vazhenin, I. G. (1987).Guidelines for the study of dispersal principles for emissions in the vicinity of industrial enterprises Dokucheva Soil Institute. Moscow, 145.

Vincet, G. (1974). Produkce ekotypu teze dfeviny, jejich pfenos a kfizeni. *Lesnictvi, 20*(N8), 717-730.

Wilcox, M. D. (1982). Selection of genetically superior Eucalyptus regna’s using family tests. *N.Z.J. Forest. Sci., 12*(3), 480-493.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Control over the Process of Thermo-Oxidative Degradation of Polymers in a Solution

Igor Anatolyevich Khaustov[1], Vitaly Ksenofontovich Bitukov[1], Sergey Germanovich Tikhomirov[1], Anatoly Anatolyevich Khvostov[1] & Alexei Petrovich Popov[1]

[1] Voronezh State University of Engineering Technologies, 19, Revolutsii ave., 394036, Russia

Correspondence: Vitaly Ksenofontovich Bitukov, Voronezh State University of Engineering Technologies, 19, Revolutsii ave., 394036, Russia.

Received: January 9, 2015 Accepted: January 19, 2015 Online Published: June 25, 2015

doi:10.5539/mas.v9n8p128 URL: http://dx.doi.org/10.5539/mas.v9n3p128

Abstract

The article discusses the process of low molecular rubber synthesis by the method of thermo-oxidative degradation as a controlled object. It was found that the control parameters affecting the degradation rate are the following: concentration of reaction initiator, temperature and mass concentration of the polymer. The main control parameters are the following: degree of degradation, quality of the resulting polymer and time of reaction. It was found that for effective control the process temperature and concentration of the initiator are to be stabilized in the vicinity of the predetermined value, and mass concentration is determined by dosage of initial components. In this case, the kinetics of the degradation process can be described as a linear function of time. I.e., the degree of degradation is linearly related to duration of reaction. Then, defining the initial quality of the polymer (its average molecular weight), the degradation process is performed till the time the predetermined value of degradation is reached.

Since the initiator is consumed during the degradation, in order to stabilize its concentration, it is proposed to add the initiator into the reaction medium. We obtained a mathematical description that describes the kinetics of the process during continuous and divided feeding of the initiator, and dependency for calculating initiator feed rate. It has been shown that if the initiator is fed in a divided manner, linearity of degradation kinetics over time is ensured, which simplifies the technology of the process.

Parameters of the approximating linear dependence of the kinetics of degradation have been determined. It has been shown that, even at the initial stage, the kinetics of degradation steadily tends to the linear asymptote. The asymptote's tilt angle depends on the constant of degradation rate (and consequently in the temperature of reaction) and on the initial concentration of the initiator. Degradation kinetics offset by the Y-axis depends on the initial molar concentration of the polymer. Thus, it is possible to control degradation rate by adjusting initial polymer concentration.

Keywords: thermal oxidative degradation of polymers, control system

1. Introduction

One of the promising methods of obtaining low molecular polymers with adjustable molecular weight is the thermo-oxidative degradation. Such polymers are used in the national economy for manufacturing adhesive compositions, anti-corrosion coatings, polymer-and-bitumen compositions, construction sealants, etc.

The process of obtaining low molecular polymers on destruction on the basis of high-molecular polymers generally consists of the following basic steps: loading components of the reaction (solvent, polymer); heating the reaction mass to the predetermined temperature and dissolution of the polymer; loading initiator; degradation to the predetermined degree of polymer degradation; loading inhibitor - end of degradation; cooling of the reaction mass; and drying and recovering resulting low molecular polymer. The main stage of the process in question is the destruction process, since it is at this stage where the main indicators of product quality reformed (average molecular weight of the polymer and the polydispersity index).

In this context, the objectives of studying the system for controlling degradation process are the following:

- adjusting the degree of polymer degradation, which is a variable and depends on what kind of polymer the technology should produce, and in what area it will be used;

- calculation of controlling actions that would allow to achieve the desired degree of degradation in the period of reaction defined in the technology.

Controlling actions that affect the behavior of the degradation process include the following parameters: temperature of the degradation process, concentration (volume ratio) of the polymer in the solution, concentration of the initiator of degradation (in moles), the ratio of decomposer and polymer concentrations in the reaction mixture, intensity of bubbling reaction mixture with atmospheric air; and feed rate of the initiator. Block diagram of thecontrolling object is shown in Figure 1.

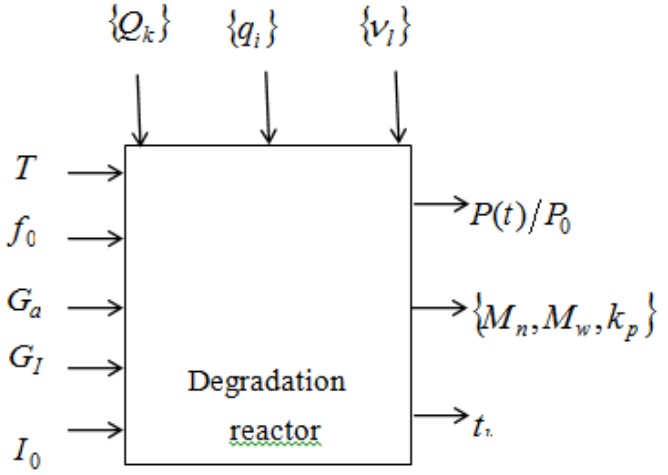


Figure 1. Block diagram of the controlling object

In Fig. 1 $P(t)/P_0$ is the degree of polymer degradation; $\{M_n, M_w, k_p\}$ are parameters of polymer quality (numerical-average and weight-average molecular weight of the polymer, polydispersity index); $P_0, P(t)$ is the initial and the current concentration of the polymer [mole/l]; I_0 is the initial concentration of the initiator [mole/l]; f_0 is mass concentration of the polymer, [kg/l]; T is the temperature of the reactor [°C]; t_k is degradation time [min]; G_a, G_I are feed rates of air and initiator into the reactor [kg/min]; $\{q_i\}$ are thermal parameters (heat capacity, heat transfer coefficient, etc.); $\{\nu_l\}$ are physical and rheological properties of the reaction mixture (viscosity, density, concentration of substances formed during secondary processes, etc.); and $\{Q_k\}$ is the set that characterizes thermal agitations (heat generation due to the reaction, mixing, temperature of environment, coolant, etc.).

Resolution of these tasks is directly related to studying kinetics of the controlled decomposition, including mathematical modeling.

The problem of studying degradation of polymers by simulation is described in many works of both domestic (Berlin et al., 1978; Bryk, 1989; Koptelov and Milekhin, 2012) and foreign researchers (Ivanova et al., 2012; Browarzik and Koch, 1996; Ziff and McGrady, 1986). Studying of the polymer degradation process in a solution by modeling is described in a relatively small number of studies (Sivalingam and Madras, 2004; Sterling et al., 2001).

2. Method

For the purpose of synthesis and studying of control systems, in work (Tikhomirov et al., 2013) kinetic equations were obtained in the form of analytic dependencies that describe the dynamics of component changes in the reaction mixture - concentration of polymer, process initiator and decomposition agent.

In order to describe the kinetics of polymer thermal oxidative degradation in the solution, the following assumption was made - in the ideal case, one mole of the initiator produces 2 moles of the decomposition agent, and reaction of one mole of the decomposition agent with one mode of the polymer produces 2 moles of the polymer, i.e., the following kinetic scheme is used:

$$\mathrm{I} \xrightarrow{k2} 2\mathrm{D}$$

$$\mathrm{D}+\mathrm{P} \xrightarrow{k1} 2\mathrm{P}$$

In modeling the process, the following assumptions were used:

1. Due to breaking of double bonds and elementary act of decomposition, one molecule of the decomposition agent is spent for formation of two polymer macro-molecules with a hydroxy end group on one of the two macro molecules of the polymer.
2. Both macro-molecules formed during the reaction of degradation belong to the same class of macromolecules as the original macro-molecule of the polymer, i.e. $P^* \equiv P$.
3. The degradation rate constant does not depend on the length of the macromolecule.
4. Exhaustion of the decomposition agent is described by a first-order reaction, whose rate is independent of temperature and pressure in the reactor.
5. The degradation rate constant is many times greater than the constant speed of deactivation of the decomposing agent.
6. In the degradation reactor, intensive stirring occurs.

The kinetics of thermal oxidative degradation of the polymer in the solution in accordance with the kinetic scheme and assumptions is described in general by the system of equations:

$$\frac{dI(t)}{dt} = -2k_2 I(t)\,. \tag{1}$$

$$\frac{dD(t)}{dt} = 2k_2 I(t) - k_1 P(t) D(t)\,. \tag{2}$$

$$\frac{dP(t)}{dt} = k_1 P(t) D(t)\,. \tag{3}$$

$$P(0) = P_0,\ D(0) = 0,\ I(0) = I_0\,, \tag{4}$$

where $D(t)$ is concentration of the decomposing agent, [mole/l]; k_2 is the constant of formation rate of the decomposing agent, [1/min]; k_1 is the constant of decomposition rate, [l/(mole·min)].

The expressions describe the kinetics of change in concentration of the initiator, the polymer and the decomposition agent in the volume of the reaction mixture during the reaction in isothermal conditions provided that the process is performed in a batch reactor with constant intensive stirring and bubbling with atmospheric air.

The influence of the reaction mass temperature on the rate of chemical degradation is described by the Arrhenius equation

$$k_j = k_{0j} \cdot \exp\left(-E_j / RT\right), \tag{5}$$

where R = 8.315 is the universal gas constant, [J/(mole·°K)]; k_{0j} is the pre-exponential constant, dimensions correspond to k_j; E_j is the energy of activation [J/mole]; $j = \overline{1,2}$.

For equation (5), estimates of activation energies were obtained for chemical reactions and pre-exponential constants: k_{01} = 7.96·1013 [l/(mole·min)]; k_{02} =2.3·1010 [1/min]; E_1 = 81.8 [KJ/mole]; E_2 = 86.4 [KJ/mole]. According to the literature, the value of activation energy E_2 = 83.7 [KJ/mole] (Geller et al., 1996), which confirms veracity of the estimates.

Adjustment of the reaction temperature is an effective way to control the rate of chemical degradation. As it follows from Fig. 2, the reaction rate increases exponentially with increasing temperature. However, at 70 °C, the reverse process, jellification is observed, in spite of introduction of antioxidants. This can be seen in Figure 3 and 4. After four hours, according to the experimental data, the degradation rate decreases, as compared to the calculated theoretical concentration values (Figure 3). This corresponds to increasing of jellification rate after 4 hours of degradation (Figure 4).

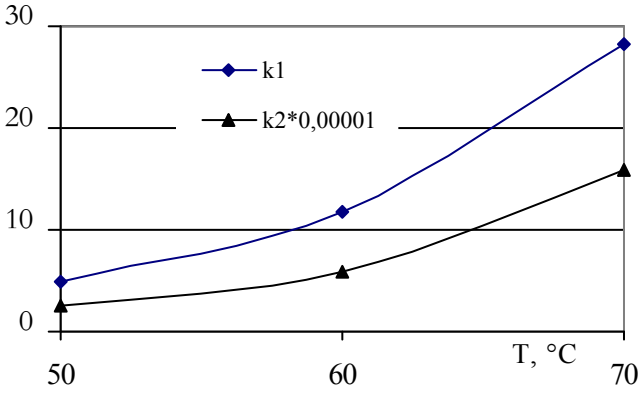


Figure 2. Influence of reaction temperature on degradation rate

The difference between experimental and calculated data at 70 °C is explained by the fact that at higher temperatures, the intensity of secondary processes increases - polymer chains recombination reaction, which leads to cross-linking of the macromolecules, and as a consequence, to reducing their concentrations. In this regard, introduction of antioxidants cannot completely prevent jellification.

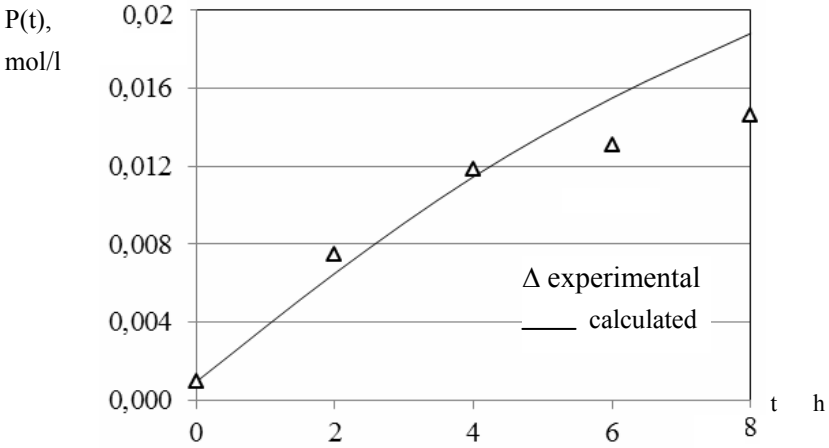


Figure 3. The dynamics of concentration change during polymer degradation, $T = 70$ °C

It is clear that at high temperatures, the rate of formation of gels increases. In this regard, for performing the chemical degradation with maximum efficiency it is reasonable to use stabilization of the maximum temperature. In this case jellification does not occur, or can be prevented.

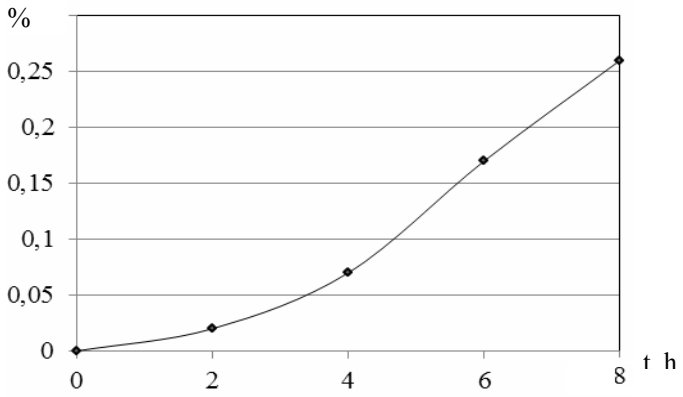


Figure 4. Dynamics of peroxide compounds formation during degradation, $T = 70$ °C

The possibility of controlling the degradation process by adjusting the rate of bubbling of the reaction mass with the atmospheric air is limited by the rate of air dissolution in the solution and by the limit of its saturation with oxygen. In this case, air supply into the reaction zone with the maximum efficiency should be ensured, i.e. $G_a > G_{\max}(t)$, where $G_{\max}(t)$ is the maximum rate of feeding air into the reactor, which ensures saturation of the solution with oxygen, [kg/min].

Increasing the mass concentration of the polymer in the solution leads to accelerated degradation process. However, at the same time, as experimental studies show, the rate of jellification increases, since with increasing concentration of the solution, the probability of polymer molecules recombination increases as well. The permissible concentration of polymer of 10% has been found experimentally; in case of a higher concentration, active jellification starts.

The influence of the initial charge of initiator and the possibility to control degradation rate by adjusting concentration of the initiator has been described in the mathematical model (1) - (3).

3. Results

3.1 Control on the Basis of Continuous Dosing of the Initiator

Suppose the initiator is continuously introduced into the reaction zone at the rate of $G_I(t)$ [kg/min], so that the concentration of the initiator in the reaction mixture is maintained constant during the entire process of degradation, i.e.

$$I(t) = I_0 = const . \tag{6}$$

Let us find the solution (1)-(3) by introducing an assumption that the volume of the reaction mass slightly increases due to additional dosing of initiator into the reactor, i.e. $V_0 >> V_I$.

In this case, the change in concentration of the reactants caused by increase in the volume is not considered, and the initiator feed rate into the reactor is calculated basing on conditions $-k_2 I(t) + \frac{G_I(t)}{V(t)M_I} = 0$, i.e.

$$G_I(t) = k_2 I_0 V_0 M_I . \tag{7}$$

where M_I is the molar mass of the initiator [kg/mole].

Then the initial system of equations is reduced to

$$\frac{dD(t)}{dt} = 2k_2 I_0 - k_1 P(t) D(t) . \tag{8}$$

$$\frac{dP(t)}{dt} = k_1 P(t) D(t) . \tag{9}$$

The resolution of a system of equations (8) - (9) using initial conditions (4) is obtained in form

$$P(t) = \frac{\varphi(t)}{k_1 \psi(t) + 1/P_0} . \tag{10}$$

$$D(t) = P_0 + 2k_2 \cdot I_0 \cdot t - P(t) . \tag{11}$$

where $\psi(t) = \int_0^t e^{k_1 \cdot P_0 \cdot \tau + 2k_1 \cdot k_2 \cdot I_0 \cdot \tau^2} d\tau$, $\varphi(t) = e^{k_1 \cdot P_0 \cdot t + 2k_1 \cdot k_2 \cdot I_0 \cdot t^2}$.

Figure 5 shows the variants of polymer degradation for different initial conditions of equations calculation (10). Curve 1 corresponds to the reaction flow with the initial starting dosage. Curve 2 - the initial concentration of the polymer is increased 10 times. Curve 3 - the initial concentration of the initiator is increased 5 times. In all 3 cases, the initial concentration of initiator is maintained constant.

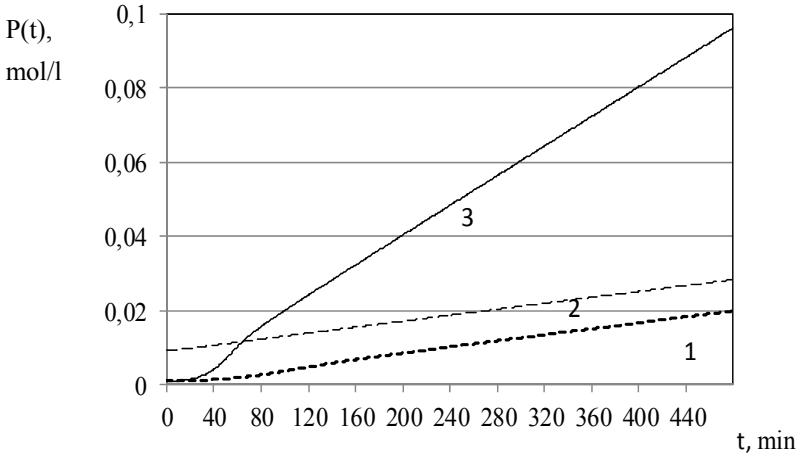


Figure 5. The kinetics of polymer degradation with $I(t) = I_0$

As can be seen from Fig. 5, if the concentration of the initiator stabilizes regardless of the original properties of the polymer (initial concentration of the polymer P_0 and, consequently, its initial average molecular weight), the function of polymer concentration to time tends to be linear. That is, changes in polymer concentration in this case are proportional to the time of degradation. Since P_0 is a constant and is determined by the quality of initial polymer, the degree of degradation is also close to the linear law of change with time.

Besides, change of the initial concentration of the polymer proportionally influences the offset of graphs along the Y-axis (curve 2), and change of the initiator concentration is proportional to the angle of curve inclination (curve 3), i.e., to degradation rate. So, the graph of changes in polymer concentration has a slant asymptote expressed by a linear function

$$P_{lin}(t) = a_0 + a_1 \cdot t\,, \tag{12}$$

where a_0, a_1 are coefficients of linear dependence, which are functions of the initial concentrations of the reactants and the reaction rate constants.

To find a_1, let us consider the finite limit

$$a_1 = \lim_{t\to\infty}\left(\frac{P(t)}{t}\right) = \lim_{t\to\infty}\left[\frac{e^{k_1\cdot P_0\cdot t + 2k_1\cdot k_2\cdot I_0\cdot t^2}}{\left(k_1\int_0^t e^{k_1\cdot P_0\cdot\tau + 2k_1\cdot k_2\cdot I_0\cdot\tau^2}d\tau + 1/P_0\right)t}\right]. \tag{13}$$

The finite limit (13) has uncertainty in form $[\infty/\infty]$, so in order to find it, the L'Hôpital's rule is used,

$$a_1 = \lim_{t\to\infty}\left(\frac{P_0 + 4\cdot k_2\cdot I_0\cdot t}{t}\right) = 4\cdot k_2\cdot I_0\,. \tag{14}$$

Similarly, we obtain a_0, which is found from the condition

$$a_0 = \lim_{t\to\infty}\left(P(t) - a_1\cdot t\right) = P_0\,. \tag{15}$$

Figure 6. shows the kinetics of degradation at various initial conditions and its approximation by expression (12). As it can be seen from the figure, the kinetics of degradation even at low t takes the form of its asymptote's graph (12).

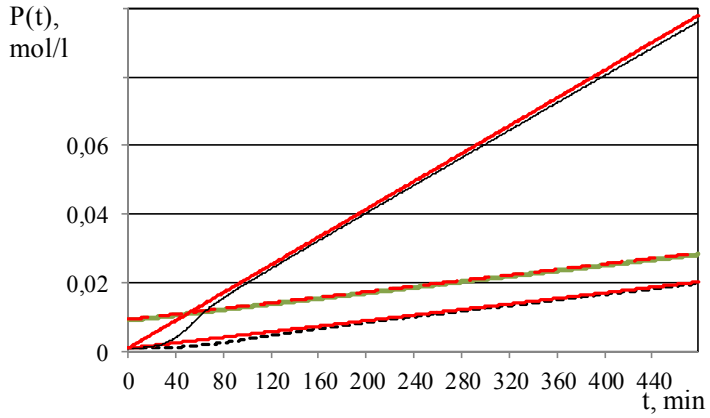


Figure 6. Linear approximation of the kinetics of polymer degradation according to (12) with $I(t) = I_0$

Calculations show that due to continuous dosing of initiator for 8 hours of degradation, the volume of reaction mass increases by 0.5 - 3.5% for the temperature of reaction between 50°C and 70°C.

These data confirm validity of using equations (8) - (9) for an approximated description of the degradation kinetics with the initiator continuously fed into the reaction medium.

3.2 Control on the Basis of Partial Addition of the Initiator

However, in industrial production conditions it is rater hard to ensure continuous addition of the initiator with high accuracy. It should be noted that due to sensitivity of the degradation kinetics to concentration of the initiator, minor deviations caused by metering accuracy can significantly affect the kinetics of degradation. For example, for the reaction mass of 1 ton with the initial starting concentrations of reaction components, initiator dosage at a rate of 7 to 50 g/min should be ensured.

Much more convenient in terms of technology of process implementation is partial addition. Addition is made at regular intervals in equal batches.

The dose is calculated according to the formula

$$V_d = I_0 V_0 M_I \left(1 - e^{-k_2 t_d}\right), \tag{16}$$

where t_d is the period of partial addition of the initiator in minutes.

Figure 7 shows comparison of calculated degradation kinetics between continuous and partial feeding of the initiator at the temperature of the reaction mass 70 °C. Curve 1 is the graph of polymer concentration in case of continuous addition of the initiator, curve 2 is the fractionality of addition of the initiator for 1 hour, curve 3 - for 2 hours, curve 4 - for 3 hours.

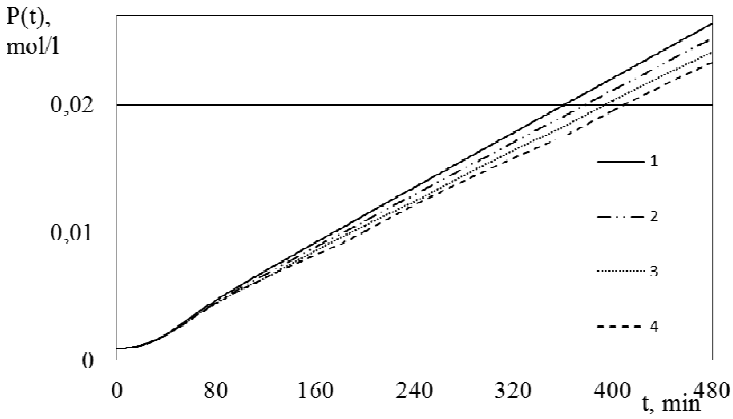


Figure 7. Comparison of degradation kinetics in case of continuous and partial addition of the initiator

As it can be seen from the figure, in case of partial addition, the linear form of the degradation reaction kinetics is maintained. Besides, the deviation of polymer concentration in case of continuous and partial addition of the initiator is insignificant. Concentration ratio of the polymer Pk_d / Pk_n with initial starting dosage is presented in Table 1 (position 1). Pk_n and Pk_d are polymer concentrations after the degradation process with continuous and fractional addition of the initiator, respectively, mole/l.

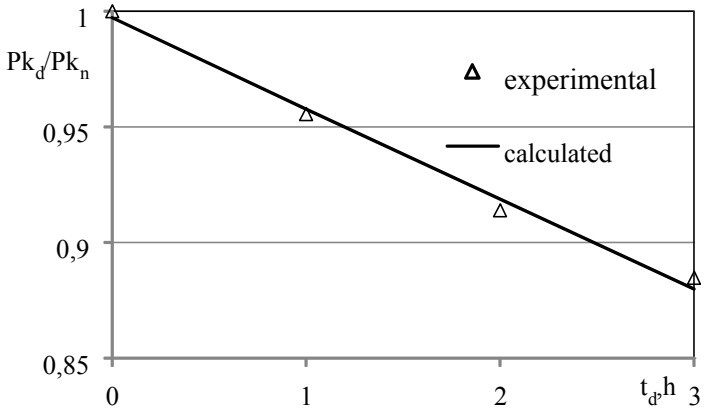


Figure 8. Dependence of ratio Pk_d / Pk_n on the period of partial addition

Figure 8 shows that dependence of the Pk_d / Pk_n ratio on the period of partial addition of the initiator can be described by a linear dependence of form

$$Pk_d / Pk_n (t_d) = b_0 + b_1 t_d \,. \tag{17}$$

To assess the influence of initial dosage of the initiator and the initial polymer concentration, numerical experiments were performed. With that, initial concentration of the initiator and polymer were decreased and increased two-fold. The results of the study are shown in Table 1.

Table 1.

Position	Terms of numerical experiment	Pk_d/Pk_n			Coefficients (17)	
		$t_d = 1$ h.	$t_d = 2$ h.	$t_d = 3$ h.	b_0	b_1
1.	$I_0 . P_0$	0.956	0.914	0.885	0.997	-0.039
2.	$2I_0 . P_0$	0.955	0.914	0.883	0.997	-0.039
3.	$0{,}5I_0 . P_0$	0.957	0.917	0.886	0.997	-0.037
4.	$I_0 . 2P_0$	0.957	0.917	0.889	0.997	-0.037
5.	$I_0 . 0{,}5P_0$	0.957	0.913	0.883	0.997	-0.035

The results show that the influence of initial concentration of the initiator and polymer on the deviation of final concentrations of the polymer in case of continuous and partial addition of the initiator can be neglected.

Influence of temperature of the reaction mass on the Pk_d/Pk_n ratio is shown in Table 2.

Table 2.

Position	Terms of numerical experiment	Pk_d/Pk_n			Coefficients (17)	
		$t_d = 1$ h.	$t_d = 2$ h.	$t_d = 3$ h.	b_0	b_1
1.	$T = 70°C, , I_0\ P_0$	0.956	0.914	0.885	0.997	-0.039
2.	$T = 60°C, , I_0\ P_0$	0.984	0.969	0.957	0.99	-0.015
3.	$T = 50\ °C, I_0\ P_0$	0.993	0.987	0.982	1	-0.006

As it can be seen in Table 2, with reaction temperature decreasing, the Pk_d/Pk_n value-increases. Obviously, the temperature has little effect on b_0, and changing b_1 is approximated by linear dependence

$$b_1(T) = c_0 + c_1 T \ . \tag{18}$$

where c_0, c_1 are the empirical coefficients.

In view of the relationships obtained above (17) - (18), the kinetic equation of degradation, depending on the conditions of partial addition of the initiator into the reaction area and the temperature of the reaction mass, can be approximated by the expression

$$P(t) = (P_0 + 4k_2 I_0)/(1 + c_0 t_d + c_1 T t_d) \tag{19}$$

Fig. 9. shows comparison of the kinetics of polymer degradation at 70 °C, under the condition of partial addition of the initiator without additional dosage. The initiator is partially added every 2 hours.

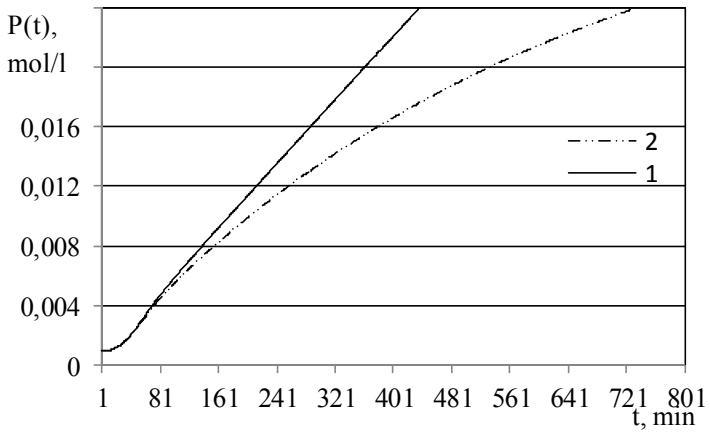


Figure 9. Comparison of the kinetics of uncontrolled degradation (single dosage of the initiator) and kinetics of degradation at fractional initiator

In Fig. 9 curve 1 describes the kinetics of degradation in case of partial addition of the initiator, curve 2 describes the kinetics of uncontrolled degradation (single dosage of initiator is used). The advantage of controlled (or continuous) degradation with partial addition of the initiator is the fact that in the same initial conditions of the reaction, it is possible to significantly reduce the time of the process until the desired degree of degradation. As it can be seen from the figure, the response time is reduced more than 1.5 times. Besides, a controlled degradation makes it possible to perform predictable degradation according to the linear law as a function of time. I.e., the control parameter is the response time that is linearly associated with the predefined degree of degradation. In this case, it becomes possible to implement software control; it is only necessary to know the original quality of the polymer - its initial average molecular weight M_n that is associated with the initial molar concentration of the polymer according to the formula $P_0 = f_0 / M_n$.

4. Discussion

The process of thermo-oxidative degradation of polymers in a solution is performed in the reactor using the periodical technology. Batch reactors are inefficient, but at the same time they are simple in design, require little auxiliary equipment, therefore, they are particularly suitable for experimental work for studying chemical kinetics. In industry they are commonly used in small-scale production and for the processing of relatively expensive chemicals. The common property of the polymer polymerization and degradation in a solution performed according to the periodic scheme in terms of production technology is their being not stationary. These processes are described by similar systems of differential equations, which make it possible to establish the relationship between the controlled settings and possible controlling influences.

An important technological task for batch processes of polymerization or degradation is the ensuring of the desired indicators of final polymer quality, such as concentrations of output products, average molecular weight, melt index, intrinsic viscosity of the polymer, degree of polymer degradation, etc. (Harish, 2011;-Bayzenberger and Sebastian, 1988).

For some polymerization processes (for example, during preparation of thermo elastoplasts), the main task of control, same as during polymer degradation, is obtaining polymers with predetermined average molecular weight M_n . Thus, the task of adjusting polydispersity K_p of obtained products is not practically set, since in the first case the narrowness of molecular weight distribution (MWD) is achieved by using highly efficient initiating system, and using polymerization mechanism of anionic polymerization (Moiseev, 1985). In the second case, MWD constantly tends to the Flory distribution, and the polydispersity index tends to value 2 (Browarzik and Koch, 1996; 15 Kaminsky and Kuznetsov, 2012).

Methods of controlling batch reactors are mainly based on changing the reaction temperature or varying concentrations of added catalyst and monomer (Bayzenberger and Sebastian, 1988). For example, in (Cott and Macchieto, 1989) temperature adjustment during exothermic reaction in a batch reactor is described. Reactors are controlled by choosing the optimal temperature profiles due to chosen criterion of quality followed by using standard control algorithms with feedback providing support for the selected mode. Common way to control the temperature in batch reactors is the use of a regulator with variable structure (Shinskey and Weinstin, 1965; Liptak, 1986). In work (Gluett et al., 1985) the process temperature was controlled using an adaptive algorithm. In (Jutan and Uppal, 1984) control was calculated according to the mathematical model of the process with estimation of the amount of heat generated in the reactor at any given moment. Based on the mathematical model of the process, the control algorithm for heating the reaction mass and further stabilization of temperature and a possibility to control its use for controlling the exothermic batch reactor was discussed in (Kershenbraun and Kittisupakorn, 1994; Marroquin and Luyben, 1973) and in (Cott and Macchieto,1989).

The values of the finished product quality are influenced by control parameter, such as time of reaction. In (Bayzenberger and Sebastian, 1988) an algorithm for resolving problems of optimal control was proposed in order to maintain the desired mode, minimizing time of reaction.

Also, the dynamics of the processes is significantly affected by the choice of initial temperature of polymerization or degradation (Bitukov et al., 2010), which allows to select initial temperature to determine future dynamics of periodic processes.

Considering the process of degradation of polymers in solution, it should be noted that in the framework of this study, practicability of the process control is shown, being based on choice of degradation time. At the same time, all other process parameters influencing the process in question are to be stabilized.

5. Conclusion

In course of system analysis of the processes of obtaining low molecular weight by degradation in a solution, in

order to find opportunities of controlling these processes, the following main results were obtained.

The main stage of the considered process where the main indicators of the finished product quality are obtained is the process of destruction. Control parameters in this step are reaction time, degree of degradation and the polydispersity index. It is expedient to control these parameters by feeding the initiator to the reaction medium. Other control parameters, such as polymer temperature and mass concentration, are to be stabilized in the vicinity of the maximum allowable value, which does not result in increasing the rate of jellification.

For the purpose of efficient control, molar concentration of the initiator in the reaction mixture should also be stabilized. For this purpose it is proposed to use additional feeding of the initiator into the reaction mass. In this case, as the estimates show, the process of destruction occurs linearly as a function of time. It is shown that implementation of additional initiator is made easier in fractionally equal doses at regular intervals, in this case, same as with continuous addition, the initiator ensures linearity of degradation kinetics over time.

Selection of the initial concentration of the initiator can accelerate or slow down the process of degradation. To implement this method of control, it is required to know the original quality of the polymer (average molecular weight). Thus, knowing the original quality of the polymer, its degree of degradation, it is possible to calculate the initial concentration of the initiator for achieving the final result within the time predetermined by technology.

The work presents theoretical results based on numerical and field experimental studies of polymer degradation kinetics in the laboratory. Definition of the possibility to use the obtained results in industrial conditions requires further experimental work. Thus, a promising direction of development of the research is testing of the obtained control methods in industrial environments.

References

Bayzenberger, D. zh. A., & Sebastian, D. Kh. (1988). *Principles of Polymerization Engineering.* Moscow: Khimiia.

Berlin, A. A., Volfson, S. A., & Enikolopian, N. S. (1978). *The Kinetics of Polymerization Processes*. Moscow: Khimiia.

Bitukov, V. K., Tikhomirov, S. G., & Khaustov, I. A., et al. (2010). Using the Methods of Determining the Polymerization Initial Temperature for Thermoplastic Elastomers Synthesis. *Bulletin of VSTA, 44*(2), 614-68.

Browarzik, D., & Koch, A. (1996). Application of Continuous Kinetics to Polymer Degradation. *Journal of Macromolecular Science: Pure and Applied Chemistry*, *33*(11), 1633-1641. http://dx.doi.org/ 10.1080/10601329608010928

Bryk, M. T. (1989). *Degradation of Filled Polymers*. Moscow: Khimiia.

Cott, B. J., & Macchieto, S. (1989). Temperature Control of Exothermic Batch Reactor Using Generic Model Control. *Industrial & Engineering Chemistry Research, 28(8),* 1177-1184. http://dx.doi.org/10.1021/ie00092a010

Geller, B. E., Geller, A. A., & Chirtulov, V. G. (1996). *Practical Guidance on the Physical Chemistry of Fiber Forming Polymers*. Moscow: Khimiia.

Gluett, W. R., Shah, S. L., & Fisher, D. G. (1985). Adaptive Control of a Batch Reactor. *Chemical Engineering Communications, 38*(1-2), 67-78. http://dx.doi.org/ 10.1080/00986448508911296

Harish, I. E. (2011). The Synergetic Method of the Synthesis of Control Systems over Chemical Batch Reactors. *Proceedings of the SFU. Technical sciences. Special Issue 94*. Retrieved October 1, 2014, from http://izv-tn.tti.sfedu.ru/wp-content/uploads/2011/6/9.pdf.

Ivanova, E. D., Dimitrov, I. V., & Georgieva, V. G. et al. (2012). Non-Isothermal Degradation Kinetics of Hybrid Copolymers Containing Thermosensitive and Polypeptide Blocks. *Open Journal of Polymer Chemistry*, (2), 91-98. http://dx.doi.org/ 10.4236/ojpchem.2012.23012

Jutan, A., & Uppal, A. (1984). Combined Feedforward-Feedback Servo Control Scheme for an Exothermic Batch Reactor. *Industrial and Engineering Chemistry Process Design and Development, 23*(3), 597-602. http://dx.doi.org/10.1021/i200026a032

Kaminsky, V. A., & Kuznetsov, A. A. (2012). Kinetics of Polymers Degradation by Lawcase. *Theoretical Bases of Chemical Technology, 46*(4), 453-457. Retrieved from http://elibrary.ru/item.asp?id=17795903

Kershenbraun, L. S., & Kittisupakorn, P. (1994). The Use of a Partially Simulated Exothermic (Parsex) Reactor for Experimental Testing of Control Algorithms. *Chemical Engineering Research & Design*, *72*(1), 55-63.

Koptelov, A. A., Milekhin, Iu. M., & Baranets, Yu. N. (2012). Simulation of Thermal Decomposition of a Polymer at Random Scissions of C-C Bonds. *Russian Journal of Physical Chemistry B, 6*(5), 626-633. http://dx.doi.org/10.1134/S1990793112050168

Liptak, B. G. (1986). Controlling and Optimizing Chemical Reactors. *Chemical Engineering,* (1), 69-81.

Marroquin, G., & Luyben, W. L. (1973). Practical Control Studies of Batch Reactors Using Realistic Mathematical Models. *Chemical Engineering Science*, *28,* 993-1003. http://dx.doi.org/10.1016/0009-2509(73)80001-6

Moiseev, V. V. (1985). *Thermoplastic elastomers*. Moscow: Khimiia.

Shinskey, F. G., & Weinstin, J. L. (1965). *Dual-Mode Control System for a Batch Exothermic Reactor.* In the Proceedings of the Twentieth Annual ISA Conference*,* Los Angeles, USA.

Sivalingam, G., & Madras, G. (2004). Thermal Degradation of Poly(vinyl acetate) and Poly(E-caprolactone) and their Mixtures in Solution. *Industrial & Engineering Chemistry Research, 43*(7), 1561-1567. http://dx.doi.org/10.1021/ie034115y

Sterling, W. J, Kim, Y. C., & McCoy, B. J. (2001). Peroxide Enhancement of Poly(r-methylstyrene) Thermal Degradation *Industrial & Engineering Chemistry Research, 40,* 1811-1821. http://dx.doi.org/10.1021/ie000971n

Tikhomirov, S. G., Khaustov, I. A., & Popov, A. P. (2013). Kinetics of polymers thermal oxidative degradation: analytical dependence obtaining. *Research Journal of International Studies, 10*(2), 95-96. Retrieved from http://research-journal.org/wp-content/uploads/2013/11/10-2-17.pdf

Ziff, R. M., & McGrady, E. D. (1986). Kinetics of Polymer Degradation. *Macromolecules*, *19,* 2513-2519. http://dx.doi.org/10.1021/ma00164a010

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Finite Component Approach for Modelling Entropy of Motor Road Pavements to Assess Their Efficiency

Mikhail Gennadjevich Goryachev[1]

[1] The Moscow state automobile and road technical university, Russian Federation

Correspondence: Mikhail Gennadjevich Goryachev, The Moscow state automobile and road technical university, 125319, Moscow, Leningradski prospect, 64, Russian Federation.

Received: January 9, 2015 Accepted: January 29, 2015 Online Published: July 25, 2015

doi:10.5539/mas.v9n8p140 URL: http://dx.doi.org/10.5539/mas.v9n3p140

Abstract

In statistical physics, the irreversible processes are characterized by the same value as entropy. Entropy represents a measure of system disorder consisting of many components. Originally ordered systems in irreversible processes lose their order, so that the system entropy increases. Therefore, one can judge by the size of entropy about the degree of order in the system. This article gives a brief description of the new physical and mathematical solution taken as a basis of the model of entropy accumulation and defects development processes within the motor road pavements for the prognosis and evaluation of their operational integrity and durability.

Keywords: road pavement, entropy, operational integrity

1. Introduction

Road pavement appears to be extremely difficult object for modelling. The model should simulate complex processes occurring in the road pavement under the conditions of automobile traffic and environment changes – elastic deformations, plastic displacements, and destruction. The present paper may be regarded to as an attempt to develop a universal physical-mathematical model of the road pavement and the process of its destruction at its operation. To this effect there are proposed new energy criteria based on physical laws of accumulation of potential strain energy of destruction of road pavement layers and soil subgrade from the work of multiply applied external loads (Goryachev, 2014a; Goryachev, 2014b; Smirnov & Aleksandrov, 2009; Smirnov, 1989). Abroad the entropy approach within the road science and practice is applied for evaluation of coating flatness (Andrei, 2000). Moreover, approaches may be very different (Chakrabarti & Boltzmann-Gibbs, 2000; Chakrabarti & Chakrabarty, 2005), but always based on the known principles of thermodynamics (Gaskell, 2003; Sandler, 1989; Martyushev & Seleznev, 2006).

For simplified demonstration Figure 1 shows a three-layer system "road pavement – subgrade" (sand, gravel, asphalt), each layer of which is represented by a set of large quantity of finite components of spherical shape.

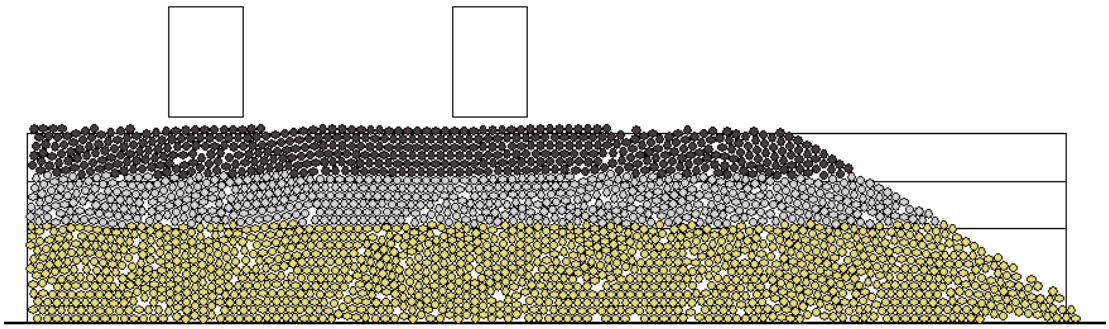


Figure 1. Model of the three-layer system of a two-lane motor road (cross section)

This model is based on the predominating principle of modern mathematics, which consists in replacing the macroscopic objects by objects of smaller sizes (preferably infinitesimal sizes, or at least sizes much smaller than

the typical size of layers themselves or automobile wheels). In accordance with this principle, the layers of road pavement are demonstrated as a totality of a large number of separate circular components (see Fig. 1). The used method of modelling is similar to the method of smoothed particles or SPH-method (Smoothed Particle Hydrodynamics), which has been recently applied in computer graphics to produce highly realistic images of different environments that are prone to fragmentation. Also modelling method is similar to the method of finite components, differing from it in that the nodes of the grid can be moved in accordance with the laws of classical dynamics, and the grid itself meanwhile changes its configuration.

2. Methodology

2.1 Model Intent

Introduction of the road pavement within the model is executed in a two-dimensional space X-Y. Provided that the axis X is located in the horizontal direction across the motor road, and the axis Y is located in the vertical direction. Elimination of the third dimension allows at a selected number of components to increase the linear dimensions of the system modelled in the directions X and Y. Otherwise, the program calculations would manifold lengthened. During the calculations, there were used up to 10000 components. Eliminated (invisible) axis Z would be located in the horizontal direction along the motor road pavement. While turning the plane along the Z direction, it is possible to search for unsafe (reference) cross-sections and perform modelling for the newly created plane X-Y.

Kinematic state of each road pavement component i is defined by four variables: the Cartesian coordinates of its centre (x_i, y_i) and the Cartesian velocity components (v_{xi}, v_{yi}).

To calculate the forces acting on the components of the road surface, used is the following method. Certain component i undergoes force action on the part of each surrounding it component j:

$$F_i = \sum_{j=1}^{N_C} \left(F_{ij}^E + F_{ij}^V \right), \text{ N}, \tag{1}$$

Where F_{ij}^E; F_{ij}^V – forces of the elastic and viscous interaction of components i and j, N;

N_c – the total number of components within a road pavement model.

To calculate the forces in each pair of components, there should be pre-calculated the distance r_{ij} between their centres $S_i(x_i, y_i)$ and $S_j(x_j, y_j)$:

$$r_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \text{ m}. \tag{2}$$

Forces between the components are calculated in the approximation of viscoelastic interaction

$$F_{xij} = \begin{cases} c\ \left(d_c - r_{ij}\right) \cdot \left(x_i - x_j\right) / r_{ij} + d \cdot \left(v_{xi} - v_{xj}\right), \ if\ r_{ij} < \alpha \cdot d_c; \\ 0, \ if\ \ r_{ij} \geq \alpha \cdot d_c; \end{cases}$$

$$F_{yij} = \begin{cases} c\ \left(d_c - r_{ij}\right) \cdot \left(y_i - y_j\right) / r_{ij} + d \cdot \left(v_{yi} - v_{yj}\right), \ if\ r_{ij} < \alpha \cdot d_c; \\ 0, \ if\ \ r_{ij} \geq \alpha \cdot d_c, \end{cases} \tag{3}$$

Where F_{xij} and F_{yij} – Cartesian components of the force F_{ij}, N;

c – stiffness coefficient of the elastic interaction of components, N/m;

d – viscosity coefficient of interaction (damping), N·s/m;

α – factor of interaction restriction.

Viscous friction coefficient d is related to the internal friction within the material, and is determined by the typical distance of sound waves attenuation within the material. The α factor is connected with the material's tensile strength (represents a ultimate tensile strain at rupture or tensile testing).

For modelling the road pavement destruction, in the model there must be followed the evolution of the component system, for which there should be calculated the trajectory of motion of each of the components. The trajectories are found by solving the equations of motion of individual components that can be recorded on the basis of Newton's second law.

$$m_c \cdot \frac{d^2 \cdot x_i}{dt^2} = \sum_{j=1}^{N_C} \left(F_{xij}^E + F_{xij}^V \right) + k_v \cdot v_{xi} ;$$

$$m_c \cdot \frac{d^2 \cdot y_i}{dt^2} = \sum_{j=1}^{N_C} \left(F_{yij}^E + F_{yij}^V \right) - m_c \cdot g + c_{c-s} \cdot r_{mii} + k_v \cdot v_{yi} , \quad (4)$$

Where m_c – mass of a road pavement component, kg;

t – time, s;

g – acceleration of gravity, m/s^2;

c_{c-s} and k_v – coefficients of stiffness and viscosity of viscoelastic interaction of the i-th component with underlying surface;

r_{mii} – distance of mutual intrusion of the i-th component of road pavement into the underlying surface, m;

v_{xi} and v_{yi} – Cartesian components of the velocity vector of the i-th component, m/s.

Set of equations of the form (4) for all N_c components describes evolution of road pavement in the course of time.

At the initial point of time, road pavement components are randomly placed by the uniform law in the geometrical field – the cross-section of the road. After the start of integration of motion equations, components, moving, get into local equilibrium positions. Thus, realized is a random dense packing of components. At the same time, the pavement model comes into a stable condition. Later within the model there is simulating the automobile wheel effect on the top layer of road pavement.

With a significant mechanical stress on the road pavement material, it loses its initial cohesion, starts to deforming and destructing, in particular, there appear microdisplacement, cracks, and voids. Within the developed model the effect of the initial cohesion loss is simulated as follows. After preparation of the model of road pavement layers (random close packing), "initial binding" of components is produced in the model. For this a two-dimensional array K (i, j) is filled in, each cell of which has the value "true" if the components i and j are initially contacting with each other, i.e. located at the distance

$$r_{ij} < \alpha \cdot d_c \quad (5)$$

and the value "false" in the opposite case.

Further, during the computer experiment, at each integration step for each bound pair i-j there is checked whether it has parted or not, under the same condition (5), and in case of inequality (5) failure, the K (i, j) array obtains the "false" value. Thus, there is a gradual disappearance of the initial cohesion of materials of the road pavement layers.

The general algorithm of the method of modelling defect accumulation in road pavements is represented by a block diagram in Figure 2 and is implemented as a computer program.

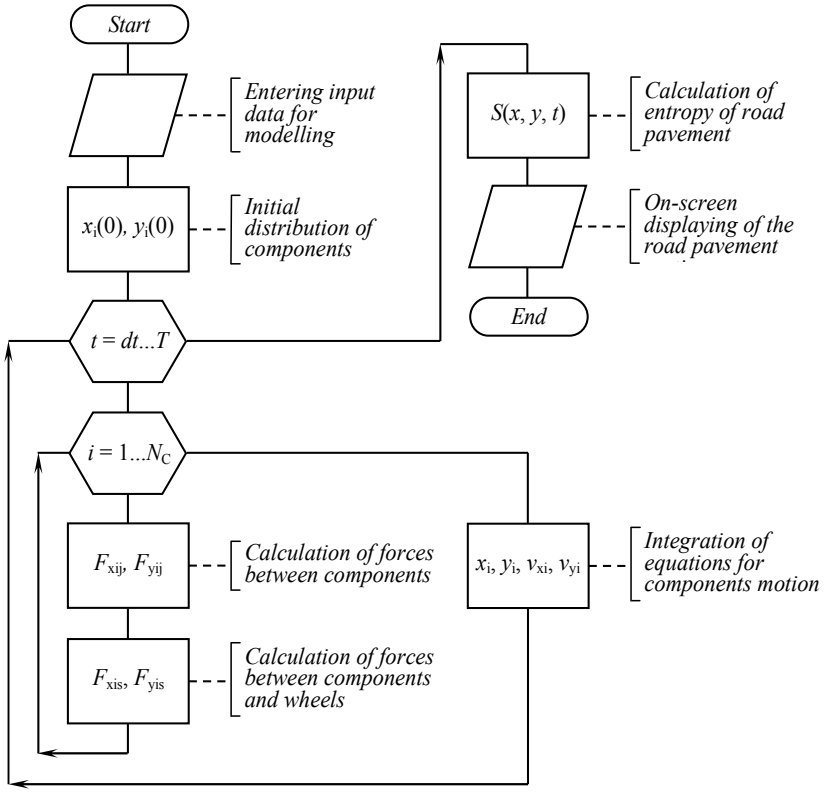


Figure 2. Diagram of the program algorithmic for modelling of defect accumulation of road pavement

The numerical integration of the equations of component motions is executed through a modified method of Euler-Cauchy. This method has the second order of accuracy in the coordinate and the first order of accuracy for speed. The method is versatile, reliable, and easily programmable. The method is implemented by the following formulas

$$x_i = x_i + v_{xi} \cdot dt + \frac{F_{xi}}{m_c} \cdot \frac{(dt)^2}{2};$$

$$y_i = y_i + v_{xi} \cdot dt + \frac{F_{yi}}{m_c} \cdot \frac{(dt)^2}{2};$$

$$v_{xi} = v_{xi} \cdot dt + \frac{F_{xi}}{m_c} \cdot dt;$$

$$v_{yi} = v_{yi} \cdot dt + \frac{F_{yi}}{m_c} \cdot dt. \quad (6)$$

Step of integration of the differential equation system amounted to $\Delta t = 0.0005$ s.

Component diameter d_c affects the adequacy of the model. For a more accurate representation of the road pavement in the model, d_c should be as small as possible. At the same time, the number of components N_c, required for simulation, sharply increases (by a cubic law) with decreasing of d_c. At increasing the number of spheres, there is also increased the number of equations describing the layers of road pavement and, correspondingly, the time for computing. Even the twice reduction of the component diameter leads to increase of the computer calculation duration by at least eight times. If you focus on the computational power of modern computers, the optimum diameter of the component is about 5 cm. This size provides sufficient specification of processes occurring in the road pavement, and does not lead to time-consuming calculations. The total number of components in the model may be several tens or even hundreds, depending on the type of problem being solved.

Mass calculation of one component of a road pavement layer m_c is executed using the tabulated value of material density and geometric considerations:

$$m_c = \rho \cdot V_c = \rho \cdot \frac{4 \cdot \pi}{3} \cdot \left(\frac{d_c}{2} \right)^3 \cdot k_f = \frac{\pi}{6} \cdot \rho \cdot d_c^3 \cdot k_f , \text{ kg}, \quad (7)$$

Where ρ – volume density of the material, kg/m^3;

V_c – component volume, m^3;

K_f – form factor necessary to account for the fact that the sphere components do not completely fill the space (unfilled pores remain between the components); k_f coefficient value depends on random packing density and is equal to approximately 1.4.

To calculate the stiffness of interaction between the two components there is used a table value of the material elasticity modulus and also the geometrical considerations concerning discretization:

$$c_n = E \cdot \frac{\pi \cdot d_c}{4} \cdot k_f , \quad (8)$$

Where E – modulus of material elasticity, Pascal.

2.2 Simulation within the Model of the Road Pavement Stressing from the Wheel Load

In the developed model, simulated a multiple (up to 2 000 000 times) passage of an automobile along the road pavement cross-section under investigation. Each passage consists in double (for two-axle automobiles) or multiple (for multi-axle automobiles) passage of wheelset over the cross-section of the road pavement (Figure 3).

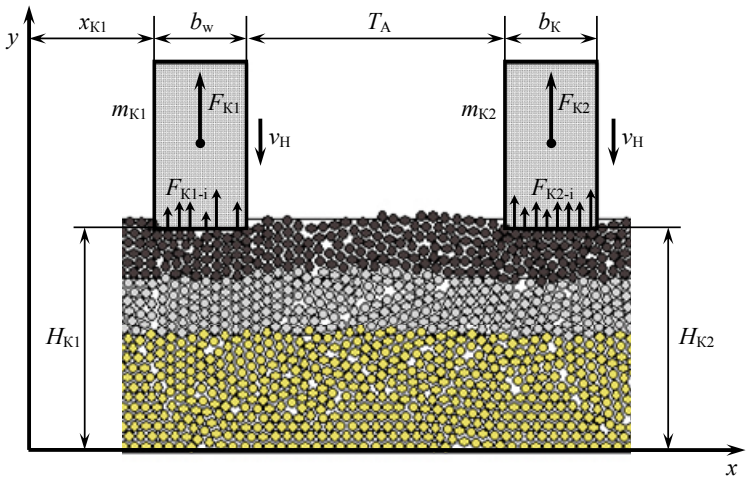


Figure 3. Design model for determination of interaction forces between wheels and road pavement components

Before the start of the simulation, there are given specific parameters for wheelsets of automobiles: wheel width b_w and track T_A. Automobile position in the transverse direction of the road is set by the variable $x_{к1}$. The model provides the possibility of automobile movement, both strictly along the track ($x_{к1}$ = const), and the possibility of accidental displacement of the track, i.e. $x_{к1}$ is considered as a random variable distributed according to a normal or uniform law.

Wheel passage is modelled as follows. Initially wheel is located at a certain distance from the road surface at height H_{ref} (in the global coordinate system XY), m. At a certain point of time $t_{нп}$ (where n – number of the wheelset passage), s, within the model there starts wheel's movement down at a constant wheel speed v_L, m/s. Loading velocity v_L depends on the automobile speed v_A, m/s, and is associated with it, in the first approximation, through the ratio

$$v_L = \frac{v_a \cdot k_н}{R_w}, \text{ m/s}, \quad (9)$$

Where $k_н$ – multifactor coefficient, depending on the type and profile of the wheel, material properties, type of tread, and air pressure in a wheel, m;

R_w – wheel radius, m.

As the downward motion of the wheel, it comes into contact with the surface of the components, wherein there is a plurality of forces between the components and the wheel $F_{к1\text{-}i}$ (or $F_{к2\text{-}i}$), which together make up surface reaction force $F_{к1}$ (or $F_{к2}$). Checking contact of the i-th component of road pavement with a wheel is carried out by the condition of entering the centre of the component (x_i, y_i) into the geometric area corresponding to the wheel, and defined by a system of inequalities:

$$\text{For left wheel} \quad \begin{cases} x_i > x_{к1}; \\ x_i < x_{к1} + b_w \cdot T_w; \\ y_i > H_{к1} - d_c/2, \end{cases} \tag{10}$$

$$\text{For right wheel} \begin{cases} x_i > x_{к1} + b_w \cdot T_w + T_A; \\ x_i < x_{к1} + 2 \cdot b_w \cdot T_w + T_A; \\ y_i > H_{к2} - d_c/2, \end{cases} \tag{11}$$

where $H_{к1}$ and $H_{к2}$ – current position of the lower points of the wheels, m.

Meanwhile, forces $F_{кk\text{-}i}$ (where k – wheel index acquiring values 1 or 2), resulting in contact between the wheel and the components of road pavement, calculated in viscoelastic approximation as follows

$$F_{кk-i} = c_{кс} \cdot \left(x_i - H_{кk} + d_c/2\right) + d_{кс} \cdot \left(v_{yi} - v_н\right), \text{ N}, \tag{12}$$

Where $c_{кс}$ – coefficient of stiffness, N/m;

$d_{КС}$ – viscous friction (damping) coefficient, N·s/m when interacting with the wheel surface material.

These coefficients are determined by the elastic moduli of materials of wheel and the top layer of the road pavement. The total reaction force on the wheel $F_{кк}$ is determined by summing the individual forces $F_{кk\text{-}i}$:

$$F_{кk} = \sum_{i=1}^{N_C} \begin{cases} c_{кс}(x_i - H_{кk} + d_c/2) + d_{кс}(v_{yi} - v_н), if \begin{cases} x_i > x_{к1} + \begin{cases} 0, if\ k = 1; \\ b_w \cdot T_w + T_A, if\ k = 2; \end{cases} \\ x_i < x_{к1} + \begin{cases} b_w \cdot T_w, if\ k = 1; \\ 2 \cdot b_w \cdot T_w + T_A, if\ k = 2; \end{cases} \\ y_i > H_{кk} - dc/2. \end{cases} \\ 0, if \begin{cases} x_i < x_{к1} + \begin{cases} 0, if\ k = 1; \\ b_w \cdot T_w + T_A, if\ k = 2; \end{cases} \\ x_i > x_{к1} + \begin{cases} b_w \cdot T_w, if\ k = 1; \\ 2 \cdot b_w \cdot T_w + T_A, if\ k = 2; \end{cases} \\ y_i < H_{кk} - d_c/2. \end{cases} \end{cases} \tag{13}$$

Contact forces between the wheel and components $F_{кk\text{-}i}$ affect both on the wheel and the components themselves, causing their displacements that extend across the entire road pavement. The force $F_{кк\text{-}i}$ is added to the horizontal by the Y-axis component F_{yi} of all forces affecting on the i-th component.

Each wheel causes formation of a flexural bowl, moves downward at a constant velocity to the point at which the reaction force $F_{кк}$ compensates weight $m_к \cdot g$ attributable to one wheel (where $m_к$ – automobile weight, attributable to one wheel). After that, in the model there begins an upward movement of the wheels at a constant velocity, for clarity substituting road pavement restoration after removal of the load, i.e. wheel's passage forward

from the considering place of loading. This restoration is equal by modulus $v_н$, and occurs until the wheel is raised to the initial height H_{ref} – elastic recovery is complete. Thus, the current velocity of the wheel $v_{кk}$, alternately takes the values that $-v_н$ (flexure under the wheel), the $v_н$ (exit from the flexural bowl). To determine the time, at which it is necessary to change the direction of the wheels on the opposite, at each integration step, there should be compared the current height of the wheels $H_{к1}$ and $H_{к2}$ with the height H_{ref}, and current forces of reaction to the wheels $F_{к1}$ and $F_{к2}$ with the weight of the automobile, attributable to one wheel $m_к \cdot g$:

Here, the first and second conditions are the conditions of motion direction change; the third is the condition of direction conservation.

This cyclic loading of wheels is made during the computer experiment until the number of passages of reference axes reaches a predetermined value N_c. The main series of computer experiments were performed at N_c equal to 250 000, 500 000, 1 000 000, 2 000 000.

$$v_{кk} = \begin{cases} v_н, if \begin{cases} v_к < 0; \\ F_{кk} > m_к g; \end{cases} \\ -v_н, if \begin{cases} v_к > 0; \\ \left[\begin{array}{l} H_{к1} > H_{ref}; \\ H_{к2} > H_{ref}. \end{array} \right. \end{cases} \\ v_{кk}, if \left[\begin{array}{l} \begin{cases} v_к < 0; \\ F_{кk} < m_к g; \end{cases} \\ \begin{cases} v_к > 0; \\ \left\{ \begin{array}{l} H_{к1} < H_{ref}; \\ H_{к2} < H_{ref}. \end{array} \right. \end{cases} \end{array} \right. \end{cases} \tag{14}$$

2.3 Simulation of Deformation and Destruction of Road Pavement within the Model

Degradation of strength and smoothness of road pavement is associated with a number of processes occurring under the influence of automobile wheels: bending, compression, plastic deformation, additional tightening, decompression, destruction of connectedness. In the developed model, road pavement is provided with a high degree of detail, so all of these processes are reproduced at the level of changes in the structure components. The ability to simulate the microstructure allows for quantitative calculations of these processes.

Bending layers of road pavement, mainly the upper ones, are simulated in a model due to the wheels effect on the covering layer and simultaneously maintain connectivity between the components that are in series (Figure 4, a).

Compression of the road pavement layers is reproduced in the model through the approximation of components to each other (the distance between the components r_{ij} becomes less than the equilibrium distance $d_i + d_j$) – Figure 4, b. Compression is expressed as stronger, as closer are the components to place of wheels contact with the surface. Approximation of components to each other gives rise to repulsive forces between them, which creates a supporting force to the wheel, and after wheels passage, the components are returned in the approximate coverage to its original state.

Plastic deformation is expressed in the model by relative movement of closely contacting components relative to each other. Taking into account the theoretical and experimental work of Zhustareva, Y.V. and field observations of Lugov, S.V., we can neglect the thinning out effect possible under displacement processes (Zhustareva, 2000; Lugov, 2004). The model is reproduced as a local plastic deformation (inelastic displacement of a part of coating material) (Figure 4, c), and "globalized" plastic deformation: for road pavement layers, the impact of the wheel in one place is squeezing the material elsewhere; for example, track (Figure 4, d).

Tightening is simulated by reduction of pore volume and other loose fragments of structure under the influence of wheel (Figure 4, d). Under the influence of strains from the wheels, loose fragments of local structure are experiencing restructuring and compacted. The initial (for example, construction), degree of packing of the particles is established through the actual coefficient of the standard tightening

$$K_y = \frac{\rho}{\rho_{st}} = \frac{S_{dg}}{S}, \quad (15)$$

Where ρ – the density of the material or soil in the layer, t/m^3;

ρ_{st} – density of the material or soil determined by standardized laboratory tests, t/m^3;

S – the area occupied by components of the layer model (particles) per unit of volume, m^2;

S_{dg} – area corresponding to the greatest degree of convergence of the layer model components per unit of volume, m^2.

A number of components is taken as a volume width.

Destruction of the road pavement reproduced cohesion is simulated in the model by removing from consideration of bonds initially present in the model (Figure 4, e). Prior to the beginning of computing experiment in the cohesive layers of the model there is performed storing of contacting components, and in the further calculations it is considered that the ultimate tensile strength between the initially cohesive components is higher than between others. Under the influence of the same wheel, at a considerable distance of cohesive components from each other, there is executed setting to zero of information concerning the cohesiveness in a separate array, cohesion is eliminated, and a pair of components, even if they will again be in contact with each other, interact only through repulsive forces.

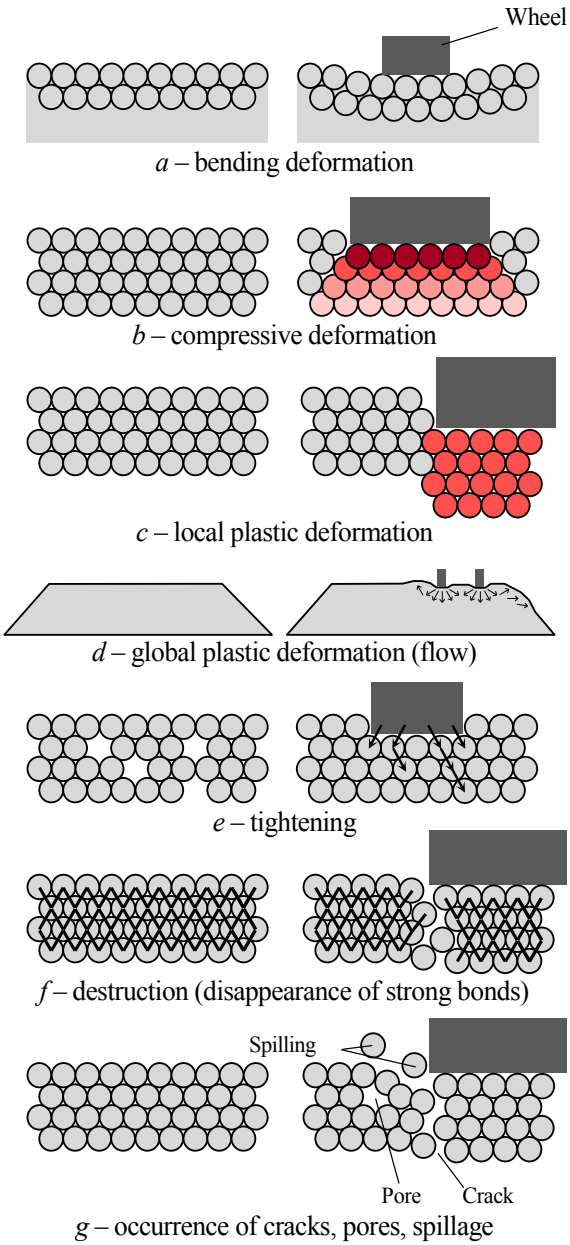


Figure 4. Simulation in the model of basic processes that occur in road pavement under loading

Cracks, pores and spillage are reproduced in a model by removing components from each other under the action of stresses propagating from the wheel (Figure 4, g). Removal of structural fragments from each other in a model causes cracks, pores, some components which are not associated with others.

These processes in the road pavement may occur simultaneously or may not be pronounced. The advantage of the developed model is the ability of quantitative checking of all mechanical and material characteristics of these processes. Furthermore, the model appears to reproduce more complex processes involving a large number of components that are difficult to be monitored and classified. Discretization of road pavement and the use of the method of particle dynamics make the model natural for this problem, so we can expect that the model is highly appropriate. Fidelity processes should increase with decreasing size of the components in the model.

Thus, the developed model simulates a wide range of processes in the road pavement that occur during the operational period. This model allows us to make calculations of basic parameters characterizing the deformation and fracture of road pavements: energy parameters (work of deformation and destruction (entropy), dissipated heat energy), the displacement of components of layers, formation of surface irregularities, including track.

In the development of physically fundamentally new models, check on adequacy can be made by comparing the

results obtained with analogues of others, but already proven approaches. Thus, to determine the residual strain in the "road pavement-subgrade" system, one can use the upgraded dependence of Mevlidinov, Z.A. (Mevlidinov, 1997) or complex solutions of Lugov, S.V. (Lugov, 2004) that meet the following essential requirements:

1. Comply with modern ideas about the processes occurring in the "road pavement-subgrade" system, to be constructed on theoretical assumptions, backed by numerous experimental studies, including different scientific areas and schools, and contain only the common parameters and characteristics that have received theoretical and experimental support.
2. Allow, at an integrated calculating base, to produce calculations of plastic defects in a defragmented manner by layers, i.e. possess, perhaps, if not absolute then sufficiently wide unification.
3. Allow to operate with a set of input parameters, including correction factors for determinations, which do not need mandatory live-update values and stability properties, laboratory testing of road construction materials and soil, i.e. possess permanent actualization and operational practicality. This rule does not apply to the minimum required information on the road pavement structure, intensity and composition of traffic.
4. Permissible error, not exceeding values and not leading to the rejection of the result, which distorts the assessment of the forecast of the road pavement state, expressed in terms of performance criteria. It is clear that the latter requirement is inherent in some convention. With the technical and economic point of view, such an error is convenient to estimate the costs of inputs, such as quantities of materials.

Outside of Russia, such models were developed by F. Finn, I. Gachwendt, S. Hushek, A. Molenaar, P. Ullidtz, B. Larsen Watanatada (Finn, 1982; Gachwendt & Poliacek, 1987; The World Bank , *1997;* Hushek, 1977; Molenaar, 1983; Watanatada, Hrall, Paterson, Dhareshwar, Bhandari & Tsunokava, 1987; Ullidtz, 1987; Ullidtz & Larsen, 1986; Grzybowska &. Salamon, 1991).

3. Results

3.1 Mechanisms of Defects Development and Nature of Changes of Operational Integrity Indicators

Computational experiments have shown that the nature of defects in the road pavement will be different depending on the physical and mechanical properties of the materials in operating conditions, loading parameters. Accordingly, with the various changes in dynamics, the efficiency parameters are also changed within the operation of road pavement.

Stage evaluation of computer experiments revealed three possible causes that lead to the loss of flatness: additional tightening, to be more precise, a displacement and a combined mechanism.

Process of additional tightening of road pavement are realized in the case of high values of the elastic modulus and internal friction coefficient of road pavement layers, low temperature and moisture content, low wheel load. Entropy under these conditions is growing according to the patterns close to exponential (Figure 5, a). In the operation of the road structure there are shown two stages: I – initial stage of quick enough additional tightening, at the same time the change in structure leads to an increase in entropy; II – stabilizing step, in which virtually no further additional tightening occurs (or entropy is increased).

The mechanism of flow-displacement is realized in the case of low values of the elastic modulus and internal friction coefficient of road pavement layers, high temperature and humidity, large wheel load (Figure 5, b). The entropy at such mechanism of operational integrity loss is developed in three steps: I – the initial stage of the flow (road pavement some time saves its configuration); II – the stage of intensive course (reconfiguration according to the nature of loading); III – step of slow flow (configuration adapted to the nature of the movement, so further it changes slightly).

In the combined mechanism of operational integrity loss, characteristic of the most real cases is implemented as a tightening and displacement forms (Figure 5, c). The first step is characterised by tightening of road pavement; the second stage shows lasting of a slow start; the third stage is a rapid steady flow; during the fourth stage it practically stops.

For the processes of additional tightening, the growth pattern of entropy with a fairly high accuracy is described by an exponential law of the form

$$S_{уд}\left(N_C\right) = S_{\max} \cdot e^{-\frac{N_c}{N_p}}, \tag{16}$$

Where S_{max} – maximum entropy, which is acquired by road pavement;

N_p – coefficient of speed of entropy growth (the number of passages in which the value of $S_{max} – S_{уд}$ differs from

S_{max} by e times).

For mechanism of flow, dependence of $S_{уд}(N_c)$ is a transition from one level to another level. Such dependence is generally analytically described by a sigmoidal function of Boltzmann (often used to describe non-equilibrium processes in physics, chemistry, economics, and biology)

$$S_{уд}(N_C) = S_2 + \frac{S_1 - S_2}{1 + e^{\frac{N_c - N_{c0}}{N_p}}} \tag{17}$$

Where S_1 and S_2 – initial and final levels of entropy;

N_{c0} – typical number of passages, at which the tendency of growth $S_{уд}$ is changed (inflection point of the sigmoid function).

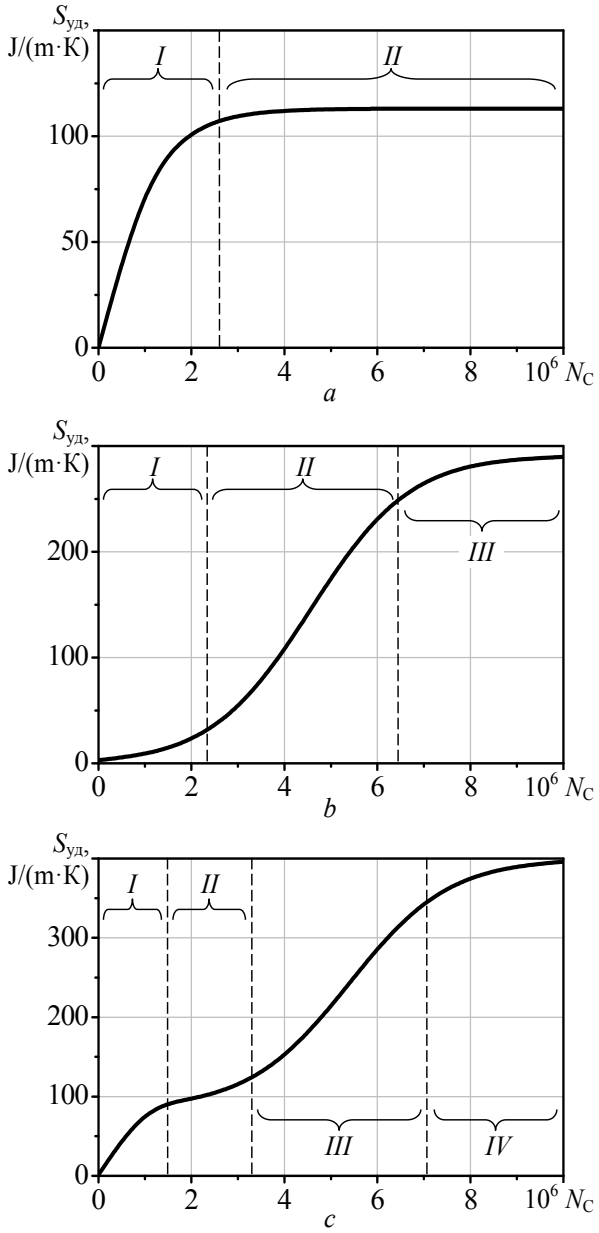


Figure 5. Nature of $S_{уд}$ entropy growth of the road pavement with increasing amounts of automobile passages for three mechanisms of operational integrity loss

a – additional tightening; b – flow-displacement; c – combined mechanism

For the combined mechanism, first occurs additional tightening within the layers of road pavement by the formula (16), then flow by the formula (17). Therefore, in general the analytical expression can be written as the sum of two terms, each of which is responsible for its own specific loss of operational integrity. Moreover, the initial level of entropy can be adopted to zero $S_1 = 0$.

$$S_{y\partial}\left(N_c\right) = S_{\max} \cdot e^{-\frac{N_c}{N_{p1}}} + S_2 \cdot \left(1 - \frac{1}{1 + e^{\frac{N_c - N_{c0}}{N_{p2}}}}\right), \tag{18}$$

Where N_{p1} and N_{p2} – coefficients of entropy speed growth, respectively, for the tightening regime and the regime of plastic deformation.

3.2 Software Implementation of the Model

To solve the above mentioned differential equations and study of the model, there has been developed a series of computer programs distinguishing themselves by detail and speed calculations, united in the "Program for modelling of road pavement distraction" in Object Pascal in an integrated programming environment Borland Delphi 7.0. Received is a certificate of official registration of the computer program No.: 2010123456.

The program is designed to study the influence of the main mechanical and geometrical characteristics of the road pavement and subgrade characteristics of traffic flow on the road on the growth of dynamics and nature of entropy of defect accumulation in the road pavement.

The basic functionality of the program:

- Conducting computer simulations of the multiple stressing of road pavement through wheel load;
- Specifying the basic parameters of materials of the road pavement layers, subgrade and traffic load;
- Displaying in the computer experiment of schematic cross-sectional images of the road, displacement maps covering components, basic output characteristics.

When you run the program, on the screen appear interface forms with lots of entry windows, in which you can specify the geometrical and mechanical parameters of the existing road pavement and the wheels of automobiles.

4. Conclusion

The developed physical and mathematical apparatus and the simulation model allow studying the relations and conversion of heat in layered systems and a special form of energy – entropy with the further determination of the degree of irreversible deformation and damage of road pavements. Operational integrity of road pavement is as higher, as smaller is the value of entropy characteristics. Representation of physical fundamentally of the new model can be tested for adequacy by comparing the results obtained with analogues of others, but already proven approaches.

References

Andrei, R. (2000). The Use of Entropy Concept in Ranking and Correlating Roughness Measurements Obtained by Different Equipment. *Second International Workshop on Artificial Intelligence and Mathematical Techniques in Pavement and Geomechanical Systems, Newark, DE*, 153-164.

Chakrabarti, C., & Boltzmann-Gibbs, D. (2000). Entropy: Axiomatic Characterization and Application. *International Journal of Mathematics and Mathematical Sciences, 23*(4), 243-251.

Chakrabarti, C., & Chakrabarty, I. (2005). Entropy: Axiomatic Characterization and Application. *International Journal of Mathematics and Mathematical Sciences, 17*, 2847-2854.

Finn, F. (1982). The use of distress prediction subsystems for design of pavement Structures. *Proceedings of 5 Int. Conference on the Structural Design of Asphalt Pavement, Delft.*

Gachwendt, I., & Poliacek, I. (1987). *Navrochovanie a posudzovanie konstrukcji vozovek cestnych komunikacji.* Bratislawa.

Gaskell, D. (2003). *Introduction to the Thermodynamics of Materials* (4th ed.). New York: Taylor & Francis Books, Inc.

Goryachev, M. (2014a). Energy states of nonrigid road pavements at simulation modeling by motor transport impact. Antiseismic construction. *Safety of constructions, 1*, 60-62.

Goryachev, M. (2014b, June). Energy criteria for assessment of nonrigid road pavement efficiency. *MADI Herald, 2*(37), 99-101.

Grzybowska, W., &. Salamon, J. (1991). Metody przewidywania glembokosce kolei ze szczegolnym uwzglednieniem metody Husheka. *Drogownictwo, 4-5*, 80-85.

Hushek, S. (1977). Evaluation of rutting due to viscous flow in asphalt pavements. *Proceedings of 4 Int. Conference on the Structural Design of Asphalt Pavements, Ann Arbor, Michigan.*

Lugov, S. (2004). Main provisions of the methodology for calculating track depth on the road pavements with asphalt covering. *Dissertation, candidate of engineering sciences, Moscow.*

Martyushev, L., & Seleznev, V. (2006). *Maximality principle of entropy production in physics and related fields.* Ekaterinburg: GOU VPO UGTU-UPI.

Mevlidinov, Z. (1997). *Justification of the main indicators, which take into account the effect of residual strains when calculating road pavements of nonrigid type.* Dissertation, candidate of engineering sciences, Moscow.

Molenaar, A. (1983). *Structural performance and design of flexible road construction and asphalt concrete overlays.* Delft University of Technology, Laboratory for Road and Railroad Research.

Sandler, S. (1989). *Chemical and Engineering Thermodynamics* (2nd ed.). Canada, Toronto: John Wiley & Sons, Inc.

Smirnov, A. (1989). Theoretical and experimental studies of the nonrigid road pavement efficiency. Dissertation, doctor of engineering sciences, Omsk.

Smirnov, A., & Aleksandrov, A. (2009). *Mechanics of road constructions. Study Guide.* Omsk: SibADI.

The World Bank (1997). *Highway Development and Management, HDM-4.* Washington DC: The World Bank.

Ullidtz, P. (1987). *Pavement analysis.* Else Vies Science Published B.V.

Ullidtz, P., & Larsen, B. (1986). Mathematical model for predicting pavement performance. *Transportation Research Record, 4*, 48-55.

Watanatada, T., Hrall, C., Paterson, W., Dhareshwar, A., Bhandari, A., & Tsunokava, K. (1987). *The Highway Design and Maintenance Standard Model, 2 volumes.* Baltimore, Maryland.

Zhustareva, Y. (2000). Influence of cohesive soil density in the subgrade working layer on residual strains of nonrigid road pavements. *Dissertation, candidate of engineering sciences, Moscow.*

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Robust Voice Activity Detection with Deep Maxout Neural Networks

Valentin Sergeyevich Mendelev[1], Tatiana Nikolaevna Prisyach[2] & Alexey Alexandrovich Prudnikov[3]

[1] St. Petersburg National Research University of Information Technologies, Mechanics & Optics, Speech Technology Center, Russia

[2] STC-innovations Limited, Russia

[3] St. Petersburg National Research University of Information Technologies, Mechanics & Optics, Russia

Correspondence: Valentin Sergeyevich Mendelev, St. Petersburg National Research University of Information Technologies, Mechanics & Optics, Speech Technology Center, 197101, St.Petersburg, Kronverkskiy pr, 49, Russia. E-mail: mendelev@speechpro.com/prisyach@speechpro.com/prudnikov@speechpro.com

Received: January 9, 2015 Accepted: March 20, 2015 Online Published: June 25, 2015

doi:10.5539/mas.v9n8p153 URL: http://dx.doi.org/10.5539/mas.v9n3p153

Abstract

Voice activity detection (VAD) under non-stationary noises is a very important task to solve when using a real-life system of automatic speech recognition, especially if a remote microphone is used. Many existing methods do not work well with noise that changes over time or with very low signal-to-noise ratio (SNR). This paper proposes a method based on deep maxout neural networks with dropout regularization. The method is effective even for very low SNR (up to -5dB). The robustness of the method is demonstrated by low FR/FA error rates on a test dataset that was recorded under conditions different from the training dataset.

Keywords: voice activity detection, maxout networks, non-stationary noise, deep learning, dropout

1. Introduction

Noise robustness of a speaker activity detector is a very important requirement for real-life use. That is especially evident when using a remote micriphone. Noise introduces substantial distortions into the speech signal. If the system is trained on clean data and used for noisy data, that leads to a significant accuracy reduction in detecting the boundaries of speaker activity. Most existing approaches require the information about statistical characteristics of the noise to be known beforehand. These methods can be divided into three categories: approaches based on a deterministic rule, statistical approaches and neural networks-based approaches.

Approaches based on a deterministic rule use a number of characteristics, such as zero crossing rate, short-time energy, autocorrelation coefficients, in order to compare acoustic features with a certain preset threshold to make a decision. In (Rabiner & Sambur, 1975) two acoustic features (log energy and zero crossing rate) are used to detect the boundaries of isolated words. This algorithm is very simple but it does not work in noisy conditions. The authors of (Savoji, 1989) first calculate the probability density function for the spectrum of each frame and then the entropy. They obtain speech and pause labels by using certain preset thresholds. This method does work for a noisy signal, but only for slowly changing noise levels, and it is not stable under low SNR. In (Krubsack & Niederjohn, 1991), a deciding rule is used based on pitch detection. A speech confidence measure is determined using a euristic procedure based on three features extracted from the autocorrelation function. In (Junqua et al., 1994) the method for detecting word boundaries is based on a time-frequency parameter which is formed from the energy in the frequency band and the log of short-time energy. The noise threshold is first calculated based on several initial frames of the input signal and then compared to the time-frequency parameter in order to determine the initial boundaries. Then the threshold rule is used to determine initial and final boundaries of words. The main drawback of approaches based on deterministic rules is that they use thresholds extracted empirically from a segment of non-speech signal. Consequently, such methods do not work for cases when noise levels change over time. They are also not effective for low SNRs.

Statistical approaches (employing hidden Markov models (HMM), Gaussian mixture models (GMM)) use maximum aposteriori probability (MAP) or maximum likelihood (ML) criteria for speech detection. It is assumed that a feature vector belongs to a certain class. Different clustering methods are used for solving the task. However, a great amount of training data for different types of background noises is needed for estimating the probability distribution. The quality of these approaches depends on the choice of probability distribution and

the possibility of estimating the parameters of noise distribution. (Atal & Rabiner, 1976) solved the problem of speech detection for clean speech using an approach based on image **recognition**. Five acoustic features were used: zero crossing rate, short-time energy, the first coefficient of the autocorrelation function, the first-order linear prediction coefficient and the residual energy of linear prediction. The model for each class was a multidimensional Gaussian distribution. The MAP criterion was used for making the decision. In (Acero et al., 1993) HMM was used for modeling speech and pause classes, and the Viterbi algorithm was used for searching. (Bhiksha & Rita, 2003) describes using non-linear likelihood obtained from a Bayesian classifier. The main drawback of statistical methods is that the distribution of acoustic features for each class must be known beforehand. (Wu and Zhang, 2011) proposed using a linear weighted combination of different statistical models as the input of the unsupervised SVM.

Neural networks-based approaches use neural networks as template classifiers. There are two advantages in using such an approach. The first is that a neural network classifier is built directly on the training data without a strict assumption about the distribution of its classes. The second is the high discrimination capability of neural networks. (Qi & Hunt, 1993) proposes a multilayer neural network for detecting voiced and non-voiced fragments and pauses. Several features are used: cepstral coefficients, zero crossing rate and mean square energy. However, this approach completely ignores context information. In (Hong & Lee, 2013) RNN is used for classifying speech and non-speech fragments under noisy conditions. The authors demonstrate the advantage of using a RNN classifier compared to a GMM classifier. They describe the efficiency of the method under changing noise levels, however they only deal with different automobile noises. The paper (Zhang & Wu, 2013) proposes a deep belief network (DBN)-based VAD. DBN is a powerful hierarchical generative model for feature extraction. Unlike traditional methods of training deep models, DBN can prevent overfitting by using a special unsupervised pretraining procedure. A DBN-based VAD first connects acoustic features in a long feature vector, which is used as a visible layer or input DBN. Then a new feature is extracted as a result of the transition of the long feature vector through multiple nonlinear hidden layers. As a result, each class of observation is predicted by the linear classifier, so the output is the softmax layer of the DBN with a new feature at the entrance.

Deep neural networks have a long history. They may describe a highly variant function using several parameters. If the training is completed successfully, they can achieve good generalization capability even with a small volume of training data.

We propose using a deep neural network with a maxout activation function and dropout regularization. Dropout technology has shown its efficiency on small training data. Using maxout improves the accuracy of model averaging with dropout. The trained neural networks are highly effective for noisy data even under low SNR and in case of training and test data mismatch.

2. The Proposed Technique

2.1 The Structure of VAD

Figure 1 shows the structure of the speaker activity detector.

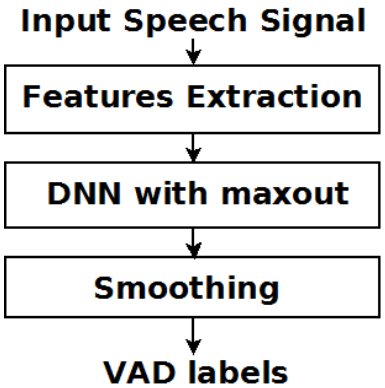


Figure 1. VAD structure.

The first stage is feature extraction. We use Fbank with context length 15.

Then the features are fed into a trained deep maxout neural network. The output of the network are the aposteriori probabilities of each frame belonging to one of the classes (speech, pause). For the correct interpretation of the speech segments and pause segments, a threshold is used for aposteriori probabilities. The threshold value may be selected automatically depending on the SNR. The threshold increases for larger values of SNR, so that VAD can separate speech from pause with greater certainty. Thus, the values of aposteriori probabilities will be close to 1 at high SNR. At low SNR, the aposteriori probabilities of speech segments may decline to 0.6. Anyway, the threshold does not fall below 0.55 in our case, since it is important to identify all the speech.

The last stage is smoothing the frame labeling. By default, fragments shorter than 1 second are smoothed.

2.2 Training Features

The choice of training features is critical for any classification task. For a speaker activity detector, good features must satisfy two conditions: 1) the distribution of speech and non-speech fragments must be different, that is, good features must not overlap for speech and noise classes; 2) the features must be robust to noise.

The following feature types are used in the literature: energy features (Rabiner & Sambur, 1975), spectrum features (Boll, 1979), cepstral features (Kinnunen et al., 2007), harmonic features (Kingsbury et al., 2002) and long-term features (Fukuda et al., 2008).

We examined and compared the following features: mel-frequency cepstral coefficients (MFCC) (Kinnunen et al., 2007) with context, filter banks (Fbank) with context and with normalization of the cepstral average, gammatone frequency cepstral coefficients (GFCC) (Shao wt al., 2009) with context. The advantage of GFCC features compared to the others is that they are more robust to noise, so they work better for speech detection.

2.3 The Training Network

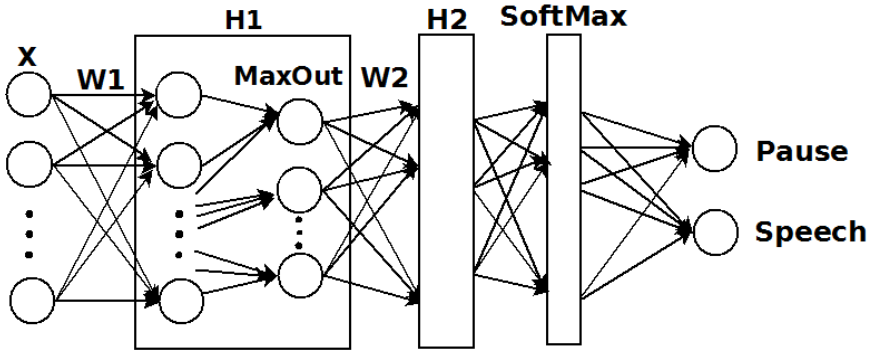


Figure 2. The neural network for VAD with two hidden layers H_1 and H_2.

2.3.1 Network Structure

Figure 2 shows the structure of the neural network used for training two classes: “speech” and “pause”. It contains two hidden layers H_1 and H_2. The structure of H_2 is analogous to H_1.

Let us examine the fully connected layer of the neural network with the input feature vector $X=[x_1, x_2, ..., x_n]^T$ and the weight matrix of the first hidden layer $W1$ with dimenstion $d \times n$, d in the experiments eqals 1200, n is the dimension of the minibatch. It is assumed that in order to increase the accuracy of network training the input features are normalized for the mean and variance. The initial weights are selected randomly in the interval $[-0.01, 0.01]$ according to uniform distribution. The output is the vector $Y1=[y1_1, y1_2, ..., y1_d]^T$, which is calculated as matrix multiplication between the input vector of the layer and the weight matrix using the activation function A. As a result,

$$Y1_{ij} = A\left(W1_{ij} * X + b1_{ij}\right) \ , \ z1_{ij} = W1_{ij} * X + b1_{ij},$$

where $W1_{ij}$ are the elements of the weight matrix of the first hidden layer, $b1_{ij}$ are the corresponding offsets, $i=1,...,n, j=1,...,d$.

After that, the neurons in the network are united into groups, each of which consititutes a maxout node. The number of groups in the experiment is 5. As the activation function A we use the maximum selected from several candidates of the maxout node. (Goodfellow et al., 2013) shows the advantage of the maxout network compared to differentiated activation functions, such as tanh (hyperbolic tangent), which consists of better

approximation of model averaging. The maxout activation function is represented as

$$Y1_r = \max_r (z1_r),$$

where $z1_r$ are the values of the neurons in the r th maxout node, $r = 1,...,R$, R is the number of neurons in the maxout node.

Dropout is used at the output of the hidden layer. Most of the literature on deep training focuses mainly on regularizing the network so as to avoid overfitting. Different regularizing methods exist (L1, L2 (Bengio, 2012), L2-prior regularizing (Liao, 2013)). Dropout (Hinton, 2012), (Wang & JaJa, 2014) is a widely used and effective regularizing method for DNNs. It makes it possible to avoid complicated coadaptations on training data. On the other hand, the dropout procedure is an efficient way of averaging models with neural networks. A good way to reduce error on the test set is to average the predictions obtained from a very large number of different networks.

The standard solution is to train many separate networks and then apply each of them to the test data, but this process is very labor-intensive both for training and for testing. Random assignment of a zero value to a neuron makes it possible to train a large number of different neural networks in reasonable time. Networks for each training vector are trained in this way, but all networks have a common weight matrix. At the output, each neuron of the layer is assigned a zero value with the probability $1-p$. Experiments show that dropout increases generalization capability of the neural network and improves results on test data. Dropout is also efficient for small training datasets. Combined with maxout, dropout makes it possible to achieve exact rather than approximate model averaging and to fully utilize its potential. The use of dropout is illustrated in Figure 3.

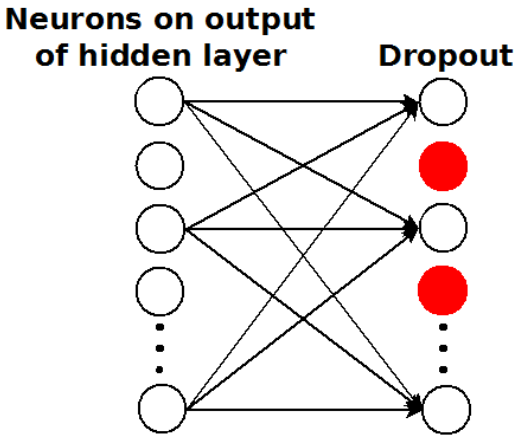


Figure 3. Regularizing using dropout.

Using dropout at the output of the layer we get

$$Y1 = M * A(W1 * X),$$

where M is the vector binary mask with the dimension d, $M_j \sim Bernoulli(p), j = 1,...,d$. So the fully connected layer with DropOut becomes a sparse layer in which the values of neurons are updated randomly during training. Each element of the mask is independent for each training feature vector and in fact establishes different connections for each new feature vector from the training dataset. In addition, the mask is also applied to offsets during training.

As with $Y1$, at the output of the second hidden layer we get the vector $Y_2 = [y2_1, y2_2, ..., y2_d]^T$.

At the output of the classifier network we use a softmax layer which normalizes the sum of output values to equal 1 and makes it possible to interpret the outputs of the neural network as aposteriori probabilities: $O = S(Y2; W_S)$, W_S is the vector of dimension parameters k, k is the number of classes (in our case, 2: speech and pause).

2.3.2 Training Conditions

According to the target function, we calculate the value of the training error E. We use the cross-entropy criterion as the target function (Golik, 2013)

$$E(D,O) = -\sum_{i=1}^{m} D_i \log(O_i), m = 1,...,k,$$

D - k - dimensional desired response of neural network.

At the final training stage we calculate the increments ΔW for the weights of each neuron for their subsequent updating. We use the standard backward propagation for that (Rojas, 1996). Weight increments are calculated starting from the softmax layer and ending with the first layer. For softmax layer

$$W_S = W_S - \eta E'_{W_S},$$

η is the training speed.

We introduce two differences from the standard backward propagation procedure. Firstly, increments for the dropout layer are calculated: $Y2 = Y2 - \eta(M * A'_{Y2})$. Here the weights that were active during the direct pass are updated according to the mask M. Secondly, such a mask is also used for the maxout layer, that is, only the weights corresponding to maximum values are updated. The increments for the weights of first hidden layer H_1 are determined in the same way.

3. Experiments and Results

3.1 Training and Testing Datasets

The training and testing data for the neural networks were taken from the speaker database recorded at Speech Technology Center (STC). The database contains recordings of phonetically rich sentences using a remote all-direction microphone under various acoustic conditions (office, home, car, street). The microphone was located at the distance of 2 to 3m from the speaker, with a 0.5m error. The experimental dataset is described in Table 1.

Table 1. Experimental dataset

Dataset	Sampling rate, Hz	Duration, hours	Men	Women	SNR, dB
STC dataset	16000	131	188	220	[-5; 20]

3.2 Experimental Results

For testing the robustness of VAD with maxout DNN we trained several DNNs with different features. All the networks had two hidden 1000-dimensional layers.

The fbankCMN_2HLx1000_L2.net DNN was trained using Fbank with context length 15 using cepstral mean normalization. L2 regularization was used during training, network configuration was fully connected.

The mfccCMN_2HLx1000_L2.net DNN was trained using MFCC, in other respects it was similar to the previous network.

The gammatoneNet_2HLx1000_L2.net DNN was trained using gammatone features with context length 15. L2 regularization was used during training, network configuration was also fully connected.

The final network, maxoutCMN_2HLx1000.net, was trained using Fbank, context length 15, using cepstral mean normalization. We used dropout regularization for training, the network configuration was described above in Section 4, maxout activation function was used.

The test data were remote microphone recordings with the total duration of 3 hours, containing different types of office and home noises, as well as street, automobile, construction noises. The test data did not match the training data: a different type of remote microphone was used, the distance was not the same (the microphone could be further away from the speaker than 3m).

Table 2 shows the results of FR/FA, where FR is the error “speech as pause” and FA is “pause as speech”.

Table 2. Comparison of FR/FA for different neural networks

Network configuration	Features	FR	FA
fbankCMN_2HLx1000_L2.net	31xFbank+CMN	8	13
mfccCMN_2HLx1000_L2.net	31xMFCC+CMN	10	12

gammatoneNet_2HLx1000_L2.net	31xGF	11	13
maxoutCMN_2HLx1000.net	31xFbank+CMN	**2.6**	**2.5**

4. Discussion

The table shows that using the maxout-activation function in combination with the dropout regularization reduces "speech as pause" and "pause as speech" errors four times. The reasons why this result is achieved are as follows. First, maxout does not use a fixed activation function, instead the function is created during training. Second, maxout is a universal approximator. Any continuous function can be approximated arbitrarily well on a compact domain by a maxout network with two maxout hidden units. Third, dropout performs model averaging, so maxout in conjunction with dropout enhances the accuracy of dropout model averaging technique and improves optimization. The maxout model benefits more from dropout than other activation functions. Through the use of such technology we can achieve greater robustness in noisy conditions.

This paper presents the results of the first experiments with DNN with maxout activation function and dropout regularization on noisy features. In the future, we are particularly interested in the following topics. We plan to perform experiments on the selection of dropout-regularization parameters and the size of maxout groups. Perhaps an increased number of hidden layers can improve the final layout. We would also like to conduct more comprehensive experiments on different ranges of SNR. The described DNN with maxout activation function and dropout regularization was trained on noise fbank. We plan to use other features with different context lengths. Using gammatone features may give the best results under very noisy data.

5. Conclusion

The paper presents a robust speaker activity detector based on DNNs with maxout activation function and dropout regularization. It is well-known that the main problem for DNN training is often the insufficient amount of training data. Our experiments show a high efficiency of speech/non-speech detection for the proposed method even in case of mismatched training and test data. The effectiveness of the method is demonstrated by lower error rates compared to standard DNNs under low SNR (up to -5 dB).

Further research will focus on the use of more robust features with different context lengths and on different SNR ranges. Experiments are planned to select optimal parameters of DNN training (selection of maxout group size, dropout-regularization parameters, number of hidden layers).

Acknowledgements

The work was financially supported by the Ministry of Education and Science of Russian Federation (Contract 14.579.21.0057, ID RFMEFI57914X0057) and by the Government of the Russian Federation, Grant 074-U01.

References

Acero, A., Crespo, C., Torre, D. L., & Torrecilla, J. C. (1993). Robust HMM-based endpoint detector. Proc. Eurospeech (pp. 1151-1154).

Atal, B. S., & Rabiner, L. R. (1976). A pattern recognition approach to Voiced-Unvoiced-Silence classification with application to speech recognition. IEEE Trans. Acoust. Speech Sig. Process, 24. (pp. 201-212). http://dx.doi.org/10.1109/TASSP.1976.1162800.

Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. arXiv report:1206.5533, Lecture Notes in Computer Science Volume 7700, Neural Networks: Tricks of the Trade Second Edition, Editors: Gregoire Montavon, Genevieve B. Orr, Klaus-Robert Muller (pp. 437-478). http://dx.doi.org/10.1007/978-3-642-35289-8_26.

Bhiksha, R., & Rita, S. (2003). Classifier-based non-linear projection for adaptive endpointing of continuous speech. Computer Speech and Language, 17 (pp. 5-26). http://dx.doi.org/10.1016/S0885-2308(02)00028-1.

Boll, S. F. (1979). Suppression of acoustic noise in speech using spectral subtraction. Acoustics, Speech and Signal Processing. IEEE Transactions on, vol. 27, no. 2, (pp. 113–120). http://dx.doi.org/10.1109/TASSP.1979.1163209

Fukuda, T., Ichikawa, O., & Nishimura, M. (2008). Phone-duration-dependent long-term dynamic features for a stochastic model-based voice activity detection. In Ninth Annual Conference of the International Speech Communication Association.

Golik, P., Doetsch, P., & Ney, H. (2013). Cross-entropy vs. squared error training: a theoretical and experimental

comparison. INTERSPEECH, ISCA. (pp. 1756-1760).

Goodfellow, I. J., Warde, F. D., Mirza, M., Courville, A., & Bengio, Y. (2013). Maxout networks. arXiv preprint arXiv:1302.4389.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. Presented at CoRR. arXiv:1207.0580v1.

Hong, W. T., & Lee, C. C. (2013). Voice Activity Detection based on Noise-Immunity Recurrent Neural Networks. *IJACT, 5*(5), 338-345. http://dx.doi.org/10.4156/ijact.vol5.issue5.41.

Junqua, J. S., Mak, B., & Reaves, B. (1994). A robust algorithm for word boundary detection in the presence of noise. IEEE Trans. *Speech and Audio Processing, 2*, 406-412. http://dx.doi.org/10.1109/89.294354.

Kingsbury, B., Saon, G., Mangu, L., Padmanabhan, M., & Sarikaya, R. (2002). Robust speech recognition in noisy environments: The 2001 IBM spine evaluation system, in Acoustics, Speech, and Signal Processing (ICASSP). *2002 IEEE International Conference on, 1*, I–53–I–56.

Kinnunen, T., Chernenko, E., Tuononen, M., Frnti, P., & Li, H. (2007). Voice activity detection using MFCC features and support vector machine. *Int. Conf. on Speech and Computer, 2*, 556–561.

Krubsack, D. A., & Niederjohn, R. J. (1991). An autocorrelation pitch detector and voicing decision with confidence measures developed for noise-corrupted speech. *IEEE Trans. Sig. Process, 39*, 319-329. http://dx.doi.org/ 10.1109/78.80814.

Liao, H. (2013). Speaker adaptation of context dependent deep neural networks, in Proc. ICASSP, 2013. (pp. 7947-7951). http://dx.doi.org/10.1109/ICASSP.2013.6639212.

Qi, Y., & Hunt, B. R. (1993). Voiced-unvoiced-silence classifications of speech using hybrid features and a network classifier. IEEE Trans. *Speech and Audio Process, 1*, 250-255. http://dx.doi.org/10.1109/89.222883.

Rabiner, L. R., & Sambur, M. R. (1975). An algorithm for determining the endpoints of isolated utterances. *Bell System Technical Journal, 54*(2), 297-315. http://dx.doi.org/10.1002/j.1538-7305.1975.tb02840.x

Rojas, R. (1996). Neural Networks: A Systematic Introduction. Springer-Verlag, New York, USA. 502.

Savoji, M. H. (1989). Robust algorithm for accurate endpointing of speech signal. *Speech Communication, 8,* 45-60. http://dx.doi.org/10.1016/0167-6393(89)90067-8.

Shao, Y., Jin, Z., Wang, D., & Srinivasan, S. (2009). An auditory based feature for robust speech recognition. In Proc. ICASSP'09, 2009. (pp. 4625–4628). http://dx.doi.org/10.1109/ICASSP.2009.4960661.

Wang, Q., & JaJa, J. (2014). From Maxout to Channel-Out: Encoding Information on Sparse Pathways. ICANN 2014. (pp. 273-280). http://dx.doi.org/10.1007/978-3-319-11179-7_35.

Wu, J., & Zhang, X. L. (2011) Maximum margin clustering based statistical VAD with multiple observation compound feature. *IEEE Signal Process. Lett., 18*(5), 283–286. http://dx.doi.org/10.1109/LSP.2011.2119482.

Zhang, X. L., & Wu, J. (2013). Deep belief networks based voice activity detection. IEEE Trans. Audio, Speech, *Lang. Process., 21*(4), 697–710. http://dx.doi.org/10.1109/TASL.2012.2229986

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Mathematical Model of Kazakhstan Economy

Orakbayev E.M.[1], Boranbayev S.N.[1], Vashenko M.P.[2] & Shananin A.A.[2]

[1] L.N. Gumilyov Eurasian National University, Kazakhstan

[2] The computer center of Russian Academy of Sciences, department of mathematical modeling of economic systems, Russia

Correspondence: Orakbayev E.M., L.N. Gumilyov Eurasian National University, 01008, Kazakhstan.

Received: January 9, 2015 Accepted: January 29, 2015 Online Published: July 15, 2015

doi:10.5539/mas.v9n8p160 URL: http://dx.doi.org/10.5539/mas.v9n3px

Abstract

This article gives a description of a model of the Republic of Kazakhstan modern economy. The model takes into account the specific character of economic brunches and the peculiarities of the competition between the domestic production and imported analogues at the domestic market. A particular attention is paid to the problem of formation of intersectional balance for the enlarged structure of economy. This article also analyzes questions of adaptation of classic approaches offered by Leontiev V.V. to the calculation of input-output coefficients for the brunches that compete with import at the domestic markets.

Keywords: mathematical model, market, exporter, importer, trade agent, currency, commercial services, production, commercial bank, national bank, imported goods, domestic goods, salary, government, population

1. Introduction

Problem of globalization and of Kazakhstan integration into world economic area became extremely urgent after the introduction of the Customs Union and Kazakhstan's possible accession to the WTO. Welfare of some brunches became dependent on processes with more open foreign trade policy which should be taken into account in short-term planning.

During the period of 2012-2013 import growth rate was much ahead of export growth rate: in 2012 the difference in growth rates was 24% and after eight months of 2013 it was 10%. This situation threatens those sectors of Kazakhstan economy that are engaged in competition with imported goods on the domestic market and that traditionally employ the major part of the population. For example share of products of animal and plants origin, of ready-made food products in import structure has increased by 0.4% for 8 months of 2013, and has decreased by 0.8% in export structure. Despite the fact that Kazakhstan agriculture, forestry and fish industry involve 25.54% of the population such changes can cause unwanted negative effects. That is the reason why the problem of agricultural adjustment (volumes and methods of its funding) is one of the most urgent at the negotiations of Kazakhstan with the WTO members.

Apparently Kazakhstan accession to the WTO will have a certain transition phase that will help economy sectors to adjust to new conditions. This makes requested those tools that give opportunity to make not only qualitative but also quantitative evaluations of influence of import rivalry and of import itself on the activity of Kazakhstan economy sectors.

Traditional methods of long-term macroeconomic framework that used to be popular in times of Soviet Union were based on the analysis and compensation of economy subjects' plans. Input-output model by V.V. Leontiev was the main mathematical tool of this adjustment. The quality of medium-term and long-term forecasts has deteriorated with the reduction of planning horizons of economy subjects. Many agents do not plan their activity for more than one year in the context of transition economy and it is impossible to base long-term forecasts on their plans. Forecasting methods that were developed during the period of planned economy turned out to be inadaptable to modern conditions. The only way to solve this problem is to develop methods of mathematical modelling that give opportunity to describe the performance of economy subject taking into account peculiarities of their activity. Countries with developed market economy possess a successful experience of forecasting based on mathematic modelling. In these countries activity of economy subjects does not undergo any rapid changes and could be studied and described by econometric approaches. But models of economic agents that are successfully applied in steady market relations could not be applied to immature market institutions of

Kazakhstan and at least require a significant adaptation. Econometric relations could be maintained only if development is stable and has no structural changes, because they do not take into account a response on the activity of economy subjects. However Kazakhstan economy hasn't reached the state that can ensure the steady tendency of the development even for medium-term perspective. Commercial banks liquidity crisis of 2008 has proved it; this crisis was accompanied by changes in strategy of bank behavior and by significant fall of production caused by scarcity of working capital. Distinctive features of Kazakhstan and other countries which are in the process of shaping modern market infrastructure are sudden changes of agents' economic strategy. In order to describe the activity of such agents it is necessary to find models that show causal relationships in economy.

The listed problems give us a sound right to claim that long-term development forecasts should be based on correctly structured economy models that would show specific features of the production and circulation and would take into consideration significant responses in economic system. These models could be formed as a part of the approach called systemic analysis of developing economy that is being worked out by scientific school of the member of the RAS A.A. Petrov (1,2) for more than 30 years. The main point of the systemic approach is in the fact that mathematic models include not only a complete description of technologies and technologic relations but also economic instruments of production and circulation arrangement typical for a particular economic system. Thus we have models that give us a self-contained description of the economy development process provided the availability of a set plan for national economic policy. Systemic approach to economy modelling in contrast to any other approaches (for example econometric or an approach based on the models of input-output balance), takes into consideration significant responses in economic system and helps to make models of indirect consequences of important social and economic decisions. Response recording helps to make system oriented qualitative evaluations of economic development programs and to study structural economic changes that could appear as the result of implementation of these programs. The evident advantage of systemic approach methods in the context of possible structural changes of economy and possible alteration of subject's part is the fact that they do not require detailed specification of economic activity. Moreover the applied methods give opportunity to describe qualitative features of agents' behavior and to analyze their possible changes. These methods consider long-term responses depending on ways of national economic adjustment. The listed advantages underline the prospect of system approach application to the problems of medium-term and long-term forecasting.

To sum up everything mentioned above we can claim that the aim of medium-term and long-term forecasting is not the creation of time sequence of macroeconomic indicators but the creation of a tool for efficient and systemic analysis of the consequences of important economic decisions with due regard to the responses. This tool could be materialized as a part of systemic approach to the analysis of a developing economy.

2. Methodology

Works of classical authors studying economic science, modern works of foreign and Kazakhstan academic economists and of specialists in the sphere of regional economy and modelling of interregional and intersectional connections represent the methodological basis of the research.

We used methods of systemic analysis, of comparative analysis and composition, of macroeconomic modelling and forecasting, of simulation modelling, methods of quantitative and qualitative studying of the economic processes under investigation and also expert evaluation methods. In this work we used reference materials, tables and schemes. The author used the following methodological approach: taking Kazakhstan input-output balance of 2008 and 2012 as a basis of forecast analysis.

3. Results

3.1 Model Description

Let us describe general model scheme with its following specifications that reflect model's features. This model distinguishes the following economic agents (or groups of agents):

A1. A group that represent the actual economy sector and service sector. The description of this group (its structure and dividing strategy) will be provided below.

A2. A group that represent national sector. This group includes the following agents:

   a. Ministry of Finance, that form a state budget;

   b. National Bank (NB).

A3. Households (population).

A4. Commercial banks.

A5. Trade agents that are responsible for import transactions and retail trading.

This model distinguishes the following markets with action mechanisms described below:

M1. The wholesale market of goods and services. The wholesale market M1 is responsible for intermediate consumption and export of A1 agents' production. M1 market participants are A1 agents and A5 trade agent.

- At the stage of intermediate consumption A1 agents are direct consumers because they purchase the required production factors. A1 agents purchase at M1 market domestic goods and services as well as imported goods and services distributed by A5 trade agent.
- Export of the production is performed by selling goods at the wholesale market to A5 trade agent who resells them at the outer market.

M2. Retail market. M2 retail market is responsible for the final consumption of A1 agent goods. Participants of this market are: A3 households, A2a government and A5 trade agent. Households and government are consumers while trade agent is a bargainer. Trade agent offers at M2 market both domestic goods purchased at A1 and imported goods.

M3. Foreign exchange market. Participants of this market are:

- Trade agent who purchases currency to supply import transactions and sells goods to supply export transactions.
- Government (Ministry of Finance), performing foreign borrowings and their maintenance.
- National Bank that carry out policy of currency rate.

Household (A3) incomes are formed by the demand of A1 agents on labor and A2a government subsidies.

Government revenues are formed by means of collection of taxes paid by A1 agents, A3 households and A4 commercial banks.

Besides M1, M2 and M3 market campaigns economic agents also provide interaction in the system of monetary circulation. This is in the responsibility of A4 commercial banks:

1. Distribute loans and take deposits from households
2. Grant loans to A1 agents who maintain them afterwards
3. Cooperate with NB that extends or absorbs liquidity of banking system.

Here is more detailed description of A1 agents. Brunches of economy are studied in the context of their cooperation with world markets. It is an attempt to single out those parts of production system and service system that interact with internal and external markets in fundamentally different ways and that have distinct responses to government policy in the sphere of foreign commerce and currency rate. The following sectors are distinguished in this model:

A1-1. "Manufacturing" industries (Sector No.1). The agent represents brunches that do not have export capacity and undergo keen competition with export. Engineering industry is an example of "manufacturing" industry. A1-1 sector is considered to be a generating one, i.e. production of this sector only is used for building new capacities.

A1-2. "Infrastructural" brunches (sector No.2). The agent represents brunches mainly oriented to domestic demand. This sector does not compete with import. Consumption of the third sector as a rule is not flexible in prices. Transportation, power industry, communication are examples of the second sector.

A1-3. "Export" brunches (sector No.3). The agent represents brunches that have high export potential. Prices to this production are formed on external markets.

A1-4. "Commerce and service" sector (sector No.4). The agent represents the sector that includes such brunches as health care and education.

As for formal criteria based on input-output balance (input-output table system) of 2011:

1. A1-3 sector included brunches which had output export percent exceeding 20% and export itself exceeded 10 billion tenge in nominal terms.
2. A1-2 sector has combined the following brunches:
    - Air transport;
    - Water supply system; sewerage system, control over the collection and distribution of wastes;
    - Land transport and transportation by pipelines;
    - Water transport;
    - Storage facilities and supporting transport activity;
    - Information;
    - Post and express activity;
    - Communication;
    - Electrical supply, supply of gas, vapor and air conditioning.

We should underline that export percent relatively to output is unusually high for these brunches and for Air transport, Water supply system, Land transport in particular (for these three brunches it is more than 10%). However if we calculate net export percent (export – import) relatively to output in order to take into consideration the brunch facilities regarding resources transfer across Kazakhstan territory we will find out that average percent of this indicator is -2.3% for the whole A1-2 sector (0.4% for three underlined brunches). This proves the hypothesis that A1-2 sector is oriented on domestic demand.

3. A1-1 sector included manufacturing branches that were not covered by sectors A1-3 and A1-2. Average ratio of import manufacturing costs to output is 10.13% for this sector.
4. Branches of commercial services has composed A1-4 sector.

3.2 Description of Markets

3.2.1 Description of the Wholesale Market

A1 manufactures and A5 trade agents interact at the wholesale market. There are two major processes at the market:

1. Exchange (purchase and sale) of goods between manufactures and trade agents to fulfill goals of satisfaction the production needs. Trade agent in this case acts as a supplier of imported goods required for the production. Key feature of this process is input-output balance that describes how many units of imported or domestic goods of any sector are required for manufacturing the production unit of every domestic manufacturer.
2. Distribution of goods, produced by all manufactures except sector 4, to a trade agent for the following resale at the retail market or export. It is considered that sector 4 ("service" sector) distributes its goods independently to population and government at the domestic market.

3.2.2 Description of the Currency Market

Purchase and sale of currency takes place at the currency market. Participants of this market are:

1. The trade agent who purchases currency to buy imported goods and sells currency to convert currency gain received from the disposal of export goods.
2. State budget that finances its deficit at the cost of foreign loans (currency sale) and that pays interest charges and pays off an external debt (currency purchase)
3. National Bank that perform currency interventions at the cost of gold and forex reserves in order to maintain the needed foreign exchange.
4. Population.

Let us take GFR for gold and forex reserves of NB, p_i^{ex} for index of prices at exported goods of i sector, p_1^{Im}

for index of prices at imported goods of i sector, Y_i for the output of i sector, S_i for the demand for domestic products of i sector, S_1^{Im} for the demand for imported products of i sector, φ for the foreign exchange (currency baskets), H for currency purchase by the population.

Currency supply

$$Ex_3(\mathrm{t}) = (\mathrm{Q}_3(\mathrm{t}) - \mathrm{S}_3(\mathrm{t})_+ \mathrm{p}_3^{ex}(\mathrm{t})) \tag{1}$$

Currency demand

$$\mathrm{Im}(\mathrm{t}) = \mathrm{C}_1^{\mathrm{Im}}\ p_1^{\mathrm{Im}}$$

$$\frac{dGFR}{dt} = (\mathrm{Q}_3(\mathrm{t}) - \mathrm{S}_3(\mathrm{t})_+ \mathrm{p}_3^{ex}(\mathrm{t})) + H_G(t) + r_{GFR}GFR -$$
$$-\delta_{GFR}Tax_3^{\varphi}(t) - \frac{1}{\varphi}(\mathrm{C}_1^{\mathrm{Im}}\ p_1^{\mathrm{Im}} + \sum_{i=1}^{3} a_{i\,\mathrm{Im}}Y_i(t) + \sum_{i=1}^{3} B_i^{\mathrm{Im}}(\mathrm{t}) + B_{Tr}^{\mathrm{Im}}(\mathrm{t}) + B_{\mathrm{Im}}^{\mathrm{Im}}(\mathrm{t}) + G_1^{\mathrm{Im}}(t)) \tag{2}$$

3.3 Description of Economic Sectors

3.3.1 Description of Manufacturing Sector

Manufacturer of I sector does its best to maximize its unreachable discounted revenue by operating maximum stock of goods. Points of goods realization are considered to be events of Poisson stream with the parameter λ. Volume of goods equal to the minimum of the current manufacturers stock $Y_1(t)$ and of maximum sales volume Y^* should be disposed during every realization point. Manufacturer produces goods to ensure the maximum stock ($\zeta(t)$ stands for the ratio of maximum stock to maximum sale volume) and then waits for the next realization point.

Demand for domestic production of the first sector:

$$S_1^{Int}(t) = \sum_{j=1}^{4} a_{1,j}^{Int}Y_j(t) + Tr_1^{Int}(t) + \sum_{j=1}^{4} B_{1,,j}(t) + B_G^{Int}(t) \tag{3}$$

The wholesale price on domestic production of the first sector is set at the wholesale market of the first sector:

$$\frac{dp_1}{dt} = \alpha p_1 \frac{S_1(\mathrm{t}) - \mathrm{Y}_1(\mathrm{t})}{\mathrm{Y}_1(\mathrm{t})}: \tag{3.1}$$

$$\frac{d\Xi_1}{dt} = Y_1(\mathrm{t}) - S_1(\mathrm{t}). \tag{3.2}$$

Points of goods realization are considered to be events of Poisson stream with the parameter λ. Volume of goods equal to the minimum of the current manufacturers stock $\Xi_1(\mathrm{t})$ and of maximum sales volume Y^* should be disposed during every realization point.

Liquidity limits:

$$M_1 = \theta_1\left(p_1 B_1(\mathrm{t}) + \varphi \mathrm{B}_1^{\mathrm{Im}}(\mathrm{t}) + \mathrm{rL}_1\right) \tag{3.3}$$

Inventory balance of the first sector:

$$\frac{dM_1}{dt} = (p_1 - y_1)(Y_1(\mathrm{t}) - \frac{d\Xi}{dt}) - \mathrm{W}_1(\mathrm{t}) - \mathrm{T}_1(\mathrm{t}) - rL_1(\mathrm{t}) - \mathrm{p}_1\ B_1(\mathrm{t}) - \varphi B_1^{\mathrm{Im}}(\mathrm{t}) - \mathrm{DIV}(\mathrm{t}) \tag{3.4}$$

Change of powers:

$$\frac{dQ_1}{dt} = g\left(\frac{B_1(\mathrm{t})}{b_1}, \frac{B_1^{\mathrm{Im}}(\mathrm{t})}{b_1^{\mathrm{Im}}}\right) - \mu_1 Q_1 \tag{3.5}$$

where μ_1 is a pace of capacity disposal.

$$\mathrm{T_1}(t) = (1-\mathrm{t_1})((\mathrm{p_1}-\mathrm{y_1})(\mathrm{Y_1(t)}-\frac{d\Xi_1}{dt}) - \mathrm{W_1(t)} - \mathrm{rL_1(t)}) \tag{3.6}$$

where t_1 is profit tax rate, w_1 is percent of salary in the profit, y_1 is manufacturing unit cost in I sector

$$\mathrm{Re}\nu_1(\mathrm{t}) = (\mathrm{p_1}-\mathrm{y_1})(\mathrm{Y_1(t)}-\frac{d\Xi_1}{dt}) \tag{3.7}$$

$$\mathrm{DIV(t)} = \nu_1(\mathrm{Re}\nu_1(\mathrm{t}) - \mathrm{W_1(t)} - \mathrm{T_1(t)} - \mathrm{rL_1(t)}) \tag{3.8}$$

$$\frac{dL_1}{dt} = U_1\,\mathrm{y}\eta - \lambda(\int_0^{(1+\zeta_0)\mathrm{Y}^*} yx\mathrm{p}(x)dx) + (1-U_1)((1+\zeta_0)\,\mathrm{Y}^* y) \tag{3.9}$$

Average characteristics of the manufacturing

$\frac{\lambda Y^*}{Q_1}$ are characteristics of manufacturing capabilities provided limits of commercial infrastructure are Y^*

Take:

$$\varsigma_0 = \frac{\tau_0 Q_1}{Y^*} \quad ;$$

$$\left(1-\frac{\Delta}{\lambda}\frac{y(\lambda+\Delta)}{p(\lambda+\Delta-r)-y(\lambda+\Delta)}\right)\left(\frac{\lambda}{\lambda+\Delta}\left(1-e^{-\frac{\lambda+\Delta}{\lambda}\varsigma_0\frac{\lambda Y^*}{\eta}}\right)+e^{-\frac{\lambda+\Delta}{\lambda}\varsigma_0\frac{\lambda Y^*}{\eta}}\left(e^{\frac{\lambda+\Delta}{\lambda}\frac{\lambda Y^*}{\eta}}-\varsigma_0\frac{\lambda Y^*}{\eta}\right)\right)=1$$

Average capacity utilization:

$$U_1 = 1-\frac{e^{-\frac{\lambda Y^*}{Q_1}(\varsigma_0+1)}}{1-\frac{\lambda Y^*}{Q_1}\varsigma_0 e^{-\frac{\lambda Y^*}{Q_1}}} \tag{3.10}$$

, where

$$\zeta_0 \in (0,1);$$

$$Q_1 U_1 = Y_1(\mathrm{t}).$$

$$\hat{\Xi}_1 = Q_1 U_1 \frac{p_1}{y_1}\left(\frac{\lambda+\Delta-\mathrm{r}}{\lambda+\Delta}\right)\left(\frac{1-R}{\lambda}\right)\left(\frac{1-U_1}{U_1}\right)\left(e^{\frac{\lambda Y^*}{R_1}\varsigma_0}\left(1-\frac{\lambda Y^*}{Q_1}\varsigma_0+e^{\frac{\lambda Y^*}{R_1}}\right)-\mathrm{r}\right) \tag{3.11}$$

Production profitability

Assume all profitability conditions are fulfilled:

1. $\lambda+\Delta>r$, $Y^*>0$
2. $p > y\frac{\lambda+\Delta}{\lambda+\Delta-r}$

Profitability of manufacturing

$$R = \frac{p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}}{p} \tag{3.12}$$

3.3.2 Description of Extractive Sector

Let's describe a representative of extractive sector of economy. And for the beginning we will consider three business lines of an agent:

- manufacturing;
- investment;
- financial.

We will assume that operational business of an agent is narrowed down to the development of existing deposits. While developing deposits the agent decides how hard it is necessary to extract resources at the current time – variable $I_3(t)$. In order to extract the required volume of resources the agent will have to invest some costs the volume of which depends on the "difficulty" – a feature of the deposit and planned production output. For the indicator of the "difficulty" we will take the cost of the first sector production used for extraction of one unit of the resource at the deposit in base-year prices – z parameter. We will assume that agent has an access to the information about density of the cumulative function curve of available oil field development and about the prices for production at the current time $\hat{I}(z,t)$. That means that $\hat{I}(z,t)$ is amount of the resource that could be extracted for z tenge at the moment of time t. Agent is always supposed to develop the deposit with the cheapest resource from the point of view of prime costs. Such minimum prime cost at the current time is marked $z_{\min}(t)$.

Resource extraction itself and expenses relating to it are separated in time. That means that after taking a decision that it is necessary to extract $I_3(t)$ units of the resource the agent should pay costs according to one diagram and to get resource according to another. Costs diagram is marked as $b_3(\tau)$ and extraction diagram is marked as $d_3(\tau)$ and it is considered that all costs and extraction take place during a certain period T_3. I.e. if it is necessary to extract $I_3(t)$ resource units with prime cost of the extraction $z_{\min}(t)$ then at the stated time $(t+\tau)$ the agent should pay the sum $p_1(t) \cdot p_1^0 \cdot I_3(t) \cdot b_3(\tau) \cdot z_{\min}(t)$, where $p_1(t)$ is an index of wholesale prices on the production of the first sector, and p_1^0 is a price for production unit in base year; thus the agent will get $I_3(t) \cdot d_3(\tau)$ of resource units.

If $z_{cp}(t)$ stands for the average cost of resource extraction, set to the moment of time t and the formula of which we will give below, then we can register combined capital costs of extracting sector that were charged off at the point of time t, – $CAPEX_3(t)$ (unit of measure – tenge of the current year) and total volume of resource production at the point of time t – $Q_3(t)$ (unit of measure – output units of the third sector):

$$CAPEX_3(t) = CAPEX_3^0(t) + p_1^0 \cdot \int_{t-T_3}^{t} p_1(x) \cdot I_3(x) \cdot b_3(t-x) \cdot z_{cp}(x) dx \quad ,$$

$$Q_3(t) = Q_3^0(t) + \int_{t-T_3}^{t} I_3(\tau) \cdot d_3(t-\tau) \cdot d\tau \, ,$$

where $Q_3^0(t)$ – is some known function that defines the agent's resource extraction and connected to the

implementation of investment projects started before the base year and requiring capital costs $CAPEX_3^0(t)$.

Then it can be put down:

$$Y_3(t) = p_3^0 \cdot Q_3(t) = p_3^0 \cdot \left(Q_3^0(t) + \int_{t-T_3}^{t} I_3(\tau) \cdot d_3(t-\tau) \cdot d\tau \right) \text{ (tenge of the base year),}$$

$$B_3(t) = B_3^0(t) + p_1^0 \cdot \int_{t-T_3}^{t} I_3(x) \cdot b_3(t-x) \cdot z_{cp}(x) dx \text{ (tenge of the base year).}$$

Agent is supposed to send the extracted resource to domestic outlet first of all and to export the remaining portion with the help of trade agents. If we mark domestic demand for the production of extracting sector as $S_3'(t)$, then we can express agent's operational profit at the point of time t, – $EBITDA_3(t)$:

$$EBITDA_3(t) = (1-n_3)\Big(\big(p_3(t) \cdot p_3^0 - tax_3^1(t) \big) \cdot \min\big(S_3'(t), Q_3(t) \big) +$$
$$+ \big(p_3^{ex}(t) \cdot p_3^{ex,0} - tax_3^1(t) - tax_3^2(t) \big) \cdot \big(Q_3(t) - S_3'(t) \big)_+ - \psi_3'(t) \cdot Q_3(t) \Big),$$

where $\psi_3'(t)$ are the agent's operational costs not including labor expenses to one output unit at the point of time t, $tax_3^1(t)$ – tax for resource extraction (tenge of the current year for the production unit of the third sector), $tax_3^2(t)$ is export tax (tenge of the current year for production unit of the third sector), n_3 is labor rate.

Then it can be put down:

operational costs of the extracting sector

$$\psi_3(t) = \left(\sum_{i=1}^{3} \big(p_i^{Int}(t) \cdot p_i^{Int,0} \cdot a_{i3}^{Int} + p_i^{\mathrm{Im}}(t) \cdot p_i^{\mathrm{Im},0} \cdot a_{i3}^{\mathrm{Im}} \big) \right) \cdot Y_3(t) \text{ (tenge of the current year)}$$

Financial result of the extracting sector

$$V_3(t) = (1-w_3)\Big(\big(p_3(t) - tax_3^1(t) / p_3^0 \big) \cdot \min\big(S_3(t), Y_3(t) \big) +$$
$$+ \Big(\big(p_3^{ex}(t) \cdot p_3^{ex,0} - tax_3^1(t) - tax_3^2(t) \big) / p_3^0 \Big) \cdot \big(Y_3(t) - S_3(t) \big)_+ - \psi_3(t) \Big) \text{ (tenge of the current year).}$$

Salary budget of the extracting sector

$$W_3(t) = \frac{w_3}{1-w_3} V_3(t) \text{ (tenge of the current year)}$$

We will assume that agent's investment activity is narrowed down to exploration of new deposits. The agent is supposed to ground the investment decisions on the analysis of the current profitability of extraction projects. By current profitability we will mean the following expression $\left(\frac{\tilde{z}(t)}{z_{\min}(t)} - 1 \right)$, where $\tilde{z}(t)$ is maximum prime cost of resource extraction which maintains investment projects profitable and which is defined in the following ratio:

$$NPV_3(z(t)) = \int_0^{T_3} \big(d(\tau) \cdot nb_3(t) - z(t) \cdot p_1(t) \cdot b(\tau) \big) \cdot e^{-\tau \cdot WACC} d\tau = 0,$$

where $WACC$ is a weighted average cost of sector's capital and $nb_3(t)$ is average "return" for the sector per

resource unit:

$nb_3(t) = V_3(t) \cdot p_3^0 \Big/ Y_3(t)$ (tenge of the current year for production unit of the third sector).

I.e. $\tilde{z}(t) = \dfrac{nb_3(t) \cdot \int_0^{T_3} d(\tau) \cdot e^{-\tau \cdot WACC} d\tau}{p_1(t) \cdot \int_0^{T_3} b(\tau) \cdot e^{-\tau \cdot WACC} d\tau}$ (tenge of the base year for the production unit of the third sector).

Thus if there is no exploration of new deposits our assumptions about the current profitability and investment activity are shaped by the following ratios:

$$I_3(t) = \hat{I}(z_{\min}(t), t) \cdot \frac{dz_{\min}(t)}{dt},$$

$$\frac{dz_{\min}(t)}{dt} = \chi \cdot \left(\frac{\tilde{z}(t)}{z_{\min}(t)} - 1 \right),$$

$$\hat{I}(z,t) = 0 \quad \text{where} \quad z < z_{\min}(t),$$

$$\hat{I}(z,t) = \rho(z) \quad \text{where} \quad z \geq z_{\min}(t),$$

where $\rho(z)$ is some function that imposes the original density of resource allocation according to extraction prime costs, χ is a constant that defines relation of the agent to projects' profitability.

Then we can define $z_{cp}(t)$ as:

$$z_{cp}(t) = \frac{z_{\min}(t) \cdot \hat{I}(z_{\min}(t), t)}{I_3(t)} \cdot \frac{dz_{\min}(t)}{dt}. \tag{3.13}$$

Now let us describe how exploration of new deposits influences the agent's operational activity. Exploration rate is supposed to be in direct ratio to agent's costs on the development of deposits with some coefficient v. Besides the exploration helps to "restore" the curve $\hat{I}(z,t)$, by adding to it new explored resources allocation of which is defined by some function $\varphi(z)$. Thus we can put down:

$$\hat{I}(z,t) = 0 \quad \text{where} \quad z < z_{\min}(t),$$

$$\frac{\partial \hat{I}(z,t)}{\partial t} = v \cdot CAPEX_3(t) \cdot \varphi(z) \quad \text{where} \quad z \geq z_{\min}(t),$$

$$\hat{I}(z,0) = \rho(z),$$

$$I_3(t) = v \cdot CAPEX_3(t) \cdot \int_0^{z_{\min}(t)} \varphi(z) dz + \hat{I}(z_{\min}(t), t) \cdot \frac{dz_{\min}(t)}{dt} \quad \text{where} \quad \frac{dz_{\min}(t)}{dt} > 0,$$

$$I_3(t) = v \cdot CAPEX_3(t) \cdot \int_0^{z_{\min}(t)} \varphi(z) dz \quad \text{where} \quad \frac{dz_{\min}(t)}{dt} \leq 0,$$

$$z_{cp}(t)=\frac{1}{I_3(t)}\cdot\left(v\cdot CAPEX_3(t)\cdot\int_0^{z_{\min}} z\cdot\varphi(z)dz+z_{\min}(t)\cdot\hat{I}(z_{\min}(t),t)\cdot\frac{dz_{\min}(t)}{dt}\right). \quad (3.14)$$

We will consider that the agent's financial activity is narrowed down to the questions of financing his/her manufacturing and investment activity and to paying dividends.

We will define financial state of extracting sector according to its account – $M_3(t)$ and to the volume of joint debt $L_3(t)$.

The agent's account is changed according to the following rule:

$$\frac{dM_3(t)}{dt}=\left(V_3(t)-r_3^dL_3(t)-T_3^p(t)\right)-\alpha_3\left(V_3(t)-r_3^dL_3(t)-T_3^p(t)\right)_+-B_3(t)p_1(t)+\frac{dL_3(t)}{dt},$$

where r_3^d is a cost of the agent's debt, α_3 is a norm of a distributed profit, $T_3^p(t)$ is a profit tax, which is defined in the following way:

$$T_3^p(t)=(1-t_3)\left(V_3(t)-r_3^dL_3(t)\right)_+,$$

where t_3 is a profit tax rate. We assume that the agent should have money at the account to ensure the payment of liability payments: $M_3(t)=\theta_3\left(V_3(t)+r_3^dL_3(t)+T_3^p(t)\right)$.

This means that account equation is an equation of agent's debt.

3.3.3 Description of Infrastructural Sector

Financial result (gross profit) of infrastructural sector is formed by means of distribution of production which has a demand in the value of $S_2(t)$ and by means of expenses connected to operational manufacturing charges $\psi_2(t)$ and costs on labor paid at the rate of w_2:

$$\begin{aligned}&V_2(t)=(1-w_2)\left(p_2(t)\min\{S_2(t),Y_2(t)\}-\psi_2(t)\right),\\&\psi_2(t)=Y_2(t)\left(\sum_{i=1}^{3}\left(p_i^{Int}(t)\cdot p_i^{Int,0}\cdot a_{i3}^{Int}+p_i^{\mathrm{Im}}(t)\cdot p_i^{\mathrm{Im},0}\cdot a_{i3}^{\mathrm{Im}}\right)\right),\\&W_2(t)=\frac{w_2}{1-w_2}V_2(t).\end{aligned} \quad (3.15)$$

Financial state of the second sector is defined by money supply of its account $M_2(t)$. Sector's account resources are refilled by means of receiving profit from primary activity and money loans. These resources are spent for investments and for payment of principal debt and percents, for dividend payment. Thus equation of the sector's account alteration is the following:

$$\frac{dM_2(t)}{dt}=\left(V_2(t)-r_2^dL_2(t)-T_2^p(t)\right)-\alpha_2\left(V_2(t)-r_2^dL_3(t)-T_2^p(t)\right)_+-B_2(t)p_1(t)+\frac{dL_2(t)}{dt}, \quad (3.16)$$

where $L_2(t)$ is the sector's joint debt, r_2^b is the interest rate on debt, $T_2^p(t)$ is profit tax, α_2 is a norm of distributed profit for the sector, $B_2(t)$ are capital investments of the sector (in real terms).

$$V_2(t)=(1-w_2)\left(p_2(t)\min\{S_2(t),Y_2(t)\}-Y_2(t)\sum_i a_{i2}p_i(t)\right) \quad (3.17)$$

Whereas

$$T_2^p(t)=(1-t_2)\left(V_2(t)-r_2^dL_2(t)\right)_+, \quad (3.18)$$

where t_2 is tax profit rate.

Process of debt servicing requires all manufacturers to maintain the certain level of the account balance:

$$M_2(t) = \theta_2 r L_2(t), \tag{3.19}$$

where θ_2 is a liquidity parameter. However manufacturers are supposed to be able only to borrow resources, i.e. $L_2(t) \geq 0$.

Here is the description of the investment activity of the second sector. We will consider that production objective of the sector is the maintenance of a certain ratio between supply and demand, i.e.

$$\frac{S_2(t)}{Y_2(t)} = const. \tag{3.20}$$

This means that output should change in accordance with the demand, dynamics of which does not depend on the sector's output:

$$\frac{dY_2(t)/dt}{Y_2(t)} = \frac{dS_2(t)/dt}{S_2(t)} = \gamma(t). \tag{3.21}$$

Taking into account that sector's output is defined by its investment activity and investments "history"

$$Y_2(t) = Y_2^0(t) + \int_{\min\{T_2^I, t\}}^{\min\{T_2, t\}} d_2(\tau) I_2(t-\tau) d\tau, \tag{3.22}$$

Then the required output growth rate is maintained by means of the involvement of the needed investments (here $d_2(\tau)$ is a "profile" according to which capacities are placed into service). Investments necessary to satisfy the increasing demand we will find in the ratio:

$$\frac{dY_2(t+T_2^I)/dt}{Y_2(t+T_2^I)} = \gamma(t)$$

If $t + T_2^I \geq T_2$, then, given that $Y_2^0(t) = 0 \quad \forall t > T_2$,

$$\frac{dY_2(t+T_2^I)}{dt} = \int_{T_2^I}^{T_2} d_2(\tau) \frac{dI_2(t+T_2^I-\tau)}{dt} d\tau = -\int_{T_2^I}^{T_2} d(\tau) \frac{dI_2(t+T_2^I-\tau)}{d\tau} d\tau =$$

$$= d_2(T_2^I) I_2(t) + \int_{T_2^I}^{T_2} \frac{dd_2(\tau)}{d\tau} I_2(t+T_2^I-\tau) d\tau \tag{3.23}$$

If $t + T_2^I < T_2$, then

$$\frac{dY_2(t+T_2^I)}{dt} = \frac{dY_2^0(t+T_2^I)}{dt} + \int_{T_2^I}^{t+T_2^I} d_2(\tau) \frac{dI_2(t+T_2^I-\tau)}{dt} d\tau + d_2(t+T_2^I) I_2(0) =$$

$$= \frac{dY_2^0(t+T_2^I)}{dt} - \int_{T_2^I}^{t+T_2^I} d_2(\tau) \frac{dI_2(t+T_2^I-\tau)}{d\tau} d\tau + d_2(t+T_2^I) I_2(0) =$$

$$= \frac{dY_2^0(t+T_2^I)}{dt} + d_2(T_2^I) I_2(t) + \int_{T_2^I}^{t+T_2^I} \frac{dd_2(\tau)}{d\tau} I_2(t+T_2^I-\tau) d\tau \tag{3.24}$$

That is why

$$\frac{dY_2(t+T_2^I)}{dt} = \frac{dY_2^0(t+T_2^I)}{dt} + d_2(T_2^I)I_2(t) + \\ + \int_{T_2^I}^{\min\{t+T_2^I, T_2\}} \frac{dd_2(\tau)}{d\tau} I_2(t+T_2^I-\tau)d\tau = \\ = \gamma(t)\left(Y_2^0(t+T_2^I) + \int_{T_2^I}^{\min\{T_2, t\}} d_2(\tau)I_2(t+T_2^I-\tau)d\tau \right) \tag{3.25}$$

$$I_2(t) = \frac{1}{d_2(T_2^I)}\left(\int_{T_2^I}^{\min\{t+T_2^I, T_2\}} I_2(t+T_2^I-\tau)\left(\gamma(t)d_2(\tau) - \frac{dd_2(\tau)}{d\tau} \right)d\tau + \right. \\ \left. + \gamma(t)Y_2^0(t+T_2^I) - \frac{dY_2^0(t+T_2^I)}{dt} \right) \tag{3.26}$$

Investment activity sets capital costs of the sector according to the following rule:

$$B_2(t) = \int_{T_2^I}^{T_2} b_2(\tau)I_2(t-\tau)d\tau \,. \tag{3.27}$$

Here $b_2(\tau)$ is a "profile" according to which costs connected to investments are made.

The sector needs to raise enough capital for financing the investment program in order to fulfill the planned investment projects. Some resources will be received as sector's undistributed profits and some capital will be borrowed. Sector's operation of undistributed profit is limited by the requirements for dividend payment.

The sector's demand for borrowed funds could be taken from the equation (3.16)by putting the expression (3.19) into it, provided fixed investment costs:

$$\frac{dL_2(t)}{dt} = \frac{1}{1-\theta_2 r_2^b}\left(p_1(t)B_2(t) - (1-t_2)\left(V_2(t) - r_2^b L_2(t)\right) + \right. \\ \left. + \alpha_2(1-t_2)\left(V_2(t) - r_2^b L_2(t)\right)_+ \right). \tag{3.28}$$

On the other hand capital supply depends on profitability of the sector's investment projects. The borrowed capital supply could be described by the following ratio:

$$\frac{dL_2(t)}{dt} = \zeta(z)p_1(t)B_2(t)\,, \tag{3.29}$$

where z is the sector's profitability evaluation, and function $\zeta(z)$ satisfies the following ratios: $0 \le \zeta(z) < 1$, $\dot{\zeta}(z) \ge 0$. Thus the market is supposed to be ready to finance some part from the required capital costs, besides the more profitability of the projects the bigger part could be financed by means of the loan.

Profitability evaluation is based at the price of investment projects given to the current point of time t, calculated according to the stream of payments for the project provided that prices on production of all brunches would not change. Thus net (i.e. except for costs) flow from payments for the project which started at the point of time t, happening at the point of time $(t+\tau)$ could be put down as

$$CF_2(t,\tau) = nb_2(t)\cdot d_2(\tau) - p_1(t)\cdot b_2(\tau)\,, \tag{3.30}$$

where $nb_2(t)$ is the return of the investment project per a production unit including operational costs, i.e.

$$nb_2(t) = (1-w_2)\left(p_2(t)\left(\min\left\{ \frac{S_2(t)}{Y_2(t)}, 1 \right\} - a_{22} \right) - \sum_{i \neq 2} a_{i2} p_i(t) \right). \quad (3.31)$$

As a matter of fact $V_2(t)/Y_2(t) = nb_2(t)$.

Provided the discount rate is r then the net cost of investment projects stated to the point of time t is defined as

$$NPV_2(r,t) = \int_0^{T_2} CF_2(t,\tau) e^{-r\tau} d\tau, \quad (3.32)$$

i.e. $NPV_2(r,t) = \int_0^{T_2} \left(nb_2(t) \cdot d_2(\tau) - p_1(t) \cdot b_2(\tau) \right) e^{-r\tau} d\tau$. Then for the evaluation of investment projects profitability we take the value

$$\hat{r} : NPV_2(\hat{r},t) = 0. \quad (3.33)$$

In order to receive the loan capital required for the fulfillment of investment program the sector should ensure a certain level of investment projects profitability, i.e. equations (3.29) and (3.28) should be integrated. Profitability conditions required for the balance of supply and demand of the borrowed funds could be defined after equating the expressions for $\frac{dL_2(t)}{dt}$ from 错误！未找到引用源。 and 错误！未找到引用源。:

$$\hat{z}(t) = \zeta^{-1}\left(\frac{p_1(t)B_2(t) - \left((1-t)(1-n_2)(1-w)V_2(t) - (1-t)r_2 L_2(t)\right)}{p_1(t)B_2(t)\left(1-\theta_2 r_2\right)} + \right.$$
$$\left. + \frac{k_N \left((1-t)(1-n_2)(1-w)V_2(t) - (1-t)r_2 L_2(t)\right)_+}{p_1(t)B_2(t)\left(1-\theta_2 r_2\right)} \right). \quad (3.34)$$

Such profitability is required for investment projects of the sector, i.e. it is necessary to follow:

$$NPV_2(\hat{z}(t),t) = 0. \quad (3.35)$$

The sector could accomplish the equation by increasing prices on its production. Target level of prices could be defined by detailing:

$$\int_0^{T_2} \left(nb_2(t) \cdot d_2(\tau) - p_1(t) \cdot b_2(\tau) \right) e^{-\hat{z}(t)\tau} d\tau = 0,$$

$$nb_2(t) = \frac{p_1(t) \int_0^{T_2} b_2(\tau) e^{-\hat{z}(t)\tau} d\tau}{\int_0^{T_2} d_2(\tau) e^{-\hat{z}(t)\tau} d\tau},$$

$$\hat{p}_2(t) = \frac{Y(t)}{S(t) - a_{22}Y(t)} \left(\sum_{i \neq 2} a_{i2} p_i(t) + \frac{p_1(t) \int_0^{T_2} b(\tau) e^{-\hat{z}(t)\tau} d\tau}{\int_0^{T_2} d(\tau) e^{-\hat{z}(t)\tau} d\tau} \right). \quad (3.36)$$

Note that level of prices $\hat{p}_2(t)$ is defined according to the assumptions that disregard the fact that $V_2(t)$

increases together with $p_2(t)$ (and vice versa). As a matter of fact when the sector needs extra borrowed funds in order to fulfill the investment program, i.e. when the right part (3.28) is positive, the evaluation $\hat{p}_2(t)$ of price level required to attract the loan capital turns out to be excessive. That is why actual development of prices for the sector's production could be put down as:

$$\frac{dp_2(t)}{dt} = \pi \cdot \left(\hat{p}_2(t) - p_2(t)\right), \tag{3.37}$$

where π is quite large parameter responsible for compliance of prices with the target level. Actual parameters of the sector $V_2(t)$ and $L_2(t)$ (from the equation (3.37)) are defined according to the price level determinate. And actual value of the sector's distributed profit is defined by the equation (3.16) and ratio (3.19)

$$DIV_2(t) = V_2(t) - (1-t)r_2 L_2(t) - T_2^p(t) - p_1(t)B_2(t) + \frac{dL_2(t)}{dt}\left(1-\theta r\right).$$

3.3.4 Description of Commercial Services Sector

Service sector is characterized by significant expenses on the maintenance of capacities and by "shadow" element of the price that sets the difference between costs on the service sector production for other economic agents and revenues coming to the sector's account.

Price on the sector's production $p_4 = \hat{p}_4 + \tilde{p}_4$, where $\tilde{p}_4$ are official rates, which is controlled by the plan in accordance with the expenses $\tilde{p}_4 = (1+\delta_4)\sum_{j=1}^{4} a_{j4} p_j$, $\hat{p}_4$ shadow element (corruptive) of the price that is defined by demand excesses Y_4 over the sector's capacity Q_4. Alteration of the "shadow" element of the price:

$$\frac{d\hat{p}_4}{dt} = \alpha \frac{Y_4 - Q_4}{Q_4}\hat{p}_4 \tag{3.38}$$

Demand for the sector's services:

$$Y_4(t) = \sum_{j=1}^{4} a_{j4} Y_j(t) + a_{4Tr}(C_1^{Int} + C_1^{Im}) + C_4(t) + G_4(t) + \hat{b}_{Tr} I_{Tr} \tag{3.39}$$

Here C_4 is household demand for the services, G_4 is demand for the government services, $\hat{b}_{Tr} I_{Tr}$ is commercial demand for the services necessary for the development of its capacity, $\sum_{j=1}^{4} a_{j4} Y_j(t) + a_{4Tr}(C_1^{Int} + C_1^{Im})$ operating demand for the services of manufacturing sectors, of commerce and importers.

Account is defined by the limits of liquidity:

$$M_4 = \theta_4 (w_4 Q_4 + \sum_{i=1}^{4} p_i a_{i4} Y_4), \tag{3.40}$$

where w_4 is official labor rate at the service sector.

Account is changed by means of

- Cash inflow for the provided services $\tilde{p}_4 Y_4$;
- Payment of financial expenses $\sum_{i=1}^{4} a_{i4} p_i Y_4$, related to rendering of services;
- Tax payments $t_4(\tilde{p}_4 - \sum_{i=1}^{4} a_{i4} p_i) Y_4$;
- Payment of salaries to the sector's employees $w_4 Q_4$;
- Payment of investments in the development of the service sector $p_1 b_4 I_4$;

- transfers $\mathrm{F}_4 = \left(-(\tilde{p}_4 - \sum_{i=1}^{4} a_{i4} p_i)\,\mathrm{Y}_4(1-\mathrm{t}_4) + \mathrm{w}_4 Q_4 \right)_+$.

Thus we come to

$$\frac{dM_4}{dt} = \left((\tilde{p}_4 - \sum_{i=1}^{4} a_{i4} p_i)\,\mathrm{Y}_4(1-\mathrm{t}_4) - \mathrm{w}_4 Q_4 \right)_+ - p_1 b_4 I_4 . \tag{3.41}$$

This equation with consideration of liquidity limits defines the volume of investments into the development of the service sector capacity $p_1 b_4 I_4$.

We will assume that capacity development is defined by the following equation

$$\frac{dQ_4}{dt} = I_4 - \mu_4 Q_4 , \tag{3.42}$$

where μ_4 is the rate of the retirement of the service sector capacities.

Revenues distributed by the household in the result of the service sector activity are the following

$$W_4(\mathrm{t}) = w_4 Q_4 + \hat{p}_4 Y_4 . \tag{3.43}$$

Taxes paid by the sector are

$$Tax_4(\mathrm{t}) = \mathrm{t}_4 (\tilde{p}_4 - \sum_{i=1}^{4} p_i a_{i4})\,\mathrm{Y}_4 \quad . \tag{3.44}$$

3.3.5 Description of the Commercial Sector

Domestic production dealer competes with the importer at the market of the first sector goods. We will assume that there is Nash equilibrium between them. The strategy of the domestic production dealer is to set a retail price $\hat{p}_1^{Int}$ to the goods of the first sector, and the importer's strategy is to set a price to imported consumer goods $\hat{p}_1^{\mathrm{Im}}$. The domestic production dealer maximizes his/her profit $\left(\hat{p}_1^{Int} - \breve{p}_1^{Int}\right) C_1^{Int}\left(\hat{p}_1^{Int}, \hat{p}_1^{\mathrm{Im}}\right)$, and importer his/her profit $\left(\hat{p}_1^{\mathrm{Im}} - \breve{p}_1^{\mathrm{Im}}\right) C_1^{\mathrm{Im}}\left(\hat{p}_1^{Int}, \hat{p}_1^{\mathrm{Im}}\right)$.

Here $\breve{p}_1^{Int} = p_1 + \sum_{i=1}^{4} a_{iTr} p_i$ is a prime cost of the production unit of the domestic production dealer, $\breve{p}_1^{\mathrm{Im}} = p^{\mathrm{Im}} + \sum_{i=1}^{4} a_{i\,\mathrm{Im}} p_i$ is an importer's prime cost, p^{Im} is a purchase value of imported goods calculated according to the exchange rate, and

$$\mathrm{C}_1^{Int}\left(\hat{p}_1^{Int}, \hat{p}_1^{\mathrm{Im}}\right) = \frac{\varphi}{\left[(\mathrm{A}_1 \left(\hat{p}_1^{Int}\right)^{\delta})^{\frac{1}{1+\delta}} + (\mathrm{A}_2 \left(\hat{p}_1^{\mathrm{Im}}\right)^{\delta})^{\frac{1}{1+\delta}} \right]} \mathrm{A}_1^{\frac{1}{1+\delta}} (\hat{p}_1^{Int})^{-\frac{1}{1+\delta}}$$

$$\mathrm{C}_1^{\mathrm{Im}}\left(\hat{p}_1^{Int}, \hat{p}_1^{\mathrm{Im}}\right) = \frac{\varphi}{\left[(\mathrm{A}_1 \left(\hat{p}_1^{Int}\right)^{\delta})^{\frac{1}{1+\delta}} + (\mathrm{A}_2 \left(\hat{p}_1^{\mathrm{Im}}\right)^{\delta})^{\frac{1}{1+\delta}} \right]} \mathrm{A}_1^{\frac{1}{1+\delta}} (\hat{p}_1^{\mathrm{Im}})^{-\frac{1}{1+\delta}}$$

according to the demand function for the imported goods depending on their prices. These functions describe the substitution of the domestic products and imported products in the structure of the household consumption assuming that they are described by the model of a representative, rational consumer who in case of budget limitation $\hat{p}_1^{Int} x + \hat{p}_1^{\mathrm{Im}} y \le \varphi,\ x \ge 0,\ y \ge 0$ maximizes the utility function

$$u(x,y)=\left[A_1x^{-\delta}+A_2y^{-\delta}\right]^{-\frac{1}{\delta}},\delta\in[-1,0], \tag{3.45}$$

where x is the consumption of the first sector domestic goods, y is the consumption of the imported goods, φ are the household expenses to the first sector goods and their imported analogs.

There is no difficulty to show that

$$C_1^{Int}(\hat{p}_1^{Int},\hat{p}_1^{Im})=\frac{\varphi}{q(\hat{p}_1^{Int},\hat{p}_1^{Im})}\frac{\partial q(\hat{p}_1^{Int},\hat{p}_1^{Im})}{\partial\hat{p}_1^{Int}},$$

$$C_1^{Im}(\hat{p}_1^{Int},\hat{p}_1^{Im})=\frac{\varphi}{q(\hat{p}_1^{Int},\hat{p}_1^{Im})}\frac{\partial q(\hat{p}_1^{Int},\hat{p}_1^{Im})}{\partial\hat{p}_1^{Im}},$$

where

$$q(\hat{p}_1^{Int},\hat{p}_1^{\mathrm{Im}})=\left[(A_1\left(\hat{p}_1^{Int}\right)^{\delta})^{\frac{1}{1+\delta}}+(A_2\left(\hat{p}_1^{\mathrm{Im}}\right)^{\delta})^{\frac{1}{1+\delta}}\right]^{1+\frac{1}{\delta}}$$

Yang conversion of the utility function $u\left(x,y\right)$. In this case the only Nash equilibrium is possible

$$\hat{p}_1^{Int}=-\frac{\breve{p}_1^{Int}}{\delta}\left((1+\delta)^{-(1+\delta)}\left(\frac{A_1\left(\breve{p}_1^{Int}\right)^{\delta}}{A_2\left(\breve{p}_1^{\mathrm{Im}}\right)^{\delta}}\right)+1\right),$$

$$\hat{p}_1^{\mathrm{Im}}=-\frac{\breve{p}_1^{\mathrm{Im}}}{\delta}\left((1+\delta)^{-(1+\delta)}\left(\frac{A_2\left(\breve{p}_1^{\mathrm{Im}}\right)^{\delta}}{A_1\left(\breve{p}_1^{Int}\right)^{\delta}}\right)+1\right).$$

3.3.6 The Commerce Sector of Domestic Goods

The commerce capacity is $Q_{Tr}(\mathrm{t})\geq \mathrm{C}_1^{Int}(\mathrm{t})$, we will consider that $Q_{Tr}(\mathrm{t})=\max_{\tau\leq t}\mathrm{C}_1^{Int}(\tau)$.

We will assume that the dealer has the capacity Q_{Tr} (for example, retail space) for the products realization. Capacities has a rate μ. The dealer injects capital $I_{Tr}(t)$ to the capacity creation. In this case the capacities will change according to equation

Equation of the capacity development:

$$\Rightarrow I_{Tr}=\frac{dQ_{Tr}}{dt}+\mu_{Tr}Q_{Tr}(\mathrm{t})$$

Take:

- $p_1^{Int}(t)$ for prices at the consumer market;
- $p_1(t)$ for prices at the products of the first sector;
- b_{Tr} for capital ratio of the capacity unit.

Loan indebtedness:

$$\frac{dL_{Tr}}{dt}=l(p_1b_{Tr}+p_4\hat{b}_{Tr})I_{Tr}, \tag{3.46}$$

Revenues distributed by the households as an activity result:

$$H_{Tr}(\mathrm{t}) = (\hat{\mathrm{p}}_1^{Int}(\mathrm{t}) - \mathrm{p}_1(\mathrm{t}) - \sum_{i=1}^{4} a_{iTr} p_i)\mathrm{C}_1^{Int}(t) - \mathrm{r}_{Tr} L_{Tr}(\mathrm{t}) - (1-l)(\mathrm{p}_1(\mathrm{t})\mathrm{b}_{Tr} + p_4 \hat{b}_{Tr}) I_{tr} - Tax_{Tr}(\mathrm{t}) \tag{3.47}$$

where $H_{Tr}(t)$ is a profit received by the dealer at the point of time t (in case if $H_{Tr}(t)$ is negative, the dealer is supposed to put up his/her capital to the business development).

$$Tax(\mathrm{t}) = t_{Tr}((\hat{\mathrm{p}}_1^{Int}(\mathrm{t}) - \mathrm{p}_1(\mathrm{t}) - \sum_{i=1}^{4} a_{iTr} p_i)\mathrm{C}_1^{Int}(\mathrm{t}) \tag{3.48}$$

3.3.7 The Commerce Sector of Imported Goods

Importer's capacity:

$$Q_{\mathrm{Im}}(\mathrm{t}) \geq \mathrm{C}_1^{\mathrm{Im}}(\mathrm{t}) \text{ , assume that } Q_{\mathrm{Im}}(\mathrm{t}) = \max_{\tau \leq t} \mathrm{C}_1^{\mathrm{Im}}(\tau) . \tag{3.49}$$

Capacity development equation:

$$\Rightarrow I_{\mathrm{Im}} = \frac{dQ_{\mathrm{Im}}}{dt} + \mu_{\mathrm{Im}} Q_{\mathrm{Im}}(\mathrm{t})$$

Loan indebtedness

$$\frac{dL_{\mathrm{Im}}}{dt} = l_{\mathrm{Im}}(p_1 b_{\mathrm{Im}} + p_4 \hat{b}_{\mathrm{Im}}) I_{\mathrm{Im}} \tag{3.50}$$

Revenues distributed by households as a result of importer's commercial activity

$$H_{\mathrm{Im}}(\mathrm{t}) = (\hat{\mathrm{p}}_1^{\mathrm{Im}}(\mathrm{t}) - \mathrm{p}^{\mathrm{Im}}(\mathrm{t}) - \sum_{i=1}^{4} a_{iTr} p_i)\mathrm{C}_1^{\mathrm{Im}}(t) - \mathrm{r}_{\mathrm{Im}} L_{\mathrm{Im}}(\mathrm{t}) - (1-l)(\mathrm{p}_1(\mathrm{t})\mathrm{b}_{\mathrm{Im}} + p_4 \hat{b}_{\mathrm{Im}}) I_{\mathrm{Im}} - Tax_{\mathrm{Im}}(\mathrm{t}) \tag{3.51}$$

$$Tax_{\mathrm{Im}}(\mathrm{t}) = t_{\mathrm{Im}}((\hat{\mathrm{p}}_1^{\mathrm{Im}}(\mathrm{t}) - \mathrm{p}^{\mathrm{Im}}(\mathrm{t}) - \sum_{i=1}^{4} a_{iTr} p_i)\mathrm{C}_1^{\mathrm{Im}}(\mathrm{t}) \tag{3.52}$$

3.3.8 Banking System

In this model the banking system is represented by two agents: National Bank and system of commercial banks.

Commercial bank

Thus we can put down the consolidated balance of commercial banks as:

$$\begin{aligned} &\mathrm{Re}\, s(\mathrm{t}) + D^H(\mathrm{t}) + \sum_{i=1}^{4} M_i(\mathrm{t}) + M_G(t) + \mathrm{L}_{CB}(\mathrm{t}) + E_B(\mathrm{t}) = \\ &M_B(\mathrm{t}) + \sum_{i=1}^{4} L_i(\mathrm{t}) + L_{Tr}(\mathrm{t}) + \mathrm{L}_{\mathrm{Im}}(\mathrm{t}) + L^H(\mathrm{t}) + L_G(t) + F_{CB}(\mathrm{t}) + \mathrm{D}_{CB}(\mathrm{t}) \end{aligned} \tag{3.53}$$

Evaluate the net balance of account deposits and credits of commercial banks in the national bank from the consolidated balance

$$\begin{aligned} &D_{CB} - \mathrm{L}_{CB} = \mathrm{Re}\, s(\mathrm{t}) + D^H(\mathrm{t}) + \sum_{i=1}^{4} M_i(\mathrm{t}) + M_G(t) + E_B(\mathrm{t}) - \\ &-(M_B(\mathrm{t}) + \sum_{i=1}^{4} L_i(\mathrm{t}) + L_{Tr}(\mathrm{t}) + \mathrm{L}_{\mathrm{Im}}(\mathrm{t}) + L^H(\mathrm{t}) + L_G(t) + F_{CB}(\mathrm{t})) \end{aligned} \tag{3.54}$$

This model supposes that activity of commercial banks system is limited by the following standards:

1. generation of mandatory reserve funds follows the requirements

$F_{CB} = \delta_B^F D^H(\mathrm{t})$, where δ_B^F is a standard of the population deposits mandatory reserve

2. limitation (standard) of commercial banks liquidity that require the liabilities are provided by liquid assets $\dfrac{M_B(\mathrm{t})}{D^H(\mathrm{t}))_+ + \sum_{i=1}^{4} M_i(\mathrm{t}) + M_G(\mathrm{t})} = \theta_B^1$

3. standard of capital adequacy $\dfrac{E_B(\mathrm{t})}{D^H(\mathrm{t}))_+ + \sum_{i=1}^{4} M_i(\mathrm{t})} = \theta_B^2$

4. creation of reserves for the impaired loans supply $\mathrm{Re}\,s(t) = h_1 L_1(\mathrm{t}) + h_H L^H(\mathrm{t})$, where h_1 is the percent of impaired loans to the first sector, and h_H is the percent of impaired loans to households

Using the described limits we get the following expression for net balance from (1)

$$D_{CB} - \mathrm{L}_{CB} = (1-\theta_B^1 - \delta_B^F + \theta_B^2)D^H + (1-\theta_B^1+\theta_B^2)\sum_{i=1}^{4} M_i(\mathrm{t}) + (1-\theta_B^1)\,\mathrm{M}_G - \\ -\left(\sum_{i=2}^{4} L_i(\mathrm{t}) + L_G(t) + L_{Tr}(\mathrm{t}) + \mathrm{L}_{\mathrm{lm}}(\mathrm{t}) + (1-h_1)\,\mathrm{L}_1) + (1-h_H)\,\mathrm{L}_H\right) \tag{3.55}$$

Financial result (change of net worth) of the commercial banks sector is reached by means of incomes from interest of commerce banks assets, liability service charges, profit tax payment at the rate t_B, salary payments, payment of dividends to employees and to owner of commercial banks $\mathrm{H}_B(\mathrm{t})$ and changes of reserves for impaired loans maintenance $\dfrac{d\,\mathrm{Re}\,s(\mathrm{t})}{dt}$.

Incomes from the interest of commercial banks asserts are formed by the following elements:

- incomes from consumer loans of the population $r_b^H(1-\mathrm{h}_H)L^H(t)$, r_b^H is the loan interest rate
- incomes from manufacturing sector loans $r_B^1(1-h_1)L_1(\mathrm{t})$, $r_B^i L_i(\mathrm{t})\ (i=2,3,4)$, where r_B^i is the interest rate on the i sector of production;
- incomes from trade loans $\mathrm{r}_{Tr}\,L_{Tr}(\mathrm{t})$, where r_{Tr} is the interest rate on trade loans;
- incomes from loans of an importer $\mathrm{r}_{\mathrm{lm}}\,L_{\mathrm{lm}}(\mathrm{t})$, where r_{lm} is the interest rate on importer loans;
- incomes from loans to the government $r_G L_G(t)$, where r_G is an interest rate on the loans to the government;
- incomes from commercial banks deposits in the National Bank $r_{CB}\,\mathrm{D}_{CB}(\mathrm{t})$, where r_{CB} is an interest rate on deposits in the National Bank.

Expenses on liabilities are formed by the following elements:

- payments of interests on the population deposit $\hat{r}_b^H D^H(t)$, where $\hat{r}_b^H$ is an interest rate on the population deposits;
- payments of interest on loans at the NB $\hat{r}_{cb}L_{cb}(t)$, where $\hat{r}_{cb}$ is an interest rate on loans at the NB.

Thus the net worth could be described by the equation:

$$\frac{dE_B(\mathrm{t})}{dt} = (1-t_B)\left(r_B^1(1-h_1)L_1(\mathrm{t}) + \sum_{i=2}^{4} r_B^i L_i(\mathrm{t}) + r_G L_G(t) + (1-h_H)\,\mathrm{r}_B^H\,L^H(\mathrm{t}) + r_{CB}\,\mathrm{D}_{CB}(\mathrm{t})\right) - \\ -(1-t_B)\left(\hat{r}_B^H D^H(\mathrm{t}) + \hat{r}_{CB}L_{CB}(\mathrm{t}) + \mathrm{r}_{Tr}\,L_{Tr}(\mathrm{t}) + \mathrm{r}_{\mathrm{lm}}\,L_{\mathrm{lm}}(\mathrm{t})\right) - \mathrm{H}_B(\mathrm{t}) - \frac{d\,\mathrm{Re}\,s(\mathrm{t})}{dt} \tag{3.56}$$

This model assumes that the interest rate on interbank loan is less than the interest rate on the National Bank loan and is bigger than the interest rate on commercial banks deposits in the National Bank. This assumption leaves us consider only net balance of deposits and loans in the National Bank. Then we will suppose that commercial banks system is either a loan debtor of the National Bank and then $L_{CB} > 0, D_{CB} = 0$, or credit provider of the National Bank and then $L_{CB} = 0, D_{CB} > 0$. Put it down

$$\tilde{r} = \begin{cases} r_{CB}, \ if \quad D_{CB} > 0, L_{CB} = 0, \\ \hat{r}_{CB}, \ if \quad D_{CB} = 0, L_{CB} > 0. \end{cases} \tag{3.57}$$

After inserting the expression for net balance of deposits and loans of commercial banks in the National Bank and the expression for the impaired loans service reserve into (3.40), we get

$$\frac{dE_B(\mathrm{t})}{dt} = (1-t_B)\left((1-h_1)\left(r_B^1 - \tilde{r}\right) L_1(r_B^1) + \sum_{i=2}^{4} \left(r_B^i - \tilde{r}\right) L_i(r_B^i) + \left(r_G - \tilde{r}\right) L_G(t) + (1-h_H)\left(\mathrm{r}_B^H - \tilde{r}\right) L^H(\mathrm{r}_B^H) \right) +$$
$$+ (1-t_B)\left(\left(\mathrm{r}_{Tr} - \tilde{r}\right) L_{Tr}(\mathrm{r}_{Tr}) + \left(\mathrm{r}_{\mathrm{Im}} - \tilde{r}\right) L_{\mathrm{Im}}(\mathrm{r}_{\mathrm{Im}}) + \left(\left(1 - \theta_{CB}^1 - \delta_B^F + \theta_{CB}^2\right)\tilde{r} - \hat{r}_B^H \right) D^H(\hat{r}_B^H) \right) +$$
$$\left(1 - \theta_{CB}^1 - \delta_B^F + \theta_{CB}^2\right)\tilde{r} \sum_{i=1}^{4} M_i + \left(1 - \theta_{CB}^1\right) M_G - \mathrm{H}_B(\mathrm{t}) - \frac{d\,\mathrm{Re}\,s(\mathrm{t})}{dt}.$$

Commercial banks maximize their incomes from operations at markets of loans and deposits by setting the interest rates. Note that the expression for the change of commercial banks net worth shows that the choice of interest rates is defined by the completion of independent tasks of revenues optimization

- from crediting the first sector manufacturers $(1-h_1)\left(r_B^1 - \tilde{r}\right) L_1(r_B^1)$ by setting the interest rate r_B^1;
- from crediting the manufacturers of i sector $\left(r_B^i - \tilde{r}\right) L_i(r_B^i)$ by setting the interest rate $r_B^i \ (i = 2,3,4)$;
- from crediting households $(1-h_H)\left(r_B^H - \hat{r}\right) L^H(r_B^H)$ by setting the interest rate r_B^H;
- from crediting trade enterprises $\left(\mathrm{r}_{Tr} - \tilde{r}\right) L_{Tr}(\mathrm{r}_{Tr})$ by setting the interest rate r_{Tr};
- from crediting the importers $\left(\mathrm{r}_{\mathrm{Im}} - \tilde{r}\right) L_{\mathrm{Im}}(\mathrm{r}_{\mathrm{Im}})$ by setting the interest rate r_{Im};
- from loans at the market of household deposits $\left(\left(1 - \theta_{CB}^1 - \delta_B^F + \theta_{CB}^2\right)\tilde{r} - \hat{r}_B^H \right) D^H(\hat{r}_B^H)$ by setting the interest rate $\hat{\mathrm{r}}_B^H$.

The interest rate $\breve{r} = Arg \max_r (r - \tilde{r}) L(r)$ meets the first order requirements

$$(\breve{r} - \tilde{r}) \frac{dL(\breve{r})}{dr} + L(\breve{r}) = 0.$$

Whence it follows that

$$\breve{r} = \tilde{r} - \left(L(\breve{r}) \Big/ \frac{dL(\breve{r})}{dr} \right).$$

Requirements for the maximum of the second order are

$$(\breve{r}-\tilde{r})\frac{d^2L(\breve{r})}{dr^2}+2\frac{dL(\breve{r})}{dr}<0.$$

Assume that the interest rate set by commercial banks changes according to the equation

$$\frac{dr}{dt}=\theta\left(\tilde{r}-\left(L(r)\Big/\frac{dL(r)}{dr}\right)-r\right). \tag{3.58}$$

Note that in case of the steady demand for loans and fixed lending rate is $\tilde{r}$ the interest rate $r=\breve{r}$ is an asymptotically stable equilibrium according to Lyapunov.

Thus we get the following equations for changes of interest rates on loans

$$\frac{dr_B^i}{dt}=\theta_{B,i}\left(\tilde{r}-\left(L\left(r_B^i\right)\Big/\frac{dL\left(r_B^i\right)}{dr}\right)-r_B^i\right)\quad (i=1,2,3,4),$$

$$\frac{dr_{Tr}}{dt}=\theta_{B,Tr}\left(\tilde{r}-\left(L\left(r_{Tr}\right)\Big/\frac{dL\left(r_{Tr}\right)}{dr}\right)-r_{Tr}\right),$$

$$\frac{dr_{\mathrm{Im}}}{dt}=\theta_{\mathrm{B,Im}}\left(\tilde{r}-\left(L\left(r_{\mathrm{Im}}\right)\Big/\frac{dL\left(r_{\mathrm{Im}}\right)}{dr}\right)-r_{\mathrm{Im}}\right).$$

Interest rate $\ddot{r}=Arg\max_r\left(\left(1-\theta_{CB}^1-\delta_B^F+\theta_{CB}^2\right)\tilde{r}-r\right)D^H(\mathrm{r})$ meets the requirements of the first order

$$\left(\left(1-\theta_{CB}^1-\delta_B^F+\theta_{CB}^2\right)\tilde{r}-\ddot{r}\right)\frac{dD^H(\ddot{\mathrm{r}})}{dr}-D^H(\ddot{r})=0.$$

Whence it follows that

$$\ddot{r}=\left(1-\theta_{CB}^1-\delta_B^F+\theta_{CB}^2\right)\tilde{r}-\frac{D^H(\ddot{r})}{\frac{dD^H(\ddot{\mathrm{r}})}{dr}}.$$

Requirements for the maximum of the second order are

$$\left(\left(1-\theta_{CB}^1-\delta_B^F+\theta_{CB}^2\right)\tilde{r}-\ddot{r}\right)\frac{d^2D^H(\ddot{\mathrm{r}})}{dr^2}-2\frac{dD^H(\ddot{r})}{dr}<0.$$

Assume that the interest rate set by commercial banks changes according to the equation

$$\frac{d\hat{r}_B^H}{dt}=\theta_{B,H}\left(\left(1-\theta_{CB}^1-\delta_B^F+\theta_{CB}^2\right)\tilde{r}-\frac{D^H\left(\hat{r}_B^H\right)}{\frac{dD^H\left(\hat{r}_B^H\right)}{dr}}-\hat{r}_B^H\right). \tag{3.59}$$

Note that in case of the steady supply of deposits and fixed lending rate $\tilde{r}$ the interest rate $r_B^H=\ddot{r}$ is an asymptotically stable equilibrium according to Lyapunov.

National Bank

Here is the description of the National Bank (NB) balance.

The NB liabilities are:

- Monetary base, issued by the NB $M(t)$. Monetary base is formed from account balances of all economic agents and from cash money.
- The NB liabilities (loan stocks of the Bank of Kazakhstan, commercial banks deposits in the NB)
- Mandatory reserve funds founded by commercial banks, $F_{cb}(t)$
- The NB reserves $\mathrm{Re}\, z_{CB}$

The NB assets are:

- The NB loans to commercial banks, $L_{cb}(t)$
- Available assets formed by the NB $A_{cb}(t)$
- Gold and forex reserves *GFR (t)*

The NB balance ratio could be put down like:

$$L_{CB} + GFR(\mathrm{t})\varphi + \mathrm{A}_{CB}(\mathrm{t}) = M_H(\mathrm{t}) + M_B(t) + \mathrm{D}_{CB}(\mathrm{t}) + \mathrm{F}_{CB}(\mathrm{t}) + \mathrm{Re}\, z_{CB} \tag{3.60}$$

The NB currency issue is supposed to happen automatically, i.e. monetary base is formed according to the agents' activity in the economy. Mandatory reserve funds are based on loans provided by commercial banks. Net-liabilities of commercial banks to the NB $\mathrm{D}_{CB}(1-\theta_{CB}) - L_{CB}$ are based on the consolidated balance of commercial banks. Available assets are proportional to deposits:

$$\mathrm{A}_{CB}(\mathrm{t}) = \theta_{CB} D_{CB}(\mathrm{t}) \quad \textbf{(5.6)}$$

$$GFR(\mathrm{t})\varphi = \mathrm{M}_H(\mathrm{t}) + M_B(t) + (\mathrm{D}_{CB}(1-\theta_{CB}) - L_{CB}) + \mathrm{F}_{CB}(\mathrm{t}) \tag{3.61}$$

$$\mathrm{Re}\, z_{CB} = \hat{r}_{CB} L_{CB} - r_{CB} D_{CB} + \frac{d\varphi}{dt} GFR \tag{3.62}$$

3.3.9 Population

Description of the Population Consumer Behavior

Population as an economic agent consumes the production of all brunches. Besides the first sector production competes with similar imported goods at the retail market of nonfood goods. Let's make an assumption that the population demand for the first sector goods is distributed between domestic and imported nonfood goods according to the maximization of the utility function $u(C_1^{Im}, C_1^{Int})$, where C_1^{Int} is population demand for domestic nonfood product of the first sector in natural units, C_1^{Int} is a demand for imported nonfood product in natural units. The demand for imported product is supposed to be met by importers in the full scale.

The population demand for domestic and imported nonfood goods results from the following expression

$$u(C_1^{Im}, C_1^{Int}) \to \max$$

Financial limits

$$p_1^{Im} C_1^{Im} + p_1^{Int} C_1^{Int} \leqslant \phi^H - p_2^H C_2 - p_3^H C_3 - p_4^H C_4 = \phi_1 .$$

where C_i is the population consumption of the i sector production, p_i^H is a price for production of i sector set for the population.

$$C_1^{Im} \geqslant 0, C_1^{Int} \geqslant 0$$

The size of population incomes:

$$\phi^H = W_G + \sum_{i=1}^{4} W_i(1-t_C) + \hat{r}_b^H D^H - r_b^H L^H + \tilde{p}_4 Y_4 + H_{Tr} \tag{3.63}$$

where W_G is the size of government welfare payments, W_i is size of salary payments in i sector, t_C is the income tax rate, $\hat{r}_b^H D^H$ are revenues from bank deposits, $r_b^H L^H$ are expenses for bank loans, $\tilde{p}_4 Y_4$ – "shadow" revenues from the forth sector, H_{Tr} are revenues/expenses of the retailers.

Volumes of loans and deposits of the population are defined by the model of the population saving behavior.

Take CES function as a utility function:

$$u(x,y) = \left[A_1 x^{-\delta} + A_2 y^{-\delta} \right]^{-\frac{1}{\delta}}, \delta \in [-1,0]$$

Solving a problem:

$$\phi_1 = q(p_1^{Im}, p_1^{Int}) u(C_1^{Im}, C_1^{Int}) \tag{3.64}$$

where

$$q(x,y) = \left[(A_1 x^{\delta})^{\frac{1}{1+\delta}} + (A_2 y^{\delta})^{\frac{1}{1+\delta}} \right]^{1+\frac{1}{\delta}}$$

$$C_1^{Im} = \frac{\phi_1}{q(p_1^{Im}, p_1^{Int})} \frac{\partial q}{\partial p_1^{Im}} \tag{3.65}$$

$$C_1^{Int} = \frac{\phi_1}{q(p_1^{Im}, p_1^{Int})} \frac{\partial q}{\partial p_1^{Int}} \tag{3.66}$$

Description of the Population Saving Behavior

Households interact with commercial banks at the market of loans and deposits. The population has opportunity to raise borrowed assets at the interest rate $r_L = r_b^H$ and to offer its own assets at the rate $r_D = \hat{r}_b^H$.

In oredr to provide the regularity of consumer expenditures $C = p_1^{Im} C_1^{Im} + p_1^{Int} C_1^{Int} + p_2^H C_2 + p_3^H C_3 + p_4^H C_4$ the household needs to have a reserve of funds $M(t) \geq \theta C(t)$. Households are granted loans so that at the final point of time they would have been able to pay the loan, i.e. $L(T) = 0$.

Assume that household forecasts exponential salary growth $S = W_G + \sum_{i=1}^{4} W_i(1-t_C) + \tilde{p}_4 Y_4 + H_{Tr}$ at the rate γ, i.e. $S(t) = Se^{\gamma t}$.

The population welfare at the point of time t is characterized by the value $x = D - L + M$. Alteration of this value is defined by its revenues received as a salary $S(t)$ and payments $D(t)r_D$ of deposits $D(t)$, and by expenses that consist of consumer expenses $C(t)$ expenses of loan obligations $L(t)r_L$. Assume that the population can immediately redistribute its money from asset of one form to another. In this case the household welfare changes according to the equation

$$\frac{dx}{dt} = Se^{\gamma t} - C + Dr_D - r_L L. \tag{3.67}$$

By redistributing its funds the household tends to maximize the benefit of consumtion discounted in time:

$$\int_0^T C(t)^\alpha e^{-\Delta t} dt \to \max .$$

Thus the distribution of the population funds is described by optimal control problem

$$J = \int_0^T (C(t))^\alpha e^{-\Delta t} dt \to \max$$
$$\dot{x} = Se^{\gamma t} - C + Dr_D - r_L L,$$
$$x = D - L + M,\ M \geq \theta C,\ D, L, M \geq 0,$$
$$x(0) = x_0\ x(T) \geq 0.$$

Note that

$$x_1 = \frac{S\theta(\Delta - \alpha r_L)}{(r_L - \gamma)(1-\alpha)(1+r_L\theta)},\quad \hat{x} = \frac{S\theta}{\gamma\theta+1},\quad \lambda(\xi) = \frac{S}{\xi - \gamma},\ \kappa(\xi) = \frac{(1+\xi\theta)(1-\alpha)}{\theta(\Delta - \alpha\xi)},\quad \rho = \frac{\Delta + \gamma(1-\alpha) + 1/\theta}{\gamma + 1/\theta},$$

$$g(x,y,\xi) = \left[\frac{x-\hat{x}}{y-\hat{x}}\right]^{-\rho}\left(\frac{\gamma\theta+1}{\xi\theta+1}y^{\alpha-1} + \int_y^x z^{\alpha-1}\left[\frac{z-\hat{x}}{y-\hat{x}}\right]^{\rho}\frac{dz}{z-\hat{x}}\right),$$

$$h(x,\tau,\xi) = \left(\lambda(\xi) + x - x\kappa(\xi)\right)e^{(\gamma-\xi)\tau} + x\kappa(\xi)e^{((\Delta-\xi)/(1-\alpha)+\gamma)\tau} - \lambda(\xi),\ b(x,\xi) = \frac{1+\xi\theta}{1+\gamma\theta}\int_0^1 \left(px + (1-p)\hat{x}\right)^{\alpha-1} p^{\frac{\theta(\Delta-\alpha\gamma)}{\gamma\theta+1}} dp .$$

The problem will be solved in the following way:

1. $\Delta - (1-\alpha)/\theta < r_D < r_L < \Delta + \gamma(1-\alpha)$

$$M(x_0) = \begin{cases} \left(x_0 + \dfrac{S}{r_L - \gamma}\right)\dfrac{\Delta - \alpha r_L}{(1-\alpha)(1+r_L\theta)}, & x_0 < x_1, \\ x_0, & x_0 \in [x_1, x_2], \\ x_2 e^{\left(\frac{\Delta - r_D}{1-\alpha} + \gamma\right)\tau_2}, & x_0 > x_2, \end{cases}$$

2. $\Delta - (1-\alpha)/\theta < r_D < \Delta + \gamma(1-\alpha) < r_L$

$$M(x_0) = \begin{cases} x_3 e^{\left(\frac{\Delta - r_L}{1-\alpha} + \gamma\right)\tau_3}, & x_0 < x_3, \\ x_0, & x_0 \in [x_3, x_4], \\ x_4 e^{\left(\frac{\Delta - r_D}{1-\alpha} + \gamma\right)\tau_4}, & x_0 > x_4, \end{cases}$$

Here x_2 is the solution of the equation $g(x_2, x_1, r_L) = \dfrac{\gamma\theta+1}{r_D\theta+1}x_2^{\alpha-1}$, x_3 is the solution of the equation $b(x_3, r_L) = x_3^{\alpha-1}$, x_4 is the solution of the equation $b(x_4, r_D) = x_4^{\alpha-1}$, τ_2 is the solution of the equation $h(x_2, \tau_2, r_D) = x_0$, τ_3 is the solution of the equation $h(x_3, \tau_3, r_L) = x_0$, τ_4 is the solution of the equation $h(x_4, \tau_4, r_D) = x_0$.

Loans are determined as $L(x_0) = (M(x_0) - x_0)_+$, deposites are determined as $D(x_0) = (x_0 - M(x_0))_+$, consumption is determined as $C(x_0) = M(x_0)/\theta$. Volume of loans is positive if $x < x_1$ or $x < x_3$, volume of deposits is positive if $x > x_2$ or $x > x_4$.

Volumes of loans ($L^H = L$) and deposits ($D^H = D$) of the population are used in the problem of the population

consumer behavior and in bank balances, cash money on hands of the population ($M_H = M$) is a part of the NB balance monetary base.

Household has two types of assets – cash money ■ and savings $[0, t_1]$. Household allocates available funds to the current consumption $r_L < \Delta + (1-\alpha)\gamma$, interest payments $[t_1, T]$ on the existing loans $\varphi \le \alpha x^{\alpha-1} / (r_D + 1/\theta)$. Household revenues are composed of salary $\varphi(T) = 0$ and dividends from the deposits $\varphi_{us}(x_0) = \frac{\alpha\theta}{\gamma\theta+1} \int_0^1 \left(px_0 + (1-p)\frac{S\theta}{\gamma\theta+1} \right)^{\alpha-1} p^{\frac{\theta(\Delta-\alpha\gamma)}{\gamma\theta+1}} dp$. Values $\varphi = \alpha x^{\alpha-1} / (r_D + 1/\theta)$ and $x = x_4$ stand for interest rate on deposits and loans correspondingly. Assume that household forecasts exponential growth of the salary $x < x_4$ with the rate $x > x_4$, i.e. $\sigma > 0$. In order to provide regularity of consumer costs dV / dx it is necessary to have monetary reserve $\frac{dV}{dx} = \alpha \int_0^\infty \left(\frac{S\theta}{\gamma\theta+1} + \left(y - \frac{S\theta}{\gamma\theta+1} \right) e^{-(\gamma+1/\theta)t} \right)^{\alpha-1} e^{-(\Delta+1/\theta+(1-\alpha)\gamma)t} dt$. Let's take value t_2 for the liquidity ratio.

The household maximizes profitability of the further consumption discounted at the coefficient $x_0 > x_{r_D}$.

Take $x_0 \ge x_2$, $t_2 = 0$, $M_{rL} = (x + w_{rL}) / \kappa(rL)$, $x(t)$, $[0, t_2]$, $x(t)$, $x < \tilde{x}$.

Statement 1. *Assume that requirements for the solution existence are met and* $a = (1 + r_L\theta)(1-\alpha) / (\theta(\Delta - \alpha r_L))$.

Then the equation $x_1 = \frac{S\theta(\Delta - \alpha r_L)}{(r_L - \gamma)(1 - \alpha + \theta(r_L - \Delta))}$, $\tau_1(x) = \frac{\theta}{\gamma\theta+1} \ln \left(\left(x - \frac{S\theta}{\gamma\theta+1} \right) / \left(x_1 - \frac{S\theta}{\gamma\theta+1} \right) \right)$. *has a positive solution* $x_1 > S\theta / (\gamma\theta + 1)$. *Besides for any* $\Delta - (1-\alpha)/\theta < r_L < \Delta + (1-\alpha)\gamma$, *more than* $\Delta - \alpha\gamma > 0$, *there is a solution* $r_L > \gamma$ *of the equation* $\tau_1(x)$.

Statement.2. *Assume that requirements for the solution existence are met and* $r_L > \gamma$. *Then the equation* $\forall$ *has a solution. Besides if* $x_0 \ge 0$ *is a solution where* T, t_1 *is a solution where* $x_0 \le x_1$ *u* $t_1 = 0$, *then* $x_0 > x_1$. *Specify the solution* $\forall$ *if* $\xi = r_L$ *as* x_3, *if* $\xi = r_D$ *as* x_4.

Statement 3. *Assume that requirements for the solution existence are met and* $[t_1, T]$. *Then* $x(T) = 0$, *there is a solution* $\hat{\varphi}(x)$ *for the equation* $x \le \tilde{x} = S / \left((r_L - \Delta) / (1-\alpha) + 1/\theta \right)$.

Statement 4. *Assume that requirements for the solution existence are met and* x. *Then* $\tilde{x}$, *there is a solution* $\dot{x}|_{x=\tilde{x}} < 0$ *for the equation* $x_0 > \tilde{x}$.

<u>Identification:</u>

Saving households:

$$T \le \ln(M_4(T_4) / x_0) / ((r_L - \Delta) / (1-\alpha) - \gamma) \tag{3.68}$$

4. Discussion

Today one of the most actively discussed national problems is the rise of Kazakhstan economy investment attrectiveness. Moreover the investments are required not only for the development of depressed manufacturing brunches but also for the relatively successful power-generating sectors. Oil complex, for example, needs investments to maintain the level of hydrocarbon production, electric energy sector has an urgent problem of renovating the old capacities.

In order to make a reasonable investment model of any econmic sector it is necessary first of all to learn how to evaluate the primary parameter that defines the investment activity or in other words the profitability of investments. Problem of evaluation of profitability and efficiency of investments is one of the major problems in the sphere of corporate finances. Works by I. Fisher (3,4) are considered to be the founders of scientific

discussion of the problem. Then works by D. Hirshleifer (5), R.M. Solow (6), D. Gale (7), R. Dorfman (8) continued the discussions on this issue. Two indicators were generaly offered for the profitability evaluation; that is NPV – net present value of cashflows connected to the project and IRR – internal rate of return at which project's NPV reduces to 0. It is necessary to mention that there were arguments pro and contra every indicator. Moreover the authors often reasoned their opinion by notorious "easiness" of usage in practice.

D.G. Cantor and S.A. Lipman in (9) have offered the statement of the problem which gave them opportunity to justify the usage of the indicator (and a unified pattern of its calculation) IRR for the evaluation of investment projects profitability and to show its connection to NPV. It is interesting that the authors haven't justified the privilege of the first indicator. In (9) there is a formular of the investor's objective that posesses common premises. This objective gives opportunity to control the speed of investment projects implementation in an attempt to receive a maximum final income. Investment rate at the widening investment horizon was accepted as the investment profitability evaluation. It turned out that it is possible to calculate this value explicitly for a wide range of investment projects and that this value corresponds to IRR. Thus the profitability evaluation was formed in a natural way as a result of an analysis of the investor's initial objective completion.

Cantor-Lipman approach turned out to be fruitful. Statements of problems offered by D.G. Cantor and S.A. Lipman are discussed in works by I.M. Sonin and E.L. Presman (10,11). (10) reasons the results of (9) by the instrument of dynamic programming. In (11) there is an interpretation of the investment polyniminal positive roots that are distinct from maximum (connected to IRR). In (12) D.G. Cantor and S.A. Lipman expand their approach to the case with several projects and reason the profitability indicator for the whole pool of investment projects. In works (13, 14) by B.Z. Belenkiy the investment activity is described as a part of linear model of Neiman-Gale and the problem of the investment projects profitability evaluation is narrowed down to the question of existance of a mainline in the model of Neiman-Gale. In (15) the model of Cantor-Lipman was expanded: they considered a situation where there is a subjective (from investor's point of view) chance on fading of the demand on investments and the investor is not allowed to take short positions. In such statement the problem was brought to the analysis of Bellman equation. In (16) there were found the ratios of model's parameters at which investor's effective behavior is a cautious strategy that excludes bankruptsy. In (17) the authors created low-bound evaluations of the capital increase rate of the investor who uses a cautious investment strategy.

Works by D.G. Cantor and S.A. Lipman and works that followed their ideas were oriented on the analysis of investments in conditions of discrete time and for that reason they dealt with investment projects which payment flows were equally distributed according to arrival time. This are quite restrictive conditions for modelling actual investment projects in applied models. Payment flows often arrive at nonmeasurable intervals. Intuitive approach, at which a uniform structure (with sufficient number of zero payment flows) approximated such payment structure, leads to a serious reduction of a time step. Thus additional difficulties in calculation appear, because it is necessary to reduce a time step in the whole model. That is why a transition to continious time is natural. Thus we will get an opportunity to use the Cantor-Lipman approach for the analysis of investments in macroeconomical models that are often formed in continious time for ease of analysis. For example the modelling of investment activity of real economic sector is a topical problem for systemic approach to modelling Kazakhstan economy (16, 17).

5. Conclusion

The strategy of mathematical modeling of the Republic of Kazakhstan was developed. Statistics of the basic economic indicators was collected and studied. The economic model was created and the procedure for settling its course was implemented. A number of analytical and prediction calculations was made.

The developed model is one of the most difficult types of mathematical models of economy; it is a model of intertemporal economic balance. This means that economic dynamics is described as a result of interaction between the stated economic agents. The most significant scientific achievement of this work is the fact that this approach provides a quite satisfactory description of economic processes even under the conditions of a crisis.

Analysis of the developed model of Kazakhstan economy has proved that in order to find a tool for medium-term and long-term forecasting it is necessary to make new clusters of the model that would take into account long-term aspects of economic agents' behavior.

That is the reason why we have begun to study the development of separate clusters of the model which to our opinion are essential for the aims of medium-term and long-term forecasting. On the basis of the research we have singled out three key problems of systemic analysis of Kazakhstan economy in long-term perspective:

- Problem of production modeling in Kazakh industries with low competitiveness in view of scarcity of working capital;
- Problem of description of investment activity in fuel and energy complex;
- Problem of medium layer formation in the process of modeling the population saving behavior.

The developed model could be used to forecast the changes of macroeconomic indicators and to study the way they are influenced by economic policy.

References

Belenkiy, V. (2002). *Economic dynamics: analysis of capital projects as a part of linear model of Neiman-Gale.* Moscow: TEMI RAN, preprint 137.

Belenkiy, V. (2007). *Optimization model of economic dynamics. Conceptual framework. One-dimensional models. Bellman's approach.* Moscow: Science.

Bikkinina, L., & Shananin, A. (n.d.). To the theory of the investment projects profitability in the context of imperfect financial market. *XLVI MIPT conference*, 136-137.

Cantor, D., & Lipman, S. (1983). Investment selection with imperfect capital markets. *Econometrics, 4*(51), 1121-1144.

Cantor, D., & Lipman, S. (1995). Optimal Investment Selection with a Multitude of Projects. *Econometrics, 5*(63), 1231-1240.

Dorfman, R. (1981). The meaning of internal rates of return. *J. of Finance, 5*(36), 1011-1021.

Fisher, I. (1907). *The rate of interest.* New York: Macmillan Co.

Fisher, I. (1930). *The theory of interest*. New York: Macmillan Co.

Gale, D. (1973). On the theory of interest. *The American mathematical monthly, 8*(80), 853-868.

Hirshleifer, J. (1958). On the theory of optimal decision. *J. of political economy, 66*, 229-239.

Petrov, A., Pospelov, I., & Shananin, A. (1996). *Experience of economy mathematic modeling.* Moscow: Energoatomizdat.

Petrov, A., Pospelov, I., & Shananin, A. (1999). *From State plan to non-efficient market. Mathematical analysis of Russian economic structures.* New York, Lewinston: The Edwin Mellen Press.

Presman, E., & Sonin, I. (2000). *Growth rate, internal rates of return and financial bubbles*. Moscow: CEMI Russian Academy of Sciences, preprint 103.

Solow, R. (1963). *Capital theory and the rate of return.* Amsterdam: North Holland Press.

Sonin, I. (1995). Growth rate, internal rates of return and turn pikes in an investment model. *Economic theory, 5*, 383-400.

Vaschenko, M. (2006). Study of Bellman equation in the problem of optimal investment. *Cl. of articles of young scientists of CM&C faculty of MSU, Moscow, 3*, 32-43.

Vaschenko, M. (2009). Evaluation of investment projects profitability in uncertain conditions. *Mathematical modeling, 3*(21), 18-30.

Website of the Agency of the Republic of Kazakhstan for Regulation and Supervision of Financial Market and Institution. (n.d.). Retrieved January 25, 2014, from http://www.afn.kz/.

Website of the National Bank of the Republic of Kazakhstan. (n.d.). Retrieved January 25, 2014, from http://www.nationalbank.kz/.

Website of the Statistical Agency of the Republic of Kazakhstan. (n.d.). Retrieved January 25, 2014, from http://www.stat.kz/.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The Information Cost Estimation as Realization of the Problem of Indistinct Mathematical Programming

Lyudmila Nikolaevna Rodionova[1], Olga Genadievna Kantor[2], Rodionov Anton Sergeevith[3] & Rukhliada Nataliia Olegovna[4]

[1] Doctor of Science, Economics, Russian Federation

[2] Cand. Sc., Physics and Mathematics, Russian Federation

[3] Ufa State Aviation Technical University, Russian Federation

[4] Cand. Sc., St.Petersburg State Polytechnical University, Russian Federation

Correspondence: Lyudmila Nikolaevna Rodionova, Doctor of Science, Economics. E-mail: rodion@ufanet.ru

Received: January 9, 2015 Accepted: January 29, 2015 Online Published: July 15, 2015

doi:10.5539/mas.v9n8p186 URL: http://dx.doi.org/10.5539/mas.v9n3p186

Abstract

All information in the market has the consumer value. It is possible to estimate the recommended cost of investments in chance of reception of exact future result on the basis of calculation of consumer cost of the full information, and it is defined as a difference between expected values of a choice at presence or absence of the full information.

The problem of definition of cost of the information a priori is connected with necessity of the account of variety of the uncertain factors characterizing, firstly, its reliability, and, secondly its utility. It is possible to estimate reliability or utility of any information on the basis of the available statistical data, or expertly. Certainly, statistical methods allow to receive plausible enough estimations, however, by no means always the researcher possesses necessary statistical base that speaks or absence of the admission to it for privacy reasons, or exclusiveness of the information. The formalization of a problem of definition of information costing as a problem of indistinct mathematical programming, proceeding from purposes of the person, making decision (PMD), and on the basis of processing of the expert data is carried out in given work

Keywords: cost information, fuzzy mathematical programming

1. Introduction

The swift growth of rates of scientific and technical revolution and computerization of all fields of activity of a society has led to allocation of some concepts, seeming ordinary several years ago. The information and derived concepts generated by it become such phenomena occupying more and more significant place in our everyday life. The gain of information quantity circulating today in the world, has transformed it from a minor resource into the factor, fatefully influencing practically on all spheres of a public life, reflecting thereby increasing information dependence of a society.

Uncertainty is the characteristic peculiar to the majority of corporate administrative decisions. Usually deal with uncertainty of time and resources in realization of this or that purpose. One of the most difficult case for the analysis is the uncertainty connected with efficiency and costs at development of stocks or working out of a new kind of production when new technologies are introduced in manufacture.

In this work the problem is not to present the exhaustive list of all kinds of the uncertainty accompanying administrative decisions. The single aim is to underline that fact that the key problem lying before managers of the company is not to operate with uncertainty, but to explain it accurately at all stages of formation of decisions.

In the management theory it is supposed that the relevant information or, at least, the accessible information is given to the manager before he starts decision-making. But such situation appears extremely seldom in practice. One of the most important components of decision-making process is knowledge of what questions should be set. For example, the doctor cannot count that results of all possible trouble-shooting tests and interrogations will be given him by then as the patient will enter for the first time into his office.

The financier makes decisions when results are not defined and based on the limited information. There is often

a choice of acceptance of strategies before him that is- the minimum profit at the minimum risks or the maximum profit at the maximum risks. For risk decrease the presence of the additional information which costs money is necessary for the inancier as a result of decision-making.

The concept of cost of the information is not defined in any Russian law. Consequently its calculation is not defined also. Cost of the state information is defined by assignment of a security label for the information of different level. It is possible to guess its cost indirectly on privacy degree. The private person or a business firm should define cost of that information which they would like to get or sell. In large firms there are security services, own or involved which estimate information cost. They are effective or not, it depends on their experience. Thus, as a rule, design procedure of cost of the information is not disclosed. All it allows to estimate efficiency of purchase and sale of the information rather conditionally.

The problem of definition of cost of the information a priori is connected with necessity of the account of variety of the uncertain factors characterizing, firstly, its reliability, and, secondly its utility. It is possible to estimate reliability or utility of any information on the basis of the available statistical data, or expertly. Certainly, statistical methods allow to receive plausible enough estimations, however, by no means always the researcher possesses necessary statistical base that speaks or absence of the admission to it for privacy reasons, or exclusiveness of the information. In such situation expert estimations remain hardly probable not the unique way, allowing to define information cost. Thus it is necessary to have in view that any expert estimations, being subjective reflexion of preferences, comprise uncertainty which is necessary to PMD for taking into consideration.

2. Methods

Let's available m various types of information $I_1,...,I_m$. Required cost of each type of information we will designate $x_1,...,x_m$ accordingly. For definiteness we will believe that the target information is estimated only. We will input following assumptions: we will consider that each information possesses

utility for PMD $\alpha_i(x_i)$ which depends on cost of the information, (the given assumption is quite logical, considering that fact that if for the most valuable information unreasonably high price will be appointed, utility of such information will repeatedly decrease);

certain subjective reliability β_i (in understanding PMD) which depends on degree of confidence PMD in utility of the information

Table 1. The Designations used for formalization of a problem of definition of cost of the information

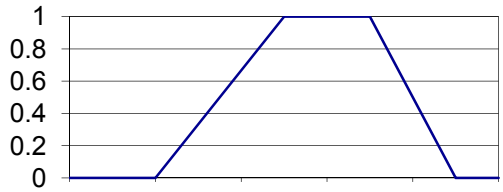
Type of information	Information cost	Subjective reliability of the information	Utility of the information
I_1	x_1	β_1	$\alpha_1(x_1)$
I_2	x_2	β_2	$\alpha_2(x_2)$
I_m	x_m	β_m	$\alpha_m(x_m)$

To set dependence of utility of the information i of the kind on its cost it is equivalent to definition of function of utility. Certainly, it is the independent problem, which methods of the decision are well enough studied. Concerning functions of utility $\alpha_i(x_i)$, $i=\overline{1,m}$ we will assume the following:

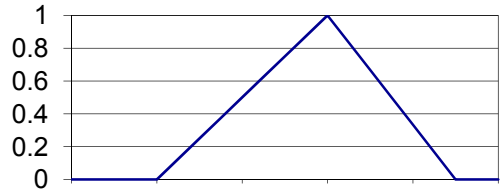
- All functions of utility $\alpha_i(x_i)$, $i=\overline{1,m}$ change in a range from 0 to 1 (1 − the maximum utility, 0 − minimum);
- Each function $\alpha_i(x_i)$ accepts the maximum value in some vicinity of point x_i^* (or only in point x_i^*), and in process of removal from it utility falls to the minimum mark (the given assumption corresponds to a situation when PMD on the basis of the information available for it defines "fair" in its understanding the price of the information i of th kind to deviate from which in the big party it is unprofitable for financial reasons, and in the smaller party - it is dangerous, in view of increase of risk of reception of the unfair information from "the unsatisfied" seller or possibility to lose the seller at all).

Let's notice that the more PMD it is assured of "the fair" price defined by it, the there should be a reduction of values of function of utility $\alpha_i(x_i)$ in process of removal from point x_i^* and on the contrary, than there is more than doubt PMD concerning "the fair" price faster, function $\alpha_i(x_i)$ in a vicinity of point x_i^* should be especially flat.

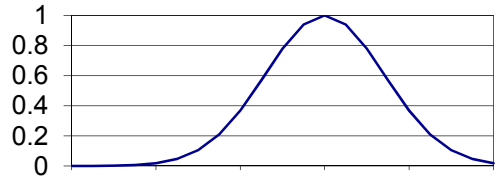
Possible variants of graphic interpretations of functions of utility $\alpha_i(x_i)$ are more low resulted.



a) trapezoid function of utility



(b) triangular function of utility



Gauss utility of function.

Figure 1. Graphic interpretations of functions of utility

In our opinion, the greatest interest is represented by Gauss functions of the utility which analytical task looks like

$$\alpha(x) = e^{-\frac{(x-a)^2}{2\sigma^2}} \quad (1)$$

It is caused by that parameters a and σ such functions suppose simple enough interpretation: parameter a represents the most probable value (population mean) of cost of information x^*, and parameter σ - expected disorder of values of cost of the information from its most probable value x^* (mean-square deviation). Both of these parameters can be defined on the basis of processing of the available statistical data, and on the basis of expert opinions. We will notice that on the basis of considered in a course of mathematical statistics of factor of a variation, on size $V = \sigma / a$ it is possible to judge Gauss "steepness" of a curve (than more factor of a variation, especially flat it is). It is obvious that the similar statement is fair and with reference to function of utility of a kind (1) (Figure 2 see).

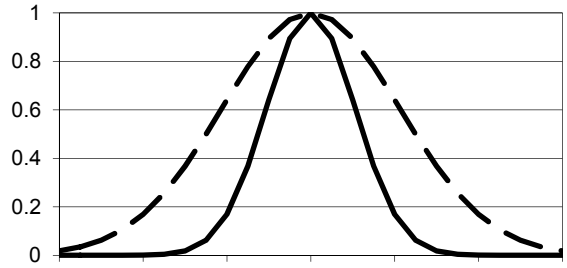


Figure 2. A type of Gauss function of utility at various values of factors of a variation

Varying parameters a and σ, PMD can correct formalization of the representation about "the fair" price. And these representations should be considered at calculation of an estimation of reliability of the information: than more PMD it is assured of "the fair" price defined by it (i.e. the less size $V = \sigma / a$), the there should be in understanding PMD a reliability of information β, and on the contrary more. Thus, size $V = \sigma / a$ can be taken as a principle definitions of subjective reliability of the information. In the present work β it is offered to count as follows:

$$\beta(V) = e^{-\left(\frac{\sigma}{a}\right)^2} \tag{2}$$

Graphic representation of function $\beta(V)$ is represented on Figure 3.

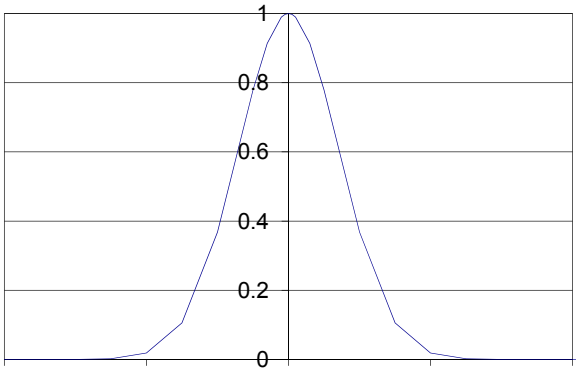


Figure 3. Graphic interpretation of function of reliability $\beta(V)$

Thus, having formalized the function of utility (having set parameters a and σ), PMD can automatically estimate and subjective reliability of the information.

Lets name the size $\beta_i \alpha_i(x_i)$ the *corrected utility* i of the information. Then size

$$\sum_{i=1}^{m} \beta_i \alpha_i(x_i) \tag{3}$$

The size will define the *total corrected utility* of all set of the got information, and

$$\frac{1}{m} \sum_{i=1}^{m} \beta_i \alpha_i(x_i) \tag{4}$$

The *average total corrected utility* of the got information. We will notice that in the total corrected utility of the information (3) unlike simple total utility $\sum_{i=1}^{m}\alpha_i(x_i)$ the factors of uncertainty inherent in process of definition PMD of the "fair" prices are considered.

3. Results

Using the entered sizes and the approach described above, it is possible to formulate or a problem of definition of cost of the got information minimizing total expenses, under condition of achievement of the set level of the average total corrected utility of the got information, or a problem of definition of cost of the got information, maximizing average total corrected utility of the got information, in the conditions of limitation of means for information acquisition.

In the first case the model will look like:

$$\sum_{i=1}^{m} x_i \to \min \tag{5}$$

$$\frac{1}{m}\sum_{i=1}^{m}\beta_i\alpha_i(x_i) \geq \gamma \tag{6}$$

$$x_i \geq 0, \quad i=\overline{1,m}, \tag{7}$$

Where γ is the set level of the average total corrected utility of the got information. We will notice that size $\frac{1}{m}\sum_{i=1}^{m}\beta_i\alpha_i(x_i)$ accepts values in a range from 0 (the worst value) to 1 (the best value). Therefore level γ has simple enough and evident interpretation: γ corresponds to the minimum comprehensible share from the best value of the average total corrected utility of the got information.

For the second of the formulated problems the model will register in a kind:

$$\frac{1}{m}\sum_{i=1}^{m}\beta_i\alpha_i(x_i) \to \max \tag{8}$$

$$\sum_{i=1}^{m} x_i \leq S \tag{9}$$

$$x_i \geq 0, \quad i=\overline{1,m} \tag{10}$$

Where *S* is the size characterizing the top border of the sum which PMD assumes to spend for information acquisition.

4. Discussion

Let's stop on a problem of an estimation *of uncertainty* of the available information. The availability of the factor of uncertainty is caused by that PMD can not possess 100 percent confidence at definition of functions of utility of the information that is in turn expressed in the task of parameters α and σ in a random way. Further it leads to uncertainty, both at an estimation of subjective reliability of the information, and at definition of its cost. Certainly, uncertainty degree varies for various situations.For practice it is important to have possibility to make a numerical estimation of uncertainty in each situation. As such measure we will use entropy which is equal to average uncertainty of all possible outcomes:

$$H(x) = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{11}$$

Where *x* is one of independent events with *n* possible conditions Pi - probability *i* of the condition. Thus size $\log_2\frac{1}{p_i}$ is called as private entropy and characterizes only i condition. For practice entropy is important that allows comparing uncertainty of various experiences to casual outcomes.

In our case event x results from definition PMD of parameters of all functions of utility $\alpha_i(x_i)$. Each of such events we will consider independent from similar others. The given assumption is fair, if PMD makes decisions, being based only on the knowledge, instead of on results of the last ways of definition of functions of utility.

Event x leads to definition of sizes of subjective reliability of information β_i, $i=\overline{1,m}$. Owing to a way of their calculation the given sizes can be interpreted as probability of approach of corresponding events $A_i =$ «i th information is authentic in understanding PMD», and sizes $(1-\beta_i)$ − as probabilities of approach of opposite events $\overline{A_i} =$ «i th information is not authentic in understanding PMD». Thus, following 2^m of variants of combinations of events is possible:

$$A_1, \overline{A_2}, \overline{A_3}, \ldots, \overline{A_m},$$

$$\overline{A_1}, A_2, \overline{A_3}, \ldots, \overline{A_m},$$

… … ….

$$A_1, A_2, \overline{A_3}, \ldots, \overline{A_m},$$

… ….

$$\overline{A_1}, \overline{A_2}, \overline{A_3}, \ldots, \overline{A_m}, \tag{12}$$

Which probabilities are accordingly equal to

$$\beta_1(1-\beta_2)(1-\beta_3)\ldots(1-\beta_m),$$

$$(1-\beta_1)\beta_2(1-\beta_3)\ldots(1-\beta_m),$$

… … ….

$$\beta_1\beta_2(1-\beta_3)\ldots(1-\beta_m),$$

… ….

$$(1-\beta_1)(1-\beta_2)(1-\beta_3)\ldots(1-\beta_m). \tag{13}$$

Events (12) form full group, therefore the entropy generated by event x, will pay off under the formula:

$$H(x) = -\sum_{\substack{i_1=\overline{0,1} \\ \ldots\ldots \\ i_m=\overline{0,1}}} (1-\beta_1)^{1-i_1}\beta_1^{i_1}\ldots(1-\beta_m)^{1-i_m}\beta_m^{i_m}\log_2\left[(1-\beta_1)^{1-i_1}\beta_1^{i_1}\ldots(1-\beta_m)^{1-i_m}\beta_m^{i_m}\right]. \tag{14}$$

It is known that entropy is maximum, if all outcomes are equiprobable. In case of described above the approach to entropy calculation its maximum value is equal

$$H_{\max} = -\sum_{i=1}^{2^m}\frac{1}{2^m}\log_2\frac{1}{2^m} = 2^m\frac{1}{2^m}\log_2 2^m = m \tag{15}$$

On the basis of the received value of expression (14) it is possible to judge degree of uncertainty of the available information formalized according to representations PMD. Thus it is possible to compare various sets of the information among themselves (more uncertain will be considered at what value of entropy above) or to compare

entropy of set of information $H(x)$ to the maximum value m. Also during research by the useful there can be an analysis of values of private entropies.

$$\log_2 \frac{1}{(1-\beta_1)^{1-i_1}\beta_1^{i_1}\ldots(1-\beta_m)^{1-i_m}\beta_m^{i_m}},$$

$$i_1 = \overline{0,1},\ \ldots,\ i_m = \overline{0,1}. \tag{16}$$

To help PMD to understand the assumptions put forward by it and to facilitate a problem of definition of exact restrictions, transformation of problems of mathematical programming (5) - (7) and (8) - (10) in problems of indistinct mathematical programming is expedient. In problems of indistinct mathematical programming probably to consider situations when about some parameters of model that they are in some set reflecting real (or desirable) possibilities is known only, and also situations in which instead of maximization (minimization) of criterion function it is enough that it reached only some set level, thus to various deviations of criterion function from the given level various degrees of an admissibility can be attributed.

Let's admit that PMD, solving a problem (5) - (7), it is not assured of correctness of unequivocal definition subjective reliable β_i, $i = \overline{1,m}$, and considers that more correct will consider their range of possible values, and concerning the average total corrected utility of got information γ can assert only that $\gamma \in \left[\gamma^{\min}, \gamma^{\max}\right]$. In this case it is possible to speak about the indistinct task of parameters β_i, $i = \overline{1,m}$ and γ, and, as consequence, about statement of a problem *of indistinct mathematical programming* (5) - (7).

Let's set a range of possible values for β_i as follows:

$$\beta_i \in \left[\beta_i^{\min}; \beta_i^{\max}\right],$$

(17)

Where $\beta_i^{\min} = \beta_i(1-\delta)$, $\beta_i^{\max} = \min\{\beta_i(1+\delta); 1\}$, $\delta \in [0;1]$ - is a constant, setting which, PMD can vary width of corridors for possible values β_i, $i = \overline{1,m}$. It is obvious that the more δ, the there is less than confidence at PMD concerning correctness β_i.

Owing to definition of functions of utility $\alpha_i(x_i)$, $i = \overline{1,m}$ a problem (5) - (7) with indistinct parameters β_i, $i = \overline{1,m}$ and γ can lead the problem with accurately certain parameters as follows:

$$\sum_{i=1}^{m} x_i \to \min \tag{18}$$

$$\frac{1}{m}\sum_{i=1}^{m} \beta_i^{\min}\alpha_i(x_i) \geq \gamma^{\min} \tag{19}$$

$$\frac{1}{m}\sum_{i=1}^{m} \beta_i^{\max}\alpha_i(x_i) \leq \gamma^{\max} \tag{20}$$

$$x_i \geq 0,\ i = \overline{1,m}. \tag{21}$$

As a result of the decision of a problem (18) - (21) at preset value δ the vector of optimum cost of information

$X^* = \{x_1^*;...;x_m^*\}$ and optimum value of criterion function $f^* = \sum_{i=1}^{m} x_i^*$ will be defined. Number $1-\delta$ can be considered as accessory and vector X^* degree to indistinct set of decisions, and simultaneously, as degree of accessory f^* to indistinct set of optimum values of criterion function of an initial problem. We will notice that if $\delta = 0$ the problem (18) - (21) is not indistinct, and, hence, decision X^* is unequivocally optimum, and f^* - minimum of all possible values of criterion function. Varying possible values δ, it is possible to receive functions of an accessory of the indistinct decision and indistinct value of criterion function [1,2,15,16].

Let's assume that, solving a problem (8) - (11), PMD do not aspire to maximize the average total corrected utility, and to aspire to that the given size has reached some comprehensible level γ^*, and let there is threshold level γ^0, less which criterion function (8) should not accept value. Then function of an accessory to indistinct function of the purpose can be defined as follows:

$$\mu_{f_0}(X) = \begin{cases} 0, & if \sum_{i=1}^{m} x_i \le \gamma^0, \\ \mu_\gamma(X), & if\ \gamma^0 \le \sum_{i=1}^{m} x_i < \gamma^*, \\ 1, & if \sum_{i=1}^{m} x_i \ge \gamma^*. \end{cases} \quad (22)$$

Where $\mu_\gamma(X)$ - the function of an accessory describing degrees of performance of the corresponding inequality from the point of view of PMD.

5. Conclusion

Function of accessory $\mu_S(X)$ for indistinct restriction (9) can be similarly defined. Indistinct conditions can be considered as set of alternatives X together with its indistinct subsets representing indistinctly formulated criteria (purpose f_0 and restriction f_1), i.e. as system (X, f_0, f_1). As a result the initial problem (8) - (11) can be formulated in the form of a performance problem of indistinctly definite purpose to which we will apply the approach of Bellman - Zade according to which is to take into consideration whenever possible all criteria in such problem means to construct function into which the purposes and restrictions are included equally [18,19,20]:

$$D = f_0 \bigcap f_1 . \quad (23)$$

The decision can be defined as an indistinct subset of universal set of possible values. The optimum corresponds to that area which elements maximize *D*. Thus, and the problem (8) - (11) can be shown to a problem of indistinct mathematical programming.

References

Altunin, A. E. (2010). Semukhin M.V Models and algorithms of decision making under uncertainty. Tyumen: TSU Publishing House, 352.

Berndt E. R. (2005). Econometrics practice. – M.: Yuniti-Dana.

Derevyanko, P. M. (2003). Elements of fuzzy logics in investment portfolio formation. Economics and infocomms in XXI century: Proceedings of the II International scientific and practical conference. November 24-29. SPb.: SPSUE Publishing House, 317-319.

Derevyanko, P. M. (2005). Comparison of indeterminate and simulation approach to modeling of enterprise activity under uncertainty. Modern problems of economics and management of national economy. *Collection of scientific articles. Issue. SPb.: SPSUE*, 289-292.

Derevyanko, P. M. (2005). Fuzzy modeling of enterprise activity and evaluation of risks of making strategic financial decision under uncertainty. Modern problems of applied informatics: I scientific and practical conference, May 23-25. Book of reports - SPb.: SPSUE, 81-83.

Doherty K. (1997). Introduction in econometrics. -M.: INFRA-M. XIV, 402 p.: ill. - (University book) Literature

reference: 384-386.

Dubois, D., & Prade, H. (1990). Possibility Theory. Applications to presentation of knowledge on informatics. Translated from French into Russian. – M. Radio and communication. 288.

Eliseeva, I. I. (2001). Econometrics: Textbook /I.I. Eliseeva et al., M.: Finances and Statistics.

Kahraman, C., Ruan, D., & Tolga, E. (2002). Capital Budgeting Techniques Using Discounted Fuzzy versus Probabilistic Cash Flows. *Information Sciences, 142*, 57-76.

Kleiner, G. B. (2001). Economical and mathematical modeling and economic theory. *Economics and Mathematical Methods, 37*(3).

Kleiner, G. B., Tambovtsev, V. L., & Kachalov, R. M. (1997). Enterprise in unstable economical environment: risks, strategy, safety. M.: Economics.

Knyazevskiy, V. S., & Zhitnikov, I. V. (2014). Analysis of time series and prediction: Learning guide. Rostov-on-Don: RSEA, 161.

Kofman, A., & Hil, A. (2012). Introduction to the theory of fuzzy sets in the management of enterprises: Per. with App. - Mn .: Higher School, 224 p.

Saati T. L. (2010). Decision making. Hierarchy Analysis Method. — M.: Radio and communication, 316.

Saati, T. L. (2008). Decision making under dependences and feedback: Analytical networks. — M.: Publishing House LKI, 360.

Sobol, I. M., & Statnikov, R. B. (1981). A choice of optimum parameters in problems with many criterias. M: The Science.

Tsarev, V. V. (2014). Evaluation of economical efficiency of investments.

Voschinin, A. P. (2014). Problems of analysis with ambiguous data – intervals and/or randomness? Interval mathematics and constraint propagation: Business meetings, 147-158.

Zade, L. (1974). The basis of the new approach to the analysis of difficult systems and decision-making processes. Mathematics today: transl. from English - M: Znanie, 5-48.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Systems of Lineaments of Magnetic and Gravity Anomalies in the Zone of Convergent Interaction of the Amur and the Eurasian Tectonic Plates

Sergey Vladimirovich Trofimenko[1,2], Nikolay Nikolaevich Grib[2], Aleksandr Ivanovich Melnikov[3] & Tatiana Vladimirovna Merkulova[1]

[1] Institute of Tectonics and Geophysics, Far Eastern Branch, Russian Academy of Sciences, Russian Federation

[2] Technical Institute (branch) North-Eastern Federal University, Russian Federation

[3] Institute of the Earth's crust, Siberian Branch, Russian Academy of Sciences, Russian Federation

Correspondence: Sergey Vladimirovich Trofimenko, Institute of Tectonics and Geophysics, Far Eastern Branch, Russian Academy of Sciences, 680000, Russian Federation.

Received: January 9, 2015 Accepted: March 20, 2015 Online Published: June 25, 2015

doi:10.5539/mas.v9n8p195 URL: http://dx.doi.org/10.5539/mas.v9n3p195

Abstract

We studied the spatial distribution of anomalies of the gravitational and the magnetic fields at the border between the Amur and the Eurasian plates. Methods of statistical analysis showed that the systems of anomalies of geophysical fields fit into the regular spatial structures and are controlled by lineaments of latitude, longitude and diagonal stretch, as well as zones of plastic-elastic flow of rocks (shear-zone). These facts allowed us to establish time of formation of the orthogonal system of geophysical fields anomalies in the strike azimuths (84-354^0). As a result of comparing diagonal system of geophysical fields anomalies with spatial distribution of the fields of contemporary seismicity and epicenters of strong earthquakes, an assumption was made that the diagonal system of lineaments fit into zones of tectonic deformation and has been controlling the epicentral fields of contemporary seismicity since the Mesozoic period of seismotectonic activation. The main result of this study is the synthesis of early scientific works and getting new ideas about the spatial distribution of anomalies of geophysical fields. Establishing the sequence of forming isolated systems of anomalies and relationships between geophysical fields and seismicity is the subject of further research.

Keywords: geophysical fields, lineaments, shear-zones, activation, crustal deformation, seismicity, earthquakes, prediction

1. Introduction

In course of geological study of the structure of the crust, the analysis of spatial distribution of geophysical fields anomalies is essential. Problems of tectonic zoning of spatial models of gravitational and magnetic fields tend to reflect major tectonic dislocations and the position of the contact zones of individual crustal blocks in the form of grouping and consistency in certain azimuths of extended linear zones - lineaments (Alaa A. Masoud and Katsuaki, 2011; Yuanyuan Li et al., 2011).

Start of theoretical studies of the lineament structures and large-scale irregularities of the earth's crust has been given in works (Hobbs, 1904; Hubbert, 1937), which are being successfully developed at present with the use of new technologies for analyzing digital models, such as geophysical fields and landforms (Anokhin, Odessa, 2001; Arellano-Baeza et al., 2006; Anokhin, Maslow, 2009; Gilmanova et al., 2012; Loiane et al., 2014). New technologies make it possible not only to track the date of lineaments formation, and, hence, geophysical fields in geological time scales (Loiane et al.,2014), but, also to track changes in the structure of lineaments due to the influence of strong earthquakes (Arellano-Baeza et al., 2006).

Study of the tectonic structure and tectonic nature of geophysical fields in the Northeast of Asia within the Aldan-Stanovoy megablock (Figure 1) have laid the foundation of work (Kazan, 1965; Grishkyan, 1968; Malyshev, 1977). Subsequent synthesis of structural, geophysical and geological data from works (Stativa et al., 2006A; Stogniy et al., 1996; Popov, Smelov, 1996;Trofimenko, 2010) revealed that in this region lineaments and the anomalies of geophysical fields that reflected them formed multidirectional orthogonal systems with

north-eastern and north-west spread (Stativa et al., 2006 B). These lineaments on the surface are recorded in the form of zones up to 15 km thick and up to several hundred kilometers length. They are predominantly spread to the north-east (45-75°) and to the northwest (300-320°), vertical or near-vertical planes placement of planes and control the position of granitisation fields.

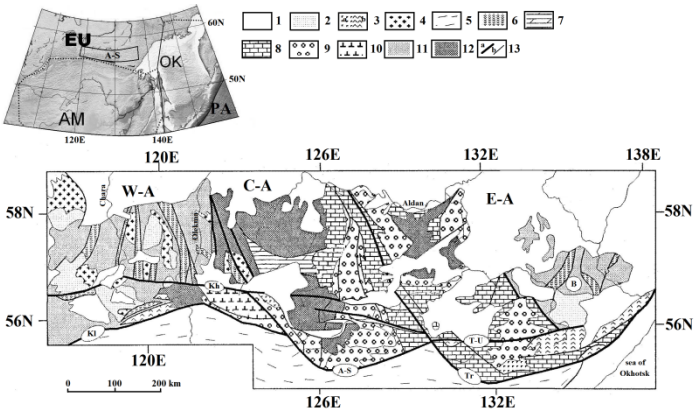


Figure 1. Block diagram of the Aldanian Shield structure and lineaments of the Aldanian-Stanovoy megablock

Designations. Overview map: EU, PA - Eurasian and Pacific plates; AM, OK - Amur and Okhotsk microplate; AS - Aldanian-Stanovoy megablock.

Main diagram. Aldanian shield blocks: WA - West Aldan; C-A - Central Aldan; E-A - East Aldan; B- Batomgsk microblock.

Numbers stand for: 1 - Riphean-Phanerozoic formations; 2 - complexes of proterozoic sediments Pt1; 3 - gabbroids (a) and anorthosite (b), Pt1; 4 - granitoids, Ar2- Pt1; 5 -Stanovoy complex of the Proterozoic age; 6 - Greenstone belts; 7 - amphibole gneissic series and gabbro-plagiogranitoids; 8 - orthoamphibolite, twopyroxene plaghio gneissic, and garnet-plaghio gneissic strata; 9 - hypersthene-gneissic, garnet-gneissic and two-pyroxene-gneissic strata; 10 - granulite-basite formation;11 - tonalite-throndhemite gneisses; 12 - granite, and charnokite and enderbite gneisses; 13 - tectonic boundaries: (a) - regional, (b) - sutures. Kl - Kalaro-Stanovaya; Kh - Khaninskaya; A-S - Amgino-Stanovaya; Tr - Tyrkandin-Dzhugdzhur; T-U - Timpton-Uchursk.

Many researchers indicate facts proving that systems of faults were formed in the Archean age and subsequently underwent repeated activation in later periods. For example, faults in the diagonal direction, according to views of authors (Kazansky, 1965; Grishkyan, 1968) were formed during the late Archean. Nevertheless, so far the problem of determining the periods of fault systems formation relate to controversial issues.

During geological development of the earth's crust, in periods of both slow tectonic and fast seismotectonic activations, anomalies of geophysical fields experience changes as a reflection of physical properties of rocks. Most intensively the processes of reformation of the structural plan of spatial anomalies should occur in seismically active regions in the areas of interaction of tectonic plates and lineaments that divide them (Heleno C. et al., 2014). Due to this fact, there is a problem of determining the spatiotemporal relationships between systems of faults with different courses and, consequently, of determining periods of formation of geophysical fields anomalies.

Thus, studying spatial structure of geophysical fields is one of scientific areas of research in the various applied problems of structural geology, tectonics, metallogeny, geodynamics and seismotectonics. Research data get a special role in seismically active zones (Joseph and William, 2002).

Preliminary data about geometric formations of lineament structures in the Aldan-Stanovoy megablock using methods of statistical analysis of geophysical fields spatial distribution are described in works (Stativa et al., 2006A; Stativa et al., 2006B; Stativa and Trofimenko, 2006; Trofimenko, 2010).

This work sets tasks of clarifying periods of formation of extended zones of geophysical anomalies by comparing these systems with spatial position of zones of elastic-viscous rocks flow (shear-zones) (Melnikov, 2008), and the tasks of finding patterns in distribution of spatial field of seismicity within the zone of interaction between the Amur and the Eurasian lithospheric plates.

2. Results of Statistical Modeling of the Geophysical Fields Spatial Structure

To study lineament structures of the Aldan-Stanovoy megablock for distribution of geophysical fields, the method of statistical analysis of the spatial distribution of crustal faults geophysical features characteristically manifested in gravitational and magnetic fields have been used. Compared with the initial maps of the gravitational and magnetic fields, diagrams of their linear elements have the advantage that they only retain orientation, length and location of fields characteristics, reflecting linear tectonic and stratigraphic contacts of rock complexes with various physical properties.

A detailed analysis of statistical distributions of azimuths of gravitational and magnetic fields anomalies showed that (Figure 2), in the graphs appear statistically significant local maxima, of which we can make a pair of mutually orthogonal systems: (3-273^0), (26-296^0) (35-305^0 *) (40-310^0) (64-334^0) (70-340^0) (84-354^0). The error in determining the preferred direction of maxima ranged from ±3^0 to ±5^0.

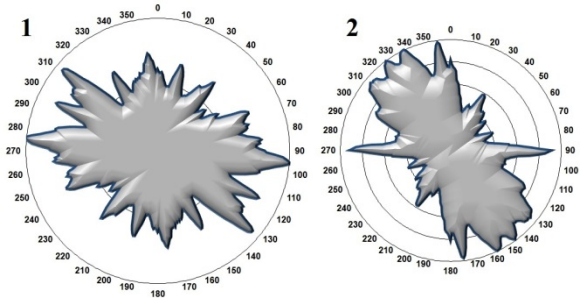


Figure 2. Statistics of distributing spreads of azimuths of anomalies of linear elements of the gravitational (1) and magnetic (2) fields

Figure 3 shows spatial position of the lineament structures built on statistical distributions of the anomalies of geophysical fields in azimuths (3-2730) and (84-3540).

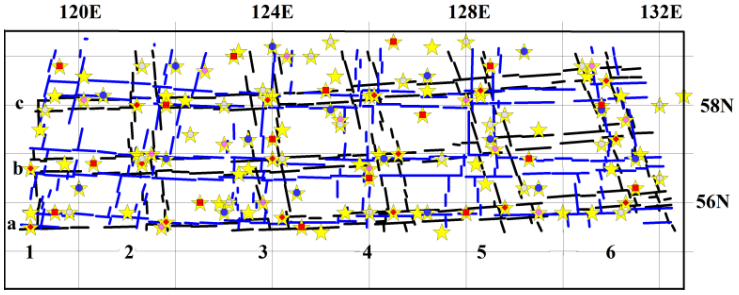


Figure 3. Spatial position of domains of system (3-2730) ±30 (blue) and a similar system (84 and 354 °) ±30 (black) of the Aldan-Stanovoy megablock

Designations. Asterisks are intersections of the lineament structures of all selected systems; digits (1-5) are numbers of meridional zones from the west to the east; (a-c) are denominations of latitudinal zones from the north to the south:

It may be noted that the structures of lineaments with azimuths (3-273^0) ± 3^0 and azimuths (84 and 354 °) ±3^0 have sides with meridional length of the domain about 120 km and latitudinal width of 120 to 180 km. In whole, in the studied territory, zones of mutually orthogonal lineaments are sustained by their spread. However, in the southern and the south-western part of the territory, starting with latitude 570, either a displacement of center lines of selected meridional zones, or partial or complete destruction of the lineaments is observed.

Similar geometric constructions made it possible to establish that system (35-305^0*) ±3^0 is similar to system (26-296^0). The average length of domain side is 150 km. In system (40-310^0) ±3^0, domain boundaries appear more regularly with a spatial period of 120 km. The inner area of domains is filled with indicators, predominantly of the north-west spread (310^0). Domains with the largest number of indicators form a regular network of mutually embedded domains of the second order. System (64-334^0) ±5^0 and a system that is similar to it (70-340^0) ±5^0 form a regular network on the whole territory with the spatial period of 120-140 km (Trofimenko, 2010).

3. Analysis of the Obtained Results and Determination of the Age of Primary Systems of Gravitational and Magnetic Anomalies

3.1 Structural-Tectonic Positions of Systems of Geophysical Fields Anomalies

The designated areas of high density of linear indicators of lineaments by the results of statistical analysis of geophysical fields anomalies spatial distribution (as models of tectonic disturbances from geophysical data) were compared with the materials of geological - geophysical and structural-tectonic exploration of the region (Popov, Smelov, 1996; Melnikov, 2008).

It was found that system (84-354^0) fits most major faults in the territory of the Aldan-Stanovoy megablock (Figure 1, 3). Comparison of the spatial position of the domains with the theoretical studies and models presented by authors (Hubbert, 1937; Kazansky, 1965; Trofimenko, 2010) made it possible to make the assumption about availability of self-similar systems domain of tectonic disturbances in the studied area.

Other orientations of faults at the regional and local levels found earlier by various authors are also confirmed by statistical analysis of the distribution of linear elements of geophysical fields anomalies. In addition to the existing models of the Aldan-Stanovoy megablock (Grishkyan, 1968; Popov, Smelov, 1996), results that indicate the presence of a domain structure for all selected systems were obtained. Moreover, the domain structure retains the principle of self-similarity. In other words, not only faults of certain orientation, but faults orthogonal to them are ranked.

Comparison of this study results with the results of simulation by authors (Gilmanova et al., 2012) of the gravitational field in the Bouguer reduction made using the method of analysis of digital fields show the correspondence of basic structural "gravitational" lineaments in azimuths 64 and 296^0. Taken together, results of these model constructions generalize findings and conclusions of previous studies.

3.2 Models of Forming Geophysical Fields Anomalies Systems

Diagonal rhombic structures (domains) are formed as transpressure under horizontal compression, theoretical model of which is considered in (Woodcock & Fisher, 1986). Figure 4A shows two systems: orthogonal (84-354^0) as a primary or basic grid of breaks and diagonal (40-310^0), which could be formed as a grid gaps in the implementation of the indenter. Within the studies area under horizontal (sublatitudinal) compression in the model (Woodcock & Fisher, 1986) indenter is the Batomgsky block located on the eastern edge of the Aldanian Shield (Figure 1). Figure 4.B-C shows a scheme of mutual position of systems (84-354^0) with systems (26 -296^0) and (35-305^0).

Regular mutual placement of the systems allowed, in addition to model (Woodcock & Fisher, 1986), to develop a model where it is possible to form structures in azimuths (26 - 296^0) and (35-305^0), in the form of a diagonal mutually orthogonal grid on the primary (main) system of first order breaks (Figure 5).

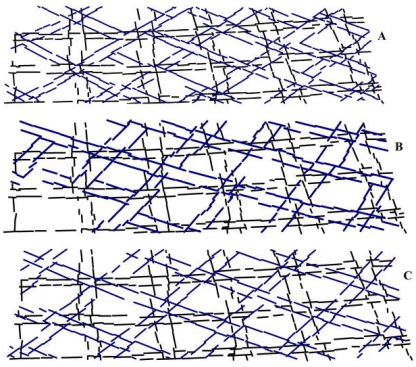


Figure 4. Scheme of mutual placement of diagonal lineament in azimuths (40-3100), (26 -2960) and (35-3050) on a primary orthogonal system of lineaments

A - relative positioning of systems (84-354^0) and (40-310^0); B - (84-354^0) and (26- 296^0); C - (84-354^0) and (35-305^0). The system of coordinates is similar to that in Figure 3.

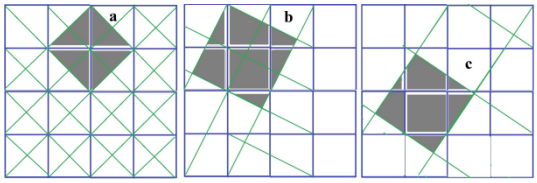


Figure 5. The model of forming system of domains in different azimuths on the primary orthogonal grid of faults

a- domain of system (45-315^0) ±30 according to (Woodcock & Fisher, 1986); b - domain of system (26-296^0) ±30; c - domain of system (32-302^0) ±30.

Systems similar by azimuths of spreads (26-296^0) - (35-305^0) and (84-354^0) - (3-273^0) are spatially combined (conjugated) by systems (Figure4.B-C). Theoretical value of angles (Figure 5) is arctg (1/2) = 26.56^0 for system (26 -296^0) and for system 35-305^0 - arctg (2/3) = 33.69^0. System (64 - 334^0), being complementary to system (26 - 296^0), since 26^0 +64^0 = 90^0, is also formed as a subordinate system.

The presented geometric models show that the presence of one (primary) orthogonal system of faults as a result of long-term variable loads on the Aldan-Stonovoy megablock makes it possible to explain the origin of all existing systems in the area of collision interaction of the Amur and the Eurasian lithospheric plates. This can be proved by the fact that 90% of intersection points of lineaments are concentrated within the orthogonal systems (84-3540) and (3-2730): latitudinal "a-c" and meridional "1-6" (Figure 3).

The developed model of mutual arrangement of systems of geophysical anomalies fit into the idea of existence of three dominant systems of crustal deformation in Northeast Asia: latitudinal-meridional (coupled systems of linear elements of fields 3 - 273^0, 84 - 354^0), north-west of the Paleozoic (coupled systems 64 - 334^0, 70 - 340^0), north-east - Mesozoic (system 40 - 310^0), corresponding, respectively, to the main periods of Archean-Proterozoic and Paleozoic-Mesozoic tectonic deformations of the Aldan-Stanovoy megablock. Mutual spatial arrangement of bimodal (3-273^0 and 84-254^0) and trimodal (26-296^0, 35-305^0 and 40-310^0) stretches of linear elements of gravity and magnetic fields may also indicate that changes in the Paleotectonic environments and related orientations of strain vectors in the territory of the Aldan-Stanovoy megablock did not occur simultaneously and sharply, but gradually and smoothly, which may find an explanation in the uneven (cyclic) rotation of the Amur plate.

3.3 Clarification of the Formative Period of System (84-3540) When Compared with Zones of Elastic-Plastic Flow

The developed model of inherited development of fault systems assumes availability of primary faults grid, which in this case is a system in azimuths (84-364^0) or (3-273^0). According to the compliance of the spatial position of these systems (Figure 3) to the main structural lineaments of the Aldan-Stanovoy megablock (Figure 1), time of their formation is not later than Archean - Proterozoic period (Gorokhov et al., 1981).

Concretization of the date of forming the basic system of anomalies in azimuths (84-354^0) was made possible by comparing the spatial position of individual elements of the system with zones of elastic-viscous flow of rocks (shear-zones).

Within the Aldanian Shield, same as in other areas of foundations of ancient platforms (Hamimi et al., 2014), zones, which are composed of rocks of adjacent blocks, slightly metamorphosed formations, anorthosite intrusions, granitoids, mylonites and blastomylonites, are mapped. They seem to ligate adjacent terranes as zones of tectonic melange (Melnikov, 2008). A detailed structural analysis of internal structure of these areas shows that they all formed using the mechanism of shear flow of rocks and are typical brittle-plastic shear-zones. All in all, within the Aldanian shield, four shear-zones are separated: Kalar, Amginsk, Tyrkandinsk and Ulkan (Figure 6).

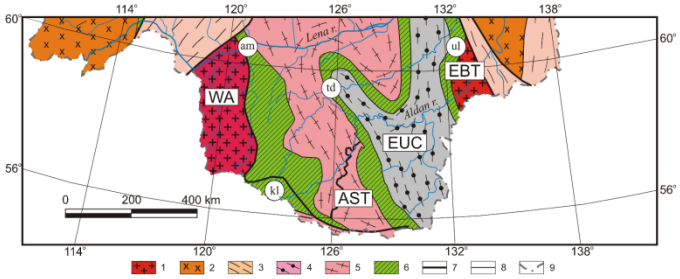


Figure 6. The structure of foundation of the southeastern part of the North Asian craton

1 - granite-greenstone terranes; 2 - tonalite-trondhjemite terranes; 3 -Proterozoic orogenic belts; 4 - granulite-paragneiss terranes; 5 -granulite-orthogneiss terranes; 6 - large zones of shear flow (shear zones); 7- large fault borders; 8 - minor faults; 9 - the south-eastern boundary of the North Asian craton. In circles are shear-zones: am - Amga, kl - Kalar, td - Tyrkandinsk, ul -Ulkan. In squares are terranes: WA - West Aldanian, AST - Sutamsky, EUC - Uchursk, EBT- Batomgsk.

The Kalar brittle-plastic shear-zone separates the West Aldanian granite-greenstone terrane located to the south of the Tyndinsk tonalite-trondhjemite-gneiss terrane and can be seen in the sub-latitudinal direction at the distance of about 650 km, 50 to 150 km wide. Various types of pegmatites and layered gabbro ultrabasite plutons 1.8 -1.9 billion years old seem to seam together plates and blocks of different age and types in the Kalar Shear Zone (Melnikov, 2008).

The Amga shear-zone separates the Central Aldanian composite terrane from the West Aldanian and Tynda terranes located to the west and south of it, and cuts off the Kalar shear-zone (Figure 6). It has an arc shape seen in the plane and can be seen at a distance of about 650 km. The width of the zone varies from several to 150 km. Rocks that connect different plates are presented here with granites and pegmatites 1.9 - 2.0 billion years old. Similar or close age have the gabbro-diorite-plagiogranites of the Ungrinsk complex, which have been metamorphosed into amphibolite facies. In general, the structure of the Amga shear-zone consists of the Archean rocks of amphibolite - epidote-amphibolite facies, Early Proterozoic ortho-gneiss and para-gneiss strata of the sub-granulite - granulite facies, Archean and Early Proterozoic fragments of greenstone structures and differentiated plutons of ultrabasic and basic rocks that have experienced widespread structural metamorphic processing in the range between 2.15 - 1.9 billion years.

Tyrkandinsk brittle-plastic shear-zone separates the East Aldanian superterrane from the Central Aldanian one in the west and from the Tynda one in the South of the composite terranes (Figure 6). It has a curved shape in plan and can be traced at a distance of about 1,650 km with its width ranging from 50 to 200 km. The age of magmatic zircons defined using the U-Pb isochron method is 1.9 billion years (Bibikova et al., 1989). The Ulkan Shear Zone by its internal structure, material composition and age of formation is no different from the Amga and the Tyrkandinsk Shear zones.

In general, formation of brittle-plastic shear-zones of the Aldan-Stanovoy megablock that connected its disparate terranes into a single continental block occurred in the age range of 2.1-1.8 billion years (Brandt et al., 1981).

Qualitative analysis revealed that the fragment of the latitudinal anomalous zone in the azimuth 840 at the latitude 56W from meridian 122E to meridian 126E (Figure 3) is spatially coincident with the Kalarsk shear-zone and meridional anomalous zones in azimuth 354^0 at longitudes 121E, 126E and 132E (Figure 3) - with the Amga, Tyrkandinsk, and Ulkan shear-zones (Figure 6).

Spatial relationships between the pulse-meridional adjacent systems (3 - 84 and 273^0 - 354^0) (Figure 3) and shear-zones of the Aldanian shield (Figure 6), composed of geological formations of detected age, leads to the conclusion that the spatial anomalies of geophysical fields in these azimuths (3 - 84 and 273^0 - 354^0) were formed in the period not later than 1.8-2.1 billion years ago.

3.4 System of Active Lineaments (40-3100), Seismicity and Prediction of Location of a Strong Earthquake

Comparative analysis of spatial distribution of earthquakes and the developed tectonics model of the Aldanian-Stanovoy megablock by statistical distribution of anomalies of geophysical fields has shown that the system of lineaments with azimuths (40 - 310^0) fits into the structure of the seismic field, and all gravitate epicenters of strong earthquakes tend to the points formed by intersecting of lineaments of the system, or to axis lines (Figure 7).

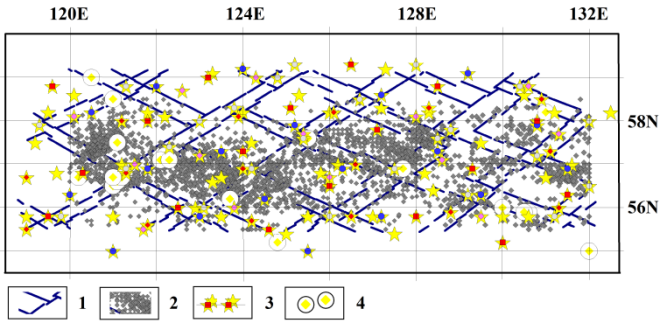


Figure 7. A fragment of the spatial structure of domains of a diagonal system of lineaments of geophysical fields with azimuths (40-3100) with epicenters of earthquakes and the points of all selected systems

Legend: 1 - lineaments of anomalies of geophysical fields in azimuth (40-310^0); 2 - dimensional field of earthquake epicenters, 3 - lineament intersection points of selected seven systems of anomalies of geophysical fields; 4 - strong earthquakes.

Stretch of concentration zones of earthquake epicenters in the central part have common azimuth 310^0, and after that, during transfer from one domain to another, consistently changes azimuth 310^0 to 40^0, from 40^0 to 310^0.

In the west of the area (121 HP, 56.5SSH) (Figure 7), there is a tectonic point of intersection of lineaments of five systems (intersection of the latitudinal structure "b" and meridional - "2", Figure 3), with azimuth (3-273^0), (26 -296^0) (40-310^0) (64-334^0) (70-340^0), respectively. This tectonic point attracts strong earthquakes with magnitude M = 6.5 and M = 7.0, with epicenters fitting into the main spatial domain structure of lineaments system (40-310^0).

In the center of the analyzed region (Figure 7, 126 VD, 56.5 SSh), (intersection of latitudinal structure "b" and meridional - "4" Figure 3) there is a point of intersection of six systems of lineaments, i. e. 6 systems converge in the center of the domain system (40-310^0), at the north-eastern border two strong earthquakes have been registered. Southeast fragment of this domain at present is aseismic, however, by analogy with the western domain, it can be considered to be one of the most likely places of tectonic stress release.

Thus, the obtained regularities suggest that the system of lineaments in azimuth (40-310^0) may be associated with the zones of active deformation of the earth's crust and currently controls the seismic process within the area of convergent interaction between the Amur and the Eurasian lithospheric plates.

4. Conclusion

This study presents the results of a statistical analysis of spatial distribution of anomalies of the gravitational and magnetic fields of convergent interaction between the Amur and the Eurasian plates. This area is a seismically active region in North-East Asia, with energy potential of individual fragments which corresponds to ten zones of seismic concussions.

Using methods of statistical analysis of azimuths of linear elements of anomalies, gravity and magnetic stages and other geophysical signs of faults, seven primary directions have been selected with azimuths (3-273^0), (26-296^0) (35-305^0) (40-310^0) (64-334^0) (70-340^0) (84-354^0), respectively.

It has been established that within the Aldanian Stanovoy megablock, these linear data fit into certain fault indicators (lineaments) or systems tightly oriented in space grids, each of which comprises two mutually orthogonal areas that form domain structures.

The length of the sides of domains for different systems is 120 to 150 km and is conditionally called a spatial period of the first order.

Comparison of systems of geophysical fields anomalies with peculiarities of tectonic plan of the area, with zones of elastic-plastic flow of rock masses (shear-zones) showed that systems (3-273^0) and (84-354^0) are spatially correlated with geological - tectonic structures of the first order, formation of which refers to period 1.8-2.1 billion years.

Systems (26-296^0) (35-305^0) and (40-310^0) may be sequentially formed on the primary grid of faults. The order of formation of these systems is the subject of further study and geological interpretation.

In general, this analysis made it possible to establish structural-tectonic position and relative time of formation of

individual systems of geophysical fields anomalies. As a result of comparing diagonal system of geophysical fields anomalies with spatial distribution of the fields of contemporary seismicity and epicenters of strong earthquakes, an assumption was made that the diagonal system of lineaments reflects zones of tectonic deformation and has been controlling the epicentral fields of contemporary seismicity since the Mesozoic period of seismotectonic activation.

Determinant result of this study is the synthesis of early scientific works and getting new ideas about geophysical models of Earth's crust in the Aldanian-Stanovoy megablock. The indicated direction of search works for studying relationship of systems of geophysical fields anomalies and seismicity is the subject of further research.

The main direction of further studies is the construction of lineament systems based on new digital elevation model (DEM) technologies and carrying out the lineament trend analysis of the geophysical fields in the western segment of the Aldan-Stanovoy megablock. Eventually, taking into consideration systematization of all data on the age determination of rocks in the study area, this will allow us to establish the time of origin of diagonal lineament systems and the succession in their formation.

Acknowledgments

This work was supported by the state task №5.1771.2014/K the Ministry of Education and Science of Russia and by Program “The Far Eastern”, under Grant of the Far Eastern Branch of RAS.

References

Alaa, A. M., & Katsuaki. K. (2011). Auto-Detection and Integration of Tectonically Significant Lineaments from SRTM DEM and Remotely-Sensed Geophysical Data. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*(6), 818-832. http://dx.doi.org/10.1016/j.isprsjprs.2011.08.003

Anokhin, V. M., & Maslov, L. A. (2009). Regularities of Orientation of Lineaments and Faults of the Russian Part of the Bottom of the Sea of Japan. *Pacific Geology, 2*, 3-16.

Anokhin, V. M., & Odesskiy, I. A. (2001). Characteristics of a Global Network of Planetary Fracturing. *Geotektonika, 5*, 3-9.

Arellano-Baeza, A. A., Zverev, A. T., & Malinnikov, V. A. (2006). Study of Changes in the Lineament Structure, Caused by Earthquakes in South America by Applying the Lineament Analysis to the Aster (Terra) Satellite Data. *Advances in Space Research, 37*(4), 690-697. http://dx.doi.org/10.1016/j.asr.2005.07.068

Bibikova, E. V., Gracheva, T. V., & Drugova, G. M. (1989). U-Pb Age of Granulitic Complex (the Aldanian Shield). *Proceedings of the USSR Academy of Sciences, 304*(4), 949-952.

Boyce, J. I., & Morris, W. A. (2002). Basement-controlled Faulting of Paleozoic Strata in Southern Ontario, Canada: New Evidence from Geophysical Lineament Mapping. *Tectonophysics, 353*(1-4), 151-171. http://dx.doi.org/10.1016/S0040-1951(02)00280-9

Brandt, S. B., Grabkin, O. V., Lepin, V. S., & Melnikov, A. I., et al. (1981). About Geochronology of the Western Part of the Aldanian Shield. *Soviet Geology, 3*, 58-67.

Gilmanova, G. Z., Shevchenko, B. F., & Rybas, O. V., et al. (2012). Linear geological Structures of the Southern Aldan-Stanovoi Shield and Eastern Central-Asian Foldbelt: The geodynamic aspect. *Russian Journal of Pacific Geology, 6*(1), 52-60. http://dx.doi.org/10.1134/S1819714012010058

Gorokhov, I. M., Duk, V. L., & Kitsul, V. I., et al. (1981). Rb-Sr Systems of Polymetamorphic Complexes of the Central Part of the Aldanine Crystalline Massif. *Bulletin of USSR Academy of Sciences, Series geological., 8*, 5-16.

Grishkyan, R. I. (1968). About the Mechanism of Formation of Late Pre-Cambrian Faults in Crystalline Strata in the Central Part of the Aldanian Shield. *Geotectonics, 4*, 136-139.

Hamimi, Z., El-Sawy, E. K., El-Fakharani, A., Matsah, M., Shujoon, A., Mohamed, K., & El-Shafei, M. (2014). Neoproterozoic Structural Evolution of the NE-trending Ad-Damm Shear Zone, Arabian Shield, Saudi Arabia. *Journal of African Earth Sciences, 99*(1), 51-63. http://dx.doi.org/10.1016/j.jafrearsci.2013.09.010

Hobbs, W. N. (1904). Lineaments of the Atlantic Border Region. *Bulletin Geological Society of America., 15*, 483-506.

Hubbert, M. K. (1937). Theory of Scale Models as Applied to the Study of Geologic Structures. *Geological Society of America Bulletin, 48*, 1459-1520.

Kazansky, V. I. (1965). About the Internal Structure of Archean Faults in the Central Aldanian Region. *Geology*

of Ore Deposits, VII(2), 63-79.

Lima, N. H. C., Ferreira, J. M., Bezerra, F. H. R., Assumpção, M, do Nascimento, A. F., Sousa, M. O. L., & Menezes, E. A. S. (2014). Earthquake Sequences in the Southern Block of the Pernambuco Lineament, NE Brazil: Stress Field and Seismotectonic Implications. *Tectonophysics, 633*, 211-220. http://dx.doi.org/10.1016/j.tecto.2014.07.010

Malyshev, Y. F. (1977). *Geophysical studies of pre-Cambrian Aldanian shield* (pp.127). Moscow: Nauka. Melnikov, A. I. (2011). *Structural Evolution of the Metamorphic Complexes of Ancient Shields* (pp. 288). Novosibirsk: GEO Publishers.

Popov, N. V., & Smelov, A. P. (1996). Metamorphic formations of the Aldanian Shield. *Geology and Geophysics, 37*(1), 148-161.

Rosen, O. M. (2002). Siberian Craton – a Fragment of a Paleoproterozoic Supercontinent. *Russian Journal of Earth Sciences., 4*(2),103-119. http://dx.doi.org/10.2205/2002ES000090

Stativa, A. S., & Trofimenko, S. V. (2006). Building Systems of the Aldan Shield Faults by Geophysical Data. The *Mining Informational and Analytical Bulletin (Scientific and Technical Journal), 17*(3), 193-198.

Stogniy, V. V., Smelov, A. P., & Stogniy, G. A. (1996). Deep Structure of the Aldanian Shield. *Geology and Geophysics, 37*(10), 148-161.

Stativa, A. S., Trofimenko, S. V., & Grib, N. N. (2006). Analysis of Existing Ideas about the Tectonics of the Aldan Shield. *The Mining Informational and Analytical Bulletin (Scientific and Technical Journal), 17*(3), 167-189.

Stativa, A. S., Trofimenko, S. V., & Imaev, V. S. (2006). Spatial Distribution of Indicators of Fault Systems in the Aldan Shield. *The Mining Informational and Analytical Bulletin (Scientific and Technical Journal), 17*(4), 188-197.

Trofimenko, S. V. (2010). Tectonic Interpretation of the Statistical Model of Distributions of Anomalies Azimuths of Gravity and Magnetic Fields of the Aldanian Shield. *Pacific Geology, 29*(3), 64-77.

Woodcock, N. H., & Fisher, M. (1986). Strike-Slip Duplexes. *Journal of Structural Geology, 8*(7), 725-735.

Yuanyuan, L., Braitenberg, C., & Yang, Yu. (2013). Interpretation of Gravity Data by the Continuous Wavelet Transform: The Case of the Chad Lineament (North-Central Africa). *Journal of Applied Geophysics, 90*, 62-70. http://dx.doi.org/10.1016/j.jappgeo.2012.12.009

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Features of Aboveground Pipeline Compensation Part Stress-Deformed Study at Permafrost

Dinar Flerovich Bikmukhametov[1], Gennady Evgenyevich Korobkov[1] & Anna Pavlovna Yanchushka[1]

[1] Ufa State Petroleum Technological University, Russian Federation

Correspondence: Dinar Flerovich Bikmukhametov, Ufa State Petroleum Technological University, Russian Federation.

Received: January 15, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p204 URL: http://dx.doi.org/10.5539/mas.v9n8p204

Abstract

The gradual shift of the Russian Federation oil production center to the areas with difficult climatic conditions, states additional challenges to the Oil production companies during development of these regions. Impossibility of underground design of oil pipelines in waterlogged areas, areas with subsurface voids unstable of different origin, permafrost zones, landslide zone forces companies to focus laying pipelines by above ground manner. The purpose of this article is to present a review of the stress-deformed state of the aboveground pipeline, its behavior during operation and pressure test stage, the forces acting on the pipeline. The study allows us to understand the state of the pipeline, to assess its stability and reliability. Analyzing the research results, it must be concluded that the decision of choice the above ground pipeline design (route geometric parameters), should be performed, guided not only meet the conditions of strength, as well as other reasonable important factors: the values of the forces acting on the fixed supports, three-dimensional position of the pipeline during hydraulic testing and operation, vibration in the pipeline, the impact of ambient temperature, wind resistance aboveground pipeline, further arising forces affecting piping supports, due to thawing of ground and soil bulging, etc.

Keywords: stress-deformed state of the pipeline, above ground pipeline, pipeline at difficult geotechnical conditions

1. Introduction

Today, Western Siberia is the leading oil-producing region of the Russian Federation. However, according to predictive estimates, about 60 % of the deposits of Western Siberia are in the stage of production decline. The reasons for this are several: the physical resources exhaustion, lack of investment in geologic exploration and ineffective production methods. At these conjunctures, today analysts predict a gradual shift the center of oil production from Western Siberia to the Eastern Siberia. Hydrocarbon fields depletion in developed areas, technical and economic constraints associated with the heavy oil extraction, bituminous sands development - all these circumstances objectively determine displacement of oil and gas development center into the new regions: the Yamal Peninsula, Eastern Siberia, Far East and shelf of Russian Federation (Abovskiy, 2005).

These regions are characterized by extremely difficult geotechnical conditions (Borodavkin et al., 1974), such as wetlands and waterlogged areas, areas with subsurface voids of different origin (area with cave formations, undermined territories in the areas of mine construction, etc.), permafrost zones, landslide zones, earthquake endangered zones, rugged terrain (Shammazov et al., 2005).

Due to the gradual shift the center of oil and gas production in these areas attention to the pipeline laying increases. Field development in such areas requires a huge amount of investments, main part of which shall be intended to implementation of technologically and structurally safe pipeline systems. Under these circumstances, evaluation of the stressed-deformed state of such pipelines is becomes extremely valuable (Vitchenko, 2008.).

During the construction and operation of oil and gas transportation pipeline systems occurs introduction their technological elements in the natural environment (Petrov and Spiridonov, 1973). This is often done in violation of the dynamic equilibrium, accompanied by activation of natural hazardous processes that have a significant negative impact on the technical condition of the pipelines often lead to the accident. Among such "rejection reactions" of natural environment include Structure subsidence or bulging, activation of permafrost, erosion, landslides and flooding on the pipeline routes all these processes belong to such "rejection reactions" of natural environment

(Borodavkin and Berezin, 1974).

Above Ground pipelines in many cases, the most economically viable and feasible from the design, technological, construction and operation of the pipeline point of view. Above Ground laying of pipelines is usually applied in complex geological and hydrogeological conditions, such as marshy and flooded areas, in areas with frequent ravines, streams and other obstacles, on the rocky soil and in permafrost conditions (Khrenov, 2005).

2. Method

This research was base on the study of stressed-deformed state of aboveground oil pipeline considering of the really existing aboveground pipeline part in the Komi Republic. The pipeline is laying in the permafrost zone, at very low temperatures and complex natural and geological conditions. The route of the pipeline passes trough the multi-year frozen soil with a temperatures -6 -7° C. The water saturation of the ground reaches 70%. Ice content is 10-20%, and the ice flat sheet deposits reach several tens of meters thickness, at some places they are cryopegs. The pipeline route crosses a large number of rivers and streams. The intensively occur frost heaving cause major deformation of the supporting part of the pipeline. Currently, the land pipeline construction in such circumstances is almost have not reliable engineering solutions to ensure provision of general and longitudinal stability of the pipeline, and therefore, there is no guarantee to ensure their reliable operation. The route passes in difficult engineering-geological, climatic and environmental conditions, in many places in marshland, in the permafrost. The combination of these conditions determines the complexity of the construction and operation of the pipeline.

There are oil, gas and water pipelines laying in one right-of-way, but for the stressed-deformed state consideration it was taken the part of the aboveground oil pipeline shown on the picture, 1between 2 fixed supports near Block Valve Station (BVS) #2.

Figure 1. Part of the aboveground pipeline with a diameter of 220 mm

To study the stressed-deformed state of the aboveground oil pipeline part, as well to analyze of the pipeline displacement study part of the pipeline has been modeled in accordance with the design and technical documentation in software "Start", represented on the Figure 2. The modeled part of the aboveground pipeline contains of: 2 fixed supports at the ends #400, #411; 6 different angle elbows #500 – 505; 10 sliding supports #401-410; 1 Emergency Shut Down Valve (ESDV) - #600.

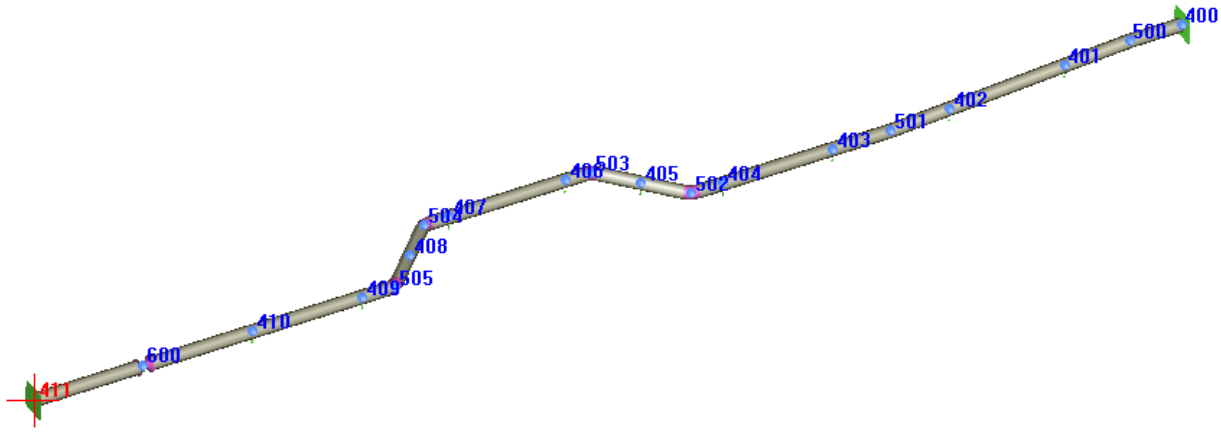


Figure 2. Modeled part of the above ground oil pipeline in «Start» software

Input data for calculation:

Installation temperature – (-30)° C;

The test liquid - water;

Test temperature – (10) ° C;

External pipe diameter – 219,1 mm;

Nominal wall thickness of 14,3 mm;

Pipe material - 10G2FBYU;

Operating Pressure – 13,9 MPa;

Design temperature – (42,5)° C;

Test pressure - 20.85 MPa.

Based on the data and operating conditions "START" software varied out the strength calculations of the pipeline. The software product is capable to simulate the behavior of the pipeline, as well as to determine the three-dimensional position of the pipeline sections during the operation and hydraulic tests stages.

The program calculates the circular, longitudinal and equivalent stress. According to the method of limiting conditions on Construction norms and rules (SNIP) 2.05.06-85 (SNiP 2.05.06-85, 1985) program calculates ultimate strength designed resistance R1 and the calculated yield strength designed resistance of R2 (Ainbinder and Kamerstein, 1982). Also, the software calculates the total forces acting on the pipeline and determines the values of pipeline displacement during operation and hydraulic tests. For this pipeline R1 = 2322,58 (kgf / sq. cm), R2 = 2350 (kg / sq. cm) (VSN-2-26-71, 1971).

3. Results

Let's consider the stressed-deformed state of the aboveground oil pipeline represented on the Figure 3. Figure 3 shows the installation position of the pipeline - in green, and red - the position of the pipeline during the operation.

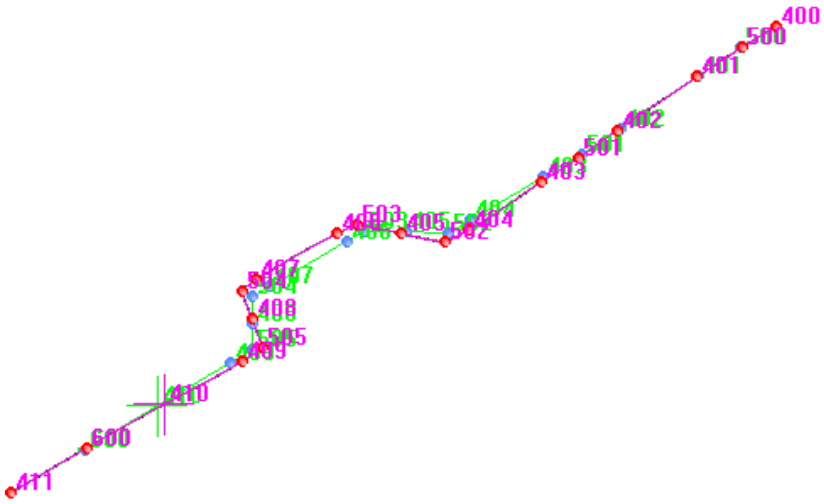


Figure 3. Aboveground pipeline behavior during its operation

As seen from the program calculations and also the graphical calculation results presentation window, during the operation stage considered pipeline would be mostly displaced at the supports #404, 406, 407, 409. Modeled aboveground pipeline has the fixed supports at the beginning and at the end of the considering part. Accumulated during the operation stresses (due to the internal pressure in the pipe, thermal expansion, etc.) are distributed along the considered part of the pipeline between the two fixed supports. Since there is the compensator in the considering part, the stresses arising in the pipeline, maximally would be compensated on it, resulting in the displacement of the pipeline in the direction of least resistance - in this case displacement from the designed axis. In many cases the result of this displacement could be derailing of the aboveground pipeline from the pipeline support.

Software calculation results confidence confirm the field measurements. Before the stat-up of the pipeline lines was marked on the sliding support and on the support beam. After the start up of the pipeline reaching the project capacity it was noted that lines was displaced relative each other as it shown on the pictures 4 and 5.

Figure 4. Displacement of the aboveground pipeline at the support #407

Figure 5. Displacement of the aboveground pipeline at the support #409.

Software calculation results shows that the displacement of the pipeline from the Y-axis at the support #407 would be 33,4 mm to the right downstream direction and displacement of the pipeline from the Y-axis at the support #409 would be 12,5 mm to the left downstream direction. Field measurements confirm «Start» software calculations.

The total equivalent thee-dimensional displacement of the considered aboveground pipeline during the operation shown on the Figure 6.

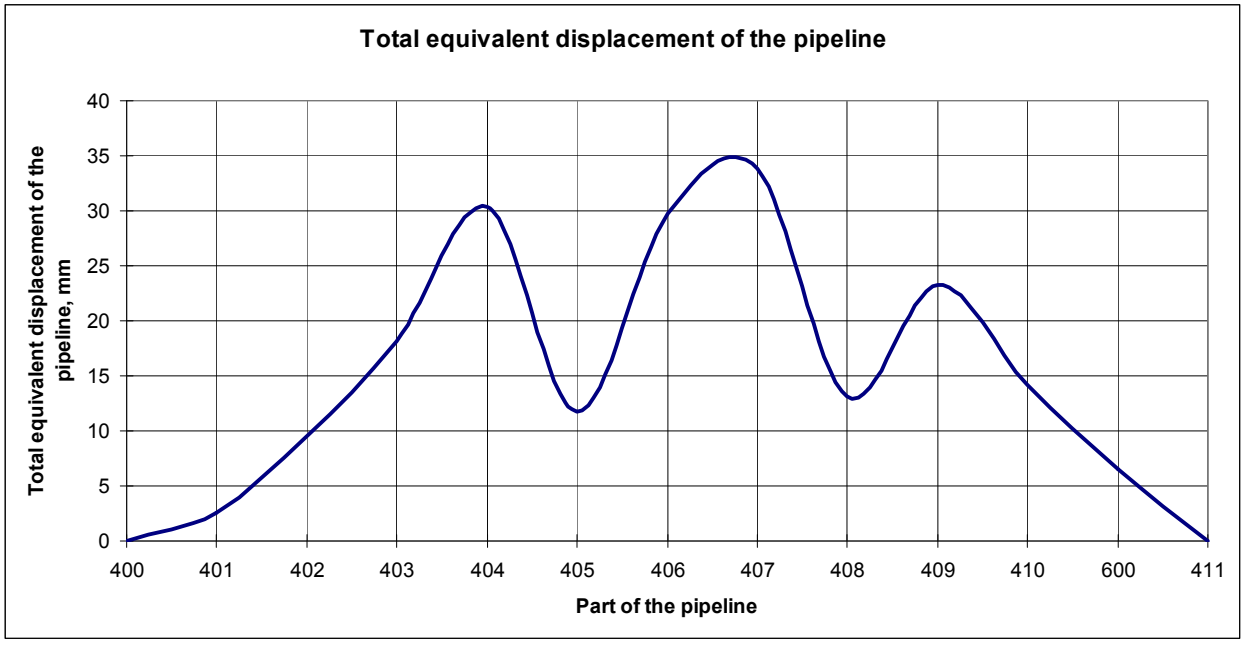


Figure 6. Diagram of the total equivalent displacement of the pipeline

Diagram shows that the largest displacements of the pipeline occur in the flexible areas - it means at compensation sites. The most displaced parts of the pipeline are at the supports #404, 406, 407, 409 – where the elbows installed a shortest distance away.

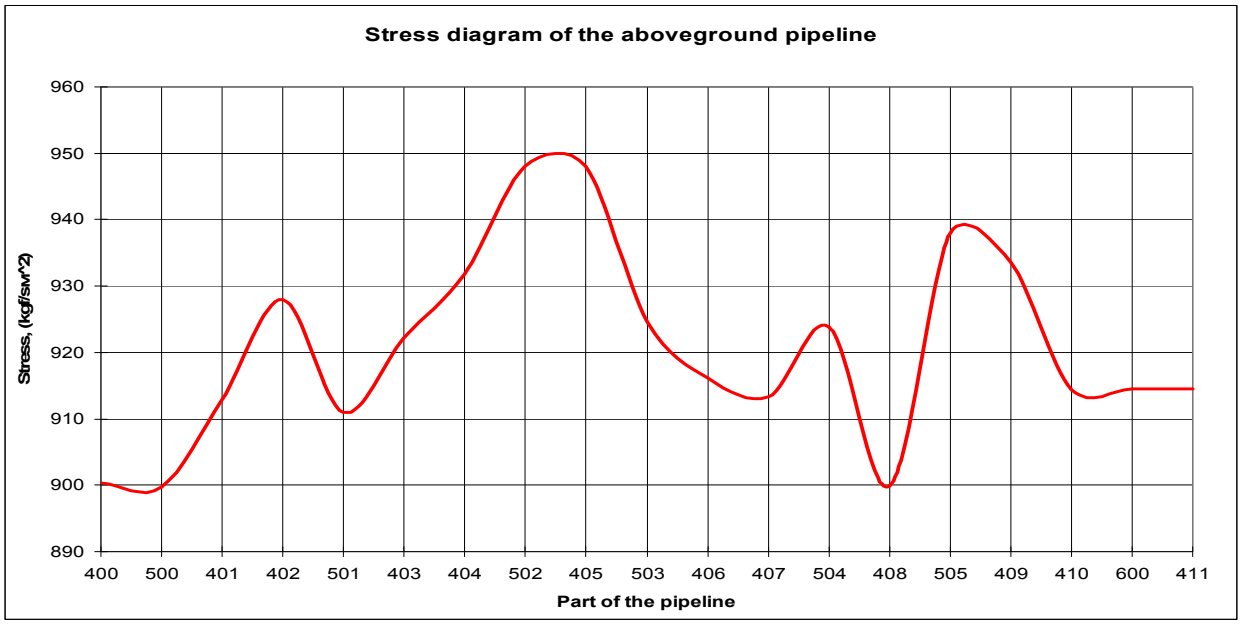


Figure 7. Diagram of the total equivalent stresses at the pipeline

Displacement of the pipeline it is the result of stresses arising during thermal expansion of the pipe sections, and this stresses are compensates at the pipeline bend points, thereby increasing the bending moment arising in a slightly curved parts. The diagram of the stress distribution along the all part of the considered aboveground pipeline shows that the maximum stresses would be located at the compensational part of the pipeline. Rising of the stresses due to the arising bending forces explain the highs of the stresses close to the elbow parts of the pipeline (see Figure 7). Based on our calculations we can say that the compensation sections of the pipeline are the most critical sections of aboveground pipeline. Design of the pipelines without compensation parts impossible due to the necessity of unloading the stresses in the pipeline occurs during the operation.

Let's consider the necessity of pipeline compensation parts, and in case of their essentiality determine compensator optimal configuration. To do this we have modeled in "Start" section of the pipeline without compensation part - with a straight section of the pipeline. Figure 8 shows this section of the pipeline without the compensation part.

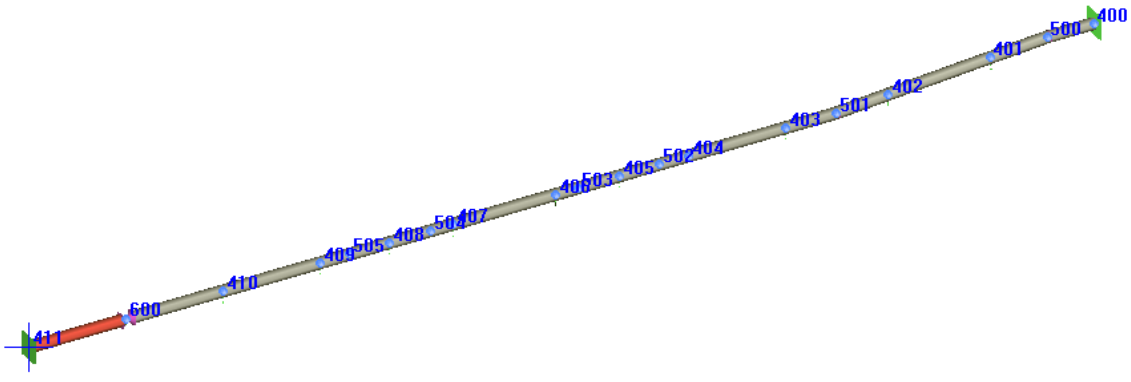


Figure 8. Modeled part of the above ground oil pipeline without compensation part in «Start» software

Then we have modeled similar sections of the pipeline with the elbows 15°, 30°, 45°, 60°, 75° and 90° at the compensation parts. First we estimated the total equivalent displacement of the pipeline sliding supports depending on the elbow angles. In addition to the displacements of the pipeline, at the same time, for a full analysis and evaluation of the given pipeline part, we have considered the stress-deformed state of the pipeline in

its areas depending on the elbow angles at the compensator. Considering the case with a pipeline without compensator part we have discovered unacceptable 120 mm displacement of the pipeline at the support #401- the closest point to the elbow, necessary to achieve the contour of the pipeline route. Moreover due to the absence of compensation stresses sections the maximum stresses from thermal expansion of the pipeline and internal pressure will be impact on to the fixed support at beginning of the pipeline, and to the bending stresses in the elbow. These stresses exceed the maximum permissible value, so such kind of design is not allowed. In the design of the pipeline section with a compensator 15° the stress value along the length of the pipeline is significantly reduced, but the pipeline displacement in the compensation section is unacceptable - more than 165 mm.

Compensator 30° for this section of the pipeline creates quite large displacement values of the pipeline in the compensation section up to 65 mm. And in case of operation characteristics change, extreme weather conditions, wind and temperature loads, this displacement might increase and become unacceptable.

Compensator 90° for this section of the pipeline meets the stress-deformed state requirements of the pipeline and its displacement. The feasibility of using this type of compensator goes into the economical rationality area, as the use of this type of compensator increases the length of the pipeline (number of tubes). Moreover in the pipeline with such configuration might arise additional hydraulic resistance and hydraulic shock. Also with such kind of elbows it is difficult to use some kind of pipeline pigs and diagnostic devices. Similar, but a lesser degree, conclusions also apply to the compensator with 75° elbows.

Analyzing the data obtained it could be concluded that the compensation sections with 45° and 60° elbows are acceptable from the values of stresses and the pipeline displacements points of view. The decision to choose the type of compensator in each case would be taken individually depending on the length of the section between the fixed supports, stress values requiring compensating, wind and temperature load, technical and economic characteristics.

4. Discussion

The issue of the structural strength of pipelines in operation stage has the high importance. This question has been studied for several decades ago with the beginning development of areas with difficult climatic conditions (Kamerstein at el., 1969). The main parameters for the aboveground pipelines strength estimation are designed resistance R1 and the calculated yield strength designed resistance of R2. Most of the calculations reduced to the sufficient pipeline wall thickness determination at the given operation parameters. Subsequently, there were works of evaluating the pipelines strength as under the static load as under dynamic. It was investigated the methods of pipeline strength and durability calculations, after alternative calculation methods have been proposed (Sokolov and Limar, 2009). The pipelines deformations at operation were also investigated and several suggestions were offered for the pipeline operation reliability increasing (Spiridonov and Sverdlov, 1971). Particular attention has been paid to the pipeline laying scheme: straight and with compensators in the form of low sinuosity pipe sections (Spiridonov and Svarts, 1974).

Some works have been dedicated to the calculation of the bearing capacity of pile supports aboveground piping that was the proposal of constructive measures to improve their reliability (Kharionovskiy, 1990). Were analyzed correspondances between high-altitude position of the pipeline support and its stress-deformed state (Bykov at el, 2012). Also, researchers suggested using of the frame pipeline crossing and its improvement based on the stress-deformed analysis (Avtakhov, 2004). One of the important parameters affecting the operation of the above-ground pipelines is wind flow. Necessary to provide aerodynamic stability of the pipeline as a whole, as well as individual elements. Criteria for assessing the aerodynamic stability of the pipeline must be in the list of obligatory check, as well as considered in the design of above-ground pipelines (Kuzmin at el, 2006). Wind flow and the fluid flowing through the pipeline could have a significant impact on the oscillation frequency of the pipeline, which affects the total stability of the system (Sokolov, 2012).

Deformations of the aboveground pipeline are not critical if they are within acceptable ranges (Vitchenko and Bereznyakov, 2008). It is important to develop a system of evaluation of the stress-deformed state aboveground the pipeline after its commissioning (Vitchenko, 2008). Evaluation of stress-deformed state of pipelines with the help of modern computer products and techniques has also been studied in modern science (Lalin and Yavarov, 2010).

5. Conclusion

Thus, analyzing the calculation results, it must be concluded that the compensation sections on the pipeline are essential but the decision of choice the compensator and pipeline route geometric parameters, should be

performed, guided not only meet the conditions of strength, as well as other reasonable important factors: the values of the forces acting on the fixed supports, three-dimensional position of the pipeline during hydraulic testing and operation, vibration in the pipeline, the impact of ambient temperature, wind resistance aboveground pipeline, further arising forces affecting piping supports, due to thawing of ground and soil bulging, etc.

It is necessary to check the reliability of the pipeline for each parameter of estimation and in case of dissatisfaction with one of the parameter appropriate adjustments should be made. Should take into account that the made corrections can positively affect on the checking to one of the criteria and negatively affect to the other. Therefore, a comprehensive and in-depth analysis is required.

Perspectives of research - an assessment of the stress-deformed state of the pipeline and its position taking into account all factors of influence: vibration, temperature, wind, etc. It is important to be able to assess the reliability of the pipeline for critical scenarios simultaneously, provide the opportunity to assess the reliability of the pipeline based on the probability of occurrence of these scenarios. The results of the study should be the main proposals to increase the reliability of the pipeline, the recommendations on basic design moments (placement of fixed supports, road crossing design, design of compensation areas), taking into account topography, soil conditions, the ambient temperature and other factors.

References

Abovskiy N. P. (2005). Construction in the northern oil and gas regions of the Krasnoyarsk Territory.: KrasGASA, Krasnoyarsk. ISBN 5-89628-136-6. pp. 228.

Ainbinder, A. B., & Kamerstein, A. G. (1982). Pipelines calculation for strength and stability. Resource book. Nedra, Moscow. pp. 341.

Avtakhov, Z. F. (2004). Improvement effectiveness of using of frame pipeline crossing. Dissertation for the degree of Doctor of Philosophy, Ufa State Petroleum Technological University, Ufa, Russian Federation. UDK 621.644.074.3(043)

Borodavkin, P. P., & Berezin, V. L. (1974). Construction of pipelines. Nedra, Moscow. UDK: 621.643.002 (075.8). pp. 407.

Borodavkin, P. P., Berezin, V. L., & Ruderman, S. U. (1974). Selection of optimized pipelines routes. Nedra, Moscow. pp. 240.

Bykov, L. I., Kotov, M. Y., & Murasov, T. T. (2012). Assessing the influence of changes of supports high-altitude position to the stress-deformed state of beam crossings. *Transport and storage of petroleum and of hydrocarbons, 1*, 7-12. Retrieved from http://elibrary.ru/item.asp?id=17787913

Kamerstein, A. G., Rogdestvenskiy, V. V., & Ruchimskiy, M. N. (1969). Pipeline stress analysis. Resource book. Nedra, Moscow. UDK 625.9(03). pp. 440.

Kharionovskiy, V. V. (1990). Strength improvement of the gas pipelines at difficult conditions. Nedra, Moscow. UDK 621.643.002.2. pp. 180.

Khrenov, N. N., (2005). Basics of integrated diagnostic of northern pipelines. Gas press, Moscow. ISBN 5-87719-008-3. pp. 606.

Kuzmin, S. V., Zakuraev, A. F., & Ivanov, V. A. (2006). Aerodynamic calculation methods of aboveground pipelines. University news. *Oil and gas, 1*(62-66). Retrieved from http://mars.arbicon.ru/index.php?mdl=content&id=38520

Lalin, V. V., & Yavarov, A. V. (2010). Modern calculation technologies of main pipelines. *Engineering and construction magazine, 3*(43-47). Saint-Petersburg. Retrieved from http://www.engstroy.spb.ru/index_2010_ 3/yavarov.html

Petrov, I. P., & Spiridonov, V. V. (1973). Aboveground pipelining. Nedra, Moscow. pp. 472.

Shammazov, A. M., Zaripov, R. M., Chichelov, V. A., & Korobkov, G. E. (2005). Calculation and ensuring the strength of the pipelines in difficult engineering-geological conditions. Inter, Moscow. ISBN 5-98761-006-0. pp. 706

SNiP 2.05.06-85 Construction norms and rules. Pipelines. (1985). Retrieved from http://xn--h1ajhf.xn--p1ai/snip/view/95

Sokolov, S. M., & Limar, O.V. (2009). Above-ground main pipelines calculation and durability test. *Oil industry. 6*(86-90). Retrieved from http://www.oil-industry.ru/archive_detail_en.php?ID=8714

Sokolov, V. G. (2012).Vibrations, static and dynamic stability of large diameter pipelines. Dissertation for the degree of Doctor of Philosophy, Saint-Petersburg State University of Architecture and Civil Engineering, Saint-Petersburg, Russian Federation. Retrieved from http://www.dissercat.com/content/kolebaniya-staticheskaya-i-dinamicheskaya-ustoichivost-truboprovodov-bolshogo-diametra

Spiridonov, V. V., & Svarts, L. E. (1974). Investigation of aboveground pipelines with low sinuosity pipe sections. *VNIIST issue, 29*(95-100).

Spiridonov, V. V., & Sverdlov, M. F. (1971). Aboveground pipelines deformability investigations. *VNIIST issue, 25*(230-245).

Vitchenko A. S. (2008). Control of pipelines stress-deformed condition at permafrost. Dissertation for the degree of Doctor of Philosophy, Engineering and technical science, Nadym, Russian Federation. Retrieved from http://www.dissercat.com/content/kontrol-deformirovannogo-sostoyaniya-nadzemnykh-truboprovodov-v-kriolitozone

Vitchenko, A. S., & Bereznyakov, A. I. (2008). Determination of the allowable deformation parameters of pipeline interconnections on the gas fields, operated in at Far North. *Science & Technology in the Gas industry, 4,* 104-109. Retrieved from http://naukaitechnika.ru/index.php?option=com_content&task=view&id=58&Itemid=19

VSN-2-26-71. (1971). Instructions of pipeline design with the longitudinal deformations compensation. Mingazprom, Moscow. pp. 258.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Mills Model Based Evaluation of Security of Software Systems

Valeriy Valentinovich Gurov[1] & Grigory Grigoryevich Novikov[1]

[1] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Russia

Correspondence: Valeriy Valentinovich Gurov, National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Russia.

Received: January 3, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p213 URL: http://dx.doi.org/10.5539/mas.v9n8p213

Abstract

The presented paper discusses a possibility for an implementation of software reliability models for an evaluation of security of software. Mills model is proposed as the most suitable one. In order to use it more effectively it is proposed to implement a division of vulnerabilities into groups. Classification of vulnerabilities based on a method of their connection with features of a process's execution on a level of interaction of system resources and operating system is presented. Security of system in a context of a specific application can be evaluated more precisely by means of assigning of a certain weight to each group of vulnerabilities. The suggestions on an evaluation of vulnerability of software taking into account a division of vulnerabilities into groups according to the presented classification are made.

Keywords: software vulnerabilities, unauthorized access, software reliability models, Mills error seeding model, software security analyses

1. Introduction

Security is an inherent property of software, which is currently being given more and more attention. The various sources give slightly different definition of that property. In the presented paper, we will adhere to (ISO/IEC 9126-93 Information technology), which defines security as a set of software properties, characterizing its ability to prevent unauthorized access, both accidental and deliberate, to software and data, as well as a degree of convenience and completeness of a detection of results of such access or actions aimed for a destruction of software and data.

Unauthorized access may be carried out if there are certain vulnerabilities in software code. The aim of a detection of vulnerabilities in software is developing of tests for a detection of security breaches in software's security system and determining an effective testing procedure (Zhang, Liu, Csallner, Kung &Yu, Lei, 2014). For quantitative evaluation of software security in a phase of software development various methods are used, including hierarchical methods of modeling, stochastic Petri networks, discrete Markov chains, Bayesian network models, etc. (Holm, Korman, & Ekstedt, 2015; Zhu, Qian, Zhang, Yanlong, Zhang, 2014; Yang, Yu, Qian, Sun, 2012; Vibhu, Saujanya, Sharma, Kishor, 2007). However, such methods are often too complicated and significantly depend on volume and accuracy of initial data, which obtainment is, in turn, a separate complicated task.

To a large extent concept of security coincides with concept of reliability of software, one of the main properties of which is a resistance to errors, i.e., the ability to sustain a certain level of quality of the functioning in cases about software errors or failures of a specific interface (IPX over IEEE Std 610.12-1990).

Issues of evaluation and provision of software's reliability are discussed for a long time and possess a large set of models and methodologies. Its presentation is given in the work (Myers, 1976). An increase of number models and a development of methods of their implementation is continuing until the present time (see, for example, the study of Lohmor, & Sagar, 2014). At the same time, analysis of software security and provision of their security, as an aim, appeared much later. However, nowadays security issues had become leading for software development of a variety of systems, ranging from simple applied problems to cloud computing (Wright, McQueen, & Wellman, 2013; Younis, Kifayat, & Merabti, 2014). Therefore, it is interesting to consider a possibility of an implementation of software reliability models in relation to an evaluation of its security.

The paper presents a procedure for an application of Mills model of software reliability for an evaluation of software security. In order to increase effectiveness of initial data preparation process and to increase accuracy of

obtained results it is proposed to use the developed classification of low-level vulnerabilities.

2. Methodology

2.1 Basic Characteristics of Mills Model

Software debugging is performed until intensity errors flow is reduced to an acceptable value for the given application, i.e., until mean time between failures is not large enough. Therefore, many reliability models focus on analysis of time of appearance of next error.

At the same time, it can be stated that, if software falls into a criminal's hands for a sufficiently long time, it will be trespassed. Therefore, it is reasonable to carry out an evaluation of its security by a degree of its resistance, i.e., by length of time during which a software can resist against unauthorized access. That parameter is determined by a number of vulnerabilities, which remained in a software after its testing.

Therefore, in that application it is unreasonable to use software reliability models, which are based on temporary characteristics of error flow. For that purpose one should use another class of models, for example, Mills model (Mills, 1972).

An implementation of Mills model in relation to an evaluation of software reliability consists of that, firstly, software is «obstructed» by a certain number of known errors. These errors are randomly made in a software, then it's presumed that for its own and introduced errors detection probability for a subsequent testing is the same and depends only on quantity. By means of testing of software for a certain period of time and sorting its own and introduced errors, one can estimate parameter N, which is an initial number of errors in a software, as well as to set confident level of that forecast.

Let's presume that, S error were introduced into a software, after which testing was initiated. Let's presume that n own errors and v introduced errors were detected during the test. Then, an evaluation for N according to maximum likelihood method will be as follows:

$$N = \frac{n}{v} * S \quad .$$

That model is simple and results obtained by means of that model are clear. However, a mechanism for an introduction of errors is currently the weakest part that model, because it is assumed that, a detection probability is identical (but unknown) both for its own and introduced errors. It results in the fact, that introduced errors must be «typical» sample errors. That, in turn, is quite hard to ensure, because software developers and persons, who introduce deliberately known errors, have usually have different programming style, skills and other aspects.

2.2 Classification of Low-Level Vulnerabilities

Simplification of «typification» of introduced errors can be achieved by means of their classification. In that case, it is easier to introduce similar errors into software, while knowing features of typical errors of a certain class.

Turning to an applicability Mills model for an evaluation of software's security, first of all, let's specify that the criterion for an absence of security vulnerabilities in software is the fact that, during testing of software no violations of security requirements for studied software in a system of its presumable application were found (Kazarin, 2012). Probability of all vulnerabilities depends on completeness of coverage of all possible conditions of a software. At the same time, it is presumed, that there are means of vulnerabilities' detection for a given run of a software. It is necessary to create a test for a detection of a vulnerability of each type, the test is formed depending on developer ideas about how a certain vulnerability appears in a certain situation, and, thus, a way of its detection.

Thus, we have the following situation. A multiplicity of vulnerabilities $W=\{w_0, w_1, ..., w_m\}$, their appearance in a software in a given test impact is unknown. A multiplicity of test impacts $T=\{t_0, t_1, ..., t_m\}$ exists, each of which is designed to detect a specific vulnerability. It is necessary to minimize a test's length, which with a certain probability detects all vulnerabilities of software. At the same time it worth noting that a detection of vulnerabilities always has a probabilistic nature.

For a solution of that problem is reasonable to divide all vulnerabilities into groups, and inside those groups by means of a certain test (presumably, it can also consist of a several test impacts) all vulnerabilities can be detected.

In that case, on the one hand, it is necessary to reduce a number of groups, in order to reduce a number of

conducted tests. However, in that case, a situation can take place, in which there is no common attributes among all vulnerabilities, selected into a certain group. As a result, within each group for each vulnerability it will be necessary to generate its own test impact. That, in turn, will not produce any advantage for a test procedure, which leads us an original option.

At the same time, an excessive increase of number of groups can lead to the fact that a total number of groups will be equal to an original number of identified vulnerabilities.

Because of that, it is necessary to develop such a classification of vulnerabilities, which would make it possible to bring together in separate groups vulnerability, which are of common appearance in a certain test impacts.

A large number of approaches for a division of vulnerabilities into classes is suggested. Such approaches are offered and used both by various companies developing anti-virus software and other companies, conducting their own analysis of their security software systems (Classification is detected objects, 2014; Classification of viruses' names according to «Doctor Web», Engle and Bishop, 2008; Knight, 2000).

However, in the discussed case, the most effective measure is use of a classification of vulnerabilities related with characteristics of a process's execution at a level of interaction of system resources (memory, processor) and the operating system -*low-level vulnerabilities* (Gurov, Gurov, Ivanov and Shustova, 2010). Areas of appearance of vulnerabilities and mechanism of their use for carrying out an attack on a computer systems are taken as a parameters for that classification.

According to the classification, vulnerabilities are divided into 6 classes. Those are buffer overflow, integer overflow, format string error, double free and race conditions. Buffer overflow vulnerability, in turn, is divided into two subclasses: stack overflow and heap overflow.

Also, data interpretation vulnerabilities can be attributed to a separate class. Low-level vulnerabilities are united by a cause of an appearance, which is an inadequate monitoring of input values.

Vulnerabilities inside classes are divided into types.

Stack overflow subclass is characterized by an overwrite of return address from a vulnerable function and is divided into 3 types:

- executable code is placed in stack;
- executable code is placed in environment variable;
- data, overflowing buffer, is placed in stack in such a way that they interpreted as input parameters of functions from system library (e.g., system( )), which is called, when a next command, which is indicated by a return address, is executed. This is «ret scheme attack» (the name is based on a name of assembler's command).

A vulnerability, which is also connected to those main types of vulnerabilities, is single-byte overflow, which appears because of an incorrect calculation of a number of elements in an array. As a result, it is impossible to rewrite return address, but it is possible to rewrite lower byte register *EBP*, which is in a case of *x*86 architecture stores an address of stack of software, which is called a function with a vulnerability. As a result, in a case of a return of a function with a vulnerability, stack frame can be modified and a next execution of ret and reti commands will pass a control to malicious code, which is the same as for another types of stack overflow (via environment variable, return to system library or putting of malicious code directly into stack).

«Heap buffer overflow» subclass is different from stack overflow in that an aim of a criminal is not return address but, in the first place, data in working memory. The subclass is divided into 4 types according to a type of rewritten data:

- -first type is rewriting of local function data, containing vulnerability. In a case of passing through borders of overflown array, adjacent set internal variables are deleted. In many cases it is enough to obstruct authorization system, which leads to authorization without a correct verification of user rights.
- -in the second type of vulnerability, which is «heap overflow», a function is rewritten. As a result, a criminal is becoming able to pass control to an arbitrary point in memory and execute introduced code.
- -the third type of vulnerability consists of rewriting of virtual functions table, which, in a case of use of gcc compiler, is placed after variables of a class, which is using a virtual function. Risk factor is the same, i.e. passing of control to arbitrary code.
- -the forth vulnerability type is based on use of mechanism of dynamic memory allotment. Blocks of allocated memory are united in doubly-connected list, and, as a result of overflow in data area of that block,

service information is rewritten. It leads to a situation, when during an execution of free() command, which is emptying an overflown block, a substitution of address of one of system functions in global offset table takes place, which allows a criminal to pass control to an arbitrary code in a case of calling of a rewritten function.

Integer overflow is a next class of vulnerabilities. Vulnerabilities of that class are related with an interpretation passing of results of arithmetic operation through limits of format (as a rule, ignoring of such overflow), and is divided into 3 types:

- overflow in a case of setting a variable's value bigger, than it is capable to enclose (e.g., in a case of setting integer type value to byte type variable a truncation of higher high-order bits takes places, digit of initial number is unchanged).
- overflow as a result of arithmetic operation, if a result is enclosed into unsigned variable (e.g., if in a case a result of an addition of two positive numbers is bigger than format, bit, which is passing out of its limits, is truncated);
- overflow in a digit in a case of executing of a arithmetic operation, if the result can be enclosed in character variable (that kind of overflow takes place, if a result, which is unsigned number with one in high-order bit can be enclosed in a character variable, in that case, that bit is interpreted as sign and the number becomes negative).

A change of a variable inside a software can lead to a situation, in which normal run of a software will be compromised, e.g. not enough space can be allocated for operations with string, which leads to buffer overflow, authorization mechanism logic will be changed, as a results, a criminal can, for example, conduct unauthorized access into a system with the highest user rights (superuser rights).

A vulnerability of string format takes place because of an absence of functions' specificators, which are conducting formatted input, or because of a specification of smaller number of specificators, as compared to a number of inputted variables. Nature of vulnerability is in the mechanism of data transfer into a function during a variable input of parameters. If a specificator is set as one of parameter's values, a function will search in memory to find data itself for processing with that specificator. As a result, a criminal can read any byte of memory or (by means of using «%n» specificator) rewrite any byte by means of a transfer of various number and specificators as a parameter of a function of a formatted input. Thus, knowing a structure of address space of an attacked process, a criminal can obtain parameters of authorization of legit users or ever rewrite return address or addresses of system functions.

The last vulnerability class, which is directly related with features of compiler's operation and placement of service information during an execution of software, is double free. Mechanism of realization of that vulnerability is close to «heap overflow» type, which is related to mechanism of dynamic memory allotment. If a software empties the same block twice, service data after a first emptying can be modified in a such way, that repeated calling of free() function leads to a situation, when a function's address in a global table offset table will be replaced by an address of introduced code.

Race conditions is a next class of vulnerabilities. That class of vulnerabilities is based on a transparency of resources of operating system for processes, each of which is working in a way like it has an exclusive time-constant accesses to OS resources. In that vulnerabilities class file race condition type, which is related with a verification of a file's existence and its following opening for witting. If a file doesn't exist, but in time between a verification and opening external process is able to create a file (which can be a link to a certain system file), an attacked application will send its output into system file, destroying of changing information in a file.

Signal race condition vulnerability can be used in a same manner, in that case a software after receiving a signal of completion stops its operation and after repeated launch continues it from «the same point» without carrying out a verification for unchanged external conditions (in particular, its temporary files). A presence of that kind of vulnerability can lead to an interruption of running of software with a vulnerability. In the same time, a criminal, depending on user rights, for example, can overwrite system file, change a system's configuration or even seize control with superuser rights.

Finally, the last class of discussed vulnerabilities is data interpretation vulnerabilities. Those vulnerabilities consist of a transfer of input data to an external interpreter without necessary verifications and restrictions. Vulnerabilities of that class can be described as follows. Logic of software implies, that data, entered into software, will be connected with some kind of statistic block and transferred to interpreter. It can be a value for a certain field of SQL-query (e.g., a software gives interpreter request «select permits from users where name =

user_input») or name of a catalog, which contents are produced by a software (shell-command «ls –al *user_input*» is formed). However, if not just number or name is passed into software, but text, which is separated by special signs (e.g., semicolon), it can be processed by interpreter as other command. In a case of processing of such a inputed fragment unauthorized reading or modification of data can take place.

2.3 Implementation of Mills Model for An Evaluation of Security of Software

In respect to an evaluation of security of software, an implementation of Mills model with a consideration of use of introduced errors will be as follwos (Gurov, Gurov, & Ivanov, 2012). Let's introduce S vulnerabilities into a software, including s_i vulnerabilities of i type ($S = \sum_{i=1}^{m} s_i$), where m is a number of types of vulnerabilities. Then, if during testing it is detected that n own vulnerabilities, including n_i vulnerabilities of i type ($n = \sum_{i=1}^{m} n_i$), and v introduced vulnerabilities, including v_i introduced i type ($v = \sum_{i=1}^{m} v_i$), then a number of own vulnerabilities of i type, which were initially in a software, is $N_i = \frac{n_i}{v_i} * s_i$, and a total number of vulnerabilities N, which were in a software before beginning of testing will be as follows: $N = \sum_{i=1}^{m} N_i = \sum_{i=1}^{m} (\frac{n_i}{v_i} * s_i)$.

Number of own vulnerabilities of i type, which remained after testing, will be following: $k_i = N_i - n_i = \frac{n_i}{v_i} * s_i - n_i$

A total number of own vulnerabilities, remained in a software after testing, will be as follows: $K = N - n = \sum_{i=1}^{m} (\frac{n_i}{v_i} * s_i) - n$

That value can be used for solving of the second problem, which is an evaluation of the model's confidence level. Let's presume that software has not more that k own vulnerabilities and introduce S vulnerabilities into it. Testing will continue until all vulnerabilities will be found. After testing is over, let's calculate number of detected own vulnerabilities, which are contained in a software n. Significance level C, which defines the model's confidence level and a probability, that the model will correctly reject a false presumption can be calculated by means of the following equations: $C = \begin{cases} 1, & \text{if } n > k \\ \frac{S}{S+k+1}, & \text{if } n \le k \end{cases}$.

Thus, if we presume, that after first step of testing in a software there is no more than 2 vulnerabilities (k = 2), than, in order to achieve significance level of C=0.9, it is necessary to introduce 27 vulnerabilities and find all of them.

However, testing before all introduced vulnerabilities will be found is a quite complicated task. For a situation, when at testing phase only j from S of introduced vulnerabilities are found ($j \le S$), for an evaluation of validity of the result the equation can be used, which was introduced in the work

(Teichroev, 1972):

$$C = \begin{cases} 1, & \text{if } n > k \\ \binom{S}{j-1} / \binom{S+k+1}{k+j} & \text{if } n \le k \end{cases},$$

where expression of $\binom{a}{b}$ type is calculated as follows:

$$\binom{a}{b} = \frac{a!}{b!*(a-b)!}.$$

By means of obtained relationship, for example, a number of vulnerabilities, that should be introduced to achieve a desired reliability, or any other parameters, necessary for subsequent analysis, can be defined.

In the discussed case, a task is complicated by a fact that it is necessary to calculate j_i, which is a number of vulnerabilities found during testing of initially introduced S_i vulnerabilities of each of m types. In this case, the reliability of an evaluation of received number of own vulnerabilities of i type, which were initially present in a software, should be evaluated individually.

$$C_i = \begin{cases} 1, & \text{if } n_i > k_i \\ \binom{S_i}{j_i - 1} / \binom{S_i + k_i + 1}{k_i + j_i}, & \text{if } n_i \le k_i \end{cases},$$

where n_i is a number of found own vulnerabilities of i-type, k_i is an estimated number of vulnerabilities of i -type.

3. Results and Discussion

In the presented paper is suggested to use Mills model for an evaluation of software reliability for an assessment of degree of security of software systems. In order to improve quality of the work of the model, it is proposed to use a division of all vulnerabilities into separate groups in a case of reduced introduction of vulnerabilities into a software. Such an approach would make it possible, on the one hand, to minimize a number of tests to verify security of software system, and, on the other hand, will allow to simplify a process of a creation of tests for vulnerabilities for each group. Classification of vulnerabilities based on a method of their connection with features of a process's execution on a level of interaction of system resources (memory, processor) and operating system is presented. That classification allows to detect typical attributes of vulnerabilities and, consequently, mechanisms for their appearances in a certain test impacts.

The expressions obtained in the presented study allow to define a number of vulnerabilities, remaining in the program after a test phase, as well as to determine a number of vulnerabilities of each type, that should be artificially to introduced into software before start of a test, to obtain a specified level of a reliability of results. Presented relationships allow to evaluate that level even for a case, when not all artificially made vulnerabilities were detected during testing, which, as a rule, is actual test conditions.

Using the obtained relationships, it is possible, in particular, to introduce vulnerabilities of a certain type before testing phase in order to obtain the most reliable results for vulnerabilities that pose the greatest treat for a given application. Level of that risk is defined either by an expert or according to data by organizations authoritative in that area (see, for example, CWE/SANS Top 25, 2011).

Thus, the method proposed in the presented study allows to eliminate the main disadvantage of Mills model, i.e. uncertainty, related with an introduction of artificial errors, while maintaining simplicity of Mills model.

Application of the proposed methodology implies the following steps:

- definition of danger level for a given application for all possible types of vulnerabilities presented in Figure 1;
- on that basis, calculation of a number of vulnerabilities of each type, which should be introduced in an initial software in order to obtain desired reliability of testing results;
- testing of software until a moment, when a number of detected vulnerabilities of each type will allow to provide specified level of security.

General issues of applicability of certain models for software reliability for an evaluation of security were discussed before in the works (Kazarin, 2012; Markov, 2011; Sharma, & Kishor, 2007). However, the approach proposed in the presented paper shows that it is reasonable to use exactly Mills model, it also allows to improve accuracy of obtained results by means of differentiated approach for a selection of initial data on a basis of the classification presented in the paper.

4. Conclusion

Currently, the security of software against unauthorized access is a topical subject. Criminals do not always want only to steal information. Often their purpose is a violation of a computing system, seizing control over it, distortion or deletion of information stored in it. Appearance of vulnerabilities can occur in a wide variety of situations and for a variety of reasons.

In a stage of a development of a software a topical problem is search for vulnerabilities and their elimination, which are based on features of interaction between processor and memory, order of an execution of applications in a processor, placement of data in memory, an organization of access of processes to resources and so on Such vulnerabilities has do not lead to an interruption of system security during normal operation, i.e. as it was planned by a programmer. They are, as a rule, are not detected during normal testing, as they do not violate logic of the operation of a software. However, in a case of use of specific input data, for example, data of a very large size, various non-standard situations can take place, which in some circumstances can lead to a system's compromise and, in particular, a criminal obtaining full access to a system. Thus, for a developer, in order not to allow presence of any vulnerabilities in a final software product, it is necessary to take additional special measures for their detection and localization.

Detection of vulnerabilities in software is a complex and challenging task, especially, because a problem of ensuring security of software is relatively new. In the presented work it is suggested to use approaches for a detection of vulnerabilities, which are analogous to those, that are used for an evaluation of software reliability for a long time. Among a large number of models of software reliability Mills model is marked out as the most suitable for that case. In order to minimize the least formalized part of the model, which is artificial introduction of inconsistencies in tested software, the classification of vulnerabilities based on principle of interaction system resources and operating system is proposed.

It is demonstrated that, in that case, an evaluation of large number of vulnerabilities, remained in a software after its testing, can be carried out, as well as recommendations for additional testing of software with a view to detect those remaining vulnerabilities, which are the most critical for that specific application of software system, are given.

Further studies in that direction are planned to comprise formalization of description of vulnerabilities of different classes and testing impacts, which are ensuring their detection, as well as verification of approaches, presented in the paper, in the widest possible range of specific software systems.

Acknowledgments

The authors would like to express gratitude to Doctor of Technical Sciences Professor Mikhail Ivanov, as a specialist in a security of software systems, for careful attention to our attempts to use non-traditional software reliability models to assess state of security of software and valuable comments made by him in that regard.

References

Classification of detected objects (2014, May 20). Securelist. Retrieved October 20, 2014, from http://www.securelist.com/ru/threats/detect?chapter=32

Classification of viruses' names according to «Doctor Web». (n.d.). *Website of Dr. Web*. Retrieved October 20, 2014, from http://vms.drweb.com/classification/

Engle, S., & Bishop, M. (2008). *A Model for Vulnerability Analysis and Classification.* Retrieved October 20, 2014, from http://www.cs.ucdavis.edu/research/tech-reports/2008/CSE-2008-5.pdf

Gurov, V., Gurov, D., & Ivanov, M. (2012). Implementation of software reliability models for an evaluation of security of software. *Security in information technologies, 1*, 88-91. ISSN 2074-7128

Gurov, V., Gurov, D., Ivanov, M., & Shustov, L. (2010). Secure programming technology and its teaching features in a case of higher educational institutions. *Remote and virtual training, 9*, 35-49. ISSN 1561-2449.

Holm, H., Korman, M., & Ekstedt, M. (2015, February). A Bayesian network model for likelihood estimations of acquirement of critical software vulnerabilities and exploits. *Information and Software Technology, 58*, 304-318. http://dx.doi.org/10.1016/j.infsof.2014.07.001

IEEE Std 610.12-1990. (1990). *IEEE Standard Glossary of Software Engineering Technology (ANSI).*

ISO/IEC 9126-93. (1993). *Information technology. Software product evaluation. Quality characteristics and guidelines for their use.*

Jason, L. (2013, May). Wright, Miles McQueen, Lawrence Wellman Analyses of two

end-user software vulnerability exposure metrics (extended version). *Information Security Technical Report, 4*(17), 173-184. ISSN 1363-4127.

Kazarin, O. (2012). *Security of computer systems' software*. Moscow: Moscow State Forest University, 2003. Retrieved October 20, 2014, from http://citforum.ru/security/articles/kazarin/2.shtml#2

Knight, E. (nd.). *Computer Vulnerabilities.* Retrieved October 20, 2014, from http://www.ussrback.com/docs/papers/general/compvuln_draft.pdf

Lohmor, S., & Sagar, B. (2014). Overview: Software Reliability Growth Models. *International Journal of Computer Science and Information Technologies – IJCSIT, 5*(4), 5545-5547.

Markov, A. (2011). Model of an evaluation and planning of software testing program according to information security requirements. *Bulletin of Bauman Moscow State Technical University. Series «Instrument engineering».* Retrieved October 20, 2014, from http://www.npo-echelon.ru/doc/software_reliability_models.pdf. ISSN 0236-3933

Mills, H. (1972). *On the Statistical Validation of Computer Programs, FSC-72-6015.* IBM Federal Systems Div., Gaithersburg, Md.

Myers, G. (1976). *Software reliability. Principles and practices*. New York, London, Sydney, Toronto: A Willey Inter science Publication, John Willey & Sons.

Teichroev, D. (1972). Survey of Languages for Stating Requirements for Computer-Based Information Systems. *Proceedings of the 1972 Fall Joint Computer Conference. Montvale, N.J.: AFIPS Press*, 1203-1224.

Top 25 Most Dangerous Software Errors. (2011). *CWE/SANS – Common Weakness Enumeration*. Retrieved October 20, 2014, from http://cwe.mitre.org/top25/

Vibhu, S., & Kishor S. (2007, April). Trivedi Quantifying software performance, reliability and security: An architecture-based approach. *Journal of Systems and Software, 4*(80), 493-509. http://dx.doi.org/10.1016/j.jss.2006.07.021

Yang, N., Yu, H., Qian, Z., & Sun, H. (2012, January). Modeling and quantitatively predicting software security based on stochastic Petri nets. *Mathematical and Computer Modeling, 1-2*(55), 102-112.

Younis, A., Kifayat, K., & Merabti, M. (2014, February). An access control model for cloud computing. *Journal of Information Security and Applications, 1*(19), 45-60. http://dx.doi.org/10.1016/j.jisa.2014.04.003

Zhang, D., Liu, D., Csallner, C., Kung, D., & Lei, Y. (2014, January). A distributed framework for demand-driven software vulnerability detection. *Journal of Systems and Software, 87*, 60-73. http://dx.doi.org/10.1016/j.jss.2013.08.033

Zhu, H., Zhang, Q., & Zhang, Y. (2014). Chapter 5 – HASARD: A Model-Based Method for Quality Analysis of Software Architecture. *Relating System Quality and Software Architecture*, 123-156.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Analysis of Effectiveness of Interphase Transfer in a Case of Purification of Biogas in Microbubbling Equipment with a Consideration of Chemisorption of Carbon Dioxide

Botagoz Myrzakhmetovna Kaldybaeva[1], Alisher Evadilloevich Khusanov[1], Darkhan Sabyrkhanovich Sabyrkhanov[1], Marat Isakovich Sataev[1] & Zakhangir Evadilloevich Khusanov[1]

[1] Auezov South Kazakhstan State University, Kazakhstan

Correspondence: Botagoz Myrzakhmetovna Kaldybaeva, Auezov South Kazakhstan State University, Kazakhstan.

Received: January 15, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p221 URL: http://dx.doi.org/10.5539/mas.v9n8p221

Abstract

The presented paper describes energy saving technology of biogas purification by microbubling method Description of a new design of microbubling equipment for gas-liquid solution and mathematical simulation of a process during chemisorption of carbon dioxide are presented. Experimental studies of mass transfer characteristics of microbubling equipment, based on model systems, were conducted with an aim to define a possibility of application of microbubling process for removal of carbon dioxide from biogas in order to obtain highly concetrated metane. As a result of the study optimal process parameters are defined, key factors affecting mass transfer characteristics of membrane microbubling method are also established. Efficiency of membrane microbubling equipment from the point of view of interphase mass transfer is assessed.

Keywords: biogas, microbubling equipment, tubular ceramic membranes, highly concentrated methane, unconventional energy sources, mathematical simulation, mass transfer, liquid, gas, microbubbles

1. Introduction

Currently, Kazakhstan has enormous potential for a development of agriculture, vast territories allows that country to become a leading country, both in livestock and in crops production. In order to achieve that goal, it is necessary to implement the most advanced technologies and innovations in a field of biogas implementation for energy needs by means of processing livestock wastes in biogas equipment, as a result the following products can be obtained: biogas, mineralized nitrogen fertilizer, methane, carbon dioxide, electricity, heat energy. Therefore, rational use of agricultural waste products is todays important problem, if it's solved, there is anenormous potential for an implementation of biomass for a production of liquid and gaseous fuel (biogas), which will allow companies to reduce expenses related with use of electric energy and fertilizers, as well environment related taxes, aimed to a reduction of a contamination of water bodies, a contamination of soil with pathogenic bacteria and helminths, which are contained in manure flows of livestock farms. Thus, it allows to reduce cost of products.

Biological processing is not something novel nowadays. It is already widely used throughout the world, but, largely, as a way to recycle a large amount of wastes of agricultural and food industries, and as an alternative source of energy, which is one of possible solutions to a deficit of energy sources (Rutz, & Janssen, 2014), (Steinhauser, 2008). This method essentially consists of anaerobic expansion of organic wastes, mainly, produced by farms, food production facilities and sewage systems. As a result of vital activity of micro-organisms, wastes are recycled with a formation of mixture of gases (so-called biogas) and their volume is reduced. Biogas, obtained by anaerobic decomposition of wastes, contains methane (≈60%(vol.)) and carbon dioxide (≈40%(vol.)). Resulting products, which are biogas and half fermented semiliquid substance, are of great value as gaseous fuel and organic fertilizer. The gas contains hydrogen sulphide, ammonia, water vapor; its calorific value is quite low – 19.5-19.8 MJ/m^3 and burning of such gas may cover only a small portion of energy needs. After purification and drying, gas must contain not less than 98 % (vol.) of CH_4 (calorific value is not less than 33.0 MJ/m^3), a concentration of N_2S should not exceed (Biogas plants in Europe: A practical handbook, 2007), (Eder, & Schulz, 2006), (Ditnerski, et al., 1991) 10^{-4}% (3-5 million^{-1}). Separation of components and

production of highly concentrated methane is a big problem, which is still not solved (Christensen, et al., 1996), (Concise Encyclopedia of Bioresource Technology, 2004), (Kimura, 1980). Cost-effectiveness of biogas equipment is of certain interest in a context of agricultural enterprises, taking into account an integrated use of biogas as energy source and fertilizer in agriculture. Nowadays, in the world there is a large interest for problems of methane fermentation of manure wastes and other organic wastes. In addition to an equipment, which includes fermentation chambers, gasholder and storage for half-fermented substance (sludge), it is also necessary to widely implement a technology of methane concentration using methods of separation of multi-component gases.

Use of well-known methods of gas separation (Ditnerski, et al., 1991), (Christensen, et al., 1996), (Concise Encyclopedia of Bioresource Technology, 2004), (Kimura, 1980) in that case is ineffective and requires a development of new methods of purification. An alternative way of solving that problem is a combination of membrane and absorptive processes.

Recently, scientific community of developed countries of the West (Atchariyawut et al, 2007), Japan (Kukizaki, & Goto, 2006 (a), (Kukizaki, & Goto, 2006b) produced a number of publications, reporting that during a dispersion of a gas through porous membranes microbubbles are formed (Rodriguez, & Rubio, 2014), (Zong, et al., 2006), (Kim, et al., 2000), (Cho, et al., 2005) with dimensions of 0.5-150 μm. As a result of such small sizes, microbubbles have a number of unique properties, such as increased contact surface of interacting phases, they can be widely applied in chemical (Loubiere, & Hebrard, 2004), (Painmanakul, et al., 2005), food and pharmaceutical (Unger, et al., 2003) industries, as well as in biotechnology, medicine and unconventional energy production.

As can be seen from the presented review, study of gas purification with a formation of microbubles is currently still cannot be considered complete, as an integrated approach, i.e. physical and mathematical simulation, as well as design of gas separating equipment, which provides formation of microbubles and its use for purification of biogas from carbon dioxide emissions with an aim to increase methane concentration, is expected to provide new and important results in that field.

The development of technology, which allows to increase concentration of methane through a development of interphase surface, using unique microbubles properties during fine dispersion of gases, will lead to an entirely new high-efficient equipment of gas-liquid contact type, including new types of reactors and fermenter. At the same time, a preparation of biogas (removal of CO_2 , H_2S and further compression for storage and distribution to consumers) using a the developed microbubling equipment can provide a significant economic effect, as compared to traditional methods, for example, absorption and adsorption. The are several variants of process organization, for each of which defined parameters are required surface of membranes, cost of compression, degree of methane extraction from original mixture in different conditions. As a result of separation concentration of methane in fuel gas reaches 98% (vol.).

The problem includes a design of microbubling equipment and effective purification technology for biogas by means of a substantial development of surface of phases in gas-liquid system. The solution of those problems lies in dispersion of biogas components through membrane in fluid, its inertia effects forming dispersion of nano- and micro-bubles, resulting in significant increase of phase interface surface for separation of highly-concentrated methane.

Due to ultra small size of generated microbubes, microbubling membrane process can be used as a basis for a development of a highly efficient biogas purification technology. In the same time, it is possible that there are no disadvantages related to limiting of loads by gas and liquid for such a technology. Unique properties of microbubles during fine dispersion of gases allows to implement that technology also in oil and gas and mining industries for a purification of associated gas and mining gas, as well as in fields of biotechnology and pharmaceutics.

Mathematical simulation of dispersion mechanics of biogas components in the developed technology of separation of highly-concentrated methane from biogas and a definition of a relationship of main mass transfer parameters, i.e. specific interphase surface, interphase flow of absorbed substances (CO_2, H_2S, NH_3 etc.), mass transfer coefficients, and speed of fluid and concentration of active part of absorbent in microbubling processes are not established so far, which makes it a topical problem.

The developed process of biogas purification with an aim to obtain highly-concentrated methane in microbubling equipment allows to eliminate unfavourable influence of bioenergetics on environment: emission of solid particles, carcinogenic and toxic substances, carbon dioxide, biogas, biospirit, impoverishment of organic components of soils, exhaustion and erosion of soils, explosion risk, large number of wastes in a form of

by-products (waste waters, distillation residues).

2. Theoretical Analysis

Analysis of features of mass transfer processes in microbubling equipment allows to conclude, that in a case of a design of an equipment by a type of shell-and-tube module, average time of presence of microbubles in apparatus is quite small. In that case, considering that the flow's structure in an equipment is close to perfect displacement model (PDM), average time of presence can be calculated as follows: $t_{pr} = \frac{l}{2\omega}$. Considering standard length of ceramic membranes of 0.8 m and fluids' speeds of 1-3 m/s, time of presence is 0.1-0.4 s, the time limit might be not enough for effective physical absorption. However, in a case of chemisorption, time of gas-liquid reaction is of 0.01 s magnitude (Astarita, 1971). Thus, from a point of view of effectiveness of mass transfer processes in membrane module, the most prospective are chemisorption processes in a case of comparatively fast chemical reaction. Moreover, membrane microbubling equipment can be used not only as absorbers, but also as chemical reactors. Thus, from a point of view of study of interphase mass transfer in microbubling contactor, it is necessary to discuss main existing chemisorption theories.

Two directions in a development of chemisorption theory can be marked out, they are based on two models of mass transfer. The first direction is based on film theory of Uitmen and is developed in works of Hatta (Hatta, 1932) and Van Krevelen (Krevelen, & Hoftijzer, 1958). The second direction is based on a penetration theory, it discuses non-stationary process of absorption during continuous renovation of interphase surface. That direction is mainly developed in works of Danckwerts (Danckwerts, 1973). It's worth mentioning, that all existing models can't be considered complete and posses a number of disadvantages and number of studies, which would allow to select one of the directions, is insufficient. At the same time, in some practical cases, results, obtained by means of different models, are virtually the same.

So far there is no data on studies of interphase mass transfer during membrane dispersion of gas. That data would allow to compare effectiveness of mass transfer equipment. Hence, the presented study is dedicated to experimental study of mass transfer in membrane microbubling contactor on an example of CO_2 absorption by CaO suspension. CaO content in the suspension is 10-20% (mass).

Study of interphase mass transfer in the presented study was carried out on an example of chemisorption of carbon dioxide by alkali solutions of various concentrations, in order to do that the method, proposed by Danckwerts and Sharma was used (Sharma, & Danckwerts, 1970).

The process comprises two consecutive reactions:

$$CO_2+OH^-=HCO^-_3 \quad (1)$$

$$HCO^-_3+OH^-=CO_3^{2-}+H_2O \quad (2)$$

Thus, total reaction is occurring according to the formula:

$$CO_2+2OH^-=CO_3^{2-}+H_2O. \quad (3)$$

In a case of sufficient excess of alkali, the second reaction is occuring almost momentarily, that's the process is limited by the first reaction, which is speed is defined by means of the expression:

$$N_R = k_2 [CO_2][OH^-]. \quad (4)$$

Thus, in a case of excess of OH^- ions, the reaction can be considered as a reaction of pseudo-first order (Sharma, & Danckwerts, 1970). According to the main mass transfer equation, molar flow of CO_2 from phase to phase can be defined as follows:

$$M = k_y(C_{0,G} - C_{i,G})F = k_x^{/}(C_{i,L} - C_{0,L})F, \quad (5)$$

where k_y, $k_x^{/}$– coefficients of mass transfer in gas and liquid phases respectively, $C_{0,G}$ – CO_2 concentration in a volume of gas phase, $C_{i,G}$ – CO_2 concentration in gas phase in interphase border, $C_{0,L}$– CO_2 concentration in a volume of liquid phase, $C_{i,L}$– CO_2 concentration in liquid phase in interphase border, F – interphase surface area. If chemical reaction is fast enough, it can be accepted that $C_{0,L}$ = 0, considering that concentrations in interphase surface are related by equilibrium – $C_{i,G} = mC_{i,L}$, the following can be obtained from expression (5):

$$M = \frac{G_{0,G}F}{\left(\frac{1}{k_y}+\frac{m}{k_x'}\right)}. \quad (6)$$

In the following, let's discuss the solution, leading to expression for mass transfer coefficient for chemisorption. The discussed reaction is very fast nonreversible chemical reaction in liquid phase of the following type (Hatta, 1932):

$$\nu_A A + \nu_b B \rightarrow \nu_d D \ , \quad (7)$$

At the same time, it's considered that reaction zone starts directly in a vicinity interphase surface and, depending on speed of reaction and concentration of components, is extended on a different depth. The process is considered stationary. The solution can be applied to reaction of first (of A $\rightarrow$ ν_dD type) and pseudo-first type.

In a case of nonreversible reaction of first order *n* by substance *A*, equation of convective diffustion will take the following form (Richardson, & Coulson, 1999):

$$D_A \frac{d^2 C_A}{dy^2} = k_P C_A^n \quad (8)$$

where reaction speed constant k_P is defined according to expression $k_P = k_{n+m} C_B^m$, in which k_{n+m}– reaction speed constant of order of $n+m$. Boundary conditions:

$$y = 0 \qquad C_A = C_{Ai}$$
$$y = \infty \qquad C_A = 0$$

By means of specifying $\frac{dC_A}{dy} = q$, expression (8) can be presented in the following form:

$$q\frac{dq}{dC_A} - \frac{k_P}{D_A} C_A^n = 0. \quad (9)$$

Integration of that expression considering boundary conditions leads to the expression:

$$\left(\frac{dC_A}{dy}\right)^2 = \frac{2}{n+1}\frac{k_P}{D_A} C_A^{n+1}. \quad (10)$$

Because $\frac{dC_A}{dy}$is negative:

$$\left(\frac{dC_A}{dy}\right) = -\sqrt{\frac{2}{n+1}}\sqrt{\frac{k_P}{D_A}} C_A^{\frac{n+1}{2}}. \quad (11)$$

Thus, interphase chemisorption speed can be defined as follows:

$$N_A' = -D_A\left(\frac{dC_A}{dy}\right)_{y=0} = \sqrt{\frac{2k_P D_A}{n+1}} C_{Ai}^{\frac{n+1}{2}}. \quad (12)$$

In a case chemical reaction of first (or pseudo-first) order is not fast enough, $C_{A0} \rightarrow 0$, and from expression (12) it follows, that chemisorption speed can be calculated as follows:

$$N_A' = \sqrt{k_P C_{Ai} D_A}, \quad (13)$$

Then, considering, that in a case of fast enough reaction C_{A0}=0, the following expressing for coefficient of mass transfer for absorption, which is accompanied by reaction of pseudo-first order (also, considering that the reaction has the first order by substance *B*):

$$k_x' = \frac{N_A'}{(C_{Ai} - C_{A0})} = \sqrt{k_p D_A} = \sqrt{k_2 C_B D_A} \quad (14)$$

In all relationships presented above it is presumed, that substance *B* is in excess and C_B = const. In the study (Krevelen, & Hoftijzer, 1958) an investigation into influence of concentration of C_B on ration of values $\frac{k_x'}{k_x}$ and $\frac{\sqrt{k_2 C_B D_A}}{k_x}$ (as C_B an average value is used). The studies were carried out using carbon dioxide – sodium hydroxide, results are presented in Figure 1.

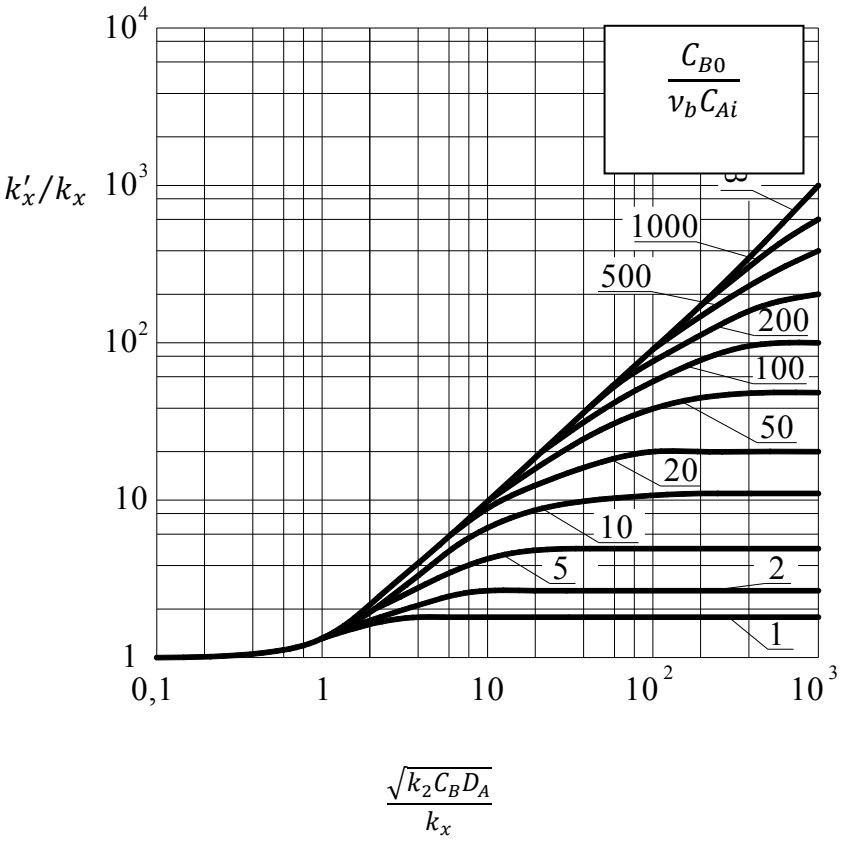


Figure 1. Relationship of value $\frac{k'_x}{k_x}$ from $\frac{\sqrt{k_2 C_B D_A}}{k_x}$ with different values $\frac{C_{B0}}{\nu_b C_{Ai}}$.

Analysis of the diagram shows, that even in a case of slow reaction, independently of concentration B, value of $\frac{k'_x}{k_x} \approx 1$and process are close to physical absorption. Increase of reaction speed leads to that value of $\frac{k'_x}{k_x}$ becomes approximately equal to $\frac{\sqrt{k_2 C_B D_A}}{k_x}$, which demonstrates a possibility of application of expression (14) for calculation of mass transfer coefficient. In that area the process is limited by speed of chemical reaction. With a further increase of speed value of $\frac{k'_x}{k_x}$ becomes constant, which is close to $\frac{C_{B0}}{\nu_b C_{Ai}}$, at the that, the process is limited by mass transfer B to reaction zone.

According to aforementioned points, expression (5) can be used with acceptable accuracy in the following range:

$$5 < \frac{\sqrt{k_2 C_B D_A}}{k_x} < \left(\frac{C_B}{\nu_b C_{Ai}}\right). \quad (15)$$

where k_x– mass transfer coefficient for physical absorption. Substitution of expression (14) in expression (6), considering that $F = aV_{work}$, produces the following:

$$M = \frac{C_{0,G} a V_{work}}{\left(\frac{1}{k_y} + \frac{m}{\sqrt{k_2 C_B D_{CO_2}}}\right)} \quad (16)$$

Expression (16) can be transformed in the following form:

$$\frac{C_{0,G}V_{work}}{M} = \frac{1}{k_y a} + \frac{m}{a\sqrt{k_2 C_B D_{CO_2}}} \quad (17)$$

In coordinates $y = \frac{C_{0,G}V_{work}}{M}$, $x = \frac{m}{\sqrt{k_2 C_B D_{CO_2}}}$ expression (17) describes line with angle of inclination tangent $\frac{1}{a}$ and crossing with coordinate axis in point $\frac{1}{k_y a}$. Thus, by means of experimental study of relationship of y from x specific interphase surface area in equipment can be defined, as well as mass transfer coefficient in gas phase, which allows to define effectiveness of interphase mass transfer in membrane microbubling equipment.

Following conclusion can be made on a basis of analysis of mass transfer with chemical reaction: existing theories regarding chemisorption, event though they can't be considered complete, provide sufficiently reliable results for a case of fast nonreversible reactions of first and pseudo-first order. At the same time, in fact, all models of mass transfer are demonstrating similar values for mass transfer coefficient for chemisorption. However, because an implementation of expression (4) requires conforming with condition (15), it is necessary to carry out additional experimental studies in order to define value of k_x. For conditions, specified by inequality (15), coefficient of mass transfer may be calculated according to expression $k'_x = \sqrt{k_P D_A} = \sqrt{k_2 C_B D_A}$, , which is conforming both with film and penetration theories. The expression shows, that in a case of fast chemical reaction, mass transfer coefficient doesn't depend on hydrodynamic conditions in flow core, but it is defined by speed of reaction and speed of diffusion of absorbed component.

3. Methodology

At the first stage of anaerobic half-fermenting of organic substances by means of biochemical decomposition (hydrolysis), first, decomposition of high-molecular compounds (carbohydrates, fats and proteic substances) into low-molecular compounds takes place (Rutz, & Janssen, 2014), (Christensen, et al., 1996), (Concise Encyclopedia of Bioresource Technology, 2007). At the second stage, with acid-forming bacteria taking part in the process, further decomposition with a formation of organic acids and their salts occurs, as well as spirits, CO_2 and H_2, then H_2S and NH_3. Final bacterial transformation of organic substances in CO_2 and CH_4 is carried out at the third stage of the process (methane fermentation). In addition, additional amount of CO_2 and CH_4 is formed in further from CO_2 and H_2. Those reaction occur simultaneously, at that, methane -forming bacteria form a significantly higher requirements for existence condition, as compared to acid-forming bacteria. For example, they require absolutely anaerobic media and longer time for reproduction. Speed and scaled of anaerobic fermentation of methane-forming bacteria depend on their metabolic activity.

For the study of interphase mass transfer during chemisorption of carbon dioxide by CaOH solution in membrane contactor, the method proposed by Sharma and Danckwerts was selected (Sharma, & Danckwerts, 1970).

Molar flow of absorbed CO_2in that study was experimentally defined by means of a change of concentration of alkali in solution, the following expression was used for calculations:

$$M = \frac{(C_B^H - C_B^K)V_L}{\nu_b}, \quad (18)$$

where C_B^H and C_B^K – initial and final alkali concentration, V_L – consumption of solution, ν_b – stoichiometric coefficient in overall equation, which is equal to 2.

Reaction speed constant of second order k_2 can be calculated using the following expression (Pohorecki, & Moniuk, 1988):

$$lg\left(\frac{k_2}{k_{2,\infty}}\right) = 0.221I - 0.016I^2. \quad (19)$$

Value of speed constant in infinitely diluted solution is defined as follows:

$$lg\left(k_{2,\infty}\right) = 11.895 - \frac{2382}{T} \quad (20)$$

Expression (20) can be used in temperature range 290-314 K. Ionic force is calculated according to the following expression (Darmana, et al., 2007):

$$I = \frac{1}{2}\sum_{j=1}^{n} C_j Z_j^2 \quad (21)$$

where C_j – concentration of Ca^+, HCO_3^-, OH^-, CO_3^{2-} ions, Z_j –valence of those ions. At the same time, it is noted (28) that, because of HCO_3^- and CO_3^{2-} ions' concentration are quite small, their values can be neglected. Thus, value of I becomes equal to concentration of alkali in a solution.

Diffusion coefficient CO_2 in alkali solution is defined according to the following equation (Bobilev, 2003):

$$D_{CO_2} = D_{CO_2}^{aq}\left(\frac{\mu_{aq}}{\mu_s}\right)^{0.85}, \quad (22)$$

where μ_{aq} – viscosity of water, μ_s – viscosity of alkali solution (was defined according to the data (Бобылев, 2003)), CO_2 diffusion coefficient in clear water is defined as follows (Versteeg, & Swaaij, 1988):

$$D_{CO_2}^{aq} = 2.35 \cdot 10^{-6} exp\left(-\frac{2119}{T}\right). \quad (23)$$

Distribution coefficient m was defined on a basis of experimental data on solubility of carbon dioxide in solutions of CaOH, presented in (28). In experiments on membrane with average pore diameter of 0.5 μm, in a case of alkali concentrations of 0.030-0.070 kmole/m^3, value of m=10.3 was used, in experiments on membrane with average pore diameter of 2.6 μm, in a case of alkali concentrations of 0.014-0.030 kmole/m^3, value of m=9.9 was used,

4. Results of Experimental Studies

The experimental studies, which were carried out in the presented work, were aimed at a definition of main mass transfer characteristics of membrane microbubling process..

From the point of view of interphase mass transfer, the objects of the study were realtionships of interphase molar flow of absorbed substance, mass transfer coefficients in gas and liquid media, as well as specific interphase surface area from speed of liquid in membrane module and from concentration of active part of absorbent. The goal of the experimennts was data for an estimation of effectiveness of interphase mass transfer in membrane microbubling contactor and its comparison with mass transfer equipment of other types.

For study of mechanisms of gas-liquid reaction in a case of microbubling the process of carbon dioxide chemisorption by CaOH solutions of various concentrations was selected. Because of small content of surfactants, physical properties of liquid phase (viscosity, density) were accepted the same as for clean water. Carbon dioxide and biogas mixture with CO_2 content of 10-40 % (vol.) was used in all experiments as gas phase. Studies were carried out with microporous membranes with internal selective layer and average pore diameter of 0.5 μm and 2.6 μm. For the membrane with 0.5 μm pores gas consumption (depending on liquid consumption) was 26.5-101.4 l/h (which was corresponding to pressure of 0.062-0.064 MPa), for the membrane with 2.6 μm pores gas consumption was 25.8-100.7 l/h (which was corresponding to pressure of 0.062-0.064 MPa). As liquid phase CaOH solutions with the following concentrations were used: 0.030; 0.040; 0.050; 0.060; 0.070 kmole/m^3 (for the memebrane with d_0 = 0.5 μm) and 0.010; 0.014; 0.018; 0.024; 0.030 kmole/m^3 (for membrane with d_0= 2.6 μm). Those alkali concentrations were selected on the condition that, from one point of view to provide sufficient effectiveness of chemical reaction, from another point of view – to ensure necessary sensitivity for measuring equipment. Consumptions of liquid during experiments were changing in range 77-300 l/h, which corresponds to speeds of 0.7-3.0 m/s. Definition of a volume of consumed CO_2 was carried out by means of measuring of final concentration of alkali in the solution, which was coming out of module, by means of electronic device for measuring of pH.

In order to carry out microbubling process memrbrane chemisorption device is designed (Figure 2), which uses membrane module made of ceramics. It consists of steel cylinder-shaped hull with 750 mm length and Ø50×3 mm diameter, inside the hull tubular ceramic membrane is installed. Direct contact between gas and liquid phases occurs in membrane module, which results in formation of fine dispersion of bubbles. Gas phase (mixture of CH_4 from CO_2 and H_2S) is fed with a required pressure inside the body from inside of membrane. Liquid phase is coming inside in ceramic membrane.

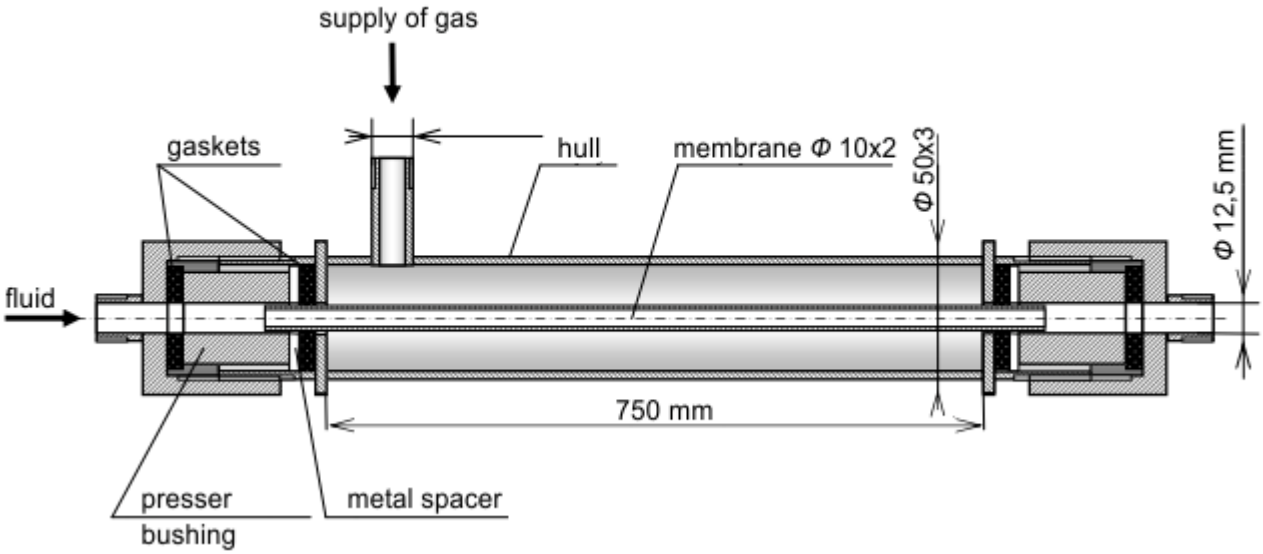


Figure 2. Membrane module for microbubling purification of biogas from CO_2

The experimental device consists of chemasorber (1) for carrying out of microbubbling process, vat (2), vortex pump (3) of first level of fire safety, closing (4,5,6,7,8) and regulating (9,10) valves, measuring and control devices (Figure 3).

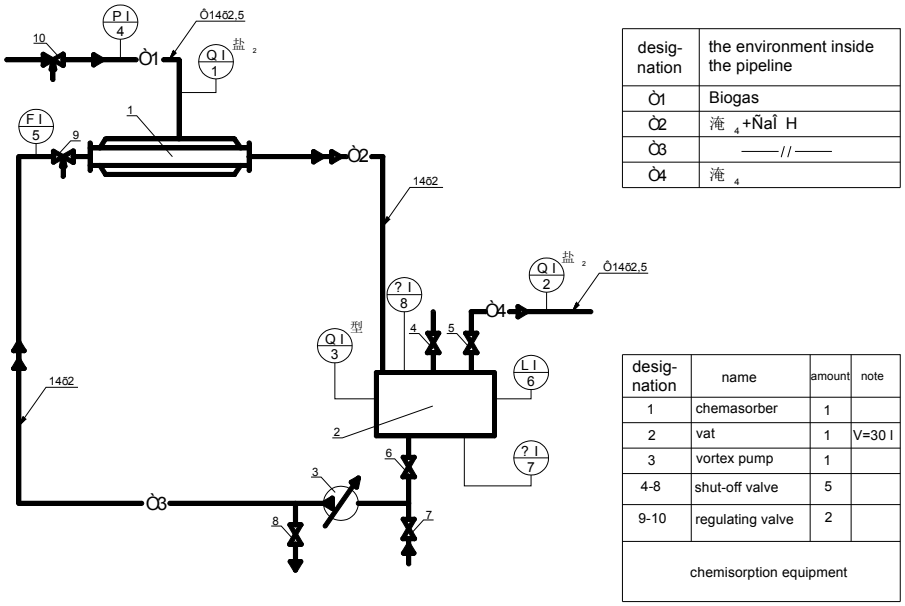
Gas mixture containing methane, CO_2, H_2S and other impurities is coming to the zone between tubes of chemasorber (1) through pipeline (T1), it is coming inside of pores of tubular ceramic microfiltering membrane and is coming out into the internal zone in a form of microbubles. Formed microbubles are continuously washed out by flow of water suspension of CaO, containing microbubles. In further, suspension together with bubbles returned to the vat (2) by means of vortex pump (3) with the following parameters: flow 0.001 m^3/s, manometric head 16 mm of water column, rotation speed 24.15 rotations/s. The selected pump is of explosion-proof version.

CaO content in the suspension is 10-20% (mass). Content of CO_2 in gas mix is measured in input (QI,1) and output (QI,2) by device for measuring gas concentration pH of media (QI,3, pH) is measured in the vat (2). Pressure is measured at input of gas into the chemasorber (1) by manometer (PI,4). Flow of suspension in at input of chemasorber (1) by means of rotameter (FI,5). Volumetric flow of the suspension with bubbles is 0.5÷1 l/s.

CaO, which is contained in suspension, before complete transformation into carbonate can purify 0.8 m^3 of gas mixture, containing 50% (vol/) of CO_2.

In the vat (2) level of suspension is also measured by water-gauge glass (LI,6), bottom pressure is measured by manometer (PI,7), pressure above gas-liquid layer is also measured (PI,8).

Filling of the vat (2) is carried out through valve (7) by the pump (3) with closed valves (5,6,8,10).



desig-nation	the environment inside the pipeline
Ò1	Biogas
Ò2	淹 4+ÑaÎ H
Ò3	——//——
Ò4	淹 4

desig-nation	name	amount	note
1	chemasorber	1	
2	vat	1	V=30 l
3	vortex pump	1	
4-8	shut-off valve	5	
9-10	regulating valve	2	
chemisorption equipment			

Figure 3. Process flow diagram of microbubling purification from CO_2 of gas mixture containing CH_4

For emptying of the vat (2) from used suspension the pump is used (3) with open valves (6,8) and not fully closed valve (10), other valves are closed.

The room where the device is operated there is an instrument for a measurement of CH_4 concentration with alarming, when critical concentration is reached.

As a results of experiments for study of interphase mass transfer, relationships between specific interphase surface area in the device and mass transfer coefficient in gas phase andspeed of liquid in channel of membrane, also value of interphase flow of absorbed substances depending on speed of liquid and concentration of active part of absorber is defined.

Figure 4 presents relationship between specific interphase surface area and speed of liquid for both used membrane. In the both cases the value of specific interphase surface is increasing with increase of speed up to values of ω approximately 2 m/s. Further increase of speed of liquid up to 3 m/s doesn't lead to a signifficant increase of interphase surface. As for value of *a* itself, for membranes with average pore diameter of 0.5 μm it is in a range of 18000-30000 m^{-1}, for average pore diameter of 2.6 μm it is in a range of 7700-19200m^{-1}.

An increase of specific interphase surface with an increase of speed of liquid can be explained by means of features of hydromechanics of microbubling process, in particular, decrease of sizes of moving microbubbles, because of increase of resistance stress of incoming flow of liquid during their growth. At the same time, because size of formed microbubbles depedns on sizes of pores of membranes, than for membrane with d_0 = 0.5 μm specific interphase surface is bigger then for a membrane with d_0 =2.6 μm.

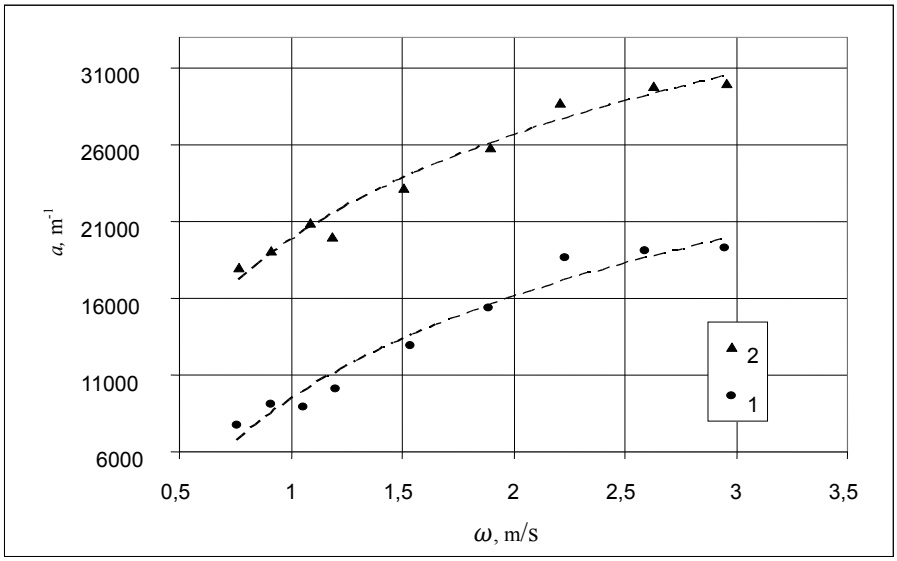


Figure 4. Relationship of specific surface of phase contact for CO2 and speed of liquid.

1) membrane with d_0 = 2.6 μm; 2) membrane with d_0 = 0.5 μm.

Relationship between interphase flow of carbond dioxide (kmole/s) from speed of liquid is presented in Figure 5.From the Figure it can be seen, that for both used membranes in all cases there is virtually linear increase of value M in the studied range of speeds (0.5-3 m/s). According to the obtained data, for a membrane with d_0=0.5 μm during chemisorption by alkali solution with 0.07 kmole/m^3 concentration M increases from $20\cdot10^{-8}$ to $70\cdot10^{-8}$ kmole/s; for a membrane with d_0=2.6 μm mass flow is also increase with an increase of speed, however, both M and speed of its increase are less as compared to the previous case: from $5\cdot10^{-8}$ to $20\cdot10^{-8}$ kmole/s during chemisorption by alkali solution with 0.07 kmole/m3 (Figure 5). Because in a case of chemisorption with significantly fast chemical reaction hydrodynamics of bottom layers of liquid doesn't seriously influence mass transfer coefficient in liquid phase, an increase of interphase flow with an increase of speed, presumably, is caused by two factors: first, increase of specific interphase surface, second, an increase of mass transfer coefficient in gas phase.

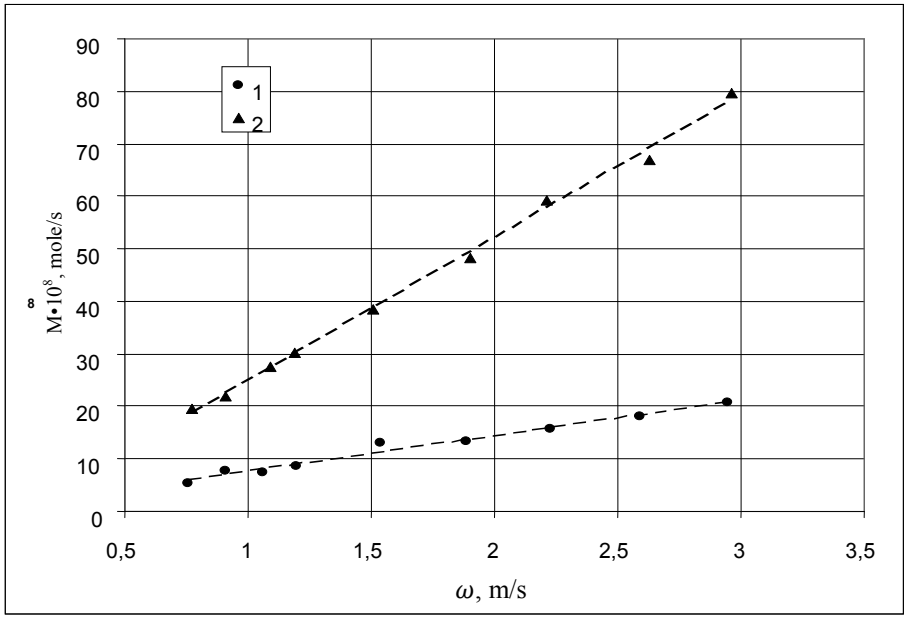


Figure 5. Relationship of interphase flow and speed of liquid

1) membrane with d_0 = 2.6 μm (C_B = 0.03 kmole/m^3); 2) membrane with d_0 = 0.5 μm (C_B = 0.07 kmole/m^3).

Let's discuss relationships of interphase flow of carbon dioxide from alkali concentration. From expression (17) with known values of a and k_y interphase flow can be described as follows:

$$M = \left(\frac{1}{k_y} + \frac{m}{a\sqrt{k_2 C_B D_{CO_2}}}\right)^{-1} C_{0,G} V_{work}. \quad (26)$$

In all cases interphase flow is increasing with an increase of concentration, which can be explained by an increase of mass transfer coefficient in liquid phase, which is calculated using the expression (14). Error in a description of experimental data is in a range 5-20%. Also, it can be seen, that experimental data, obtained with a membrane with pore sizes of 0.5 μm conform to the expression (26) better, then data for a membrane with pores of 2.6 μm size. Presumably, it is explained by higher error of the method in the second case, as in experiments with a membrane with d_0 = 2.6 μm alkali solutions of lower concentration were used (lower alkali concentrations were selected in order to ensure necessary sensitivity of measuring instruments, because for that membrane due to lower pressure volumetric concentration of CO_2 in incoming gas mixture is lower).

The fact that obtained experimental data is in good agreement with the expression (26) may be a good indirect proof of correctness of use of the selected methodology for mass transfer studies.

It is worth mentioning, that in a case of experiments with a membrane with 2.6 μm pores, in a case of small concentrations of alkali value of k'_x/k_x is close to the lower boundary of inequality, which can be an indicator of errors in that zone, which are related with that a condition of pseudo-first order of reaction is not satisfied.

5. Analysis of Adequacy of the Obtained Data

Further, let's compare membrane microbubbling contactor with other types of devices.

Let's discuss microbubbling layer on a plate and gas-liquid dispersion in membrane contactor, having the same volume and gas content. Specific contact surface in plate-type device depends on design of plate, speed of gas and physical properties of gas-liquid system. For example, mesh plates with comparatively low speeds of gas in zone of cellular foam contact surface is approximately 200 m^{-1} (Ramm, 1976).

Further increase of gas speed increases contact surface, reaching 800-100 m^{-1}. For concave type of plates a doesn't seriously depend on speed of gas and reaches approximately 1000 m^{-1} for air-water system and bigger values (up to 4000 m^{-1}) for systems with organic liquids. Thus, for place type equipment value of specific area of phase contact is in a range of 200-4000 m^{-1} and in average is 1000 m^{-1}. From the results of conducted experiments it follows, that contact surface in membrane microbubbling devices depends on membrane characteristics, speed of liquid, gas consumption and, in our case, was in a range from 8000 m^{-1} to 30000 m^{-1}. It leads to conclusion, that specific interphase surface in microbubbling device, depending on process conditions, will be 8-30 times higher than in bubbling layer on a plate.

Analysis of main mechanisms of mass transfer with chemical reaction allows to conclude, that during chemisorption mass transfer coefficient in liquid phase both in plate type and membrane type membrane microbubbling device can be calculated using expression (14), thus, values of k'_x will be comparable.

Calculations of mass-transfer coefficient, which were carried out on a basis of experimental data, shows, that for conditions, studied in the presented paper, mass-transfer coefficient in membrane contactor (calculated for gas phase) has values of $1.6 \cdot 10^{-5}$ – $4.5 \cdot 10^{-5}$ m/s for a membrane with 0.5 μm pores and $2.3 \cdot 1^{-5}$ – $4.5 \cdot 10^{-5}$ for a membrane with 2.6 μm pores. Mass transfer coefficients in plate-type device, calculated on a condition of equal k'_x and value $k_y = 1 \cdot 10^{-3}$ m/s are $4.6 \cdot 10^{-5}$ – $7.5 \cdot 10^{-5}$ m/s and $2.9 \cdot 10^{-5}$ – $5.1 \cdot 10^{-5}$ m/s respectively. Therefore, mass transfer coefficients in plate-type device in the discussed conditions will be 1.1-2.9 times higher than in membrane contactor.

Transformation of expression (26) gives the following:

$$M = \left(\frac{1}{k_y} + \frac{m}{\sqrt{k_2 C_B D_{CO_2}}}\right)^{-1} a C_{0,G} V_{work} = k_y a C_{0,G} V_{work}, \quad (27)$$

it can be concluded, that because value of k_y membrane device is in average 1.5 times smaller and specific interphase surface is in 8-30 times higher, than in a case of the same working volumes amount of absorbed substances will increase in 5-20 times. It allows to conclude that in order to reach the desired level of absorption, membrane microbubbling device must have 5-20 smaller working volume than plate type device. At that, the following recommendations can be given. Mass transfer coefficient obtained during experiments with both membranes are quite close, but in the same time values of specific interphase surface for a membrane with d_0=0.5 μm are in 2-2.5 times bigger. Thus, it can be presumed, that implementation of microfiltering membranes

with pores less than 1 μm will give more significant effect in decreasing sizes of a devices and, consequently, will reduce capital spendings. At that, the most optimal range of speeds of liquid, both from point of view of mass transfer coefficients and specific interphase surface is a range 1.5-2.5 m/s. At the same time, in a case of use of membrane with 0.5 μm pores necessary gas pressure is three times higher, as compared to a membrane with 2.6 μm pores. It can lead to significant energy expenses for gas blow off, especially with their expenses. Thus, for design of membrane device for a specific process technical and economic analysis should be carried out in order to find a compromise decision between a selection of type of membranes with smaller pores to reduce sizes of device from one side, and a selection of membranes with bigger pore to reduce necessary pressure from another side.

The following is also should be noted: in the majority of industrial plate-type absorbers gas expenses significantly exceed liquid expenses (Purification of process gases, 1977). However, in membrane contactor that ratio of gas and liquid expenses is intolerable, because it will lead to gas content in dispersion of more than 0.5. Presumably, it will lead to disturbance of microbubbles formation and change of process parameters (it worse reminding, that in the presented work mass transfer studies were carried out with gas content of 25%). Thus, design of membrane microbubbling device on a basis of gas loads, comparable with plate-type devices, would require a significant increase of use liquid absorber which will increase operation expenses and sizes of the device. In that case use of microbubbling contactor would be unreasonable.

That's why the most effective area of application of microbubbling method is processes that don't require high gas loads. Purification of biogas from carbon dioxide is an example of that kind of process.

Use of membrane devices in a process of purification of biological methane from CO_2 is extensively studied nowadays (Atchariyawut, et al., 2007), (Al-Marzouqi, et al., 2008), (Yan, et al., 2007), (Zhang, 2006), (Li, & Chen, 2009), (Mansourizadeh, & Ismail, 2009). Generally those are devices on a basis of hollow fiber membranes, comprising porous polymer fibers. Absorption in that kind of devices is carried out due to contact of gas through pores with liquid, flowing inside fibers, in a case there is no bubbling, pores are filled with gas (if membrane surface can't been wet) or with liquid (if membrane can be wet) and contact surface in that case depends on a number of fibers and porosity of membrane. Main advantages of that devices are small sizes, high selectivity of operation and absence of operational restrictions, which are characteristic for plate-type cap-type devices. That's a comparison of membrane contactors with hollow fiber membranes and microbubbling device from a point of effectiveness of interphase mass transfer is of a big interest.

Calculations of mass-transfer coefficient, which were carried out on a basis of experimental data, shows, that for conditions, studied in the presented paper, mass-transfer coefficient in membrane contactor (calculated for gas phase) has values of $1.6 \cdot 10^{-5} – 4.5 \cdot 10^{-5}$ m/s for a membrane with 0 μm pores and $2 \cdot 3^{-5} – 4.5 \cdot 10^{-5}$ for a membrane with 2.6 μm pores.

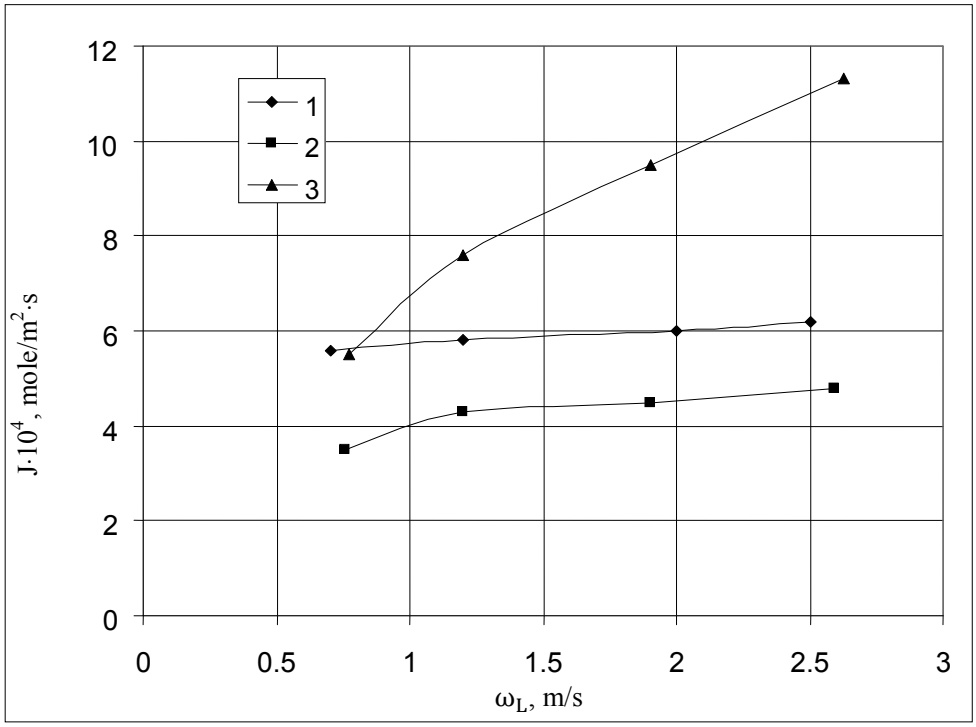


Figure 6. Comparison between specific interphase flow in the microbubbling device and membrane contactor with hollow fiber membranes during chemisorption of CO_2 water solution of CaO (C_B = 0.1 kmole/m^3)

1) Hollow fiber membrane with d_0 = 0.2 μm (reference data); 2) ceramic membrane with d_0 = 2.6 μm; 3) ceramic membrane withd_0 = 0.5 μm.

The Figure 6 shows experimental data on specific interphase flow of CO_2 in microbubbling device and in membrane contactor with polymer hollow fiber membranes with comparable conditions. As it can be seen from the Figure, in a case of use of membranes with a diameter 2.6 μm has specific flow is $3.5 \cdot 10^{-4}$– $4.8 \cdot 10^{-4}$ mole/m2·s, which is very close to the data for hollow fiber membrane ($5.6 \cdot 10^{-4}$ –$6.2 \cdot 10^{-4}$ mole/m^2·s). In the case with the membrane with 0.5 μm pores specific flow varies from $5.6 \cdot 10^{-4}$ to $11.3\ 10^{-4}$mole/m^2·s and exceeds values for hollow fiber membrane, especially at speeds of liquid of 1.5-2.5 m/s. That allows to conclude that in a case of use of microporous membranes microbubbling device may have advantages in terms of efficiency of mass transfer as compared to devices with hollow fiber membranes. In addition, the majority of polymer hollow fiber membranes lose their performance characteristics at elevated temperatures, in contrast to ceramic membranes (Koonaphapdeelert, & Zhentao, 2009), which indicates another important advantage of microbubbling device.

On a basis of the obtained data the following practical recommendations can be made. In order to obtain the largest phase interface surface, membranes made from ceramics and glass with a little value of roughness and average pore size equal to 0.2-3.0 μm should be implemented in a design of mass transferring microbubling equipment. The optimal range of liquid's speed for the process is 1.5-2.5 m/s. Ratio of flows of gas and liquid in the equipment must provide that gas content of a dispersion do not exceed 30%; in the opposite case mechanism of a formation of microbubbles can be compromised, which leads to a decrease of specific phase interface surface in the equipment. During a selection of sizes of membrane pores a compromise solution must be found, which considers both an increase of phase interface surface, leading to a decrease of a working volume of the equipment, and an increase of pressure in the equipment, which leads to increased energy expenses. Chemical absorbent shall be used for higher efficiency of mass transfer.

The results obtained in the presented study can be practically applied in the following fields:

- Definition of specific phase interface surface in membrane microbubling equipment;
- Definition of sizes of microbubbles in membrane floatation equipment;

- Definition of main mass transfer parameters in membrane equipment with gas-liquid contact (absorbers, chemical reactors, fermenters, etc.). Membrane microbubling equipment can be applied for purification of various multi component gas mixtures (e.g., purification of exhaust gases from carbon dioxide), for various chemical reactions (hydrogenation, oxidation) and for purification of waste waters from various contaminants (organic colorants, oils, iron oxides, etc.).

6. Conclusion

The developed technology biogas purification from carbon dioxide gas is proposed. Mass transfer in finely dispersed gas-liquid systems is experimentally investigated. Fine gas-liquid dispersions were created by means of microbubbling method through tubular microfiltering ceramic membranes using incoming liquid. Processes of absorption and chemisorption during microbubbling in mobile liquid phase are studied. Mass transfer coefficients are identified, specific interphase surfaces and interphase flows in microbubbling device during absorption of carbon dioxide from its mixtures with methane and water suspension of CaO are identified. The adequacy of model is proved by experiments of the authors and comparisons with reference experimental data. It is demonstrated, that value specific interphase surface during microbubbling is 8-30 times higher, than in the normal bubbling, which leads to a significant reduction in working volume of a device with the same efficiency. A comparison with membrane hollow fiber contactors shows that using of ceramic membranes allows to substantially increase value specific interphase flow in the microbubbling device and it is comparable or higher than in a case of hollow fiber contactor.

As it was noted, all presented conclusions are based on the provisions of film model of substance's transfer. However, penetration model also can be used for a description of membrane microbubbling method. Therefore, a comparison of relationships obtained by film and penetrations model is of big interest. Existing theories regarding chemisorption, even though they can't be considered complete, provide sufficiently reliable results for a case of fast nonreversible reactions of first and pseudo-first order. At the same time, in fact, all models of mass transfer are demonstrating similar values for mass transfer coefficient for chemisorption. However, because of disadvantages characteristic to existing models of transfer, for accurate prediction of k_x' it is necessary to have experimental data on k_x. For the verification of the conditions on pseudo-first order reaction (condition 15) it is planned to conduct experiments on physical absorption of pure CO_2 by water, on the same membranes and in the same speed range of liquid and gas consumptions as in experiments on chemisorption. On the basis of the obtained data it will possible to calculate values of mass transfer coefficient in liquid phase during physical absorption k_x .In addition the study physical absorption of carbon dioxide may allow a verification of a satisfaction of the condition (15).

Acknowledgments

The presented paper has been created with a help of grant the Ministry of Education and Science of the Republic of Kazakhstan in a state program "Grant funding for research", for sub priority: "The renewable sources of energy (wind, and hydropower, biofuels and photoelectricity)", on the topic "The development of technologies and simulation of process microbubbling purification of biogas with a view to obtain highly-concentrated methane from renewable sources of energy". We would like to express gratitude to our colleagues from D. Mendeleev University of Chemical Technology of Russia, doctors of technical sciences, Professors E.A. Dmitriev, A.M.Trushin, candidate to doctor of technical sciences, assistant professor I.K. Kuznetsova.

References

Al-Marzouqi, M., El-Naas, M., Marzouk, S., Al-Zarooni, M., Abdullatif, N., & Faiz, R. (2008). Modeling of CO_2 absorption in membrane contactors. *Sep.andPurif. Tech., 3*(59), 286-295. 293. http://dx.doi.org/10.1016/j.seppur.2007.06.020

Astarita, D. (1971). *Mass transfer with chemical reaction.* Moscow: Chemistry.

Atchariyawut, S., Jiraratananon, R., & Wang, R. (2007). Separation of CO_2 from CH_4 by using gas-liquid membrane contacting process.*Journal of membrane science, 304*, 163-172. http://dx.doi.org/10.1016/j.memsci.2007.07.030

Biogas plants in Europe: A practical handbook. (2007). Springer.

Bobilev, V. (2003). *Physical properties the most well-known chemical substances.* Moscow: D. Mendeleev University of Chemical Technology of Russia.

Cho, S., Kim, J. Y., Chun, J., & Kim, J. D. (2005). Ultrasonic formation of nanobubbles and their zeta-potentials in aqueous electrolyte and surfactant solutions. *Colloids and surfaces A:Physicochem. Eng. Aspects, 269*,

28-34. http://dx.doi.org/10.1016/j.colsurfa.2005.06.063

Christensen, T., Christensen, T., Cossu, R., & Stegmann, R. (1996). *Landfilling of Waste: Biogas* (1st ed.). Publisher: Taylor & Francis.

Concise Encyclopedia of Bioresource Technology. (2004). CRCPress. http://dx.doi.org/ 10.5860/choice.42-3775

Danckwerts, P. (1973). *Gas-liquid reaction.* Moscow: Chemistry.

Darmana, D., Henket, R., Deen, N., & Kuipers, J. (2007). Detailed modeling of hydrodynamics, mass transfer and chemical reactions in a bubble column using a discrete bubble model: chemisorption of CO_2 into NaOH solution, numerical and experimental study. *Chem. Eng. Sci., 62*, 2556-2567. http://dx.doi.org/10.1016/j.cec.2007.01.065

Deublein, D., & Steinhauser, A. (2008). *Biogas from Waste and Renewable Resources*. Publishing House: Wiley. http://dx.doi.org/ 10.1002/9783527621705

Ditnerski, Yu., Brikov, V., & Kaigramanov, G. (1991). *Membrane separation of gases*. Moscow: Chemistry.

Eder, B., & Schulz, H. (2006). Biogas Praxis. Translated in Russian language. Biogas installations. Practical guide.

Hatta, S. (1932). On the absorption velocity of gases by liquids.*Tech. Repts. Tohoku Imp. Univ., 10*, 119-128.

Hobler, T. (1964). *Mass transfer and adsorption.* Moscow: Chemistry.

Kim, J., Song, M., & Kim, J. (2000). Zeta potential of nanobubbles generated by ultrasonication in aqueous alkyl polyglycoside solutions. *J. Colloids Interf. Sci., 223*, 285-291. http://dx.doi.org/10.1006/jcis.1999.6663

Kimura, S., & Walmet, G. (1980). Fuel Gas Purification with Permselective Membranes Separat. *Sci. and Technol., 4*(15), 1115-1133. http://dx.doi.org/10.1080/01496398008076290

Koonaphapdeelert, S., Zhentao, W., & Li, K. (2009). Carbon dioxide stripping in ceramic hollow fiber membrane contactors.*Chem. Eng. Sci., 64*, 1-12. http://dx.doi.org/10.1016/j.cec.2008.09.010

Krevelen, D., & Hoftijzer, P. (1958). Micro- and macro- kinetics: general introduction to the symposium. *Chem. Eng. Sci., 1-2*(8), 5-17. http://dx.doi.org/10.1016/0009-2509(58)80032-9

Kukizaki, M., & Goto, M. (2006). Size control of nanobubbles generated from SPG membranes.*Journal of membrane science, 281*, 386-396. http://dx.doi.org/10.1016/j.memsci.2006.04.007

Kukizaki, M., & Goto, M. (2006). Spontaneous formation behavior of uniform-sized microbubbles from SPG membranes in the absence of water-phase flow. *Colloids and surfaces A: Physicochem. Eng. Aspects,* 14097. http://dx.doi.org/10.1016/j.memsci.2006.04.007

Li., J., & Chen, B. (2005). Review of CO_2 absorption using chemical solvents in hollow fibers membrane contactors.*Separation Purification Tech., 41*, 109-122. http://dx.doi.org/10.1016/j.seppur.2004.09.008

Loubiere, K., & Hebrard, G. (2004). Influence of Liquid surface tension (surfactants) on bubble formation at rigid and flexible orifices.*Chemical Engineering and Processing, 43*, 1361-1369. http://dx.doi.org/10.1016/j.cep.2004.03.009

Maalej, S., Benadda, B., & Otterbein, M. (2003). Interfacial area and volumetric mass transfer coefficient in a bubble reactor at elevated pressure.*Chem. Eng. Sci., 58*, 2365-2376. http://dx.doi.org/10.1016/s0009-2509(03)00085-x

Mansourizadeh, A., & Ismail, A. (2009). Hollow fiber gas-liquid membrane contactors for acid gas capture: a review.*Journal of Hazardous Materials, 171*, 38-53. http://dx.doi.org/10.1016/j.jhazmat.2009.06.026

Painmanakul, P., Loubiere, K., & Hebrard, G. (2005). Effect of surfactants on liquid-side mass transfer coefficients. *Chemical Engineering science, 60*, 6480-6491. http://dx.doi.org/10.1016/j.cec.2005.04.053

Pohorecki, R., & Moniuk, W. (1988). Kinetics of reaction between carbon dioxide and hydroxyl ions in aqueous electrolyte solution.*Chem. Eng. Sci., 43*, 1677-1684. http://dx.doi.org/10.1016/0009-2509(88)85159-5

Ramm, V. (1976). *Gas absorption*. Moscow: Chemistry.

Richardson, J., & Coulson, J. (1999). *Chemical Engineering* (6th ed.). Vol. 1. Oxford: Butterworth-Heinemann.

Rodriguez, R., & Rubio, J. (2003). New basis for measuring bubbles size distribution. Minerals Engineering, *8*(16), 757–765. http://dx.doi.org/10.1016/S0892-6875(03)00181-X

Rutz, D., & Janssen, R. (nd.). *Biofuel Technology Handbook. Dipl.-Ing. WIP Renewable Energies, Sylvenstein str.*

Munchen, Germany. Retrieved December 26, 2014, from www.wip-munich.de

Semenova, T., & Leitesa, I. (1977). *Purification of process gases* (2nd ed). Moscow: Chemistry.

Sharma, M., & Danckwerts, P. (1970). Chemical methods of measuring interfacial area and mass transfer coefficient in two-fluid systems. *British Chemical Eng., 15*(4), 522.

Unger, E., Matsunaga, T., Schumann, P., & Zutshi, R. (2003, April). Microbubbles in molecular imaging and therapy. *Medicamundi, 47*(1), 58-65.

Versteeg, G., & Swaaij, W. (1988). Solubility and diffusivity of acid gases (CO_2 and N_2O) in aqueous alkaloamin solutions. *Journal of Chem. Eng. Data, 33*, 29-34. http://dx.doi.org/ 10.1021/je00051a011

Yan, S., Fang, M., Zhang, W., Wang, S., Xu, Z., Luo, Z., & Cen, K. (2007). Experimental study on the separation of CO_2 from flue gas using hollow fiber membrane contactors without wetting. *Fuel Processing Tech., 88*, 501-511. http://dx.doi.org/10.1016/j. fuproc.2006.12.007

Zhang, H., Wang, R., Liang, D., & Tay, J. (2006). Modeling and experimental study of CO_2 absorption in a hollow fiber membrane contactors. *Journal of membrane science, 279*, 301-307. http://dx.doi.org/10.1016/j.memsci.2005.12.017

Zong, Y., Wan, M., Wang, S., & Zhang, G. (2006). Optimal design and experimental investigation of surfactant encapsulated microbubbles. *Ultrasonics, 44*, 119-122. http://dx.doi.org/10.1016/j.ultras.2006.06.005

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Development of a Complex Catalytic Conversion System for Internal Combustion Engines Fueled with Natural Gas

Vladislav Anatolievich Luksho[1], Andrey Victorovich Kozlov[1], Vladimir Ivanovich Panchishny[1] & Alexey Stanislavovich Terenchenko[1]

[1] Federal State Unitary Enterprise Central Scientific Research Automobile and Automotive Institute "NAMI" (FSUE «NAMI»), National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Russia

Correspondence: Vladislav Anatolievich Luksho, Federal State Unitary Enterprise Central Scientific Research Automobile and Automotive Institute "NAMI" (FSUE «NAMI»), National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Russia.

Received: January 15, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p237 URL: http://dx.doi.org/10.5539/mas.v9n8p237

Abstract

The paper related to developing of a new gas engines with high energy efficiency and meeting future emission standards. It is necessary to develop complex exhaust gas aftertreatment systems to treat the toxic components efficiently when the engine runs on stoichiometric and lean mixtures. It is proposed to use new combination of three-way catalyst for working on stoichiometric mixtures and a selective catalytic reduction system for NO_x aftertreatment on lean mixtures. Experimental studies have shown that efficient (over 90%) conversion of gas engine exhaust components takes place in the range of air excess ratio from 0.99 to 1.01. Theoretical studies have shown that the highest efficiency of nitrogen oxides reduction is achieved in the temperature range of 400...500°C and reaches over 97%.

Keywords: natural gas, internal combustion engines, aftertreatment system, nitrogen oxides, three-way catalyst, selective catalytic reduction

1. Introduction

Depletion of oil reserves, pollution of air with harmful substances and global warming make it necessary to look for solutions to these global problems. One of the most efficient and comprehensive ways to solve them is to replace petroleum-based fuels for internal combustion engines by alternative fuels. The most promising approach, according to many scientists and experts, is the use of natural gas as fuel both in compressed and liquefied form (Luksho, 2014; Bakhmutov and Karpukhin 2012).

The main advantage of natural gas is its better environmental footprint. The exhaust gases of natural gas powered engines contain less harmful substances, compared with petrol and diesel engines. They produce less carbon monoxide by 50-70%, less non-methane organic gases and nitrogen oxides by 70-80% and less carbon dioxide by 15-20% (Luksho et al., 2011; Luksho, 2010).

There are certain other advantages of using natural gas as motor fuel. In particular, after desulfurization methane contains almost no sulfur, whose combustion products can be harmful to the health of people and other living organisms, acidify soil and water, destroy the monuments, building facades, etc. Therefore, a comprehensive assessment of the most harmful components of exhaust gases, created by engines running on natural gas, shows that such engines are significantly less dangerous than gasoline or diesel ones.

However, transition to gas fuel does not fully solve environmental problems exposed by tightened environmental standards. These problems are mainly related to two components, namely methane and nitrogen oxides, although the transition to advanced (Euro 5 and Euro 6) standards can make it difficult to avoid violating the maximum levels of other regulated substances.

In particular, when converting a diesel engine to methane fuel, scientists at NAMI and other researchers reported a very high level of methane emissions, which is several times greater than current standards. Therefore, even Euro-3 required the use of catalytic converters with high content of precious metals, with significantly (three or more times) greater concentration thereof than in the conventional oxidation converters for gasoline and diesel engines.

Transition from Euro-4 to Euro-6 implies stricter requirements for emissions of hydrocarbons, including 2.2 - 3.4 times stricter requirements for methane, depending on the type of engine, regulated exhaust component and normalized test cycle (Table 1). For example, for CH_4 it corresponds to emission reduction from 1.1 g/kWh (ETC - cycle) to 0.5 g/kWh during transition from Euro-4 to Euro-6. The task of satisfaction to the standards for nitrogen oxides is even more difficult. In accordance with the norms for complying with future requirements quoted in the table 1, the toxicity levels caused by nitrogen oxides should be reduced by more than 8 times during transition from Euro-4 to Euro-6. Standard values for other toxic components are significantly reduced too. It is necessary to take into account the effects of inevitable aging, which affects components of the system. As a result, the initial levels of treatment should be set with a significant (20-30%) margin.

Table 1. Emission standards for toxic emissions

Regulatory requirements	ETC cycle, g/kWh				
	CO	NMHC	CH_4[a]	NO_x	PM
Euro-4	4.0	0.55	1.1	3.5	0.03
Euro-5	4.0	0.55	1.1	2.0	0.03
Euro-6	4.0	0.16	0.5	0.4	0.01

a - only for natural gas fueled engines.

One of important problems in converting diesel engines to natural gas is an efficiency drop, caused, in particular, by lower compression ratio. Studies conducted at NAMI have shown possibility of substantial (10...15%) improvement of fuel economy at partial load conditions when running with stoichiometric mixtures, using the Miller cycle and maintaining the same geometric compression ratio as in the diesel prototype engine.

Further increase of the gas engine's efficiency is possible through leaning of the combustion mixture at partial loads, however, it creates a problem of increased NOx emissions, because three-way catalyst is not efficient for NO_x at lean-burn operation.

Thus, to develop new gas engines with high energy efficiency and meeting future emission standards, it is necessary to develop complex exhaust gas aftertreatment systems to treat the toxic components efficiently when the engine runs on stoichiometric and lean mixtures. This paper discusses issues related to creation of such complex systems.

2. Complex Catalytic Conversion System

Meeting the future requirements on the toxic emissions of a natural gas engine is possible by using an integrated system of catalytic conversion, able to work efficiently with different air-fuel ratios.

Table 2 shows the influence of air excess ratio λ in the mixture on the choice of technologies that may be used for efficient treatment of exhaust gases. Engine operation mode and air excess ratio λ determine the exhaust composition of the natural gas fueled ICE's and, consequently, the type or set of catalysts, which can be used in the process of aftertreatment.

The exhaust gas treatment technologies, used for ICEs running on natural gas, may be divided into oxidation, reduction, oxidation-reduction and complex conversion systems, depending on the concentrations of toxic components, oxygen and temperature.

In the reviewed version of a gas engine operating with both stoichiometric and lean mixtures a complex system of catalytic conversion is necessary.

The system shall consist of a three-way catalyst to ensure treatment of gases when the engine runs with stoichiometric mixtures and a selective catalytic reduction (SCR) system of nitrogen oxides for lean-burn operation. A scheme of an integrated system is shown in Figure 1.

Table 2. Effect of air excess ratio (λ) on the choice of exhaust gas aftertreatment technology

λ	Process type	Toxic components	Main reactions
λ<1 ("rich" mixture)	oxidation/reduction	NO_x, CO, HC	$2CO + 2NO \rightarrow 2CO_2 + N_2$ $HC + NO \rightarrow CO_2 + H_2O + N_2$ $2CO + O_2 \rightarrow 2CO_2$ $HC + O_2 \rightarrow CO_2 + H_2O$

λ=1 stoichiometric mixture	oxidation/reduction (TWC)	NO_x, CO, HC	$2CO + 2NO \rightarrow 2CO_2 + N_2$ $HC + NO \rightarrow CO_2 + H_2O + N_2$ $2CO + O_2 \rightarrow 2CO_2$ $HC + O_2 \rightarrow CO_2 + H_2O$
	Oxidation	CO, HC	$2CO + O_2 \rightarrow 2CO_2$ $HC + O_2 \rightarrow CO_2 + H_2O$ $NO + O_2 \rightarrow NO_2$
λ>1("lean" mixture)	Selective catalytic reduction of NOx (SCR)	NO_x	$2CO + 2NO \rightarrow 2\,CO_2 + N_2$ $HC + NO \rightarrow CO_2 + H_2O + N_2$ $2H_2 + 2NO \rightarrow 2\,H_2O + N_2$
	Partial oxidation	NO_x, CO, HC	$2HC + O_2 \rightarrow CO + H_2$
	Reforming (partial conversion)	NO_x, CO, HC	$CO + H_2O \rightarrow CO_2 + H_2$ $2HC + 4H_2O \rightarrow 2CO_2 + 5H_2$

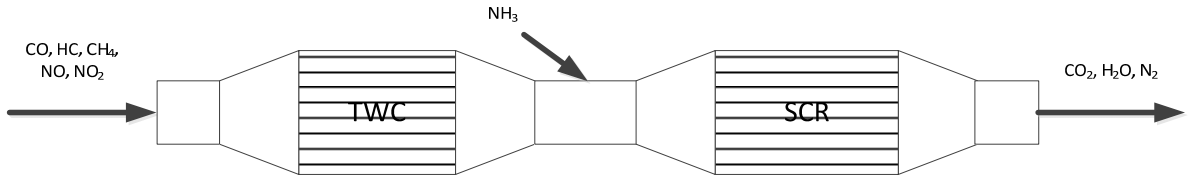


Figure 1. An integrated system of exhaust gas aftertreatment

The choice of catalysts is the question of utmost importance when creating a catalytic conversion system. The main parameters needed when choosing a catalyst include clear definition of the composition, temperature and flow rate of the exhaust gas, which, in turn, are determined by design choices, engine operation mode, fuel and air supply systems of the engine. The following subsections present the results of studies and analysis of the most efficient catalytic coatings for use in a complex catalytic conversion system of a gas engine. Three-way catalyst was experimentally tested on motor test bench. Selective catalytic reduction system was investigated by computer simulation technique.

3. Three-Way Catalytic Converter

The worldwide community have accumulated a lot of experience in creation of catalysts for three-way conversion (TWC) of CO, NOx and hydrocarbons in exhaust gases. It have been established that simultaneous conversion of these components is possible only when the oxygen concentration is close to stoichiometric or lower. These conditions are successfully implemented in modern engines with electronically controlled fuel ratio. The use of TWC conversion systems for gas engines has its own peculiarities associated with the need to oxidize relatively large (compared to other hydrocarbons) amount of methane.

Methane is the most difficult compound to oxidize among light paraffins.

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O,\ \Delta H^o_{298} = 805{,}0\ \text{kJ/mol}.$$

Inactivity of methane is due to high strength and low polarity of the C-H bond and stiffness of the tetrahedral structure, which complicates activation of the CH4 molecule. The energy gap of the C-H bond in CH_4 is $423{,}2 \times 10^3$ kJ/mol, $398{,}0 \times 10^3$ in propane and $322{,}0 \times 10^3$ in propylene. Higher strength of the C-H bond in methane complicates its combustion process and is a major cause of increased hydrocarbons emission, which is usually significantly greater than permissible level.

Therefore, to meet the requirements of modern standards for emissions of CO and hydrocarbons, a catalytic converter is required, which provides an effective afterburning of methane and other hydrocarbons at 80-90% on average at the most typical exhaust temperatures of a gas engine (350-500°C).

The most efficient catalysts for complex conversion processes are Pt-Pd-Rh-on-alumina catalysts with modifying

additives, a mixture of Ce, Zr, La oxides is successfully used. The most commonly used ratio of Pt:Pd:Rh is 1:5:1 and total concentration of active ingredients is 1.4÷2.1 grams per liter of catalyst. When using this catalyst for complex conversion of emissions from ICE running on natural gas, in $\lambda = 1$ mode or methane oxidation ($\lambda>1$) the concentration of Pt-Pd-Rh is increased to 4 or more grams per liter of catalyst.

NAMI conducted comparative laboratory tests of Pt, Pd, and Pt-Rh catalysts in a bifunctional mode on gas mixture containing CO, NO and propane or methane. Tests showed that Pt-Rh catalyst at $\lambda=1$ is active in simultaneous oxidation of HC, CO and reduction of NO by 70-80% at temperatures of 300-320 °C in the presence of propane and at 380-400 °C, if the mixture contains methane. The catalysts are ordered by their activity as follows: Pt-Rh, Pt, Pd. Increasing the concentration of active components in the catalyst from 1.2 to 2.2 g/l leads to an increase in activity of the catalyst, which is especially pronounced at temperatures more than 450°C. At $\lambda>1$ the catalysts based on platinum group metals work only in deep oxidation mode and do not initiate reduction of nitrogen oxides.

A positive effect is demonstrated by layering of active components in the catalys carrier, which allows not only to improve resistance of catalysts to poisoning, but also regulate the temperature of the three-way conversion process. The optimal choice is when the outer layer consists of palladium with stabilizing dioxide additives of, e.g. cerium, while rhodium and platinum remain in the inner layer. This arrangement makes it possible to reduce the light-off temperature of the reaction of hydrocarbons' and CO oxidation at the surface and considerably reduces platinum and rhodium sintering, which are protected by the upper layers of the catalyst and also increases both the role of rhodium during reduction of NO_x, and the ability of cerium oxide to accumulate oxygen (Gandhi et al., 2003; Kinnunen, 2011; Kašpar et al., 2003).

Oxidation of CH_4 with a "rich" mixture ($\lambda=0.988\div0.995$) runs easier than in bifunctional mode ($\lambda =1$), for example, the methane conversion rate at 360°C with a "rich" mixture is 90% versus 65% at $\lambda=1$. The increase of CH_4 contents after catalyst in the temperature range of 280-400°C is higher and the difference in effectiveness of the catalyst decreases as the temperatures go higher. Consequently, this catalyst is best used at $\lambda=0.988\div0.995$, when both NO_x reduction and CH_4 oxidation run at maximum rate (Wit et al., 2000; Nellen and Boulouchos, 2000; Gélin and Primet, 2002).

NAMI has conducted research of three-way catalyst's efficiency for aftertreatment of exhaust gases of a 6-cylinder gas engine running in Miller cycle. In the Miller cycle, the intake valve is left open longer than it would be in an Otto cycle engine. The engine has cylinders with bore and stroke 105x128 mm, nominal power of 180 kW at speed of 2300 RPM. The studies were conducted at 30...70% of the rated load and engine speed of 1400...1800 RPM. Air excess ratio was changed by variation of fuel supply rate at constant throttle position and it was calculated as ratio of real air consumption to theoretically required to full combustion of expended fuel. The results of measuring concentrations of toxic components before and after the catalyst according to the excess air ratio are shown in Figure 2.

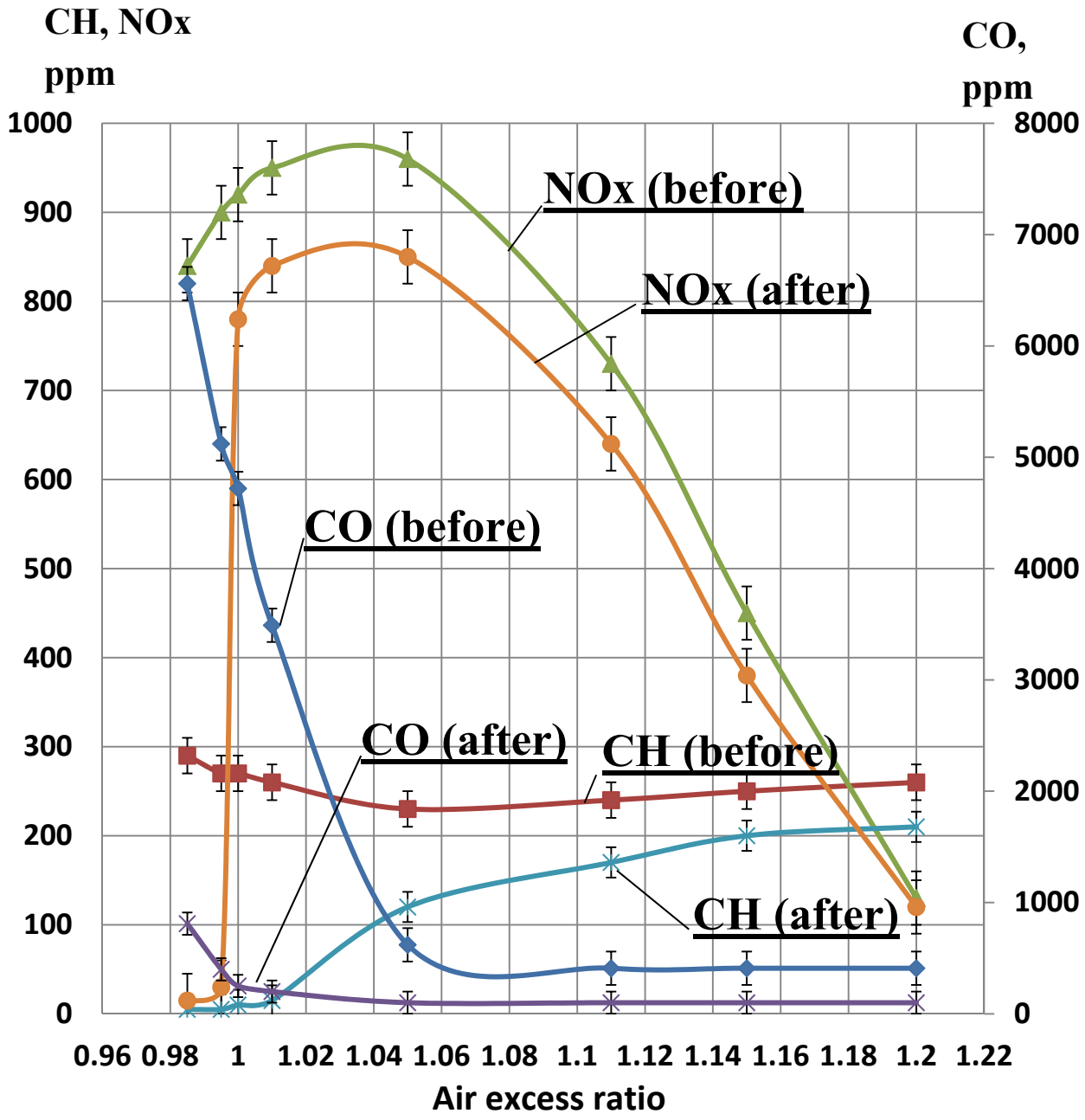


Figure 2. Concentration of toxic components in the exhaust gases before and after catalyst depending on the air excess ratio

As the experiments showed, the highest conversion efficiency (over 90%) for all three components was achieved in the air excess ratio range from 0.99 to 1.01. When the mixture got leaner to λ>1.09, there was a sharp decrease in the efficiency of nitrogen oxides reduction.

Thus, it is shown that a TWC converter is a complex multifunctional system, which must have a layered structure comprising a metallic or ceramic matrix, a secondary carrier and an active phase, able to simultaneously ensure oxidation of CO, HC and NO_x reduction at oxygen concentration below or near stoichiometric in a wide range of temperatures and presence of small amounts of poisons (SO_2, H_2S, P $\leq$ 10 ppm). The active phase should include: 1 - active components (platinum group metals Pt, Pd, Rh or a mixture thereof); 2 - oxygen accumulator (CeO_2, ZrO_2 or mixture thereof); 3 - stabilizers and a carrier (RZE, SZE or a mixture

thereof); 4 - poison sorbing components.

4. Selective Catalytic Reduction of Nitrogen Oxides

In the cases, where nitrogen oxides pose the biggest problem in the overall balance of harmful engine emissions, the systems based on selective reduction of NO_x become the most widely spread, and the preferred reductant is carbamide (urea). The most common way of urea application is 32.5% solution in demineralized water under a trade name "AdBlue". The use of selective reducing agents containing amino groups in its structure allows for deeper reduction of nitrogen oxides in exhaust gases. However, it creates problems of filling and carrying on-board an additional reagent, possible interactions between construction materials and amino compounds, own toxicity of ammonia in case of incomplete reaction and possibility of secondary nitrogen oxide formation during oxidation of unreacted ammonia in the catalyst.

Studies of the processes of nitrogen oxides reduction with hydrocarbons, in particular by gaseous methane, are presented in (Hamill et al., 2014; Jing et al., 2009), however the efficiency of such systems is not sufficient to meet Euro-6 level norms. In this regard, the use of nitrogen oxide reduction systems utilizing urea seems to be a more promising approach.

An outline of the NH_3-SCR process is shown in Figure 3.

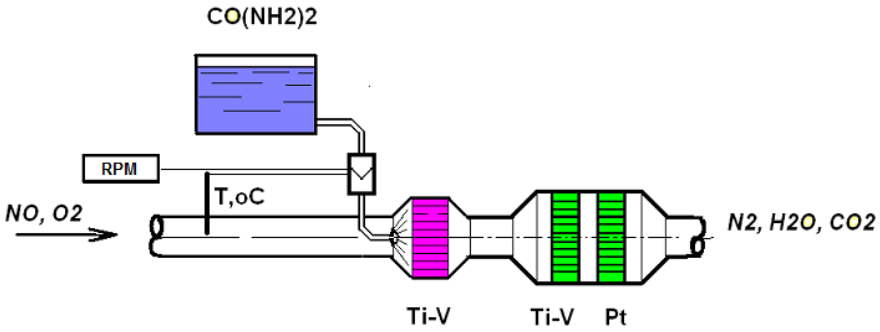


Figure 3. Scheme of the NH_3-SCR process

Selective reduction of nitrogen oxides using urea as a source of ammonia is sufficiently studied and the most efficient catalysts based on titanium oxide with additives of V_2O_5, WO_3, MoO_3, Fe_2O_3, or a combination mentioned above, are identified (Caton and Siebers, 1989; Long and Yang, 1999; Komatsu et al., 1994; Gieshoff et al., 2000; Uddin et al., 1995; Koebel et al., 2001). NAMI has conducted work to create a system and a catalyst on a block metallic carrier (BMN) to implement selective reduction of NO_x for a diesel engine. It has been found that the efficiency of vanadium-titanium catalysts on BMN in the reaction of nitrogen oxides reduction with ammonia in the presence of excess oxygen depends on several parameters, among which the most important are: composition, substrate structure, volume velocity, catalyst technology and exhaust gas temperature (Gieshoff et al., 2000).

It is possible that this catalytic system could also be used for an engine running on natural gas. In addition to the catalyst for selective reduction of nitrogen oxides the NH_3-SCR system includes catalysts of urea hydrolysis and final reduction of, e.g. ammonia. This technology has a number of thermal limitations, because hydrolysis of urea at T>400-500^oC may produce crystalline compounds poorly soluble in water, such as triurea and carbamyl urea, which are incapable of hydrolysis. In addition, the number of active V^{4+} sites at these temperatures in V-Ti composition decreases, V^{5+} and higher phases appear, which oxidize ammonia at the stage of NO_x and nitrogen reduction, significantly reducing the catalyst's efficiency.

Therefore, the main attention, as studies show, must be focused on the method to form the active surface of the catalyst and its stabilization, i.e. the method of catalyst preparation and its operating conditions. These problems and lack of experience with NH3-SCR system in vehicles powered by natural gas, do not exclude the possibility of its modernization and future use.

Alternatives to the V-Ti composite for high-temperature NO_x reduction, according to the authors (Uddin et al., 1995; Koebel et al., 2001; Weitkamp and Gläser, 2003), may be CeTi, FeTiOx, Cu/zeolite catalysts, which do not form volatile compounds and allow for the process to reduce nitrogen oxides at temperatures 50-100°C higher

than vanadium contacts. In addition to NH_3-SCR and hydrolysis catalysts, the system comprises an ammonia oxidation stage, because the Euro-6 standard provides for the restriction of ammonia at 10 ppm.

A number of computational studies has been performed to investigate the effectiveness of the selective catalyst in a gas engine running on lean fuel mixtures.

The computational studies used a mathematical model describing the basic chemical processes occurring in the catalytic conversion system. The reactions of nitrogen oxides reduction (Winkler et al., 2003; Wurzenberger and Wanker, 2005) and conversion of ammonia are as follows:

reduction of nitrogen oxides:

$$4NH_3 + 4NO + O_2 = 4N_2 + 6H_2O; \quad (1)$$

combined reduction of nitrogen oxide and dioxide:

$$4NH_3 + 2NO + 2NO_2 = 4N_2 + 6H_2O; \quad (2)$$

reduction of nitrogen dioxide:

$$8NH_3 + 6NO_2 = 7N_2 + 12H_2O; \quad (3)$$

hydrolysis of isocyanic acid:

$$HNCO + H_2O = NH_3 + CO_2; \quad (4)$$

oxidation of ammonia:

$$4NH_3 + 3O_2 = 2N_2 + 6H_2O. \quad (5)$$

The formulas for calculating reaction rates are as follows:

$$r_1 = K_1 e^{\frac{-E_1}{RT_s}} c_{NO} \frac{K_1' e^{\frac{-E_1'}{RT_s}} c_{NH_3}}{1 + K_1' e^{\frac{-E_1'}{RT_s}} c_{NH_3}}; \quad (6)$$

$$r_2 = K_2 e^{\frac{-E_2}{RT_s}} c_{NO} c_{NO_2} \frac{K_2' e^{\frac{-E_2'}{RT_s}} c_{NH_3}}{1 + K_2' e^{\frac{-E_2'}{RT_s}} c_{NH_3}}; \quad (7)$$

$$r_3 = K_3 e^{\frac{-E_3}{RT_s}} c_{NO} \frac{K_3' e^{\frac{-E_3'}{RT_s}} c_{NH_3}}{1 + K_3' e^{\frac{-E_3'}{RT_s}} c_{NH_3}}; \quad (8)$$

$$r_4 = K_4 e^{\frac{-E_4}{RT_s}} c_{HNCO} c_{H_2O}; \quad (9)$$

$$r_5 = K_5 e^{\frac{-E_5}{RT_s}} c_{NH_3}, \quad (10)$$

where K_i is pre-exponential factor of the i-th reaction; E_i is activation energy of the i-th reaction, kJ/kmol; c_n is molar concentration of the n-th component, kmol/m^3; R is the universal gas constant, kJ/(kmol*K); T_s is the temperature of the catalyst substrate, K.

Basing on the data from the literature (Winkler et al., 2003; Wurzenberger and Wanker, 2005; Beeckmann and Hegedus, 1991; Schaub et al., 2003; Tang et al., 2008; Riyandwita and Bae, 2011) the values K and E were determined, which are presented in the Table 3.

Table 3. Values of the pre-exponential factor and activation energy for calculation of reaction rates

Pre-exponential factor	Value, unit	Activation energy	Value, kJ/kmol
K_1	3200 m/s	E_1	62571
K'_1	$1*10^{-14}$ m^3/kmol	E'_1	-283699
K_2	$5*10^{12}$ m^4/(kmol*s)	E_2	72548
K'_2	$3*10^{-16}$ m^3/kmol	E'_2	-293260
K_3	3000 m/c	E_3	68391
K'_3	$1*10^{-14}$ m^3/kmol	E'_3	-268318
K_4	$1*10^{8}$ m^4/(kmol*s)	E_4	49285
K_5	1000 m/s	E_5	80862

Figure 4 shows the calculation results of nitrogen oxide reduction efficiency in exhaust gases in accordance with the above equations (6)-(10), depending on the temperature and composition of the combustible mixture. Increasing the excess air ratio lowers the reduction efficiency (by 8...17%). The best efficiency is achieved at exhaust gas temperatures of 400...500 °C.

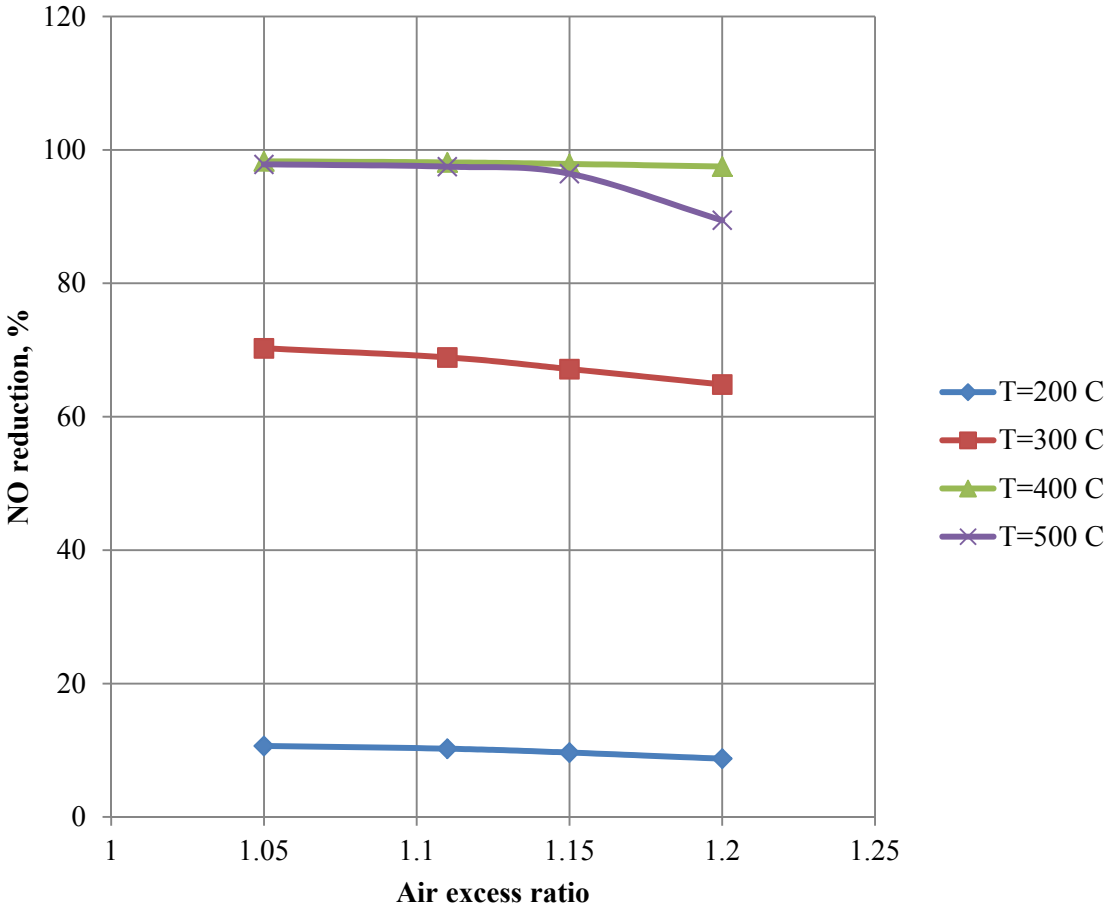


Figure 4. Calculation results of efficiency evaluation of nitrogen oxide reduction in exhaust gases depending on composition of combustion mixture

5. Conclusions

Fulfillment of the Euro-5 and Euro-6 environmental requirements is a complex task, which can be accomplished primarily through optimization of gas engine design and operation, its air and fuel supply systems, as well as the use of complex catalytic conversion systems for aftertreatment toxic exhaust conponents.

It is proposed to use for gas engine new combination of three-way catalyst for working on stoichiometric mixtures and selective catalytic reduction system for NO_x aftertreatment on lean mixtures.

In order to achieve the requirements on emissions of harmful substances in exhaust gases, when the engine runs with stoichiometric mixture, it is advisable to use a special multifunctional three-way catalyst, specifically designed for complex conversion of exhaust gases produced by a natural gas-powered ICE. A number of catalysts for engine exhaust aftertreatment were identified, which are able to oxidize methane and reduce nitrogen oxides efficiently in a wide temperature range. The key components of three-way catalysts for exhaust gases of natural gas engines are precious metals with Ce, Ce-Zr and other additives to improve efficiency and thermal stability of the contact. Experimental studies have shown that efficient (over 90%) conversion of gas engine exhaust components takes place in the range of excess air ratio from 0.99 to 1.01.

Aftertreatment of toxic components of a gas engine with $\lambda>1$ requires the use of catalytic converters, which use a modular design concept allowing to build systems that ensure conversion of nitrogen oxides and methane in the presence of excess oxygen. Such systems combine components with oxidation and reduction functions in accordance with the temperature, concentration and toxic components to oxygen ratio. A good solution to reduce NO_x emission is an SCR converter using urea. Theoretical studies have shown that the highest efficiency of nitrogen oxides reduction with ammonia is achieved in the temperature range of 400...500 °C and reaches over 97%. This research is performed, a new project with the Ministry of Education and Science of the Russian Federation (the unique identifier for Applied Scientific Research (project) RFMEFI62414X0005).

References

Bakhmutov, S. V., & Karpukhin, K. E., (2012). Greencars: the directions of realization and achived results. *Scientific magazine of Automotive engineers, 6*, 51–54.

Beeckmann, J. W., & Hegedus, L. L. (1991). Design of monolith catalysts for power plant NOx emission control. *Ind. Eng. Chem. Res, 30,* 969-978. http://dx.doi.org/ 10.1021/ie00053a020

Caton, J. A., & Siebers, D. L. (1989). Comparison of nitric oxide removal by cyanuric acid and ammonia. *Combust. Sci. And Tech., 65*, 277. http://dx.doi.org/10.1080/00102208908924054

Gandhi, S., Graham, G. W., & McCabe, R. W. (2003). Automotive exhaust catalysis. *Journal of Catalysis, 216,* 433–442. http://dx.doi.org/10.1016/S0021-9517(02)00067-2

Gélin, P., & Primet, M. (2002). Complete oxidation of methane at low temperature over noble metal based catalysts: A review. *Applied Catalysis B: Environmental, 39*, 1–37. http://dx.doi.org/10.1016/S0926-3373(02)00076-0

Gieshoff, J. et al. (2000). Improved SCR Systems for Heavy Duty Applications. SAE technical paper, 2000-01-0189, 12. http://dx.doi.org/ 10.4271/2000-01-0189

Hamill C., Burch, R., Goguet, A., Rooney, D., Driss, H., Petrov, L., & Daous, M. (2014). Evaluation and mechanistic investigation of a AuPd alloy catalyst for the hydrocarbon selective catalytic reduction (HC-SCR) of NOx. *Applied Catalysis B: Environmental, 147*, 864-870. http://dx.doi.org/10.1016/j.apcatb.2013.09.047

Jing G., Li, J., Yang, D., & Hao, J. (2009). Promotional mechanism of tungstation on selective catalytic reduction of NOx by methane over In/WO3/ZrO2. *Applied Catalysis B: Environmental, 91*, 1–2, 123-134. http://dx.doi.org/10.1016/j.apcatb.2009.05.015

Kašpar, J., Fornasiero, P., & Hickey, N. (2003). Automotive catalytic converters: current status and some perspectives. *Catalysis Today, 77,* 419–449. http://dx.doi.org/10.1016/S0920-5861(02)00384-X

Kinnunen, T. (2011). Emission catalyst solutions for Euro 5 and beyond. New Delhi, 48.

Koebel, M., Elsener, M., & Madia, G. (2001). Recent Advances in the Development of Urea- SCR for Automotive Applications. SAE technical paper, 2001-01-3625, 16. http://dx.doi.org/ 10.4271/2001-01-3625

Komatsu, T., Nunokawa, M., Moon, I. S., Takahara, T., Namba, S., & Yashima, T. (1994). Kinetic studies of reduction of nitric oxide with ammonia on Cu(sup 2+)-exchanged zeolites. *J. of Catalysis, 148*(3), 427-432. http://dx.doi.org/10.1006/jcat.1994.1229

Long, R. Q., & Yang, R. T. (1999). Selective catalytic reduction of nitrogen oxides by ammonia over Fe3+-exchanged TiO2-pillared clay catalysts. *J. of Catalysis, 186*(2), 254-268. http://dx.doi.org/10.1006/jcat.1999.2558

Luksho, V. A. (2010). Conversion of a diesel engine into a gas-diesel with variable thermodynamic cycle. *Alternative Fuel Transport, 6*, 4-50.

Luksho, V. A. (2014). Increasing the fuel efficiency of vehicles with gas engines. *Proceedings of NAMI, 257,*

124-138.

Luksho, V. A., Kozlov, A. V., & Terenchenko, A. S. (2011). Performance evaluation of natural gas as a motor fuel in the full life cycle. *Alternative Fuel Transport, 3,* 4-9.

Nellen, C., & Boulouchos, K. (2000). Natural Gas Engines for Cogeneration: Highest Efficiency and Near-Zero-Emissions through Turbocharging, EGR and 3-Way Catalytic Converter. SAE technical paper, 2000-01-2825, 12. http://dx.doi.org/ 10.4271/2000-01-2825

Riyandwita, B. W., & Bae, M. (2011). Three-Dimensional Simulation with Porous Medium as the Washcoat Layer for an SCR Monolith Reactor. SAE technical paper, 2011-01-1240, 14. http://dx.doi.org/ 10.4271/2011-01-1240

Schaub, G., Unruh, D., Wang, J., & Turek, T. (2003). Kinetic analysis of selective catalytic NOx reduction (SCR) in a catalytic filter. Chemical Engineering and Processing: *Process Intensification, 42*(5), 365-371. http://dx.doi.org/10.1016/S0255-2701(02)00056-9

Tang, W., Wahiduzzaman, S., Wenzel, S., Leonard, A., & Morel, T. (2008). Development of a Quasi-Steady Approach Based Simulation Tool for System Level Exhaust Aftertreatment Modeling. SAE technical paper, 2008-01-0866, pp:17. http://dx.doi.org/ 10.4271/2008-01-0866

Uddin, A., Md. Azhar, Komatsu, T., & Yashima, T. (1995). Selective catalytic reduction of nitric oxide with ammonia on MFI-type ferrisilicate. *J. Chem. Soc. Far. Trans, 91*, 3275-3282. http://dx.doi.org/ 10.1039/FT9959103275

Weitkamp, J., & Gläser, R. (2003). Katalyse. Berlin, 74.

Winkler, C., Floerchinger, P., Patil, M. D., Gieshoff, J., Spurk, P., & Pfeifer, M. (2003). Modeling of SCR DeNOx Catalyst - Looking at the Impact of Substrate Attributes. SAE technical paper, 2003-01-0845, 14. http://dx.doi.org/10.4271/2003-01-0845

Wit, J., Johansen, K., Hansen, P. L., Rossen, H., & Rasmussen, N. B. (2000). Catalytic emission control with Respect to CH4 and CO for highly efficient gas fueled decentralised heat and power production. 5 International Conference on Furnaces and Boilers, 9.

Wurzenberger, J. C., & Wanker, R. (2005). Multi-Scale SCR Modeling, 1D Kinetic Analysis and 3D System Simulation. SAE technical paper, 2005-01-0948, 18. http://dx.doi.org/ 10.4271/2005-01-0948

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Determination of Weber-Ampere Characteristics of Electric Devices Using Solution of Inverse Problem of Harmonic Balance

Anton Mikhailovich Lankin[1], Mikhail Vladimirovich Lankin[1], Nikolay Ivanovich Gorbatenk[1] &
Danil VadimovichShaykhutdinov[1]

[1] Platov South-Russian State Polytechnic University (NPI), Russia

Correspondence: Lankin Anton Mikhailovich, Platov South-Russian State Polytechnic University (NPI), Russia.

Received: January 15, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p247 URL: http://dx.doi.org/10.5539/mas.v9n8p247

Abstract

This work is devoted to development of measurement of weber--ampere characteristic of electric devices consisting of a magnetic core and a coil. In this method the coil of electric device is energized with sinusoidal voltage of known amplitude and frequency, and harmonic amplitudes of passing current are measured. On the basis of these data the inverse problem of harmonic balance is solved and the approximation coefficients of the equation describing the required weber--ampere characteristic are determined. The influence of the degree of approximating expression on measurement error is studied. The obtained results have demonstrated possibility of application of the proposed method to measurements of weber--ampere characteristics of electric devices. The measurement error of weber--ampere characteristic does not exceed 3 %.

Keywords: weber – ampere characteristic, electric device, measurement methods, harmonic balance, weber - ampere characteristic of operating cycle

1. Introduction

The miniaturization trend and attempts to decrease materials consumption of up-to-date electric devices present increased requirements to quality of their individual elements, which can be estimated on the basis of their mechanical, electrical and magnetic properties. Each of the aforementioned groups of indices require for application of dedicated methods and tools of testing of electric devices. This situation promotes searching for integral indices of operation of electric devices, which can provide conclusions about quality of separate parts and operational characteristics of overall product.

All electric devices (electric magnets, electromagnetic relays, motors) include moving and static parts of magnetic core and at least one operating coil (Korotyeyev, Zhuikov, Kasperek, 2010). Within operation of an electric device current flows through operating coil, working magnetic flow is generated in the static part of magnetic core, put in motion its moving part. The value of magnetic flow is determined by design and mutual position of part of magnetic core and operating coil, as well as the number of turns in operating coil and flowing current. Within operating cycle of electric device the moving part of magnetic core travels with regard to the static part, which also varies magnetic flow. All this suggests that the integral characteristic, providing information not only about operating parameters of electric device but also about quality of its individual parts, is weber - ampere characteristic of operating cycle. Weber--ampere characteristic of operating cycle is the magnetic flow passing through operating coil of electric device as a function of current flowing via the coil. Herewith, the moving part pf magnetic core of electric device executes typical operating movement. In order to obtain weber--ampere characteristic of an electric device it is necessary to apply external varying magnetic field to magnetic core and to measure magnetic flow in its cross section by means of special sensor (Singh, 2003; ASTM International, 1970; Kiefer, 1969). In order to obtain weber--ampere characteristic of operating cycle electric device in assembly is tested, thus, it is impossible to use sensors of magnetic field.

Now let us consider measurement tools of weber--ampere characteristic, where the source of internal magnetic field is the coil of electric device, and the flow is determined indirectly.

For instance, substitution method, where the circuit of sinusoidal voltage source is alternatively connected either with the coil applied on magnetic core of tested electric device, or known variable inductance L_0 and non-reactive resistance r. In addition, ammeter and variable inductance L and capacitor C are connected in series

with the circuit of sinusoidal voltage source. Initially, with known value of output voltage U of the sinusoidal voltage source it is connected with the coil of tested electric device and by means of variable inductance L and capacitor C, the circuit is adjusted to resonance which is detected by maximum value of current I in magnetizing circuit. Then the circuit, instead of winding of tested device, is connected with known variable inductance L_0 and non-reactive resistance r and by means thereof the circuit is again adjusted to resonance. Herewith, by variation of non-reactive resistance r, the same value of current I in the circuit is obtained (the values L and C are not modified). Within resonance the inductance of winding L_x, applied to the tested device, equals to know variable inductance L_0. When it is known, inductance B or flow of magnetic field Φ in the tested device are calculated as follows:

$$\Phi = \frac{L_0 \cdot I \cdot \sqrt{2}}{4\pi w} \tag{1}$$

where I is the current measured by ammeter, w is the number of coils of the tested device.

Another values of output voltage of the sinusoidal voltage source U are set, and measuring the values of current I in magnetizing circuit the magnetic flow Φ is calculated, thus obtaining the required number of points of weber--ampere characteristic.

A drawback of this method is that acceptable accuracy is provided only with sinusoidal variation of magnetizing current, in the region of small magnetic fields, and within saturation of the tested device, when current becomes non-sinusoidal, the calculation by Eq. (1) results in high error.

Another method of measurement of weber--ampere characteristic of electric device involves connection of the coil of electric device to the sinusoidal voltage source, then the voltage and current are measured on the coil of electric device, and the coordinates of weber--ampere characteristic are calculated as the integral of difference between the voltage supplied to the coil of electric device and voltage drop at its active resistance:

$$\Phi = \frac{1}{k}\int (U - IR)dt \tag{2}$$

where U is the voltage at the output of sinusoidal voltage source, I is the current flowing through the coil of electric device; R is the active component of circuit resistance of coil of electric device; k is the coefficient determined by the number of windings of coil of electric device, length of median line and surface area of transversal cross section of magnetic core of electric device. The active component of resistance R of coil of electric device, including the active portion of resistance of the coil of electric device, resistance of current instrument shunt and output resistance of sinusoidal voltage source are preliminary detected and then applied as a constant within calculation of magnetic flow, however, during the measurements as a consequence of current passing through the coil of electric device it is heated and its active resistance R increases. Hence, Eq. (2) works incorrectly, which adds significant and accumulating within integration error to the measurements.

A drawback of the method is that it is based on integration, and the coil active resistance varies within its heating. The necessity to eliminate these factors initiated several works, which more or less successfully solve this problem Roller, 2012; Sakhavova , Shirokov, and Yanvarev, 2013).

Application of the existing methods for measurement of weber--ampere characteristic of operating cycle of electric device has some drawbacks, thus, it is required to develop a method for measurement of weber--ampere characteristic of operating cycle of electric device with sufficient accuracy and without the aforementioned drawbacks.

2. Methodology

2.1 The Essence of Harmonic Balance

The problems of harmonic balance are applied for investigation into various non-linear circuits (Bessonov, 1978; Song Xing, Suting Chen, Zhanming Wei, Jingming Xia, 2013). Direct problem of harmonic balance consists of determination of shape of current passing through non-linear element at its known volt--ampere characteristic, amplitude and shape of feeding voltage. The essence of the method is based on decomposition of periodical functions of voltage applied to non-linear element and current flowing through it into Fourier series. In general case the unknown variables in non-linear electric circuit are non-sinusoidal and contain infinite spectrum of harmonics. The expected solution can be presented as the sum of main and several higher harmonics. Substituting this sum into non-linear differential equation compiled for the required variable and setting equal the coefficients preceding harmonics in the obtained expression (sinusoidal and non-sinusoidal functions) of the equal frequencies in its left- and right-hand sides, we obtain a set of n algebraic equations, where n is the number

of considered harmonics. Solving the set of equations we obtain the unknown variables. The developed method is based on inverse problem of harmonic balance, which consists of determination of unknown volt--ampere characteristic and non-linear element qith known shape of current flowing through it, shape and amplitude of feeding voltage.

2.2 Computation Algorithm of Direct and Inverse Problems of Harmonic Balance

Let us consider solution of direct problem of harmonic balance for the case of non-linear circuit consisting orf sinusoidal voltage source and non-linear resistance (Figure 1).

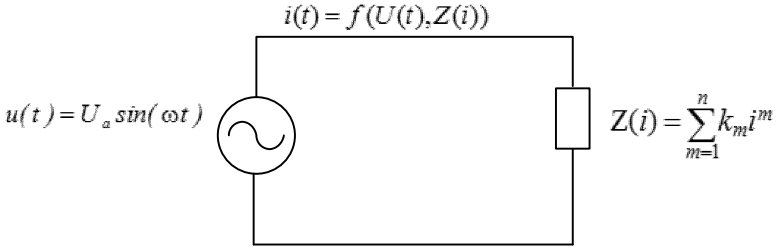


Figure 1. Schematic presentation of problems of harmonic balance

Solution of direct problem of harmonic balance includes the following steps:

1. Equations of circuit (Figure 1) are written for instant values.

$$u(t) = i(t) \cdot Z(i), \tag{3}$$

where $u(t) = U_a \sin(\omega t)$, $i(t)$ are the instant values of voltage and current; $Z(i)$ is the total resistance, U_a is the voltage amplitude, ω is the angular frequency.

2. Expression of analytic approximation of non-linear dependence of total resistance of non-linear element on current is written:

$$Z(i) = \sum_{m=0}^{n} k_{2m+1} i^{2m+1}, \tag{4}$$

where k_{2m+1} are the coefficients of approximating expression.

3. On the basis of preliminary analysis of circuit and non-linear characteristic (4) the equation of required value in the form of finite Fourier series is set (Tolstov, 2012) with unknown at this stage amplitudes of current harmonics I_{2m+1}:

$$i(t) = \sum_{m=0}^{n} I_{2m+1} \sin\left((2m+1)\omega t\right), \tag{5}$$

where I_{2m+1} are the amplitudes of current harmonics. Equation (5) contains no even harmonics of current and components of cosine portion of decomposition into Fourier series due to symmetry of voltage--ampere characteristic with regard to origin of coordinates.

4. Equations (4) and (5) are substituted into equation of circuit (3), trigonometric transformations are carried out (Lopez, 1994) and the terms in the obtained equations are grouped according to numbers of harmonics:

$$U_a \sin(\omega t) = \sum_{m=0}^{n} k_{2m+1} \left(I_{2m+1} \sin\left((2m+1)\omega t\right)\right)^{2m+1} \sum_{m=0}^{n} I_{2m+1} \sin\left((2m+1)\omega t\right). \tag{6}$$

A set of non-linear equations is compiled by substitution of various arguments (ωt) into Eq. (6). The obtained set of equations is solved with regard to unknown amplitudes of current harmonics I_{2m+1}.

In certain cases it is necessary to determine not the current shape but analytic approximation of non-linear dependence of total resistance of non-linear element on current, then we apply the solution of inverse problem of harmonic balance. It differs in that the coefficients k_{2m+1} in Eq. (4) are unknown, and the amplitudes of current harmonics I_{2m+1} are known. therefore, the set compiled of Eq. (6) is solved with regard to k_{2m+1} .

2.3 Application of Inverse Problem of Harmonic Balance for Determination of Weber--Ampere Characteristic of Electric Device

Solution of direct problem of harmonic balance makes it possible to determine the shape of current $i(t)$, flowing through the coil of electric device, given in the form of non-linear inductance (Enns, McGuire, 2000; Enns, 2010) by means of decomposition into Fourier series:

$$i(t) = \sum_{m=0}^{n} I_{(2m+1)} \sin((2m+1)\omega t), \quad (7)$$

where $I_{(2m+1)}$ is the amplitude of the $(2m+1)$-th current harmonic. Herewith, the shape and amplitude U_a of voltage applied to non-linear inductance are known:

$$u(t)=U_a \sin(\omega t), \quad (8)$$

as well as weber--ampere characteristic of non-linear inductance set by approximating equation:

$$\Phi(i)=\sum_{m=0}^{n} k_{(2m+1)}\ i^{2m+1}, \quad (9)$$

where Φ is the value of magnetic flow through non-linear inductance, $k_{(2m+1)}$ are the coefficients of approximating expression of weber--ampere characteristic, $m=\overline{(0,n)}$, $(n+1)$ is the number of components in the approximating expression, i is the current intensity through non-linear inductance.

The inverse problem of harmonic balance for determination of weber--ampere characteristic of electric device is formed as follows. There exists non-linear inductance with unknown weber--ampere characteristic, the regularities of variation of voltage (8) applied to non-linear inductance and flowing through it current (7) are known. It is required to determine the coefficients $k_{(2m+1)}$ of approximating expression of weber--ampere characteristic (9).

Now let us write the equation of circuit with non-linear inductance involving active resistance R:

$$u(t)=Ri+\frac{d\Phi}{dt}.$$

Let us rewrite it with consideration for known regularities of variation of current (7) and voltage (8):

$$U_a \sin\omega t=R\sum_{m=0}^{n} I_{(2m+1)}\sin((2m+1)\omega t)+\frac{d\sum_{m=0}^{n} k_{(2m+1)}(I_{(2m+1)}\sin((2m+1)t\omega))^{2m+1}}{dt} \quad (10)$$

On the basis of known degree $(2n+1)$ of approximating expression of weber--ampere characteristic we set the $(n+1)$ value of sine function argument in Eq. (10). The argument value is taken from the interval $]0;\pi/2[$.

Therefore, we obtain a set of $(n+1)$ linear equations. The obtained set of equations contains known voltage amplitude U_a, amplitudes of current harmonics $I_{(2m+1)}$, values of active resistance R and circular frequency of flowing current ω, since these parameters of circuit with non-linear inductance can be measured within testing of electric devices. Solving this set of equation we obtain coefficients $k_{(2m+1)}$ of Eq. (9) approximating weber--ampere characteristic.

3. Results Discussion

3.1 Simulation Experiment

In order to implement the model it was necessary to perform certain requirements to the applied package of general-circuit simulation: possibility to set the required weber--ampere characteristic by points, plotting of spectra of periodic functions (current spectra), possibility of step-by-step reproduction of circuit, plotting of several function in one plot. All aforementioned requirements are satisfied by the MicroCap simulator (Vester, 2009; Amelina, Amelin, 2007).

A peculiar feature of numerous electric devices is the existence of non-magnetic gap in their magnetic cores. The model includes weber--ampere characteristic of magnetic core of material 3100V, for which two variants of non-magnetic gap are preset: 0.4 mm and 0.9 mm, which leads to various slopes of weber--ampere characteristic.

The model of electric device contains connected in series non-linear inductance with 95 windings and active resistance R of 0.15 Ω and 0.1 Ω for the aforementioned non-magnetic gaps, respectively. The electric device is connected to sinusoidal voltage source with frequency of 50 Hz and amplitude U_a of 1.65 V and 1.7 V for the aforementioned non-magnetic gaps, respectively.

Figure 2 illustrates the currents of в coils of electric device with selected non-magnetic gaps.

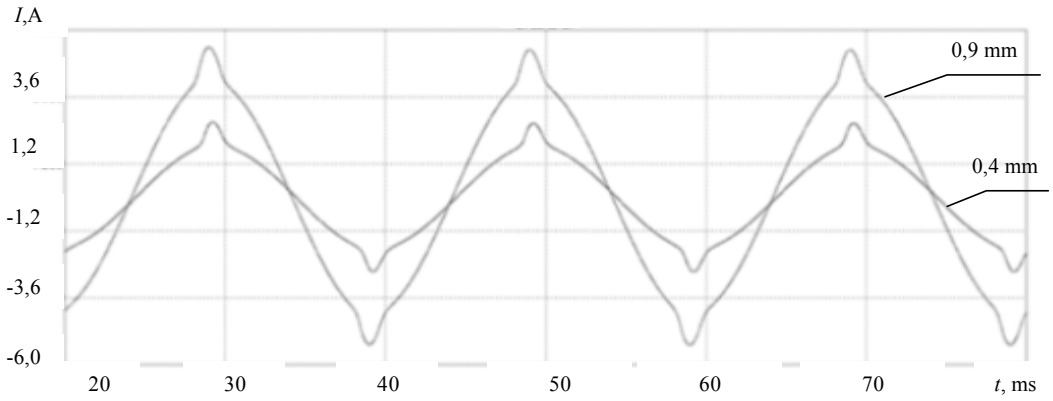


Figure 2. Currents in the coils of electric device

In Figure 2 the currents flowing through the coil of electric device are of non-sinusoidal shape, which can be attribute to the existence of non-linear element in the circuit.

Figures 3 and 4 illustrates the current spectra for selected non-magnetic gaps.

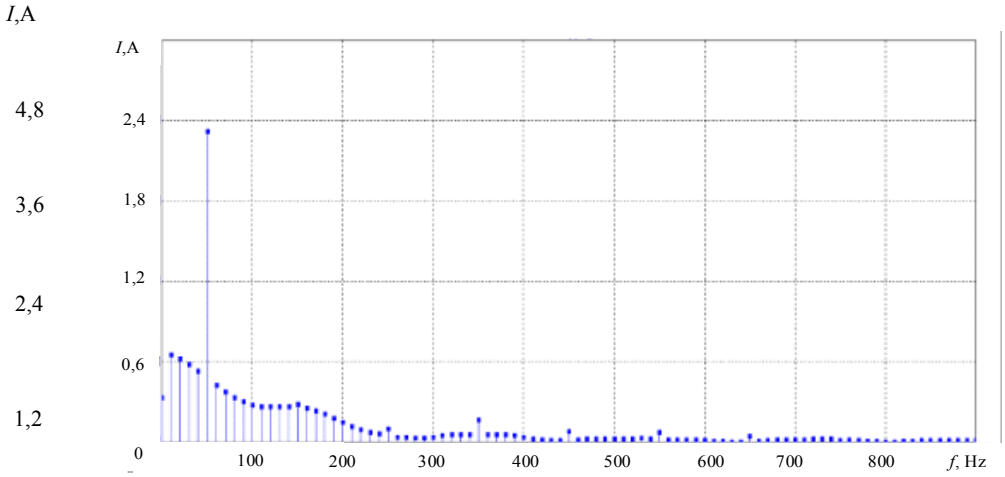


Figure 3. Current spectrum for non-magnetic gap of 0.4 mm

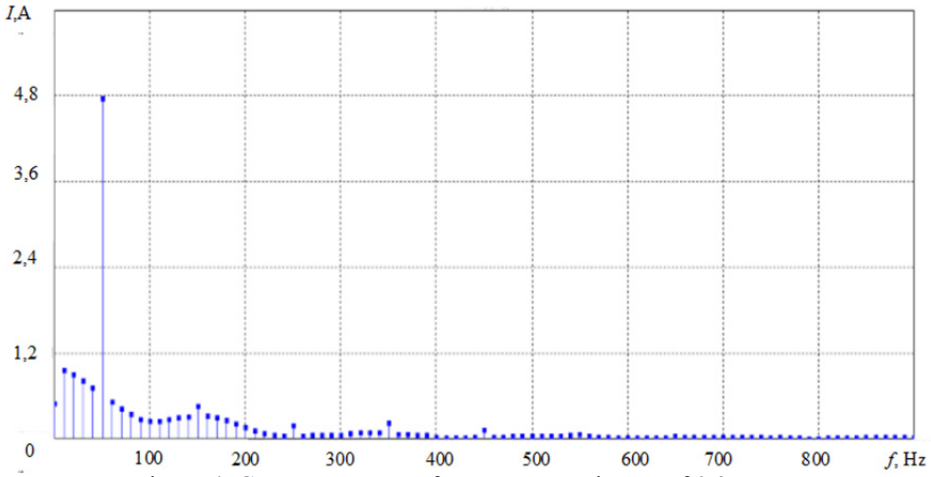


Figure 4. Current spectrum for non-magnetic gap of 0.9 mm

In Figs. 3 and 4 it can be seen that in addition to the first harmonic it is possible to determine higher harmonics from the first to the thirteenth ones. Herewith, the harmonics of current flowing through electric device with the gap of 0.9 mm have higher amplitude than the those through electric device with the gap of 0.4 mm.

Determination of calculated weber--ampere characteristics is carried out by means of the proposed method on the basis of solution of inverse problem of harmonic balance, the results are illustrate din Figure 5, and the number of considered harmonics is given in parenthesis.

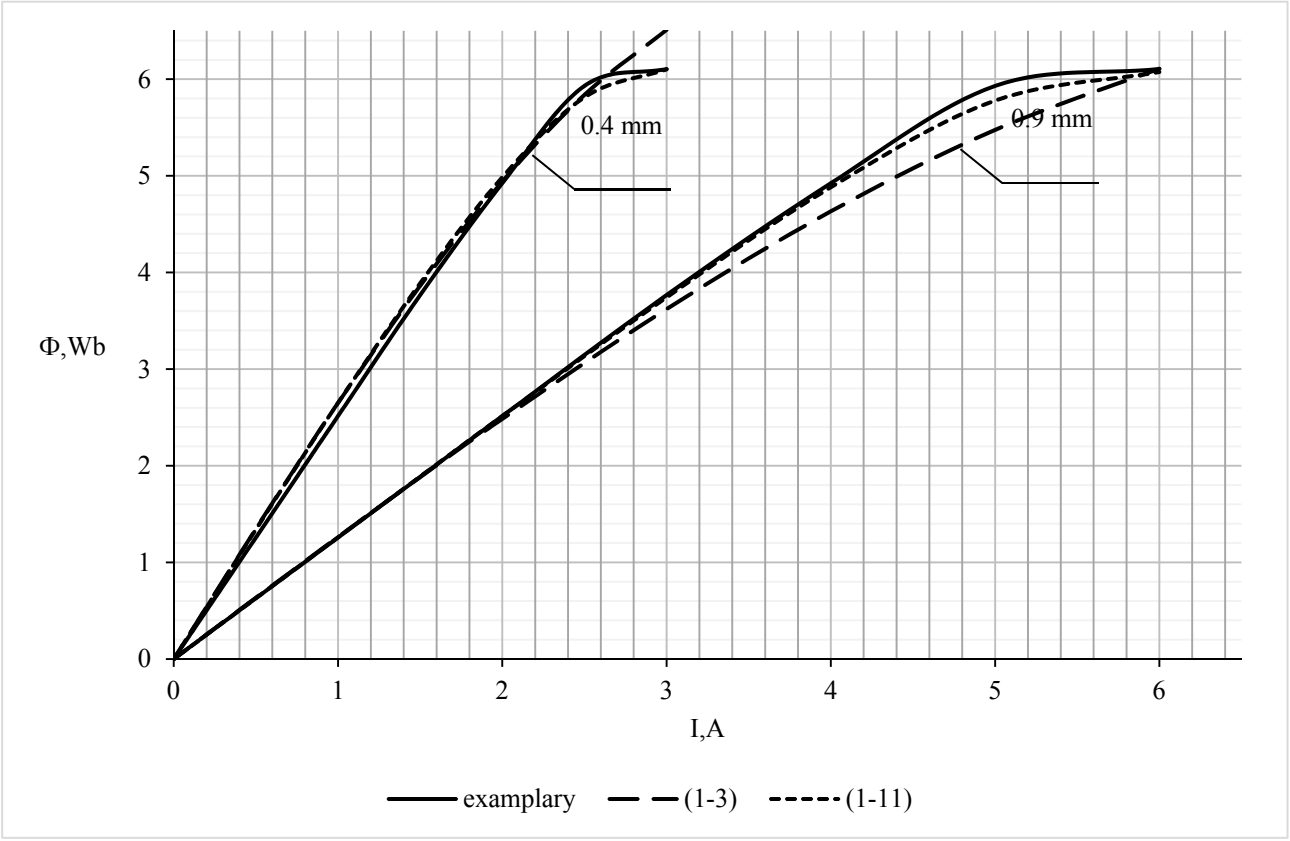


Figure 5. Weber--ampere characteristics of и electric device at preset values of magnetic core gap

3.2 Full-Scale Experiment

Now let us demonstrate the principle of determination of weber--ampere characteristic of electric device on the basis of full-scale experiment. The following electric devices were tested: AC electromagnetic realy RPU-1, asynchronous electric motor KD 1 – 2, and toroidal transformer STT - 12A.

The specifications of RPU-1 are as follows:

Operating coil has 5000 windings;

Coil-winding wire: PEV – 2;

Dimensions (mm): 69x70x37.3;

Weight (kg): 0.25.

The specifications of asynchronous electric motor KD 1 – 2 are as follows:

Supply voltage (V): 127 ;

Operating coil has 3200 windings;

Coil-winding wire: PEV - 2

Power consumption (W): 13;

Dimensions (mm): 4x74x67.5;

Weight (kg): 0.6.

The specifications of toroidal transformer STT - 12A are as follows:

Outer diameter (mm): 75;

Inner diameter (mm): 40;

Grade of magnetic core steel: E320;

Coil-winding wire: PEV – 2;

Operating coil has 750 windings;

Weight (kg): 1.1.

The experiments were performed with the following equipment: laboratory autotransformer Wusley TDGC 8A, two Fluke 289 RMS multimeters, Tektronix 2024b digital oscillograph, PC with the following characteristics:

Wusley TDGC autotransformer:

Power (kW): 2;

Input voltage (V): 220;

Output voltage (V): 0 – 260;

Fluke 289 RMS multimeter:

Basic error (%): 0.025;

Voltage measurement range: 1 μV – 1000 V;

Current measurement range: 10 nA– 10 A;

Resistance measurement range: 0.01 Ω – 500 MΩ;

Memory: 10000 counts;

USB interface;

Tektronix 2024b digital oscillograph:

Transmission range (MHz): 200;

Memory per channel: 2500 samples;

Maximum discretization frequency: 2 GHz;

Maximum input voltage: 300 V;

Input impedance: 1 MΩ;

Number of channels: 4;

Fast Fourier transform;

USB interface;

Memory: 4 oscillograms.

Flowchart of the assembly is illustrated in Figure 6.

Sinusoidal voltage of preset amplitude is supplied from the output of laboratory autotransformer L to electric device, which is measured by means of multimeter M. Shunt R_0 is connected in series with the coil of electric device. Current shape and harmonics are measured by means of digital oscillograph O. The data acquired by the oscillograph and multimeter are transferred to personal computer PC, in order to compute weber--ampere characteristic of electric device.

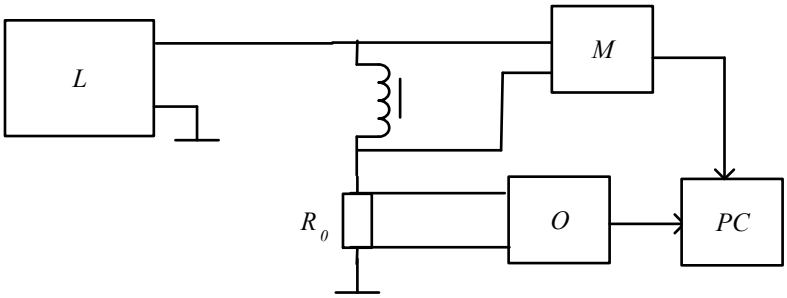


Figure 6. Flowchart of the experimental assembly

Reference weber--ampere characteristics of electric devices were measured by the ammeter--voltmeter method (Komarov, Pokrovskii, Sergeev, and Shikhin, 1984). Measuring coils are applied onto the magnetic cores of electric devices. A coil of electric device is connected to sinusoidal voltage source: laboratory autotransformer. Voltage at the measuring coil and current in the circuit of coil and electric device are measured by multimeters. The obtained results are applied for plotting of weber--ampere characteristic of electric device.

The obtained weber--ampere characteristics of electric devices $\Phi_n(I_n)$, normalized by the values of maximum voltage on the measuring coil and maximum current in the coil of electric devices are illustrated in Figure 7.

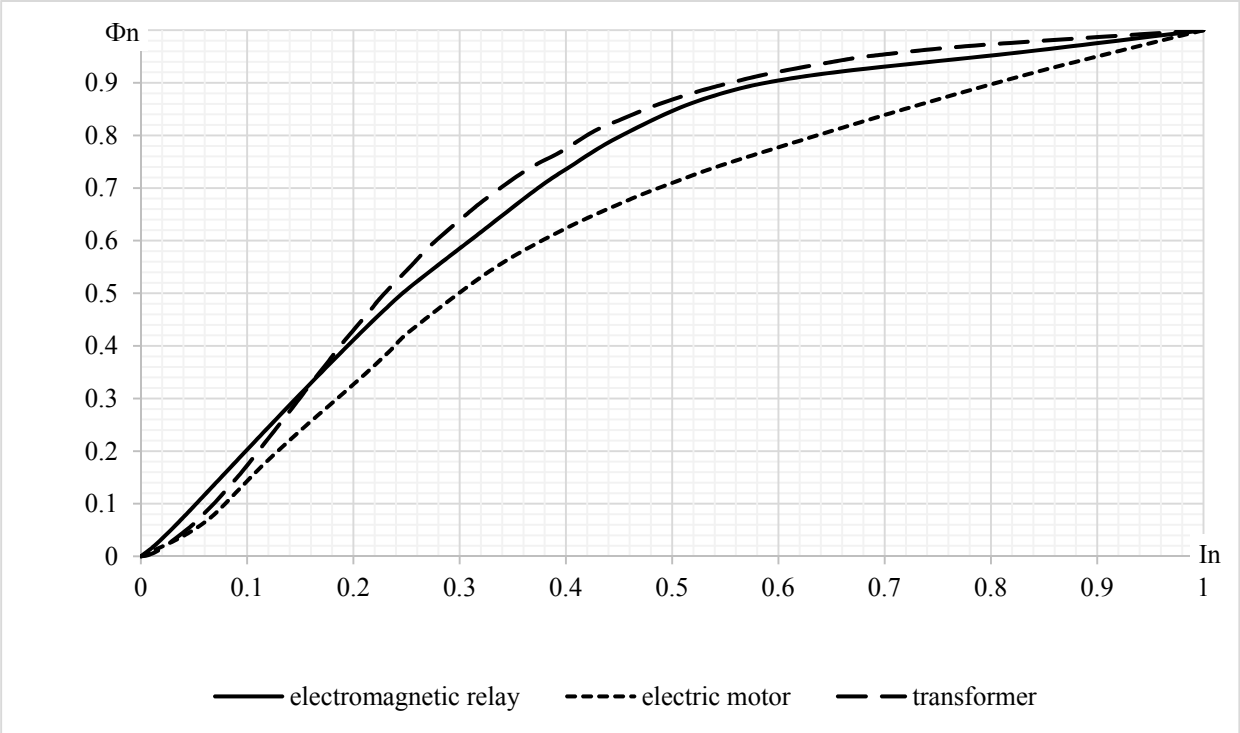


Figure 7. Reference normalized weber--ampere characteristics of electric devices

Calculated weber--ampere characteristics of electric devices are determined by the proposed method based on the solution of inverse problem of harmonic balance. With this aim the amplitude and frequency of the applied voltage, active resistance of coils of electric devices, and the amplitudes of current harmonics in the form of spectrograms are measured.

The results are summarized in Table 1.

Table 1. Results of preliminary measurements

Parameter	U_a, В	R, Ом	I_1, А	I_3, А	I_5,А	I_7, А	I_9, А
Electromagnetic relay	260	488	0.49	0.012	0.00049	0.00042	0.00005
Electric motor	97	64.5	1.47	0.092	0.0018	0.0014	0.00015
Transformer	200	512	0.38	0.004	0.00038	0.00004	0.000039

Figure 8 illustrates the shapes of current in coils of electric devices.

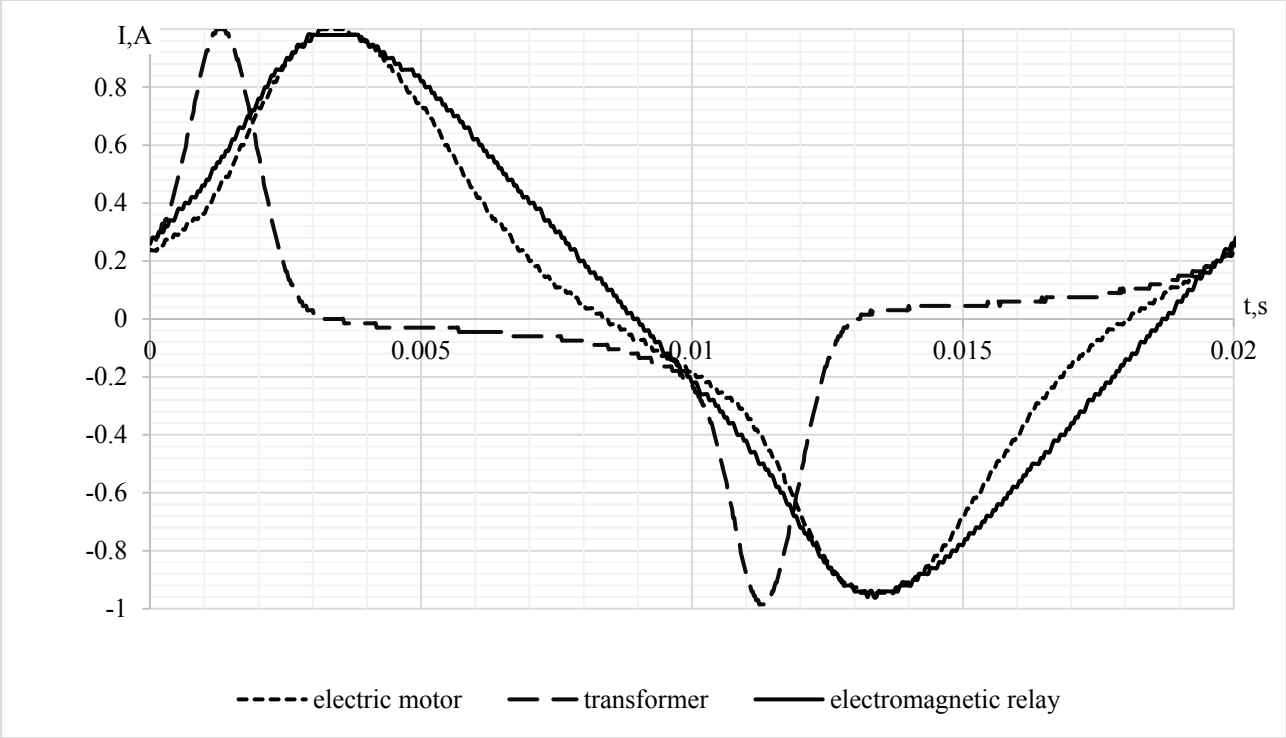


Figure 8. Shapes of current within testing of electric devices

In Figure 8 it can be seen that the current flowing through the coils of electric devices is of various degree of unsinusoidality. The shape of current of toroidal transformer has the highest distortion.

Figures 9--11 illustrate spectrograms of currents in coils of electric devices. The calculations are based on the current harmonics with attenuation not higher than 40 dB.

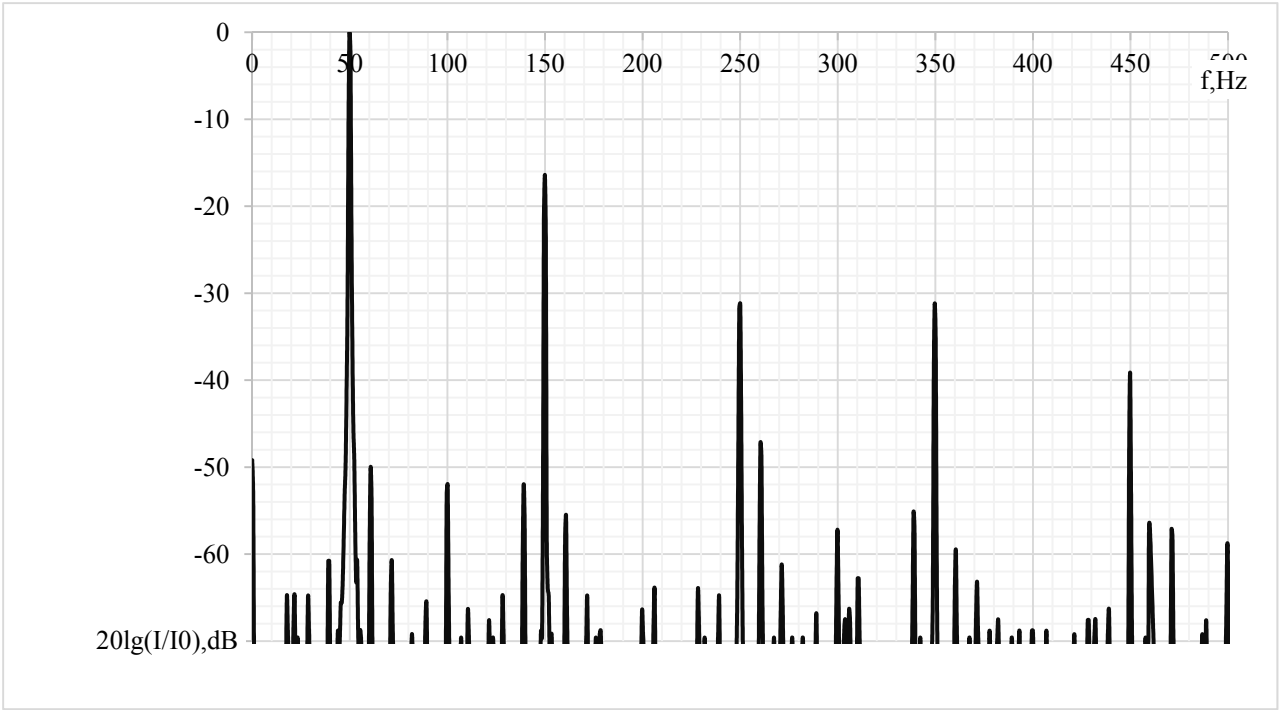


Figure 9. Current spectrum within testing of electromagnetic relay

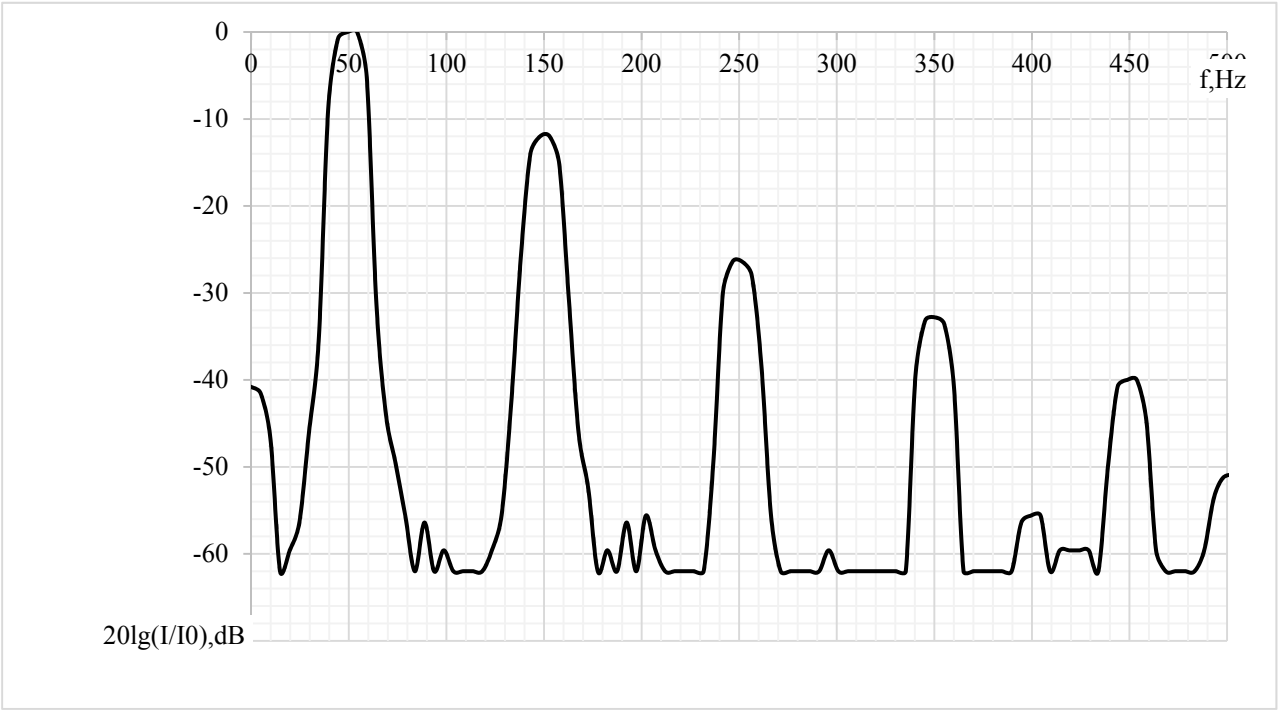


Figure 10. Current spectrum within testing of electric motor

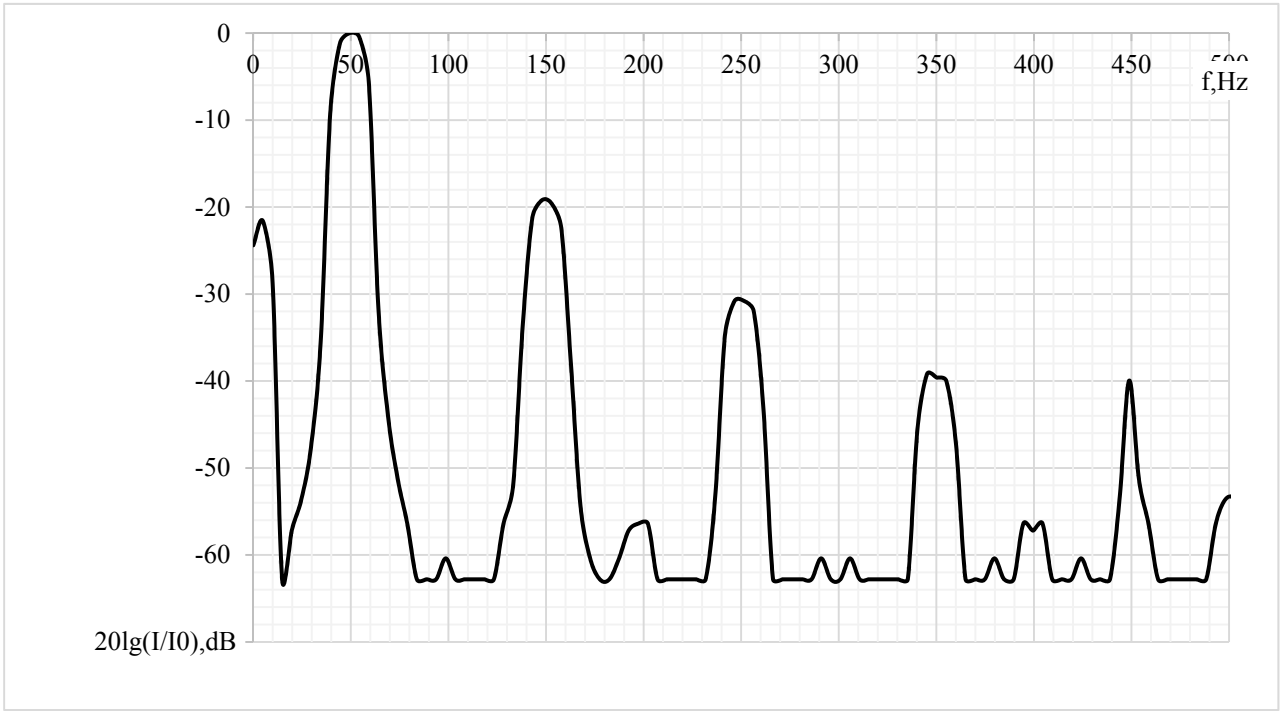


Figure 11. Current spectrum within testing of transformer

The obtained results with the use of the first and the third and from the first to the ninth current harmonics are illustrated in Figs. 12--16.

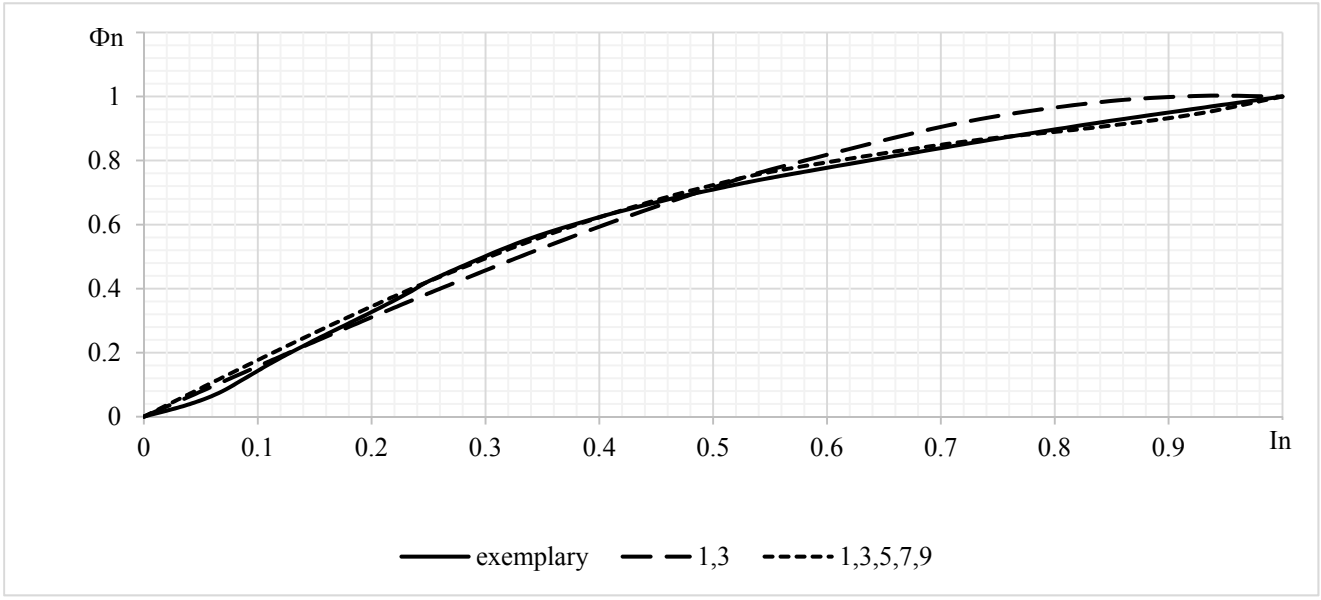


Figure 12. Weber--ampere characteristic of electric motor

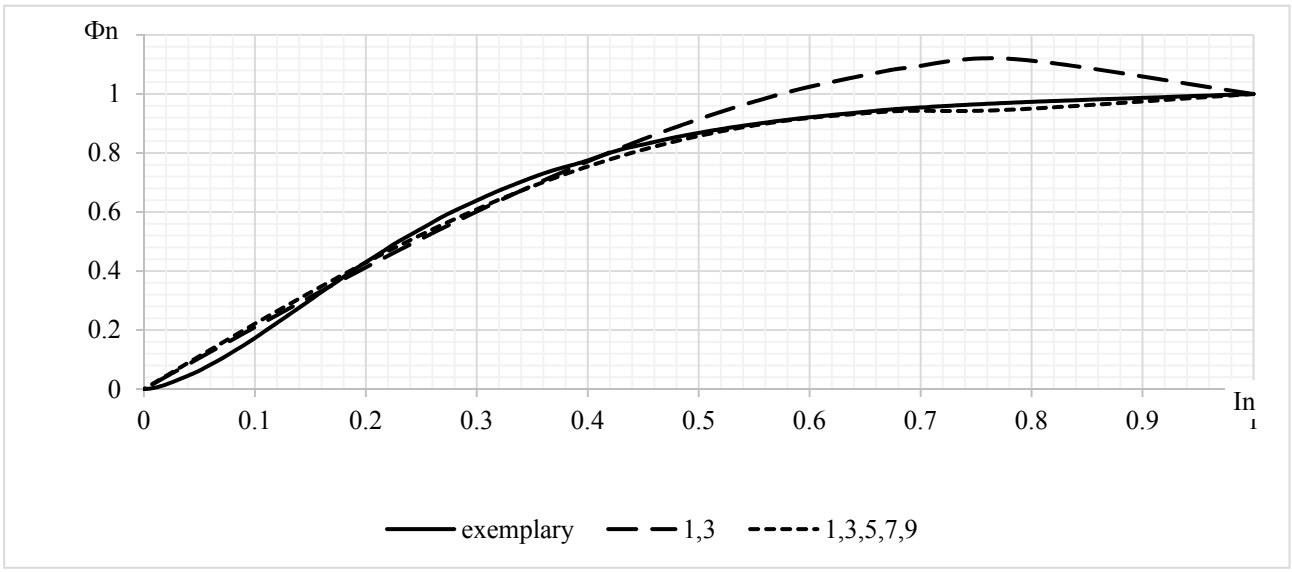


Figure 13. Weber--ampere characteristic of toroidal transformer

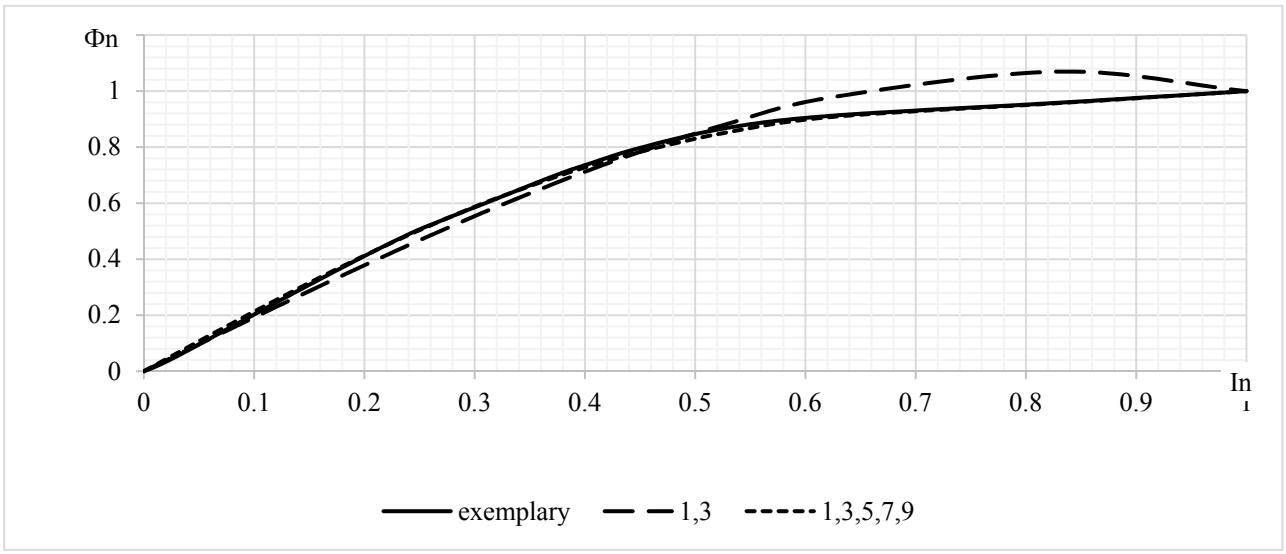


Figure 14. Weber--ampere characteristic of electromagnetic relay

In order to analyze the obtained results it is necessary to calculate the error of determination of weber--ampere characteristics of electric devices. The method of detection of this error (Antonov, Petrov, and Shchelkin,1986; Golt, 2008) is illustrated in Figure 15.

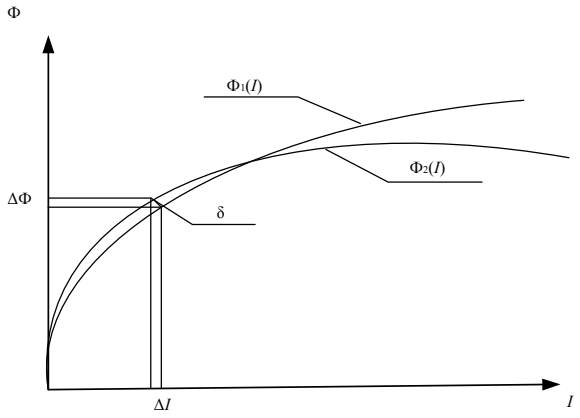


Figure 15. Method of detection of measurement error of weber--ampere characteristic of electric device

A set of points for calculation of the error is determined in reference weber--ampere characteristic $\Phi_1(I)$. Then perpendicular is plotted from the reference weber--ampere characteristic $\Phi_1(I)$ to the obtained weber--ampere characteristic $\Phi_2(I)$, absolute errors for magnetic flow Φ and current I are determined as projections to appropriate axes. Estimated relative errors of determination of weber--ampere characteristic $\Phi(I)$ for current I and flow Φ, as well as total error are determined as follows:

$$\delta_I = \frac{\Delta I}{I}, \quad \delta_\Phi = \frac{\Delta \Phi}{\Phi}, \quad \delta = \sqrt{\delta_I^{\,2} + \delta_\Phi^{\,2}}$$

The results of calculation of measurement error of weber--ampere characteristic of electric device with different non-magnetic gaps: 0.4 mm (weber-- ampere characteristic 1) and 0.9 mm (weber-- ampere characteristic 2), obtained by various number of current harmonics are summarized in Table 2.

Table 2. Errors of determination of weber--ampere characteristics of electric devices with non-magnetic gap

Maximum considered current harmonic			3	5	7	9	11	13
Error, %	weber--ampere characteristic, gap 0.4 mm	δ_I	5.3	2.8	1.3	1.4	1.1	0.9
		δ_Φ	4.5	2.4	2.8	2.4	2.1	1.8
		δ	7.0	3.7	3.1	2.8	2.4	2.0
	weber--ampere characteristic, gap 0.9 mm	δ_I	6.2	5.9	5.6	4.8	4.3	4.1
		δ_Φ	2.1	1.3	0.9	0.9	0.8	0.8
		δ	6.5	6.0	5.7	4.9	4.4	4.2

Table 2 summarizes the results of measurement error of weber--ampere characteristics of electric devices in full-scale experiment.

Table 3. Errors of determination of weber--ampere characteristics of electric devices

Maximum considered current harmonic			Electric motor	Transformer	Electromagnetic relay
Error, %	3	δ_I	21	9	18
		δ_Φ	14	7	12
		δ	25	11	22
	9	δ_I	2.5	2.1	1.2
		δ_Φ	1.8	1.5	0.9
		δ	3.1	2.6	1.5

In Table 3 it can be seen that the use of the first and the third current harmonics provides deviation from

reference characteristic in overall interval of the weber--ampere characteristics, and the use of the current harmonics from the first to the ninth makes it possible to achieve acceptable accuracy of det3ermination of weber--ampere characteristics of electric devices.

The obtained results demonstrate that the successful application of the method considerably depends on correct selection of the number of harmonics. In order to determine the required number of current harmonics involved in the calculation of coefficients k_m of approximating equation of weber--ampere characteristic (9) we apply the known approach (Lvovsky, 1988; Mason, Mukhopadhyay & Jayasundera, 2013; Freund, Mohr, Wilson, 2010).

Successively increasing the number of current harmonics we control the variation of residual dispersion:

$$S_{\text{res}}^2 = \frac{\sum_{i=1}^{n} (\Phi_i - \overline{\Phi}_i)^2}{n - q - 1},$$

where n is the number of points, where the compared reference values of flow $\overline{\Phi}_i$ and calculated Φ_i were measured, q is the number of the used current harmonics.

Termination of addition of current harmonics was determined by verification of hypothesis about the absence of difference between residual dispersions S_{res}^2. With this aim the Fischer criterion was applied, according to which the calculated Fischer criterion is determined as follows:

$$F_{\text{p}} = \frac{S_{\text{res(j)}}^2}{S_{\text{res(j+1)}}^2}$$

and then is compared with the critical value of Fischer criterion for the following number of degrees of freedom $f_1 = n - 1;\ f_2 = n - q - 1$.

When the condition $F_{\text{p}} \le F_{\text{кр}}$ is satisfied, the addition of current harmonics terminates.

The results of determination of the required number of current harmonics at confidence level α = 5 % are summarized in Table 4.

Table 4. Determination of number of current harmonics

Parameter	2p+1	$S_{\text{ост}}^2$	p	f_1	f_2	F_{p}	$F_{\text{кр}}$
weber - ampere characteristic, gap 0.4 mm	**3**	0.091	2	18	16	-	-
	5	0.022	3	18	15	4.09	2.39
	7	0.0081	4	18	14	2.75	2.44
	9	0.0031	5	18	13	2.67	2.51
	11	0.0012	6	18	12	3.00	2.6
	13	0.00072	7	18	11	1.43	2.71
weber - ampere characteristic, gap 0.9 mm	**3**	0.063	2	18	16	-	-
	5	0.021	3	18	15	3.00	2.39
	7	0.0084	4	18	14	2.50	2.44
	9	0.0058	5	18	13	1.45	2.51

From Table 4 it follows that for the first weber--ampere characteristic it is required to apply from the first to the eleventh odd current harmonics, and for the second weber--ampere characteristic from the first to the seventh harmonics.

5. Conclusions

The work analyzes known methods of measurements of weber--ampere characteristic of electric devices, their unsuitability for determination of weber--ampere characteristic of operating cycle is revealed. In order to obtain the set purpose it is proposed to apply solution of inverse problem of harmonic balance. Mathematical analysis of solution of inverse problem of harmonic analysis has been performed, simulation and full-scale experiments have been carried out using mathematical model of the Micro-Cap simulator and three types of electric devices

(electromagnetic relay, electric motor, and toroidal-core transformer). The experimental results make it possible to state that the proposed method of measurement of weber--ampere characteristic based on the solution of inverse problem of harmonic balance facilitates obtaining of characteristic of electric device with the error not exceeding 3 %. This method can be applied for testing of electric devices both within their fabrication and operation.

Acknowledgments

The work was supported by the Project No. 1.2690.2014/K titled Methods of solution of inverse problems of diagniostics of complex systems (in engineering and medicien) based on full-scal and simulation experiments, performed in the scope of designing portion of Federal Program on the premises of Center of Collective Usage "Diagnostics and poer saving electric equipment", Platov South-Russian State Polytechnic University (NPI).

References

Amelina, M. A., & Amelin, S. A. (2007). Micro-Cap 8 Simulator. Moscow, Goryachaya Liniya, Telekom.

Antonov, V. G., Petrov, L. M., & Shchelkin, A. P. (1986). Measurement tools of magnetic parameters of materials. Leningrad: Energoatomizdat.

Bessonov, L. A. (1978). Theoretical foundations of electric engineering: Electric circuits. Textbook for students of electric engineering and instrument building specialties of higher schools. Moscow, Vysshaya Shkola.

Direct-current Magnetic Measurements for Soft Magnetic Materials. ASTM International, 1970.

Frank, W. (2012). Roller Electric and Magnetic Measurements and Measuring Instruments. HardPress.

Georgi, P. (2012). Tolstov Fourier Series. Courier Dover Publications. http://dx.doi.org/ 10.1115/1.3630087

Igor, K., & Valerii, Z. (2010). Radoslaw Kasperek Electrotechnical Systems. CRC Press. http://dx.doi.org/ 10.1201/9781420087109

Joachim Vester Simulation elektronischer Schaltungen mit MICRO-CAP. Springer-Verlag, 2009. http://dx.doi.org/ 10.1007/978-3-8348-9333-8

Kiefer, I. I. (1969). Testing of ferromagnetic materials. Moscow, Energia, 272-273.

Komarov, E. V., Pokrovskii, A. D., Sergeev, V. G., & Shikhin, A. Ya. (1984). Testing of magnetic materials and systems. Moscow, Energoatomizdat.

Lvovsky, E. N. (1988). Statistic methods of development of empiric equations. textbook for higher schools. Moscow: Vysshaya Shkola,.

Mason, A. S. C., Mukhopadhyay, K. P., & Jayasundera, N. B. (2013). Sensing Technology: Current Status and Future Trends. – Springer Science & Business Media. http://dx.doi.org/ 10.1007/978-3-319-02318-2

Michael, C. (2008). Golt Magnetic and Dielectric Properties of Magneto-dielectric Materials. ProQuest. 150 c.

Richard, H. (2010). Enns It's a Nonlinear World. Springer Science & Business Media. http://dx.doi.org/ 10.1007/978-0-387-75340-9

Richard, H. Enns, & George, C. (2000). McGuire Nonlinear Physics with Maple for Scientists and Engineers. Springer Science & Business Media. http://dx.doi.org/10.1007/978-1-4612-1322-2

Robert, J. (1994). Lopez Maple via Calculus. Springer Science & Business Media. http://dx.doi.org/10.1007/978-1-4612-0267-7

Rudolf, J., Freund, D. M., & William, J. W. (2010). Statistical Methods – Academic Press.

Sakhavova, A. A., Shirokov, K. M., & Yanvarev, S. G. (2013). Application of indirect determination of weber-ampere characteristics in automated system of sensorless diagnostics of electromagnetic mechanisms. Modern problems of science and education. Retrieved from http://www.science-education.ru/111-10234

Singh, S. K. (2003). Industrial Instrumentation & Control. Tata McGraw-Hill Education.

Song, X., Suting, C., Zhanming, W., & Jingming, X. (2013). Unifying Electrical Engineering and Electronics Engineering. Springer Science & Business Media. http://dx.doi.org/ 10.1007/978-1-4614-4981-2

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The Use of Wireless Sensor Technologies for Condition Monitoring of Modern Aircraft Structures

Maksim Vladimirovich Sergievskiy[1] & Sergey Nikolaevich Syroezhkin[1]

[1] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia

Correspondence: Maksim Vladimirovich Sergievskiy, National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409, Moscow, Russia.

Received: January 3, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p262 URL: http://dx.doi.org/10.5539/mas.v9n8p262

Abstract

Monitoring of the aircraft structures' during the pre-flight testing is a critical task of the aerospace industry. One of the most promising solutions, not yet widely applied, is continuous monitoring of aircraft structures using wireless sensor network technology. The brief summary of the proposed system is the following: special sensors (devices which detect potential deformations), send signals to the local motes (autonomous computing device equipped with a wireless transmitter). Information from motes is gathered by routers which then transfer the aggregated information to the datacenter. Motes are used for collection and primary processing of the data from sensors, which then sent to the routers, and, ultimately, gateway of corporate network. Applications of corporate network control and define flexible patterns for processing of the information received from sensors. As an example of hardware components for preprocessing and data transmission the system described in this study uses motes and routers from MEMSIC, for measuring the level of deformation - sensors from Tokyo Sokki Kenkyujo. This network structure allows to centralize data collection modes in the process of testing; implement continuous data collection at a defined frequency; process and display data in real-time.

Keywords: Wireless Sensor Networks (WSN), monitoring of aircraft constructions, mote, protocol stack, router, cluster topology, network lifetime, MoteWorks platform

1. Introduction

1.1 Internet of Things and WSN

Internet of Things (IoT) – is an environment, which provides an interface for a variety of objects, equipped with computing and communication resources. In this case, Internet is understood broadly, not only as a method of transmitting data within a network using specific protocols and technologies, but also as a set of different services.

In Gartner's IT-glossary IoT is defined as "the network of physical objects that contain embedded technology to communicate or interact internally or with the external environment". Obviously, such a definition of IoT is very close to the definition of wireless sensor networks (WSN) (Gartner , 2014). However, the concept of IoT is broader than WSN concept, because the IoT brings together, firstly, animate and inanimate objects, and, secondly, the objects that do not have an active communication tools (Hart & Martinez, 2006).

WSN consists of a number of small computational and communication devices - motes. Mote is a board, which size is usually less than one cubic inch, which consists of: a processor, memory (Flash and RAM), digital-to-analog and analog-to-digital converters, RF transceiver, power supply and sensors (Cordeiro & Agrawal, 2006).

The sensors are connected to the board via the digital and analog connectors and can be very diverse: temperature, pressure, humidity, light, vibration, electromagnetic, chemical and others. The set of installed on mote sensors depends on the functions performed by the WSN. As the power source nodes use small batteries; typically, rechargeable batteries are not applied, since its charging is impossible in most cases. Motes serve for the collection, basic processing and transmission of the data.

Gateway is a special network's node, a sufficiently powerful computer, which performs the main functional processing of collected data. To retrieve data from the motes, node is equipped with a radio transmitter. Because of the limited range of the transmitter, gateway does not receive information directly from every network's mote, it can obtain data just from the motes that are located close to it. The problem of obtaining sensors data by gateway is solved as follows: motes can communicate with each other using transceivers operating in the radio bands. They

send, information collected by motes sensors, as well as service information (status of devices and the results of data transmission).

Information is transferred from one mote to another by chain, and finally nearest to gateway motes send all the collected data to it. If some motes fail, the sensor network operation should continue after reconfiguration and changing its structure. But in this case, of course, the number of motes, sources of information, decreases.

To perform the functions a special operating system is installed on each node. Currently one of the most common operating system for WSN is TinyOS (Sergievskiy & Syroezhkin, 2008; Crossbow Technology, 2014). It is event-driven real-time operating system, designed to work in limited computing resources environment. This OS, as well as a number of others (UC Berkeley, 2013; Kosachev & Ponomarenko, 2006), has special libraries and protocols that allow motes automatically establish communications with neighboring items to form a sensor network of a given topology. The controlling program of wireless sensor node describes all its operation logic, including such basic things as protocols of interaction with hardware, the primary processing of sensory data, and more complex processes, such as routing algorithms, encryption and data transmission.

The most important factor when working with wireless sensor networks is a limited capacity of the battery, mounted on mote (IEEE Std 802.15.4a, 2007; Shen, Harte, Popovici, O'Flynn, & Atk, 2009). It should be taken into account, that it is impossible to replace the battery in many cases. For this reason, firstly, motes should perform only simple primary processing, aimed at reducing the amount of transmitted information, and secondly, perhaps the most important, the number of data transmission and reception cycles should be minimized. To solve this issue special communication protocols were developed. The most common of which are the protocols of the ZigBee Alliance [], which were created to coordinate activities of wireless sensor technologies. In general, to produce a standard, including protocol stack for wireless sensor networks, ZigBee used the previously developed the IEEE 802.15.4 standard, which specifies physical and medium access layers for wireless data transmission over short distances (up to 75 m) with low power consumption, but high degree of reliability.

1.2 The Example of Monitoring System

The main purpose of the use of sensor networks is to control the state of objects during its operation. The critical components and systems can be equipped with a wireless data collection systems and diagnostic tools, for analysis and decision-making (Barbar´an, D´ıaz, Esteve, & Rubio, 2007; Mainwaring, Polastre, Szewcyk, & Culler, 2002).

Here is one example of such solution for monitoring of the critical object's state. We are talking about a wireless sensor network "RadMote" (Barbar´an, D´ıaz, Esteve, & Rubio, 2007), established under the project "Secure Middleware for Embedded Peer-to-Peer Systems", which is designed for staff safety of nuclear power stations in Spain[]. The network and the monitoring system, created on its basis, perform environment's state monitoring on object, nuclear power plant. For this purpose, many motes, equipped with radiation sensors, are installed inside and outside the station. At any time new nodes can be added upon request. Motes organized in a mesh network that allows you to transmit sensory data between each other. Sensor nodes are programmed to recognize the risk situation, when the indications of radiation sensors exceed a threshold value, and to send alarms. Finally, the data should come on intermediate nodes, equipped with more powerful radio transmitters. Once the intermediate node gets the data, it immediately passes data to PDAs, which are used by the station employees. Thus, each employee of the station can monitor the radiation level in real time.

2. Problem Statement and Methodic

Now let's discuss a solution of another important task - the control of modern aircraft structures' state in the process of testing and operation. This problem is relevant for the aircraft industry, because hidden defects caused by mechanical damage can lead to emergencies and accidents. Different manufacturers solve the task of improving the reliability and safety of the aircraft differently. They perform more preflight tests, analyze results and make estimates.

One of the most promising ways to solve this problem, not widespread yet, is the continuous monitoring of aircraft constructions in the process of testing and operation with the output on the dashboard of the necessary information. There are all preconditions to populate this approach:

- wide range of compact sensors (strain, temperature, pressure, vibration, location, etc.) with different mounting options;

- wireless sensor technology to collect information from sensors and send it to consumers; wide capabilities of control of information gathering modes;

- data processing techniques, implemented by custom applications.

The main aim of this work is to create a wireless system for monitoring of aircraft structures' deformation level during static and dynamic tests.

3. Results

3.1 System Architecture

The structure of the monitoring systems in most cases is the following: there are sensors connected to motes on the lower level. Several sensors can be connected to one mote; it is possible that the sensors are located on the mote's board - it depends on the tasks, which sensors perform. In our case, sensors are mounted directly on a controlled structure.

Motes are only used for collecting and converting the primary sensor data, which they send to each other and, eventually, to a special unit - gateway (base station) that has a wire connection to the server (the network). The network may contain multiple gateways; it increases its reliability and performance. Controlling of information collection modes and main processing of the obtained data is performed by applications on the corporate network. This approach allows to get rid of a large number of cable connections, since before wire connections were widely used in the airspace industry (Chalermek Intanagonwiwat & Ramesh Govindan, Deborah, 2000; Haenselmann, 2006).

In case, when sensors are very tight, a star topology is considered to as the most effective. In this case, motes will transmit information to special routers that communicate with the base station. Routers are not equipped with sensors.

Thus, the network consists of a set of clusters, each formed by a router and motes with sensors, polled by router. Routers forward data to each other and, finally, the data comes to the gateway. The reliability of such systems is very high: even failure of a large number of motes or routers will not affect the operation of the system. Additionally, routers may be located outside of the deformable structure (Jangra, Richa, Swati, & Priyanka, 2010).

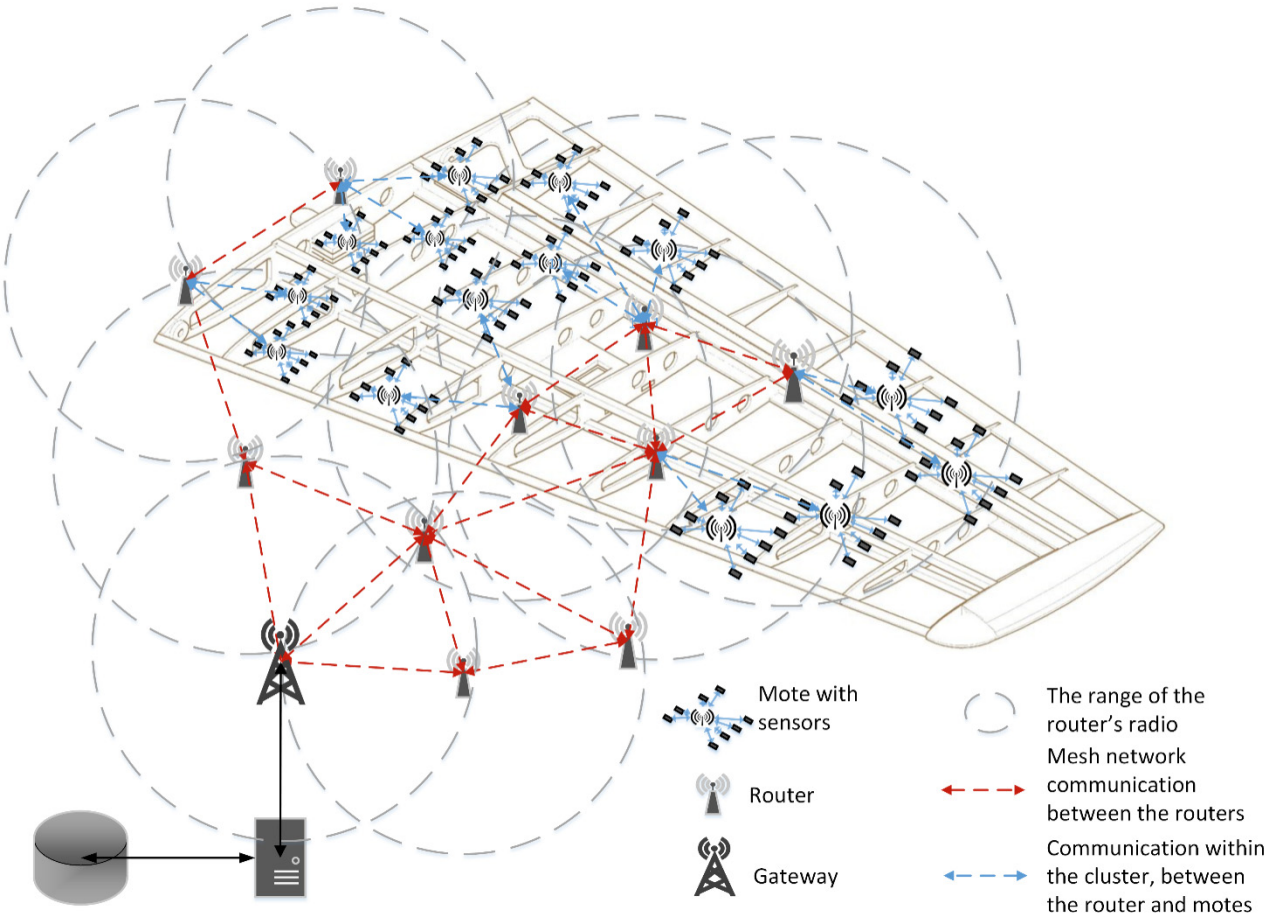


Figure 1. Structure of monitoring system

Figure 1 shows a monitoring system consisting of a gateway router and sensor motes. Each mote has several sensors, connected to it by wires. To save energy motes are unable to receive and retransmit data from other motes, they only can send to router their own data.

Some routers have a direct link with the gateway, while others are transferring data by a longer chain, using other routers. This approach has both advantages and disadvantages. On one hand, it allows the use of only one gateway, but on the other, it increases utilization of some of the routers. Since in this domain network topology usually varies little during operation, these routers will fail earlier than others. Knowing the characteristics of a particular network topology makes it possible to increase the number of routers and efficiently balance the load between them, extending the network lifetime and avoiding potential loss of data (Joel K., 2008).

The described network architecture can be attributed to the hybrid type, where some nodes (routers) form a mesh network, and the router and its motes actually use the cluster topology.

An important task that WSN developer is faced is to increase the operating time of the network without having to replace the batteries. The most energy-consuming operations for wireless node is sending and receiving data. In the process of deformation level monitoring of aircraft constructions during static and dynamic tests measuring frequency reaches 200 Hz, that is, WSN should transmit a quite large data stream (Kay & Mattern, 2011; Nakamura, Loureiro, & Aleja, 2007).

Let's determine the amount of payload data that should be transmitted to the gateway over the network, consisting of N motes with M sensors on each, if one deformation level measurement takes two bytes of memory.

$$Q = 2 * N * M(B)$$

If the required measurement frequency of deformation is H Hz, the flow of information transmitted over the network as a whole, will be equal to:

$$S = Q * H = 2 * N * M * H(Bps)$$

For example, for a network consisting of 100 motes with 10 sensors on each, and a measurement frequency of 10 Hz, the flow of information will be 20 KB/s. This estimate does not take into account the fact that the overhead information, accompanying each message transmitted over the network, increases the total network traffic in about 7-10 times. Thus, the actual flow in this case is equal to 140-200 KB / s. Note, this is an estimate of the total data stream that arrives at the gateway, for routers and, especially, motes values of local flows will be significantly lower. Continuous processing and transmission of the data stream may trigger certain network components (routers or motes) failure in some cases. In order to maximize the network lifetime, you can apply several solutions related to the choice of its functioning modes. Let's discuss them in the next paragraph.

The process of developing of wireless deformation level monitoring system includes not only determining the network structure and configuration, but the solution of two interconnected problems - choice of hardware components and software creation.

3.2 Hardware Components

To measure the deformation level of aircraft structures we use strain gauges model "FLA-5-17" by Tokyo Sokki Kenkyujo Company. Basic characteristics of the sensors are shown in the Table 1.

Table 1. Basic characteristics of the sensors

Applicable specimen	Metal, Glass, Ceramics	Backing	Epoxy
Operational temperature(°C)	-20~+80°C	Element	Cu-Ni
Temperature compensation range(°C)	+10~+80°C	Strain limit	5% (50000×10^{-6}strain)
Bonding adhesive	CN, P-2, EB-2	Fatigue life at room temperature	1×10^{6} ($\pm1500\times10^{-6}$strain)

As a hardware platform for primary data processing and transmission, we use motes IRIS (Figure 2) from MEMSIC. They have a wide range of features including built-in processor, ADC, multi-channel radio transmitter, RAM, and interface for connecting sensors (Crossbow Technology, 2007).

Figure 2. Mote IRIS

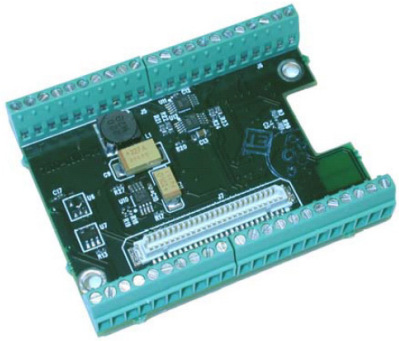


Figure 3. MDA300CA board

Special data captures and interface board are used in this WSN for strain monitoring. Both of these cards connect to the mote through the 51-pin interface. To connect strain gauges FLA-5-17 to wireless motes IRIS a special data capturing board MDA300CA (Figure 3) is used. This Board is attached to the IRIS Mote's extension connector and provides a set of digital and analog inputs for connecting a variety of sensors.

Every strain gauge connects to a data capturing board through any differential pair of channels and Ai + Ai-, which are used for connecting external sensors with a gain of 100. Thus, a single board can connect up to 10 sensors of the strain level. The dynamic range of channels is ±12.5 mV. The result obtained after analog-to-digital conversion of the sensor readings, can be translated into a voltage V (mV) according to the following formula:

$$V = 12.5 * (АЦП / 2048 - 1)$$

For routers, we use the same IRIS motes, but they are not equipped with sensors. Sensor readings received by mote, should be transferred wirelessly to the server for further processing.

Figure 4. MIB520 board

Special multi-purpose USB-interface board MIB520 (Figure 4) of MEMSIC is the link between the wireless sensor network and the server. This board (gateway) with a transmitter module, integrated with IRIS mote, on one hand is able to exchange messages with the wireless network nodes, and on the other to transmit and receive data via USB-port of an ordinary computer.

3.3 Software Components

For development of software components of condition monitoring system, we use a special platform - MoteWorks of Crossbow Company (Crossbow Technology, 2007). This platform is based on the TinyOS operating system. Developers of the MoteWorks platform considered features of such systems and created the product, which allows create effective and reliable decisions.

The monitoring system software has two groups of components: network and server. Network components include a control application for motes and special software, to ensure the normal functioning of the network. Control application implements all algorithm required for monitoring and data transmission. For the data transfer, application uses stack of network protocols XMesh of MoteWorks platform. The stack of protocols XMesh supports multilink transmission of messages and conforms to the standard 802.15.4 (Linear Technology, 2013).

Server software components of the monitoring system are running on the server special services, with main tasks to 1) collect, 2) store and 3) process the readings obtained by the network sensor. Server components uses XServe platform of MoteWorks as a basis, which collects incoming data on the interface card, converts them to a format specified by XML- template, and stores in a special database.

4. Discussion

While monitoring of the aircraft structures state, experts rarely analyze obtained sensor data in the real-time, often they study them after experiments. Thus, often there is no need to deliver sensors data on gateway in real-time, and hence, a better approach from the standpoint of energy saving, would be an accumulation of a certain amount of data at the node, its compression, and only then transmission to the network. This operating mode allows you to switch the radio transmitter in the sleep mode, and thus save energy. In addition, it will allow solve another problem - avoiding conflicts arising from the simultaneous transmission of data from different routers.

Dependency of energy consumption on data transmission distance is elevated, which means that in addition to the transmission time, the distance between transmitter and receiver and the transmitter's power also determines energy consumption. To reduce the cost of energy, required to data transmission, it is important to design the network in such a way to prevent the transmission over a long distances. Furthermore, the configuration of the domain allows us to pre-arrange the location of nodes, and hence it becomes possible to calculate the transmitter power that is required for reliable operation of the network.

In practice it is desirable to use minimum transmitter power, as this will not only reduce the energy consumption for data transmission, but also reduce the number of possible collisions (Tiwari, Ballal, & L. Lewis, 2007). Apart from that, there is a flexibility to adjust the number of active sensors, taking already available data into the account. At present, the maximum amount of sensors connected to a single mote can't exceed ten.

As an alternative it is also possible to look at WSN with variable technology. But it will most likely lead to increasing probability of collisions, as well as higher energy consumption and declining time efficiency (due to regular reconfiguration of the network requirement).

5. Conclusions

In conclusion we highlight the key advantages of the use of wireless sensor networks for controlling and monitoring aircraft structures:

- Lack of cables or a minimum number of cables;

- very small form factor of motes that basically allows to embed them in the body of various aircraft structures or connect to various devices;

- reliability of individual elements, and the whole system in the process of information collection and transmission;

- minimum energy consumption for receiving and transmitting radio signals, which allows you to use compact batteries;

- simplicity of a sensor network self-organization;

- scalability, allowing to ensure the deployment of networks with large number of nodes, which is important for a

detailed analysis of aircraft structures.

The use of aircraft structures monitoring system, created on the base of wireless sensor network technology, will allow to organize the centralized management of the data collection modes during experiments; to carry out continuous data collection from sensors of various type with the set frequency of measurements; quickly to process and display data in real time.

It is worth saying a few words about future prospects of this research. In this study we identified a number of areas which could improve efficiency of the monitoring system of aircraft structures deformation:

- to optimize the operation patterns of motes and routers, minimal power consumption without compromising reliable of transmission and processing of data from sensors;

- to analyze additional conditions of aircraft structures, for example the level of vibrations;

- to apply more advanced hardware and software platforms, that allows to build in motes in the design of object.

- to offer methods of using these type of systems, based on WSN, in the day-to-day aviation.

References

Barbar´an, J., D´ıaz, M., Esteve, I., & Rubio, B. (2007). RadMote: A Mobile Framework for Radiation Monitoring in Nuclear Power Plants. *International Journal of Electronics, Circuits and Systems, 1*(2).

Chalermek, I., & Ramesh, G. D. (2000). Directed diffusion: A scalable and robust communication paradigm for sensor networks. MobiCom, 56-67. http://dx.doi.org/10.1145/345910.345920

Cordeiro, C., & Agrawal, D. (2006). Ad hoc & sensor networks: theory and applications. Hackensack, NJ: World Scientific Pub. Co. http://dx.doi.org/10.1142/6044

Crossbow Technology. (2007). XMesh User's Manual. Crossbow Technology.

Crossbow Technology. (2014). Retrieved from Crossbow Technology. Retrieved from www.xbow.com

Gartner . (2014, 10). Retrieved from Gartner IT Glossary. Retrieved from http://blogs.gartner.com/it-glossary

Haenselmann, T. (2006, 04 05). GFDL Wireless Sensor Network textbook. Retrieved from http://pi4.informatik.uni-mannheim.de/~haensel/sn_book

Hart, J., & Martinez, K. (2006). Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews, 177-191. http://doi:10.1016/j.earscirev.2006.05.001

IEEE Std 802.15.4a. (2007). IEEE-SA Standards Board. IEEE Standard for Information technology — Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs).

Jangra, A., Richa, Swati, & Priyanka. (2010). Wireless Sensor Network (WSN): Architectural Design issues and Challenges. *International Journal on Computer Science and Engineering, 2*(9), 2010. ISSN: 2347-3878.

Joel K., Y. (2008, 08 25). Clearing up the mesh about wireless networking topologies. Part 1. Retrieved from http://www.embedded.com/design/connectivity/4007667/Clearing-up-the-mesh-about-wireless-networking-topologies-Part-1

Kay, R., & Mattern, F. (2011). The Design Space of Wireless Sensor Networks. IEEE Wireless Communications, 54-61. The Design Space of Wireless Sensor Networks.

Kosachev, A. S., & Ponomarenko, V. N. (2006). Real time operating systems. Preprint, Institute for System Programming.

Linear Technology. (2013). Retrieved from Dust Networks. Retrieved from www.dustnetworks.com

Mainwaring, A., Polastre, J., Szewcyk, R., & Culler, D. (2002). Wireless Sensor Networks for Habitat Monitoring. ACM International Workshop on Wireless Sensor Networks and Applications. Retrieved from http://www.intel-research.net/Publications/Berkeley/120520021024_43.pdf http://dx.doi.org/10.1145/570738.570751

Nakamura, E., Loureiro, A., & Aleja. (2007). Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys, 39*(3), 9. http://doi.acm.org/10.1145/1267070.1267073

Sergievskiy, M. V., & Syroezhkin, S. N. (2008). Wireless sensor networks. ComputerPress. Retrieved from http://www.compress.ru/article.aspx?id=17950&iid=831

Shen, C., Harte, S., Popovici, E., O'Flynn, B., & Atk, R. (2009). Automated protocol selection for energy

efficient WSN applications. *Electronics Letters, 45*(21), 1098-1099. http://dx.doi.org/10.1049/el.2009.1217

Tiwari, A., Ballal, P., & L. Lewis, F. (2007). Energy-efficient wireless sensor network design and implementation for condition-based maintenance. ACM Transactions on Sensor Networks (TOSN), 1-es. http://dx.doi.org/10.1145/1210669.1210670

UC Berkeley. (2013). TinyOS Open WSN project. Retrieved from http://www.tinyos.net

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Justification of the Technology for Preventing Scale in the Downhole Equipment

LiliyaAl'bertovnaShangaraeva[1]

[1] National mineral resources university "University of Mines", Russia

Correspondence: LiliyaAl'bertovnaShangaraeva, National mineral resources university "University of Mines", 21st line, 2, St-Petersburg, 199106, Russia.

Received: January 15, 2015 Accepted: February 13, 2015 Online Published: July 25, 2015

doi:10.5539/mas.v9n8p270 URL: http://dx.doi.org/10.5539/mas.v9n8p270

Abstract

The formation of scale in the downhole equipment and the bottomhole formation zoneis often a serious cause of energy losses, it reduce the productivity of the wells, leading to the current unplanned and costly workovers. Fight with scale in the well operation is one of the most important ways to increase the efficiency of oil production. The technology for preventing scale in the downhole equipment was investigated in the present work.

Keywords: Scale, downhole equipment, inhibitor, formation squeeze treatment, phosphonic acid, bifluoride of ammonia, bottomhole formation zone, adsorption-desorption characteristics

1. Introduction

One of the major problems in the oil and gas industry is to enhance oil recovery in complicated geological and technological conditions. This is due to the assumption of the most of high-yield oil fields in the final stage of the development and the growing proportion of hard to recover reserves to their total amount. Nowadays process of oil extraction is significantly complicates with high water cut production that leads to scale in the downhole equipment (Muslimov, R., 2005)

An integral part of the development and exploration of oil fields is a water factor. As a result of flooding is the formation of scale. Accumulation of scale in the downhole equipment leads to reduced productivity, premature failure of the pumping equipment, current unscheduled and costly workovers and as a consequence, a significant deterioration of the technical and economic indices of oil and gas companies (Zdolnik, S., &Akimov, O., 2009).

Water is of paramount importance in the oilfield, as the scale will occur only when the water is used. Water is a good solvent for many materials and can carry large amounts of scale mineral.

All natural water contains dissolved components resulting from contact with the mineral phases in the natural environment.This leads to the fact that the complex fluids, which are rich in ions, some of which are at the limit of saturation for certain mineral phase, is form (Krabrtri, M., et al., 2002).

Calcium carbonate, calcium and barium sulfate are major components of most oilfield scale. In the oil wells deposits of pure scale are rare. Typically they are a mixture of one or more basic inorganic components with corrosion products, sediments and particles of sand, or covered with asphalt resins paraffinic substances. Wells treatment cannot be successful without the removal of the organic component of the scale (Zdolnik, S., & Akimov, O., 2009).

The mechanism of scale formation consists of several stages. The first stage of development begins with a saturated solution of unstable formation of clusters of atoms. Then formed primary crystal nucleus, when atomic clusters transform into small crystals-buds. These crystals gradually grow due to adsorption of ions on the surface of the crystal imperfections, increasing their size, are combined together into larger aggregates. For some time large crystals formed in the solution or aggregates that they may no longer be kept in suspension in the solution is released and the solid phase (precipitate) (Krabrtri, M., et al., 2002).

Crystal growth occurs at the initiation of certain physico-chemical reactions at the existing boundary between a solid and a liquid. Portions of such reactions are various surface defects such as roughness surfaces of the tubes, the perforations, joints and seams in tubing and pipelines (Kashchavtsev, V., 2005).

Scale occur in three forms: as a thin scale or friable flakes in a layered form, in crystalline form. The deposits of the first type have a loose structure, permeable and can be easily removed.

Radioactive barium scale are the most hard-to-out of scale. Barium is often found in highly mineralized formation waters of oil fields, where its concentration often reaches 0,15-0,5 gpl. Even in the presence of low concentrations of sulfate ion barite ($BaSO_4$) precipitates. These scale are less soluble. For example, in distilled water at 25 ° C dissolves all 0,0023gpl of barite, which is almost 900 times less than the solubility of gypsum (Antipin, Y., &Islanova, G., 1999, Rizkalla E. 1983).

Noted that the barite scale, selected from oil-field equipment and tubing have increased radioactivity due to the presence of radioactive isotopes of radium, which are associated in groundwater with barium. Effect of pressure on the solubility of $BaS0_4$ understudied. In general, there is a slight increase in solubility with increasing pressure is particularly pronounced in solutions with salinity less than 30-50 gpl. With the increase in salinity influence of pressure effect is negligible (Symeopoulos B. &Koutsoukos P., 1999). Thermobaric conditions while moving upward flow of liquid through the wellbore produce little effect on the change in water solubility of barite.

The main reasons for scale in the wells are mixing incompatible waters resulting from the operation of several productive layers simultaneously or in wells operated with one layer with behind-the-casing flows from related horizons (Larichev T. et. al., 2006). Often the cause is a breakdown of the technical state of production casing and packer leaks, especially in fields on the later stages.

2. Method

The scale inhibition in downhole equipment does not have a unique solution, every case of scale formation has its own individual approach. Predicting possible complications associated with scale has particular importance in addressing such a complex multisided problem. For the successful solution requires constant well monitoring and monitoring of physico-chemical processes in it. Great help in this matter could have a map of changes in the produced water composition in various components of: chlorides, sulfates, barites and mineralization.

Casing leak of the production string that leads to scale formation in the wells can be detected as the resultbased onfollowing main characteristics. The main idea is to track changes in the produced water composition in the dynamics. So reducing the content of barium ions, then the appearance of sulfates, after complete disappearance of the barium shows may indicate leakage of production casing and imminent failure of the pump due to scale. An additional feature is the presence of barium sulphate in the suspended solids.

For most wells, was confined to Permian strata, is typical the content of barium ion in the range of 0,1– 0,8 gpl, and a low content of sulfate ions in the range of 0 – 0,25 gpl. Well under the influence of freshwater or wastewater downloadscontains thebarium ions in the range of 0 to 0,1gpl. For waters of the upper horizons (except Tournaisian) is characterized by a high content of sulfate ions (from 1 to 12 gpl). If produced water from the Devonian wells has a sulfate ion concentration greater than 0,3gpl, this wells require additional study in case of casing leak of the production string. The high content of sulfate ions without leakage may be the result of the activities of the enhanced oil recovery or geological and engineering operations (Shangareeva, L. &Petukhov, 2013, Shangareeva, L. &Petukhov, 2011).

Nowdays the most widespread in oil-field practice is the reagent method of preventing scaling. The wide range of chemical compounds which show the effective inhibiting properties against scale (Tomson, N. et. al., 2003, Glushenko, V. et. al., 2008). But, despite the huge contribution of scientists in the development of this direction, the development of chemical compositions remains a challenge (Guskova, I., 2011).

One of the most effective method is the injection of an aqueous solution of the scale inhibitor in the bottomhole formation zone for its subsequent long enough removal in the wellbore during its operation (Shaidakov, V., 2012). Existing methods of selection reagent for specific conditions are based on the selection of the brand of the inhibitor and on its capacity. The adsorption-desorption characteristics of the inhibitor are not taken into account. However, the amount of inhibitors adsorption on the formation and its subsequent desorption depends on it. They affect the efficiency and duration of effect of the compositions. At the same time, it is known that reservoirs have different wettability and different absorption capacity. In this regard, inhibiting composition has to meet strict requirements. It should adsorb as quickly and broadly as possible on the surface of the formation during injection, and should very slowly desorb from formations during well operation. Thus, the longer will be the process of removal of the reagent, the longer and more effectively will be preventing the scale. Selection of the scale inhibitor, taking into account its adsorption and desorption ability will provide the lowest removal reagent and increase the time and efficiency to prevent the scale (Ibragimov, N. et al., 2003).

As an inhibitor of the growth of salt crystals is widely known use of phosphonic acids. However, an aqueous solution of these acids has a low adsorption-desorption ability and quickly taken out from the reservoir by fluid flow (Yarkeeva, N., 2008).

To improve the sorption properties of the inhibitors recommended their injection with weak acid solutions (Gimatutdinov, S., 1983). So treatment of the bottomhole formation zone is carried out by the mixture of phosphonic acids with hydrochloric acid. Due to the acidic reagent there is rejection of oil film from the surface of the reservoir by reducing the interfacial tension in the system "formation-oil-inhibiting solution". Hydrochloric acid reacts with carbonate minerals and thus the adsorption area of the inhibitor on reservoir increases.

The major volume of oil is produced from sandstone reservoirs with carbonate concretion. Therefore, scale inhibitor should reduce the surface tension at the boundary of the “formation-inhibitor solution” and react with both silicate, aluminosilicate and carbonate concretions. It enhances the adsorption-desorption properties of the inhibitor.

Composition, which contains phosphonicand hydrofluoric acids and surfactants (Antipin Y. et. al., 1993) is known. Hydrofluoric acid in the composition interacts with silicate and aluminosilicate minerals. But this acid is expensive and dangerous to handle. So it will be more sophisticated to use a composition containing bifluoride of ammonia and hydrochloric acid, which form hydrofluoric acid as a result of reaction. In addition, if the rock contains carbonates of calcium and magnesium, they can form insoluble fluorides of calcium and magnesium during their reaction with fluoric acid. Appending a hydrochloric acid to the composition will prevent this. In laboratory was developed effective composition that meets these requirements.

Inhibitory composition development was carried out in laboratory conditions at three stages. The first was selection the most effective combination of components as a result of determining the interfacial tension at the boundary “oil – inhibitor solution”. Then the inhibiting ability of the composition is studied. The third step was determination the adsorption-desorption properties of scale inhibitor. Evaluation of the adsorption and desorption ability of inhibitor compositions were performed when filtering inhibiting compositions on the FDES-645 (Formation damage evaluation system). The developed composition represents multicomponent structure for receiving the greatest inhibiting action in relation to inorganic salts.

3. Results

Analyzing the obtained data, it is possible to note that increase of concentration of reagents reduces the size of aninterfacial tension, and hydrochloric acid has a greater influence on it. The optimal concentrations are 0,5 − 2% of a bifluoride of ammoniaand 7 − 10% of hydrochloric acid. Addition ofphosphonicacid to the compositioncan further reducetheinterfacial tension. This effect is noticeable when the content of acid in the amount of 3 %.

The assessment of efficiency of the inhibitors by the ability to prevent the precipitation was carried out according to the NACE standard TM 0374-2007 in the liquid solution of the averaged model of the formation water. Experiment results revealed that the developed chemical compositions have the necessary protective effect.

The use of chemicals to prevent scale in wells associated with the use of chemically aggressive fluids. Injection of corrosive agents in the bottomhole formation zone not only promotes downhole equipment failure, but also brings additional harm from the products which are formed as a result of exchange reactions and can worse the filtration properties of the bottomhole formation zone. All this leads to a significant increase in operating costs of oil and gas companies that have a negative impact on the oil production cost. Therefore it is necessary to consider the corrosiveness of injected compositions.

All considered chemical compositions showed acceptable corrosion rate in relation to the steel and these reagents can be considered as reagents to prevent scaling in the wells.

To determine the adsorption-desorption characteristics of inhibitory compositions was carried out filtration tests on core samples.

In the laboratory were studied three inhibitory compositions. Considering that the initial concentration of organophosphorus reagents in inhibiting compositions are different, it's possible to compare the dynamics of the relative concentrations of the solutions. While the relative concentration is determined by the ratio of their current concentration to the original.

Figure 1 shows the results of relative concentration of organophosphorus reagents determinationin the inhibiting compositions when the adsorption process.Laboratory studies have shown that the maximum adsorption is achieved by injection the nine pore volumes for inhibiting compositions # 1 and # 2, and ten pore volumes for

composition # 3.Adsorption proceeds faster when applying the inhibiting compositions # 1 and # 2.

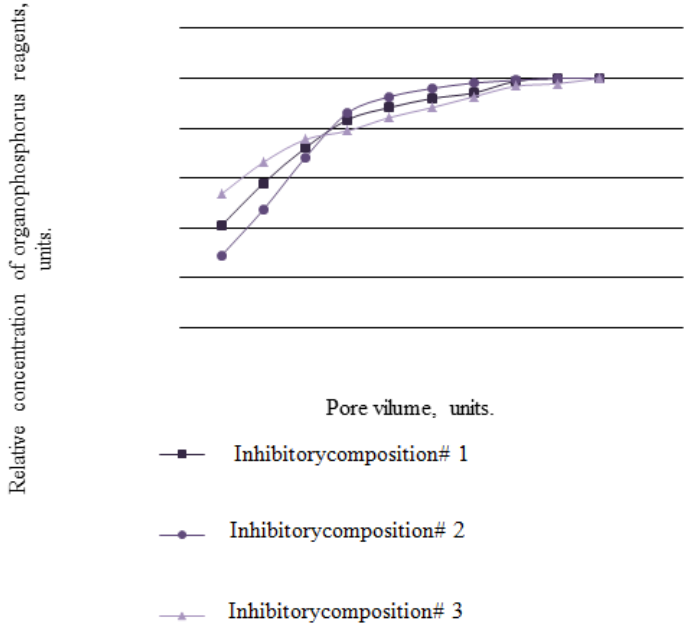


Figure 1. Changing relative concentration of organophosphorus reagentsin the process of adsorption

After the core was left for 24 hours through it was injected formation water to displace inhibitory composition. Figures 2 and 3 show the results of the relative concentrations of organophosphorus inhibitors determination during desorption.

According to figures 2 when using inhibitory composition # 3 necessary subtraction of inhibitor, which is sufficient for effective protection against scale, is achieved by injection through a core sample 30 pore volumes of water. When using the inhibitory composition # 1 effective protection against scaleis achieved by injection 46 pore volumes of water, and composition # 2 - 42 pore volumes of water.

Analysis of data shows that significant portion of the free (not adsorbed) inhibitor is taken out when injection the first two pore volumes. Considering, however, that in the composition # 1 is used concentration of organophosphorus reagents higher than that the composition # 2, the losses will be much more when applying the composition # 1.

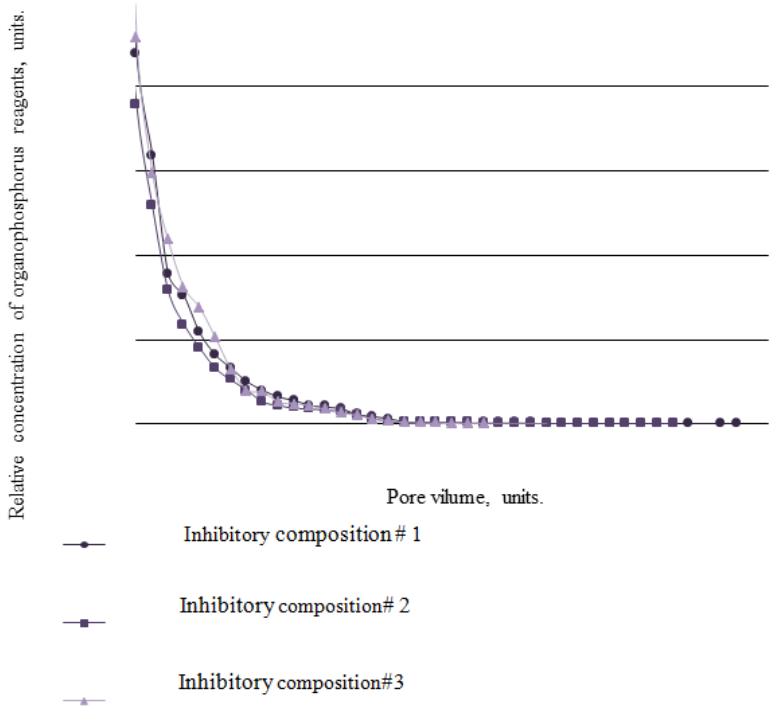


Figure 2. Changing relative concentration of organophosphorus reagents in the process of desorption

4. Discussion

Efficiency of the developed composition is explained by the mechanism of influence on the reservoir of acid additives included in the composition.

The fluoric acid which is formed as a result of interaction of hydrochloric acid with bifluoride of ammonia reacts with the quartz and kaolin which are a part of terrigenous collectors according to the following scheme:

$$SiO_2 + 4HF = SiF_4 + 2H_2O, \quad (3)$$

$$H_4Al_2Si_2O_9 + 14HF = 2AlF_3 + 2SiF_4 + 9H_2O. \quad (4)$$

Further there is an interaction of the formed fluoric SiF4 silicon to water:

$$2SiF_4 + 4H_2O = Si(OH)_4 + 2H_2SiF_6. \quad (5)$$

Hydrofluosilicic acid formed as a result of reaction remains in solution and the hydrate of silicon oxide can turn into a gel in process of decrease in acidity of solution (Amiri M. &Moghadasi J., 2010).The turned-out gel in turn can very easily block pores, thereby reducing the permeability of the reservoir. Mix of fluoric and hydrochloric acids is usedto prevent this when treatment of terrigenous reservoirs. Hydrochloric acid is necessary to maintain the high pH environment and retention of silicic acid in solution (Petrukhin O., 1992). Besides, if reservoir contains carbonates of calcium and magnesium during their reaction with fluoric acid almost hardly soluble fluorides of calcium and magnesium are formed. It is interfered by hydrochloric acid.

Also hydrochloric acid is capable to change wettability of reservoir, clearing its surface of the oil film. This achieves a smooth and more complete adsorption of the scale inhibitor.

Due to the attack by fluoric acid of a silicate reservoir surface the increase area of it surface and the extent of chemical composition adsorption on this surface is provided. The inhibiting composition will get into structure of silica sols, which cover the contact surface of the inhibitor with silicate minerals. Due to this it will desorb slower from the reservoir, as films of silica sols being a part of inhibitor composition are more resistant to leaching. Which leads to the fact that the time of desorption raises, and from this we can conclude that effectiveness of the scale inhibitor increases.

5. Conclusions

Thus, application of the developed inhibitor allows to increase desorption from formation in 1.6 times compared

with the analogues. Wide range of effectiveness of the examined compounds is explained by the different mechanism of action on the formation of specific additives. Laboratory experiments have shown that desorption from the formation with the developed inhibitory compositions longer than the existing counterparts.

Besides development of the inhibitor compositions technological ways of their injection in a well are of great importance for the prevention of scale. One of the most effective methods is the intermittent injection of the chemical composition in the bottomhole formation zone.

According to the general principles of inhibitory protection of the downhole equipment and productive formation a preparatory work is necessary just before the implementation of the technology of the prevention of scale in the chosen well.

Originally manufactured sampling of products produced from the well to determine the degree of over-saturation ions sulfate and barium. Chemical analysis of water is carried out in the laboratory. To determine the productivity index is carried out pressure transient analysis. Existence of scale and their settling zone in a well are established by calipering. Operation of bottomhole cleaning, borehole cleaning, bottomhole treatment, cleaning of downhole equipment with an organic solvent in case of scaling and decreasing the productivity index.

Formation squeeze treatment of inhibitory composition is most preferable, as it can adsorbed in the bottomhole formation zone and held on the surface of the reservoir. In the process of fluid filtration through the bottomhole formation zone gradual desorption process proceeds, the scale inhibitor is released and e comes to a well with the rformation fluid, providing conditions to prevent scale.

Before carrying out the operation of formation squeeze treatment the candidate wells is selected. Candidate wells have to meet the following main criteria:

- the well is complicated by scale;
- the low mean time to failure of downhole equipment;
- the formation of scale below the pump suction and, therefore, of traditional methods of protection against scaling;
- the high level of solids.

To prepare the surface of the reservoir, remove already formed scale it is recommended to combine formation squeeze treatment with a small volume of acid treatment of wells. This technique, at the same time, allows to increase the permeability of the bottomhole formation zone and facilitates delivery of the inhibitor into the formation.

The advantages of the technology include the fact that the protection extends to the bottomhole formation zone of the well, the production casing to the level of the pump, downhole equipment, tubing and terrestrial communications. Another advantage is the length of the period of protection between treatments and no need for construction and operation of the dosers.

References

Amiri, M., & Moghadasi, J. (2010). Prediction the amount of barium sulfate scale formation in Siri oilfield using OLI scaleChem Software. *Asian Journal of Scientific Research ed., 3*, 230-239. http://dx.doi.org/10.3923/ajsr.2010.230.239

Antipin, Y., & Islanova, G. (1999). Use gel-forming technologies for increase of scale prevention efficiency in the wells. *Development and Exploration of Oil and Gas Fields ed.,* pp. 67-77. Ufa: Ufimsky state oil technical university.

Gimatutdinov, S. (1983). Reference guide on design of development and exploration of oil fields. Oil production (p. 455). Moskow: Nedra.

Glushenko, V., Silin, M., & Ptashko, O. (2008). Oilfield chemicals: Complications in the system reservoir - well – OPPU (p. 328). Moskow: Max Press.

Guskova, I. A. (2011). Design and development of technological solutions to the problem formation of organic deposits in the changed conditions of technogenic deposits of oil (Doctoral dissertation, Tatnipineft, Bugulma, Republic of Tatarstan, Russia).

Ibragimov, N., Khafizov, A., & Shaydakov, V. (2003). Complications in oil production (p. 254). Ufa: Monograph.

Kashchavtsev, V. (2005). Role of interstitial waters in the process of precipitation of scale during oil production. *Oil, Gas and Business ed.,* 42-45.

Krabrtri, M., Eslinger, D., & Fletcher, F. (2002). Fighting scale – removal and prevention. *Oilfield review ed.,* 52-73.

Larichev, T., Sotnikova, L., & Sechkarev B. (2006). Bulk crystallization in inorganic systems. (p.364). Kemerovo: Kemerovo State University.

Muslimov, R. (2005). Up-to-date methods of enhanced oil recovery. Design, optimization and evaluation of the performance (p. 205). Kazan: Academy of Sciences RT.

Patent. 1804469 (RF) Composition for preventing scale / Y.V. Antipin, A.Sh. Husniyarov, G.A. Shamaev: UGNTU. No. 4878526/03; published 23.03.93.

Petrukhin, O. (1992). Analytical Chemistry. Chemical methods of analysis. (p. 489). Moskow: Khimiya.

Rizkalla, E. (1983). Kinetics of Crystallisation of barium sulphate. *J. Chem. Soc. Faraday Trans ed., 79*, 1857-1867.

Shaidakov, V. (2012). About the dosing of chemicals into the well. *Oil. Gas. Novation ed., 11*, 39-42.

Shangareeva, L., & Petukhov, A. (2011). Study of the deposition of salts in downhole equipment and the prediction of their education (Proceedings of the Conference. *Rassohinskie Reading.* for Graduate Students, 112-115. Uhta: Uhtinsky state technical university.

Shangareeva, L., & Petukhov, A. (2013). Prediction of formation of scale in oil wells. *Modern Problems of Science and Education.* Retrieved from http://science-education.ru/113-11180

Symeopoulos, B., & Koutsoukos, P. (1999). Radioanalytical monitoring of the formation of barium sulfate in aqueous solutions. *Journal of Radoanalytlcal and Nuclear Chemistry ed., 173*(1), 23-36.

Tomson, N., Watson, M., & Fu, G. (2003). *Mechanisms of Mineral Scale Inhibition, 18*, 192-199. SPEPF. http://dx.doi.org/10.2118/84958-PA

Yarkeeva, N. R. (2008). Improving the efficiency of prevention of scaling in wells at a late stage of development of deposits. (PhD thesis, Ufimsky state oil technical university, Ufa, Russia).

Zdolnik, S., & Akimov, O. (2009). Manage scaling - a pledge improve the efficiency of oil production *Engineering Practice ed., Pilot*, 66-69.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Modeling of Us Dollar to Euro Rate Dependence on USA GDP Dynamics

Shkodinsky S. V.[1], & Prodchenko I. A.[2]

[1] Doctor of economics, professor, department head at the State Research Institute of System Analysis of the Accounts Chamber of the Russian Federation, Moscow, Russia

[2] Ph.D., associate professor, senior researcher at the State Research Institute of System Analysis of the Accounts Chamber of the Russian Federation, Moscow, Russia

Correspondence: Shkodinsky S. V., Doctor of economics, professor, department head at the State Research Institute of System Analysis of the Accounts Chamber of the Russian Federation, Moscow, Russia. E-mail: sh-serg@bk.ru

Received: January 3, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p277 URL: http://dx.doi.org/10.5539/mas.v9n8p277

Abstract

The article analyzes the changes in the USD, to evaluate the influence of gross domestic product in the United States the value of the United States dollar, the influence of gross domestic product of the United States on the value of the United States dollar against the euro, the models of the linear dependence of the USD against the European currency, raschetan coefficient of pair of linear correlation, characterizing the degree of dependence of the United States dollar against the euro on the dynamics of the index of real GDP in the United States of a linear model.

Keywords: modeling, exchange rate, exchange rate changes, the gross domestic product, the United States, the euro, the index, a linear function

1. Introduction

It is known that the exchange climate of the state takes important effect on the economic, investment and social processes in the society (Aleksandrovich JA, 2005). And one of the significant challenges of the exchange policy is to implement the concerted actions of different states in the interest of macroeconomic regulation and ensuring the currency markets stability (Ivanter A., Peresetsky A., 1999). Many economists try to reveal regularities of exchange rates formation, which in its turn forms the basis for decision-making to overcome the crisis developments in the economy and for financial losses minimizing (Balatsky EV, 2005; OA Gulyaev, 2008; P. Kryukov A. 2011; Panilov MA, 2009; Bogoviz AV, 2013).

However, the problem of forecasting and modeling of exchange rates also is a complex multifactorial problem to which the focused attention of scientists and specialists is paid. Thus, in the P.A. Kryukov's work "The methodology of the exchange rate dynamics modeling" an overview of modern empirical research of currency market analyzing and forecasting methods and description of a new methodological approach to the modeling of exchange rate dynamics from the perspective of solving the problem of the classifivcation of its conditions by methods of factor scaling and logistic regression is provided (Kryukov, P.A., 2011). The study of E. Balatsky "The exchange rates formation factors: pluralism of patterns, theories and conceptions", which presents an analysis of the analytical material accumulated in this area, is also interesting (Balatsky E., 2010).

The urgency of the chosen theme of the scientific article is conditioned, firstly, by the manifestation of the violent fluctuations in the US dollar against major world currencies, especially agains the backdrop of the negative effects of the global financial and economic crisis; secondly, by the lack of a mechanism for determining the degree of influence of the US gross domestic product dynamics on the US dollar value against major world currencies; thirdly, by the need to improve the practice of analysis of the factors affecting the change in the US dollar value against major world currencies (P.A. Kryukov, V. Kryukov, 2011).

2. Method. Gross Domestic Product of the United States is the main indicator, which reflects the condition of the national economy (Buglaev V.B., Liventsev N.N., 1998). It represents the market value of the goods and services produced during a certain period, including the income of foreign corporations and non-residents working in the

USA, and excluding the revenues of American citizens and companies, received abroad.

Within the period from January 1, 1999 to April 1, 2011 gross domestic product of the United States increased by 2 934.3 billion US dollars (27.9 percent).

The diagram of the US gross domestic product dynamics for the period of from January 1, 1999 to April 1, 2011 is shown in the Figure1.

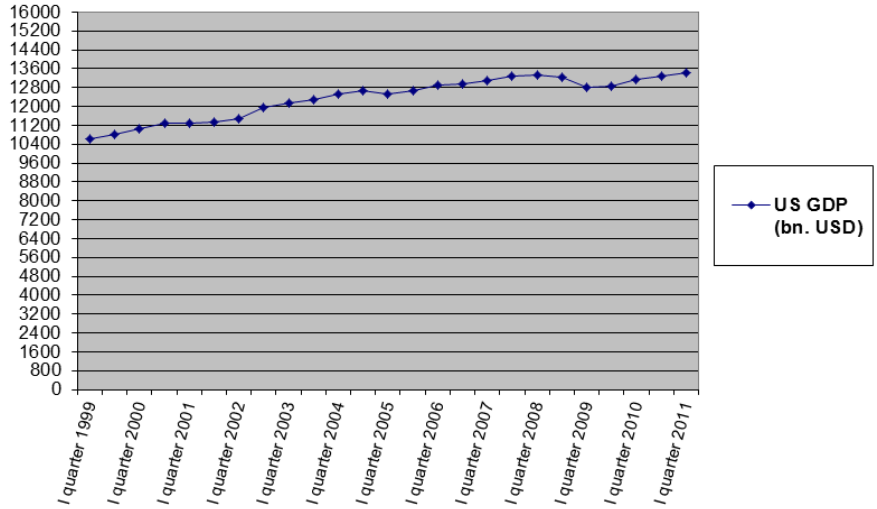


Figure 1. The US gross domestic product dynamics for the period of from January 1, 1999 to April 1, 2011

For the IV quarter of 1998 gross domestic product of the United States was 10 507.6 bln US dollars. For the I. quarter of 2011the US gross domestic product was 13 441.9 bln US dollars.

The dynamics of US gross domestic product (hereinafter - GDP) during the concerned period is presented in the Table 1.

Table 1. US gross domestic product change dynamics for the period of from January 1, 1999 to April 1, 2011 (Note 1)

Period	Value bln. of USD	Change	Publication date
I quarter 11	13441.9	61.2	26.05.2011
IV quarter 10	13380.7	102.2	25.03.2011
III quarter 10	13278.5	83.6	22.12.2010
II quarter10	13194.9	56.1	30.09.2010
I quarter 10	13138.8	119.8	25.06.2010
IV quarter 09	13019	158.2	26.03.2010
III quarter 09	12860.8	50.8	22.12.2009
II quarter 09	12810	-22.6	30.09.2009
I quarter 09	12832.6	-161.1	25.06.2009
IV quarter 08	12993.7	-229.8	26.03.2009
III quarter 08	13223.5	-135.5	23.12.2008
II quarter 08	13359	19.8	26.09.2008
I quarter 08	13339.2	-24.3	29.05.2008
IV quarter 07	13363.5	95.0	27.03.2008
III quarter 07	13268.5	74.4	30.01.2008
II quarter 07	13194.1	104.8	30.09.2007
I quarter 07	13089.3	28.6	29.06.2007
IV quarter 06	13060.7	94.8	30.03.2007
III quarter 06	12965.9	3.4	29.12.2006
II quarter 06	12962.5	46.6	29.09.2006

I quarter 06	12915.9	167.2	30.06.2006
IV quarter 05	12748.7	65.5	31.03.2006
III quarter 05	12683.2	95.7	30.12.2005
II quarter 05	12587.5	53.4	30.09.2005
I quarter 05	12534.1	123.8	30.06.2005
IV quarter 04	12410.3	106.8	31.03.2005
III quarter 04	12303.5	89.7	31.12.2004
II quarter 04	12213.8	86.2	30.09.2004
I quarter 04	12127.6	84.8	30.06.2004
IV quarter 03	12042.8	107.3	31.03.2004
III quarter 03	11935.5	196.8	31.12.2003
II quarter 03	11738.7	92.9	30.09.2003
I quarter 03	11645.8	47.0	30.06.2003
IV quarter 02	11598.8	2.4	31.03.2003
III quarter 02	11596.4	57.6	31.12.2002
II quarter 02	11538.8	60.9	30.09.2002
I quarter 02	11477.9	97.8	28.06.2002
IV quarter 01	11380.1	40.0	29.03.2002
III quarter 01	11340.1	-31.2	31.12.2001
II quarter 01	11371.3	74.1	28.09.2001
I quarter 01	11297.2	-37.3	29.06.2001
IV quarter 00	11334.5	66.6	30.03.2001
III quarter 00	11267.9	9.4	29.12.2000
II quarter 00	11258.5	215.5	29.09.2000
I quarter 00	11043.0	28.7	30.06.2000
IV quarter 99	11014.3	194.4	31.03.2000
III quarter 99	10819.9	135.9	31.12.1999
II quarter 99	10684	82.8	30.09.1999
I quarter 99	10601.2	93.6	30.06.1999
IV quarter 98	10507.6	178.8	31.03.1999

As of January 1, 1999, the exchange rate of US dollar to euro was 0.8572 euro for 1 US dollar at a value of GDP of the Uniated States for the IV quarter 1998 of USD 10 507.6 billion.

For the IV quarter of 1999, GDP of the United States reached the value of USD 11 014.3 bln, while the the rate of US dollar to euro was 0.992647 euro for 1 US dollar as of January 1, 2000.

As of January 1, 2001, the rate of US dollar to euro was 1.051138 euro for 1 US dollar at the value USA GDP for the IV quarter of 2000 of USD 11 334.5 bln.

For the IV quarter of 2001, GDP of the United States increased insignificantly and reached the value of USD 11 380.1 billion, while the rate of US dollar to euro was 1.132245 euro per 1 US dollar as of January 1, 2002.

Then, with a slight increase of GDP of the United States a decline of the rate US dollar to euro was observed. For the IV quarter of 2002 the GDP of the United States reached the value of USD 11 598.8 billion, while the rate of US dollar to euro fell to 0.955293 euros for 1 US dollar as of January 1, 2003. As of January 1, 2004, the rate of US dollar to euro fell to 0.793967 euros for 1 US dollar, while the value of GDP of the United States for the IV quarter 2003 was USD 12 042.8 billion.

For the IV quarter of 2004 the GDP of the United States reached the value of USD 12 410.3 billion, while the rate of US dollar to euro continued to fall to EUR 0.733299 per 1 US dollar as of January 1, 2005. As of January 1, 2006, the rate of US dollar to euro was sterengthened to 0.829531 euros per 1 US dollar, while the GDP of the United States for the IV quarter 2005 was USD 12 748.7 billlion.

For the IV quarter of 2006, of the GDP of the United States reached the value of USD 13 060.7, while the rate of US dollar to euro fell to 0.766872 euros per 1 US dollar as of January 1, 2007.

As of January 1, 2008, the rate of US dollar to euro fell to the record lowest value of USD 0.679302 euros per 1 US dollar, while the GDP of the United States for the IV quarter of 2007, was USD 13 363.5 billion. For the IV quarter of 2008 GDP of the United States for the first time during the concerned period fell to USD 12 993.7

billion, while the rate of US dollar to euro was 0.709471 euro per 1 US dollar as of January 1, 2009.

As of January 1, 2010, the rate of US dollar to euro fell to 0.694541 euros per 1 US dollar, while the GDP of the United States for the IV quarter 2009 was USD 13 3019.0 billion. For the IV quarter of 2010, the GDP of the United States reached the value USD 13 380.7 billion, while the rate of US dollar to euro was 0.749625 euro per 1 US dollar as of 1 January 2011.

Dynamics of change in the GDP of USA and the rate of US dollar to euro during the period of January 1, 1999 to April 1, 2011 is presented in the Table 2.

Table 2. Dynamics of change in the GDP of USA and the rate of US dollar to euro during the period of January 1,1999 to April 1, 2011

n	date	USD/EUR	GDP of USA in bln. of $ (GDPi)
1	01.01.1999	0.857202	10507.6
2	01.04.1999	0.932099	10601.2
3	01.07.1999	0.967626	10684.0
4	01.10.1999	0.93961	10819.9
5	01.01.2000	0.992647	11014.3
6	01.04.2000	1.043796	11043.0
7	01.07.2000	1.049775	11258.5
8	03.10.2000	1.13445	11267.9
9	04.01.2001	1.051138	11334.5
10	03.04.2001	1.140309	11297.2
11	03.07.2001	1.180567	11371.3
12	02.10.2001	1.099365	11340.1
13	01.01.2002	1.132245	11380.1
14	02.04.2002	1.144033	11477.9
15	02.07.2002	1.005632	11538.8
16	01.10.2002	1.017398	11596.4
17	01.01.2003	0.955293	11598.8
18	01.04.2003	0.923447	11645.8
19	01.07.2003	0.874968	11738.7
20	01.10.2003	0.859846	11935.5
21	01.01.2004	0.793967	12042.8
22	01.04.2004	0.81686	12127.6
23	01.07.2004	0.827814	12213.8
24	01.10.2004	0.81103	12303.5
25	01.01.2005	0.733299	12410.3
26	01.04.2005	0.773156	12534.1
27	01.07.2005	0.826788	12587.5
28	01.10.2005	0.831739	12683.2
29	11.01.2006	0.829531	12748.7
30	01.04.2006	0.823724	12915.9
31	01.07.2006	0.786905	12962.5
32	03.10.2006	0.788892	12965.9
33	10.01.2007	0.766872	13060.7
34	03.04.2007	0.748896	13089.3
35	03.07.2007	0.738552	13194.1
36	02.10.2007	0.701901	13268.5
37	10.01.2008	0.679302	13363.5
38	01.04.2008	0.633713	13339.2
39	01.07.2008	0.633112	13359.0
40	01.10.2008	0.69512	13223.5
41	01.01.2009	0.709471	12993.7
42	01.04.2009	0.755173	12832.6
43	01.07.2009	0.707814	12810.0

44	01.10.2009	0.683761	12860.8
45	01.01.2010	0.694541	13019.0
46	01.04.2010	0.745379	13138.8
47	01.07.2010	0.817996	13194.9
48	01.10.2010	0.736322	13278.5
49	01.01.2011	0.749625	13380.7
50	01.04.2011	0.706065	13441.9

In the Figure 2 the diagram of the dynamics of changes in exchange ratios of the US dollar to euro during the studied period.

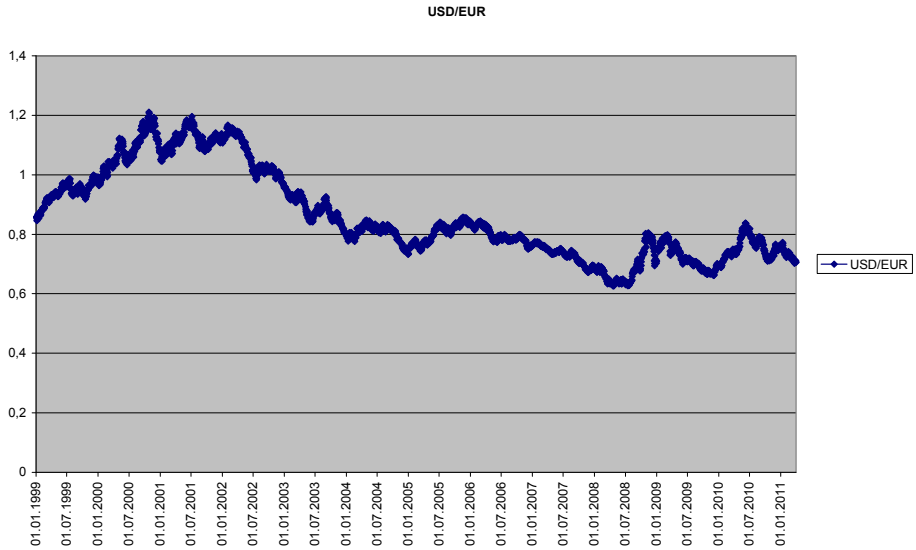


Figure 2. The diagram of the dynamics of changes of the exchange ratios of the US dollar to Euro for the period of from January 1, 1999 to April 1, 2011

The diagram of dynamics of GDP in the United Sates and exchange ratios of US dollar to euro in the concerned period is presented in the Figure 3.

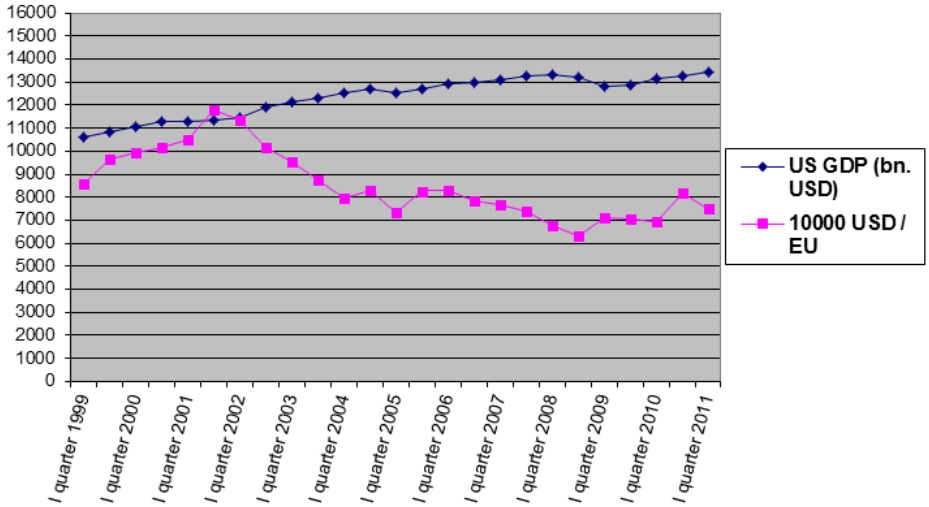


Figure 3. The diagram of dynamics of GDP in the United Sates and exchange ratios of US dollar to euro fo the period of from January 1, 1999 to April 1, 2011

Let's consider the dependence of the exchange ratios of US dollar to Euro on the index of the real GDP of the

United States. Nominal GDP (absolute) is expressed in current prices of the given year. Real GDP (as adjusted for inflation) is expressed in the prices of the previous or any other base year. To what extent the growth of GDP is determined by the growth of production, rather than by growth of prices is considered in the real GDP (Perron P., 1997).

The diagram of index dynamecs of real GDP of the United States and exchange ratios of US dollar to Euro in the concerned period is presented in the Figure 4.

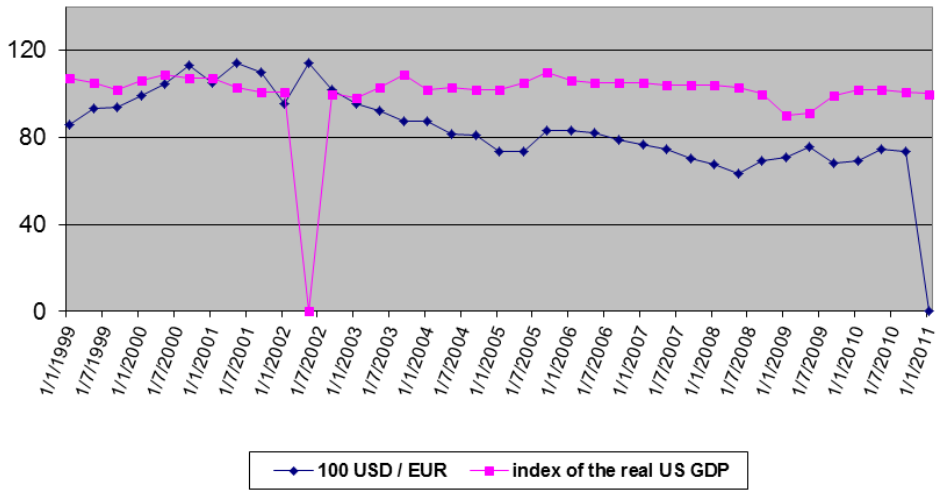


Figure 4. The diagram of index dynamecs of real GDP of the United States and exchange ratios of US dollar to Euro for the period of from January 1, 1999 to April 1, 2011

We noted above that the gross domestic product, GDP, is a general index of amount of added values created during a certain period by all manufacturers acting in the territory of the country. GDP is a generalizing indicator of the power of the economy (or vice versa, its weakness during recessions) (Pugel T.A., Lindert P.H, 2003). Its connection with the exchange rate is always obvious: the stronger GDP grows, the stronger domestic currency is. The higher the GDP is, the better the economic situation. It is one of the main indicators for the currency markets. Reaction on the publication of not only growth indicators of the main economies, but their corrected (specified) values can be very significant.

According to the quantitative theory, the growth of the country national income as a result of new goods sales leads to increased demand for transactions with the national currency (Kravtsov MA, Miksjuk SF, 2005).

National income is not an independent component, which can be changed on its own. What causes the change of national income has a great impact on the exchange rate (Keynes, 1999). Increase of the offers of goods and servicest to foreign markets made by a country leads to an increase in the national currency. The increase of national income due to additional expences of government may not cause an increase of the national currency, especially if the additional money will be used to increase imports (Fedoseev V.V, Garmash A.N & others, 2005).

Thus, much depends on what causes the change in the national income: an increase in the ability of the goods supply (currency rate increases) or an increase in domestic demand (the exchange rate falls) (Litinskii D.S., 2003).

The proposed mechanism for determination of the extent of influence of the USA GDP dynamics on the value of the rate of US dollar to Euro consists of two stages. At the first stage the model of linear dependence of the rate of US dollar to Euro on the dynamics of the GDP of the United States. At the second stage the extent of influence the the USA GDP dynamics on the value of the rate of US dollar to Euro by calculating the coefficient of pair linear correlation characterizing the closeness of dependency of the rate of US dollar to Euro on the dynamics of GDP of the United States on the model developed of the the first stage (Shelobaev S.I., 2005).

To define the mechanism for determination of the extent of the influence of the dynamics of the USA GDP on the value of the rate of US dollar to Euro, let's consider the statistical series of USA GDP and the exchange ratios of the US dollar (USD) to Euro during the analyzed period, on the basis of data obtained from the official website of Bank of Russia www.cbr. ru, as well as from the site www.quote.ru and www.prime-tass.ru.

One of the mathematical methods serving to align different statistical series is the method of least squares. It reduces to the fact that the existing dependence is modeled by some production function, the parameters of which are determined at the same time (Malihin V.I., 1999).

The main requirement of the least squares method consists in the fact that the sum of squared deviations of the calculated (theoretical) levels of the actual values of the dependent factor should be the lowest (minimum) (Johansen S., 2000; Oskorbin N.M., 1989).

We present the dependence of the US dollar rate on the dynamics of the GDP of the United States in the form of the production function in the following form:

$$Yri = f\ (GDPi) \tag{1}$$

where: Yri is a calculated value of the US dollar rate;

GDPi is the GDP of the United States.

After we denote the sum of squared deviations of the calculated (theoretical) levels of the actual values of the dependent indication by S, let's write the main requirement of this method in mathematical form.

$$S = \sum (f\ (GDPi) - Yi)^2 \rightarrow \min \tag{2}$$

where: Yi is the actual meaning of the rate of US dollar;

Since GDPi and Yi are known values, then the sum of squares of the specified values of deviations depends on the parameters a0, a1 which are determined. Therefore we can write that the S sum is a function of the known parameters:

$$S = f\ (a0,\ a1) \tag{3}$$

At the point of extremum of the differentiable function its first derivative is equal to zero. The function (3) is differentiable. S is minimal, that is extremal. Therefore, we can take the partial derivatives sums of S under certain parameters and equate them to zero.

$$\frac{\delta S}{\delta a0} = 0;\ \frac{\delta S}{\delta a1} = 0 \tag{4}$$

As a result the two equations with two determinate parameters will be obtained. The range of equations (4) constitutes the system of so-called normal equations. While solving the system of normal equations unknown parameters of the production function are found. Systems of these equations are solved by any known method (method of substitution, addition, by G. Cramer's method (Muraviev D.G., 2006), etc.). Available parameters are recorded in the the sought-for function production function (Grishin A.F., Kotov-Darty S.F., Yagunov V.N., 2005).

Since the increase of the GDP in the United States causes the growth of the the national currency rate (unless other rate forming factors counteract more) when determining the dependence of the exchange ratios of the US dollar to the European currency, we use a linear model of dependence of the rate of US dollar on the size of the GDP of the United States.

Linear function, which models the variation of the currency rate (Yi), depending on the dynamics of GDP GBPi, has the form:

$$Yi = a0 + a1*GBPi \tag{5}$$

Then the main requirement of the least squares method, according to which the sum of the squares of deviations between the calculated and actual levels shall be minimal (Bendat J., Peirsol A., 1989), can be presented in the following form:

$$S = \sum (a0 + a1*GBPi - Yi)^2 \rightarrow \min \tag{6}$$

In equation (5) two parameters shall be determined: a0 и a1. Therefore, the partial derivatives of the S sum are written (6), under the parameters a0 and a1 and they are equated to zero.

$$\frac{\delta S}{\delta a0} = 2\sum (a0 + a1 * GBP_i - Y_i) * 1 = 0$$

$$\frac{\delta S}{\delta a1} = 2\sum(a0 + a1 * GBP_i - Y_i) * 1 = 0 \quad (7)$$

after trimming the both equation by 2, opening the brackets and making a term-by-term summation, we will obtain:

$$\sum a0 + \sum a1*GBPi - \sum Yi = 0,$$

$$\sum a0 *GBPi + \sum a1*GBPi^2 - \sum Yi *GBPi = 0 \quad (8)$$

Since Yi and GBPi are known amounts, the ΣY_i and ΣGBP_i are transferred to the right side of the equations. Since a0 and a1 are the parameters, which are in this case constant numbers, they are taken beyond the summation sign.

Therefore:

$$\sum a0 = n*a0, \text{ where n is the number of observations;}$$

$$\sum a0 *GBPi = a0 \sum GBPi;$$

$$\sum a1*GBPi^2 = a1 \sum GBPi^2.$$

Ultimately, the system of normal equations takes the form of:

$$n*a0 + a1 \sum GBPi = \sum Yi$$

$$a0 \sum GBPi + a1 \sum GBPi^2 = \sum Yi *GBPi \quad (9)$$

Let's solve this system by the determinations (by Kramer's methos). The determinant of the system (9) has the following form:

$$\Delta = \begin{vmatrix} n & \sum GBPi \\ \sum GBPi & \sum GBPi^2 \end{vmatrix} = n \sum GBPi^2 - (\sum GBPi)^2 \quad (10)$$

The determinants for evaluation of a0 and a1 are:

$$\Delta_{a0} = \begin{vmatrix} \sum Yi & \sum GBPi \\ \sum Yi *GBPi & \sum GBPi^2 \end{vmatrix} = \sum GBPi^2 \sum Yi - \sum GBPi \sum Yi *GBPi$$

$$\Delta_{a1} = \begin{vmatrix} n & \sum Yi \\ \sum GBPi & \sum Yi *GBPi \end{vmatrix} = n \sum Yi *GBPi - \sum GBPi \sum Yi \quad (11)$$

As a result the required paramenters will be:

$$a0 = \frac{\Delta_{a0}}{\Delta} = \frac{\sum GBPi^2 \sum Yi - \sum GBPi \sum Yi *GBPi}{n \sum GBPi^2 - (\sum GBPi)^2}$$

$$a1 = \frac{\Delta_{a1}}{\Delta} = \frac{n \sum Yi *GBPi - \sum GBPi \sum Yi}{n \sum GBPi^2 - (\sum GBPi)^2} \quad (12)$$

3. Results

As a result we will obtain:

$$a0 = \frac{\Delta a0}{\Delta} = \frac{\sum GBP_i^2 \sum Y_i - \sum GBP_i \sum Y_i * GBP_i}{n \sum GBP_i^2 - (\sum GBP_i)^2} =$$

$$= \frac{7550028534{,}30 * 42{,}8487 - 602288{,}4 * 519632{,}3027}{50 * 7550028534{,}3 - (602288{,}4)^2} = 0{,}7146$$

$$a1 = \frac{\Delta a1}{\Delta} = \frac{n \sum Y_i * GBP_i - \sum GBP_i \sum Y_i}{n \sum GBP_i^2 - (\sum GBP_i)^2} =$$

$$= \frac{50 * 519632{,}302705 - 602288{,}4 * 42{,}8487}{50 * 7550028534{,}30 - (602288{,}4)^2} = 1{,}1816E\text{-}05$$

The function, modeling dependence of the value of the rate of US Dollar to Euro $(Yi)^{Euro}$ on dynamics of GDP of the United states has the following form:

$$Yi^{Eur} = a0 + a1 * GBP_i = 0{,}7146 + 1{,}1816E - 0{,}5 * GBP_i \quad (13)$$

As a result we will obtain:

$$a0 = \frac{\Delta a0}{\Delta} = \frac{\sum GBP_i^2 \sum Y_i - \sum GBP_i \sum Y_i * GBP_i}{n \sum GBP_i^2 - (\sum GBP_i)^2} =$$

$$= \frac{524343{,}67 * 42{,}8487 - 5012{,}5 * 4390{,}3718}{50 * 524343{,}67 - (5012{,}5)^2} = 0{,}4219$$

$$a1 = \frac{\Delta a1}{\Delta} = \frac{n \sum Y_i * GBP_i - \sum GBP_i \sum Y_i}{n \sum GBP_i^2 - (\sum GBP_i)^2} =$$

$$= \frac{50 * 4390{,}3718 - 5012{,}5 * 42{,}8487}{50 * 524343{,}67 - (5012{,}5)^2} = 0{,}0043$$

The function, modeling dependence of the value of the rate of US Dollar to Euro $(Yi)^{Euro}$ on dynamics of index of the rial GDP of the United states has the following form:

$$Yi^{Eur} = a0 + a1 * GBP_i = 0{,}4219 + 0{,}0043 * GBP_i \quad (14)$$

The simplest system of correlation relationship is a linear relationship between two signs the pair linear correlation. Its practical importance consists in the fact that there is a system in which among the factors influencing the resultant sign one important factor stands out which basically determines the variation of resultant sign (Hamilton J.D., 1994). To determine the dependence of the rate of US dollar to the main world currencies on the dynamics of the USA GDP we will calculate the coefficient of the pair linear correlation with the following formula:

$$cor = \frac{n \sum Y_i * GBP_i - \sum GBP_i \sum Y_i}{\sqrt{(n \sum GBP_i^2 - (\sum GBP_i)^2) * (n \sum Y_i^2 - (\sum Y_i)^2)}} \quad (15)$$

Linear correlation coefficient is within the range from of from 0 to 1. When the cor = 0, any dependence of (Y_i) of the indicator on the factor sign (GBP_i) (i = 1,2, ... n) is absent. When K = 1, there is a functional dependence. When 0.3 <= cor <= 0.7 - the relationship is average. When the linear correlation coefficient is less than 0.3 the closeness of the relationship is considered weak, when Cor is > 0.7 – it is strong (Gulyaeva O.A., 2008).

By using the data given in the table 3, let's calculate the coefficient of pair linear correlation of dependence of the rate of the US dollar to Euro(Yi) on the dynamics of GDP ^{Eur} (GBPi):

$$cor = \frac{n \sum Y_i * GBP_i - \sum GBP_i \sum Y_i}{\sqrt{(n \sum GBP_i^2 - (\sum GBP_i)^2) * (n \sum Y_i^2 - (\sum Y_i)^2)}} =$$

$$= \frac{50 * 519632{,}3027 - 602288{,}4 * 42{,}8487}{\sqrt{(50 * 7550028534{,}3 - 602288{,}4^2) * (50 * 37{,}85 - 42{,}8487^2)}} = 0{,}1911$$

The coefficient of pair linear correlation cor < 0.3 (cor = 0,1911).

By using the data given in the table 4, let's calculate the pair linear correlation coefficient of dependence of the value of the US dollar to Euro rate (Yi) on the dynamics of index of real GDP ^{Eur} (GBPi):

$$cor = \frac{n \sum Y_i * GBP_i - \sum GBP_i \sum Y_i}{\sqrt{(n \sum GBP_i^2 - (\sum GBP_i)^2) * (n \sum Y_i^2 - (\sum Y_i)^2)}} =$$

$$= \frac{50 * 4390{,}3718 - 5012{,}5 * 42{,}8487}{\sqrt{(50 * 524343{,}67 - 5012{,}5^2) * (50 * 37{,}85 - 42{,}8487^2)}} = 0{,}6041$$

The coefficient of pair linear correlation 0.3 < cor < 0.7 (cor = 0,6041).

4. Discussion

It is important to say that scientists and specialists actively discus the problems of modeling of currency rates on the basis of the complexity of modern macroeconomic sistuatsii (Panilov M.A., 2008; Gulyaeva O.A., 2008). The presented approaches deserve attantion because they enable to use the developed models not only taking into account the particular domestic economic situation, but also in the condition of the global financial and economic crisis. Moreover, they can form the basis of predicting the dynamics of currency exchange rates, exports, imports, and the inflow/outflow of the capital.

5. Conclusion

Thus, the present study analyzed the influence of the dynamics of the GDP of the United States and the index of real GDP of the United States of the value of the US dollar to the euro rate, developed the mechanism of determination of extent of influence of USA GDP on the value of the US dollar to Euro rate during the period of from January 1, 1999 to April 1, 2011.

The mechanism of determination of the extent of influence of the dynamics of the USA GDP on the value of the rate of US dollar to European currency was developed.

On the basis of the number values of the currency exchange rate of the US dollar to Euro, as well as the number values of the dynamics of the USA GDP and an index of real USA GDP, the coefficients of equation of dependence of the US dollar rate on the value of the USA GDP and the index of real GDP of the Unites States were calculates, the linear function that models the dependence of the value of the dollar US to the Euro rate was developed.

The article the extent of influence of the dynamics of the USA GDP and an index of real USA GDP on the value of the US dollar to Euro rate was determined, as well as the parameters of linear functions, modeling the dependence of the US dollar to Euro rate on the dynamics of the USA GDP and an index of real USA GDP were calculated.

This made it possible to develop the model of the linear dependence of the value of US dollar to Euro rate on the dynamics of the USA GDP and an index of real USA GDP and to calculate the coefficient of pair linear correlation characterizing the closeness of their dependence.

However, it should be noted that further improvement of the tools of forecasting and analytical assessments of the dynamics of currency exchange rates in order to develop regulatory measures of currency policy (Shvajko P.,

2002), providing reduction of macroeconomic instability at the national and supranational levels is of significant practical importance. To solve these problems it is necessary to continue the development and substantiation of mathematical economic models that take into account the widest possible range of macroeconomic aggregates and make it possible to quantify the dynamics of equilibrium currency exchange rates of the main world currencies depending on changes in fundamental factors. In addition, it is necessary to provide such modeling of the rate dollar to other national currencies as well.

Acknoledgements. The authors would like to thank the staff of the State Research Institute of System Analysis of the Accounts Chamber of the Russian Federation and personally the director - Vladimir Ivanovich Shchedrov for their help and support to prepair this article.

References

Aleksandrovich, J. A. (2005). Econometric analisis of the main macroeconomic indicators time series. *The economic bulletin of Belarusian economics ministery researth institute, 3*, 3-23.

Balatsky, E. (2010). Factors of exchange rates formation: the pluralism of patterns, theories and conceptions. Retrieved from http:// capital-rus.ru/articles/article/180954

Balatsky, E.V. (2005). The model the dynamics of the "dollar / euro" rate. *The Economist, 9*, 75-81.

Bendat, J., & Peirsol, A. (1989). Application analysis of random data (p. 540), M.: Mir.

Bogoviz, A. V. (2013). Modeling the Labor Process One of the Tasks of Strengthening of Positive Trends in the Economic Growth of the Industrial Enterprises of the Region. World Applied Sciences Journal, 8(25), 1222-1225.

Buglaev, V. B., & Liventsev, N. N. (1998). International Economic Relationships (p. 159). M.: Finance and Statistics.

Fedoseyev, V. V., Garmash, A. N., & Orlova, I. V. (2005). Economic and mathematical methods and applied models: education guidance for HEIs. M.: UNITY.

Grishin, A. F., Kotov-Darty, S. F., & Ygunov, V. N. (2005). Statistical models in economics (p. 344). Rostov-on-Don: PHOENIX.

Gulyaeva, O. A. (2008). Currency risk management on the basis of prognostic analysis of exchange rates by fractal methods (p. 27).

Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.

Ivanter, A., & Peresetsky, A. (1999). The development of the state bond market. Working paper series, 99/06, May.

Johansen, S. (2000). Modelling of Cointegration in the Vector Autoregressive Model. *Economic Modelling, 17,* 359-373. http://dx.doi.org/10.1016/S0264-9993(99)00043-7

Keynes, J. M. (1999). The General Theory of Employment, Interest and Money (p.351). M. Helios APB.

Kravtsov, M. A., & Miksjuk, S. F. (2005). Mathematical modeling of the macroeconomic processes. Science works of Belarusian economics ministery researth institute.

Kryukov, P. A. (2011). Methodology of modeling of the dynamics of the exchange rate. Economics, management, finance: Materials of the international. scientific. conf. Perm: Mercury.

Kryukov, P. A., & Kryukova, V. V. (2011). Forecasting of the exchange rate on the basis the factor scaling. *Bulletin of the Kuzbass State Tech. Unic, 1*, 118-127.

Litinsky, D. S. (2003). Statistical forecasting for building effective trading strategies in the currency market. PhD thesis. Moscow.

Malihin, V. I. (1999). Financial Mathematics (p. 247). M.: UNITY. International economic statistics. *Electronic Resource.* Retrieved from http://www.statinfo.biz/HTML/M128F6963A483 5L1.aspx

Muraviev, D. G. (2006). Mathematical methods of development and evaluation of trade strategies in the interbank currency market Forex. PhD thesis. Samara.

Oskorbin, N. M. (1989). Some Aspects of Optimization of Decision Support System. Volume of Abstracts International School: Seminar Optimization Methods and Their Applications. Baika.

Panilov, M. A. (2008). Development of the new complex approach to equilibrium exchahge rate analisis. Internet Journal ATiSO. Retrieved from http://www.e-rej.ru/Articles/2008/Panilov/pdf.

Panilov, M. A. (2009). Economic modeling of the dynamics of the exchange rate. PhD thesis. Moscow.

Perron, P. (1997). Further evidence on breaking trend function in macroeconomic variables. *Journal of econometrics, 80*, 355-385. http://dx.doi.org/10.1016/S0304-4076(97)00049-3

Pugel, T. A., & Lindert, P. H. (2003). International Economics (p. 799). M.: Business and Service.

Shelobaev, S. I. (2005). Economic and mathematical methods and models (p. 285). M.: UNITY.

Shvajko, P. (2002). Econometric models for analysis and forecasting primary market capacity of the state obligations. *The Economic Bulletin (ECOVEST), 2*, 111-153.

Note

Note 1. International economical statistics. Digital resourse. Access mode: http://www.statinfo.biz/HTML/M128F6963A4835L1.aspx

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

A Two-Stage Method for Considering Cardinality in Portfolio Optimization of Mutual Funds

Amir Alimi[1]

[1] Department of management, Ferdowsi university of Mashhad, Iran

Correspondence: Amir Alimi, Department of management, Ferdowsi university of Mashhad, Mashhad, Iran. E-mail: amiralimi2005@yahoo.com

Received: January 17, 2015 Accepted: February 3, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p289 URL: http://dx.doi.org/10.5539/mas.v9n8p289

Abstract

Mutual funds are important financial institutes. There are several methods for performance evaluation of mutual funds such as portfolio optimization. Portfolio optimization has a basic model that has completed up to now. One of completions is adding cardinality constraint to the model. Considering cardinality in portfolio optimization model makes it an integer programming problem that solving it, is hard and makes the efficient frontier discontinuous. In current study in first stage we rank the mutual funds with VIKOR method and based on 5 characteristics: rate of return, variance, semivariance, Treynor ratio and Sharpe ratio. In second stage according to cardinality level best ranked mutual funds are chosen. A mean-semivariance portfolio optimization model is written using chosen funds. This model is solved using fuzzy technique programming and efficient frontier is obtained. Real data from NASDAQ based on 92 mutual funds are used to illustrate the effectiveness of proposed methodology. Results show that the efficient frontier obtained from our methodology is continuous and near to unconstrained efficient frontier.

Keywords: mutual fund, VIKOR method, mean-semivariance, cardinality, fuzzy programming technique

1. Introduction

The role of mutual funds in financial markets is undeniable. Mutual funds are financial institutes that help investors to have an appropriate portfolio. Performance evaluation of mutual funds has become a serious subject in recent years. Researchers have applied new methods to choose superior funds. Murthi et al. (1997) used Data Envelopment Analysis (DEA) as a non- parametric method and Jenson alpha and Sharpe as indices. Basso & funari (2001) used DEA method too. They defined their model based on Jenson alpha, Treynor, Sharpe and semivariance indices. Chang et al. (2010) evaluated mutual funds using extended TOPSIS method and based on Jenson alpha, Sharpe and Treynor indices and Information ratio. In mutual funds portfolio optimization problem Xia et al. (2001) considered transaction cost and Chen & Huang (2009) used cluster analysis in their model. In this article cardinality subject is considered in portfolio optimization of mutual funds. In this study a two-stage method is defined. In first stage VIKOR method is used.

The VIKOR method is a new and famous method that placed on Multi Attribute Decision Making (MADM) methods. Opriovic (1998) used this method for the first time and opriovic & Tzeng (2002) extended it. In first stage of this study mutual funds are ranked based on VIKOR method using 5 characteristics: rate of return, variance, semivariance, Treynor index and Sharpe index. In second stage a multi-objective portfolio optimization model is used.

Markowitz (1952, 1959) proposed portfolio selection problem in a mathematical model. There are many studies that have done to improve and complete this model. For example Benati & Rizzi (2007) used value at risk (VaR) as a measure of risk. Lin & Liu (2008) wrote a portfolio optimization model with minimum transaction lots and solved it with Genetic Algorithm. Gupta et al. (2008) entered liquidity and minimal and maximal fraction of the asset that should be invested in an asset in their model. They solved their model using fuzzy mathematical programming. Jana et al. (2009) proposed a possibilistic model with transaction cost and entropy function in objective functions. Chang et al. (2009) considered different risk measures in their model such as variance, semivariance, variance with skewness and absolute deviation, and solved it using genetic algorithm. An important subject that has investigated in many portfolio optimization studies is cardinality.

Cardinality means the number of assets that should be held in optimized portfolio. When the problem is large scale capital allocation based on the results of problem is difficult. For example, when we have 100 assets results show that we should allocate lower than 1 percent of our capital in some of assets. In this situation portfolio management is very difficult. To avoid this problem researchers define a constraint named cardinality. Cardinality is considered as a constraint in many researches such as: Gupta et al. (2008), Chang et al. (2009), Branke et al. (2009), Golmakani & Fazel (2011), Yang et al. (2011), Anagnostopoulos & Mamanis (2011) and Woodside-Oriakhi et al. (2011). Anagnostopoulos & Mamanis (2010) proposed cardinality as an objective function that should be minimized. Chang et al. (2000) investigated the differences between basic portfolio optimization models and cardinality constrained models. They illustrated that when there is a cardinality constraint the efficient frontier is discontinuous. In their study 2 weaknesses are mentioned for basic Markowitz's model: 1- the assumption of normality for returns and 2- constraints such as cardinality that make the model an integer programming problem. In current study to avoid the first weakness semivariance is considered as a downside risk measure and to avoid the second, cardinality is considered before optimization.

Considering cardinality in portfolio optimization model change it to an integer programming problem with discrete variables. Because solving this type of problems and obtaining efficient frontier is difficult and time-consuming, in this study a two-stage method is used. In first stage mutual funds are ranked based on 5 characteristics and with VIKOR method. In second stage according to cardinality level best mutual funds based on VIKOR ranking are chosen. A mean- semivariance portfolio optimization model is applied for these mutual funds. Fuzzy technique programming solved this model and efficient frontier is obtained.

This paper is organized as follows: In section 2 the first stage of our methodology is described. In this section the VIKOR method and its attributes are discussed completely. In section 3 the basic portfolio optimization problem, mean-semivariance model and cardinality constraint are presented and fuzzy technique programming is considered as our solving method. To illustrate the proposed method a numerical example is solved in section 4. Section 5 presents our conclusions.

2. First Stage: Ranking Mutual Funds Using VIKOR

Decision making based on more than one attribute is a serious problem. There are many methods named MADM methods that consider and solve this type of problems. One of these methods is VIKOR method.

2.1 The VIKOR Method

The VIKOR method is a famous and efficient method that Opricovic (1998) and Opricovic and Tzeng (2002) proposed it for the first time. The VIKOR term has a Serbian etymon and is the abbreviation of this phrase: VlseKriterijumska Optimizacija I Kompromisno Resenje. It means multi criteria optimization and compromise solution (sanayei et al, 2010). The VIKOR method is an extension of $L_p-\text{metric}$ method. The $L_p-\text{metric}$ method tries to minimize the distinction between the ideal solutions and objective functions. This method minimizes the function that is shown in Eq. (1).

$$L_{p,j}=\left\{\sum_{i=1}^{n}\left[\frac{w_i(f_i^*-f_{ij})}{f_i^*-f_i^-}\right]^p\right\}^{1/p},\ j=1,2,\dots,J,\qquad 1\le P\le\infty \tag{1}$$

Where J is the number of alternatives, n is the number of criterions, f_{ij} is the value of ith criterion function for jth alternative, w_i is the weight of ith criteria, f_i^* is the best f_{ij} and f_i^- is the worst f_{ij} (Opricovic and Tzeng, 2004).

The VIKOR method considers $L_{1,j}$ and $L_{\infty,j}$ functions and follows these steps:

Step 1: determine f_i^* and f_i^- as Eqs. (2), (3).

$$f_i^*=\begin{Bmatrix}\max\limits_j f_{ij}, & \text{for benefit criteria}\\ \min\limits_j f_{ij}, & \text{for cost criteria}\end{Bmatrix}\ j=1,2,\dots,J \tag{2}$$

$$f_i^-=\begin{Bmatrix}\max\limits_j f_{ij}, & \text{for cost criteria}\\ \min\limits_j f_{ij}, & \text{for benefit criteria}\end{Bmatrix}\ j=1,2,\dots,J \tag{3}$$

Step 2: compute values S_j and R_j for j=1, 2,…, J by Eqs. (4), (5). (S_j is defined based on $L_{1,j}$ function and R_j

is defined based on $L_{\infty,j}$ function).

$$S_j = \sum_{i=1}^{n} \frac{w_i(f_i^* - f_{ij})}{f_i^* - f_i^-} \tag{4}$$

$$R_j = \max_i \left[\frac{w_i(f_i^* - f_{ij})}{f_i^* - f_i^-} \right] \tag{5}$$

S^-, S^*, R^- and R^* values can be computed by Eqs. (6), (7).

$$S^* = \min_j S_j \ , \qquad S^- = \max_j S_j \tag{6}$$

$$R^* = \min_j R_j \ , \qquad R^- = \max_j R_j \tag{7}$$

Step 3: compute the values Q_j for j=1, 2, … , J by Eq. (8).

$$Q_j = \frac{v(S_j - S^*)}{S^- - S^*} + \frac{(1-v)(R_j - R^*)}{R^- - R^*} \tag{8}$$

Where v is defined as the maximum group utility and is set to 0.5 in many researches.

Step 4: rank the alternatives using S_j, R_j and Q_j values in decreasing order.

Step 5: If the alternative a' has two following conditions propose it as a compromise solution:

Acceptable advantage: $Q(a'') - Q(a') \geq \frac{1}{J-1}$

Acceptable stability in decision-making: The alternative a' is ranked the best by S or/and R.

Where alternative a' has the minimum value of Q ($Q(a') = \min_j Q_j$) and a'' is ranked the second alternative by Q.

If one of the conditions is not satisfied then a set of compromise solutions is proposed that consist of:

Alternatives a' and a'' are compromise solutions if only the condition (b) is not satisfied. Alternatives a', a'', …, a^M are proposed as compromise solutions if the condition (a) is not satisfied. a^M is ranked the Mth by Q and is determined by Eq. (9) for maximum M.

$$Q(a^M) - Q(a') < \frac{1}{J-1} \tag{9}$$

However, VIKOR is used in many researches as an efficient MADM method. For instance Sanayei et al. (2010) used this method for supplier selection problem. San Cristobal (2011) used VIKOR in the selection of a renewable energy project. Peng et al. (2011) proposed VIKOR for evaluating classification algorithms of financial risk prediction. In current study mutual funds are evaluated using VIKOR and with 5 attributes.

2.2 Attributes

2.2.1 Rate of Return

The net asset value (NAV) is defined as the current value of fund's assets minus its liabilities divided by available number of current shares (Mobius, 2007). Rate of return based on net asset value defined as:

$$R_{i,t} = \frac{NAV_{i,t} - NAV_{i,t-1}}{NAV_{i,t-1}} * 100\% \tag{10}$$

In Eq. (10) i is the number of mutual funds, $R_{i,t}$ is the rate of return of ith fund at time t and $NAV_{i,t}$ is the net asset value of ith fund at time t.

2.2.2 Variance

$$\sigma_i^2 = \frac{\sum_{i=1}^{T}(R_{i,t} - \bar{R}_i)}{T} \tag{11}$$

Where $\bar{R}_i$ is the average return rate of T months.

2.2.3 Semivariance

$$SV_i = \sum_{t=1\,;\,R_{i,t}<\bar{R}_i}^{T} (R_{i,t} - \bar{R}_i) \Big/ T \tag{12}$$

2.2.4 Treynor Index

For performance evaluation of portfolios, Treynor (1965) proposed an index like Eq. (13).

$$TR_i = \frac{\bar{R}_i - R_f}{\beta_i} \tag{13}$$

Where TR_i is the Treynor index, $\bar{R}_i$ is the average rate of return, R_f is the return of risk-free asset and β_i is the measurement of systematic risk and calculated as Eq. (14):

$$\beta_i = \frac{cov(R_i, R_m)}{\sigma_m^2} \tag{14}$$

Where R_m is the return of market index, $COV(R_i, R_m)$ is the covariance between the return of mutual fund i and the return of the market index and σ_m^2 is the variance of R_m.

2.2.5 Sharpe Index

Sharpe (1966) proposed a ratio for performance evaluation of mutual funds.

$$SR_i = \frac{\bar{R}_i - R_f}{\sigma_i} \tag{15}$$

In Eq. (15) SR_i is the Sharpe index and σ_i is the standard deviation and called total risk.

3. Second Stage: Portfolio Optimization

Portfolio optimization is an old problem that decides to allocate investors' wealth among several assets. This problem has been proposed as a mathematical model since the first study of Markowitz in 1952. The model presented in the researches of Markowitz (1952, 1959) is known as basic model.

3.1 Basic Model

In financial markets each investor likes to obtain more return and less risk. So that in basic portfolio optimization model there are two objective functions and one constraint. One of the objective functions maximizes the return and the other minimizes the risk. This model is shown in Eq. (16).

$$Maximize \quad R_p = \sum_{i=1}^{N} R_i . x_i \tag{16}$$

$$Minimize \quad \sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i . x_j . \sigma_{ij}$$

$$Subject\ to \quad \sum_{i=1}^{N} x_i = 1$$

$$x_i \geq 0 \qquad i = 1, 2, \ldots, N$$

Where N is the number of assets, x_i is the proportion invested in asset i, R_i is the rate of return of ith asset, R_p is the rate of return of portfolio, σ_p^2 is the variance of portfolio and σ_{ij} is the covariance between the return rates of asset i and asset j.

The basic model considers normal distribution for rate of return so that variance is an appropriate risk measure. To solve basic model we can divide it into two models and solve one of them. One model maximizes the return and considers the risk as a constraint and the other minimizes the risk and considers the return as a constraint. Assuming different values for return or risk constraint results efficient frontier.

3.2 Mean-Semivariance Model

The assumption of normal distribution is not always accurate in practice. So that downside risk measures are defined to accurate portfolio optimization problems. Calculation of risk in downside risk measures is based on this theory that when the expected return is higher than mean we have a non-risk situation and when is lower we have risk. There are several studies that consider downside risk measures such as: Vercher et al. (2007), Estrada (2007), Benati & Rizzi (2007), Gupta et al. (2008), Chang et al. (2009) and Yang et al. (2011). In this study we consider semivariance as a risk measure and rewrite the basic portfolio optimization model using this measure.

$$Maximize \quad R_p = \sum_{i=1}^{N} R_i . x_i \tag{17}$$

$$Minimize \quad SV_p = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i . x_j . SV_{ij}$$

$$Subject\ to \quad \sum_{i=1}^{N} x_i = 1$$

$$x_i \geq 0 \qquad i = 1, 2, \ldots, N$$

In Eq. (17) SV_p is defined as semivariance of portfolio and SV_{ij} is cosemivariance (semicovariance) between the return rates of asset i and asset j. cosemivariance can be defined as Eq. (18) (Estrada, 2007). In Eq. 17 the second objective function is non-linear but it is convex and we have a convex optimization model.

$$SV_{ij} = E\{min[(R_i - \bar{R}_i), 0] . min[(R_j - \bar{R}_j), 0]\} \tag{18}$$

3.3 Cardinality Constraint

Researchers have changed the basic portfolio optimization model by adding objective functions or constraints to make it more accurate. Added objective functions are usually entropy function or liquidity. A common constraint that is used in models is cardinality constraint as in Chang et al. (2000), Gupta et al. (2008), Branke et al. (2009), Chang et al. (2009), Yang et al. (2011) and Golmakani & Fazel (2011) studies. Cardinality constraint fixes the number of assets in optimized portfolio. In large-scale problems portfolio management is difficult and investors like to have finite assets in their portfolio. For defining cardinality in portfolio optimization models a constraint like Eq. (19) adds to model.

$$\sum_{i=1}^{N} z_i = K \tag{19}$$

Where K is cardinality level and z_i is a zero-one variable. $z_i = 1$ if ith asset is held and $z_i = 0$ otherwise. This constraint makes the model an integer programming problem. The problem has at least two objective functions (return and risk) and because of risk function is non-linear. So we can call this model a multi-objective non-linear integer programming. In large-scale problems solving this model is very hard and forces us to use metaheuristics. On the other hand, in cardinality constrained problems the efficient frontier is discontinuous because of integer variables (Change et al, 2000). In this study, according to cardinality level, best mutual funds are chosen. For example when K=6, we choose six funds that ranked the best based on VIKOR ranking. Then these funds are optimized for determining the proportions of capital that should be invested on.

3.4 Fuzzy Programming Technique

Fuzzy programming is an efficient technique for solving multi-objective non-linear problems that is proposed by Zimmermann (1978). Jana et al. (2009) showed that this technique can be applied in portfolio optimization problem. Fuzzy programming technique follows these steps for solving the problem that is shown in Eq. (17):

Step 1: consider the problem as two single objective problems and minimize and maximize the problems with constraint. These solutions are named ideal solutions. R_p^* is defined as maximum value of return objective function and R_p^-, SV_p^* and SV_p^- are defined likewise.

Step 2: compute the intervals of objective functions. ΔR_p is defined as interval of return objective function and

ΔSV_p is defined as interval of risk objective function. These definitions are shown in Eqs. (20), (21).

$$\Delta R_p = R_p^* - R_p^- \tag{20}$$

$$\Delta SV_p = SV_p^* - SV_p^- \tag{21}$$

Step 3: define membership functions of objectives as Eqs. (22), (23).

$$\mu(R_p) = \begin{cases} 0 & R_p \le R_p^- \\ \dfrac{R_p - R_p^-}{\Delta R_p} & R_p^- < R_p < R_p^* \\ 1 & R_p \ge R_p^* \end{cases} \tag{22}$$

$$\mu(SV_p) = \begin{cases} 0 & SV_p \le SV_p^- \\ \dfrac{SV_p^* - SV_p}{\Delta SV_p} & SV_p^- < SV_p < SV_p^* \\ 1 & SV_p \ge SV_p^* \end{cases} \tag{23}$$

Step 4: define α_R as the percentage that return objective function is neared to its optimized solution or minimum value of its membership function and define α_{SV} likewise. Now we can write Eqs. (24), (25).

$$\alpha_R \le \mu(R_p) \rightarrow \alpha_R \le \frac{R_p - R_p^-}{\Delta R_p} \rightarrow R_p \ge R_p^- + \alpha_R.\Delta R_p \tag{24}$$

$$\alpha_{SV} \le \mu(SV_p) \rightarrow \alpha_{SV} \le \frac{SV_p^* - SV_p}{\Delta SV_p} \rightarrow SV_p \le SV_p^* - \alpha_{SV}.\Delta SV_p \tag{25}$$

Now we can rewrite the model that is proposed in Eq. (17) as a single objective model that operation research softwares can solve it easily.

$$\begin{aligned} Maximize \quad & L = w_R.\alpha_R + w_{SV}.\alpha_{SV} \\ Subject\ to: \quad & R_p \ge R_p^- + \alpha_R.\Delta R_p \\ & SV_p \le SV_p^* - \alpha_{SV}.\Delta SV_p \\ & \sum_{i=1}^{N} x_i = 1 \\ & x_i \ge 0, \quad i = 1, 2, \dots, N \end{aligned} \tag{26}$$

Where L is the percent that objective functions near to their optimized solution additionally, w_R is the weight of return objective function and w_{SV} is the weight of risk objective function.

4. Numerical Example

In this section we show our proposed methodology in a computational example. There are 92 mutual funds in our sample from NASDAQ. In first stage, these funds are studied between years 2007-2009 and in monthly data. Rate of return, variance, semivariance, Treynor index and Sharpe index are extracted from 36 data. We have a MADM problem with 92 alternatives and 5 attributes. The VIKOR method is used for ranking of funds. We consider equal weights (1/3) for return, risk and portfolio indices. Because we have two measures for risk and two for portfolio indices the weight of variance, semivariance, Treynor and Sharpe index is assumed 1/6. Ranking of best funds and their attributes is shown in table 1 for top ten funds.

Table 1. Ranking of funds and their attributes

Fund number	Return rate	Variance	Semi variance	Treynor index	Sharpe index
41	0.93	220.77	99.24	14.55	0.66
2	0.27	55.01	21.04	0.64	0.13
75	0.57	201.46	117.57	8.97	0.43
5	0.8	331.68	163.85	6.8	0.38

92	0.83	486.41	212.65	6.74	0.4
12	0.64	314.49	154.11	6.04	0.32
19	0.13	46.68	22.01	-6.18	-0.08
34	0.42	163.61	86.69	3.92	0.2
66	0.78	1120.03	685.35	7.25	0.45
91	0.66	448.48	195.56	5.18	0.34

All of the funds are ranked and compromise solutions are obtained. Then, 4 levels considered for cardinality: K=2, 4, 6, 8. when K=2, funds number 41 and 2 are chosen for portfolio optimization and when K=4 funds number 41, 2, 75 and 5 etc. From attributes return rate and semivariance are remained for next stage. In second stage a mean-semivariance portfolio optimization model is written for chosen funds according to cardinality level. For example for K=4 we have a model like the model is shown in Eq. (27). For making this model a cosemivariance matrix is computed.

$$\begin{aligned} &Maximize \quad L = w_R.\alpha_R + w_{SV}.\alpha_{SV} \\ &Subject\ to: \ 0.93x_1 + 0.27x_2 + 0.57x_3 + 0.8x_4 \geq 0.27 + 0.66\alpha_R \\ &99.24x_1^2 + 21.04x_2^2 + 10.49x_1x_2 + 117.57x_3^2 + 163.85x_4^2 + 87.41x_1x_3 + 45.67x_1x_4 \\ &+ 7.74x_2x_3 + 48.14x_2x_4 + 225.72x_3x_4 \leq 163.85 - \ 145.9\alpha_{SV} \end{aligned} \tag{27}$$

$$\sum_{i=1}^{4} x_i = 1$$

$$x_i \geq 0, \quad i = 1, 2, 3, 4$$

Lingo 11 software can solve this problem. In all problems $w_R + w_{SV}$ is assumed equal one. w_R is started from zero to one with steps equal 0.1. The results of solving this problem is shown in table 2.

Table 2. Computed values for return and risk of portfolio and investment proportions

		R_p	SV_p	x_1	x_2	x_3	x_4
$w_R = 0$	$w_{SV} = 1$	0.364	17.667	0.099	0.812	0.089	0
$w_R = 0.1$	$w_{SV} = 0.9$	0.413	18.31	0.175	0.74	0.085	0
$w_R = 0.2$	$w_{SV} = 0.8$	0.473	20.769	0.269	0.651	0.08	0
$w_R = 0.3$	$w_{SV} = 0.7$	0.606	29.426	0.413	0.469	0	0.118
$w_R = 0.4$	$w_{SV} = 0.6$	0.747	46.875	0.56	0.236	0	0.204
$w_R = 0.5$	$w_{SV} = 0.5$	0.892	73.349	0.715	0	0	0.285
$w_R = 0.6$	$w_{SV} = 0.4$	0.896	74.538	0.748	0	0	0.252
$w_R = 0.7$	$w_{SV} = 0.3$	0.903	77.57	0.803	0	0	0.197
$w_R = 0.8$	$w_{SV} = 0.2$	0.918	87.59	0.913	0	0	0.087
$w_R = 0.9$	$w_{SV} = 0.1$	0.93	99.238	1	0	0	0
$w_R = 1$	$w_{SV} = 0$	0.93	99.238	1	0	0	0

Considering 11 levels for w_R and w_{SV} results an efficient frontier. Efficient frontiers for K=4 and other cardinality levels are shown in Figs. 1-4. Each efficient frontier is compared with unconstrained efficient frontier and all of the efficient frontiers are compared in Fig. 5.

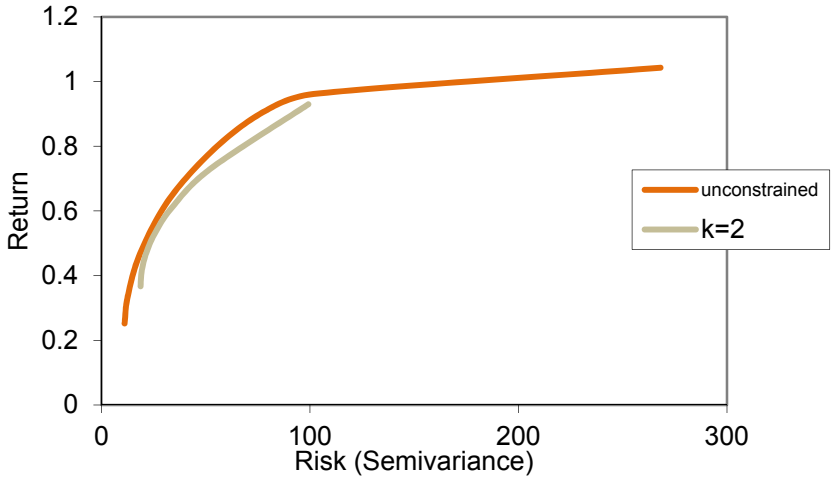


Figure 1. Efficient frontier considering cardinality level=2

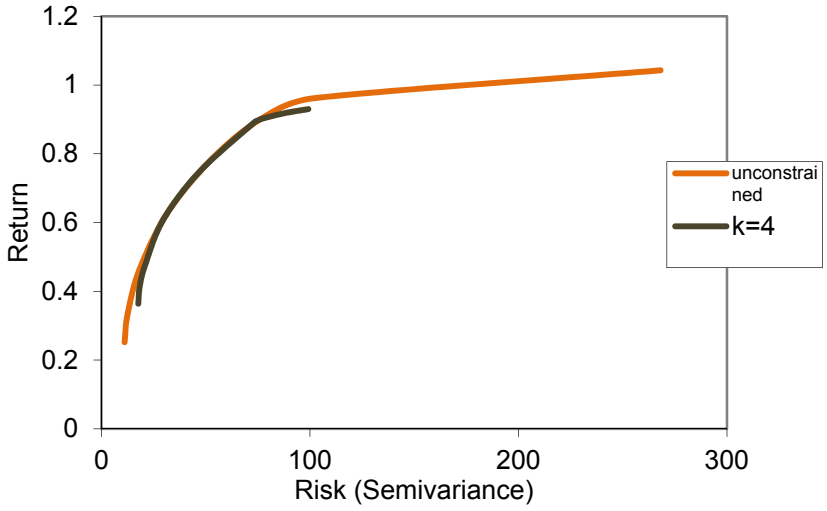


Figure 2. Efficient frontier considering cardinality level=4

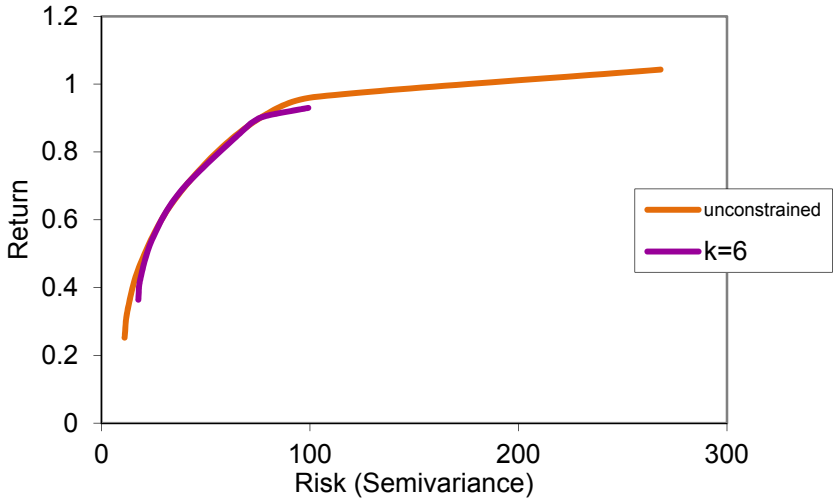


Figure 3. Efficient frontier considering cardinality level=6

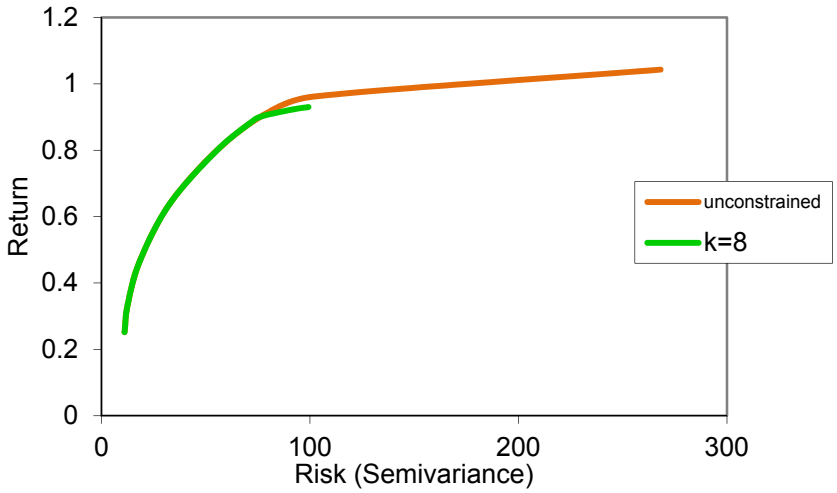


Figure 4. Efficient frontier considering cardinality level=8

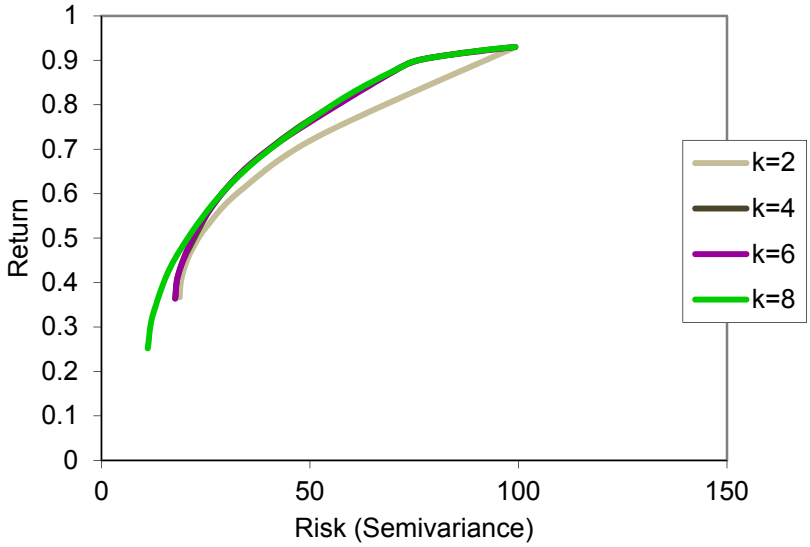


Figure 5. Efficient frontiers considering different cardinality levels

As shown in Figures 1-4 increasing cardinality level leads to near constrained and unconstrained solutions. When k=2 the difference between constrained efficient frontier and unconstrained is considerable but when we increase cardinality level this difference is decreased. As shown in Fig. 4 when k=8 both of efficient frontiers are matched in almost every risk level but unconstrained efficient frontier dominated constrained efficient frontier and it is not surprising because in unconstrained solution there are 92 funds and in constrained solution there are only 8 funds. It is important that using proposed methodology leads to similar results between unconstrained solutions and cardinality constrained solutions. In other words, we approximately obtain the unconstrained solutions with 8 funds and this is the contribution of this study. The small difference between constrained and unconstrained efficient frontier in Fig. 4 is because of one fund that was not in our portfolio in k=8. This fund has appropriate return rate but other characteristics such as its Treynor index and Sharpe index is not appropriate enough to place in best mutual funds portfolio. In table 3 the characteristics of this fund (number 40) and the fund that ranked the best in VIKOR ranking (number 41) are compared. On the other words this difference is because of effectiveness of Treynor index and Sharpe index in our methodology. Fig. 5 shows the comparison of different cardinality levels. As shown in this figure k=8 efficient frontier dominates other efficient frontiers and k=4 and k=6 efficient frontiers are matched approximately.

Table 3. A comparison between two funds

Fund number	Return rate	variance	semivariance	Treynor index	Sharpe index
41	0.93	220.77	99.24	14.55	0.66
40	1.04	442.33	210.05	-3.04	-0.05

5. Conclusion

Mutual funds decide to play the role of diversification as an institute in financial markets. In current study a new two-stage methodology for performance evaluation of mutual funds is proposed. In first stage 5 characteristics are measured for all of mutual funds. The characteristics are return rate, variance, semivariance, Treynor index and Sharpe index. The VIKOR method is used as a MADM method for ranking these funds. The funds are ranked based on 5 characteristics and considering equal weights for return, risk and portfolio indices. In second stage the best funds are chosen considering cardinality level. When cardinality level is equal 2 the first and second funds according to VIKOR ranking are used. 4 levels are assumed as cardinality levels. Then a multi-objective portfolio optimization model is written. In this model semivariance is used as a downside risk measure. This model is solved using fuzzy technique programming and efficient frontiers are obtained.

Results show that when cardinality level increases we can have appropriate solutions. In this study these solutions are observed in cardinality level equal 8. In this cardinality level the cardinality constrained efficient frontier and unconstrained efficient frontier are matched approximately. Because in models that consider cardinality constraint, the efficient frontier is discontinuous, obtaining continuous efficient frontier was our success.

In this study we have considered cardinality constraint without using integer programming or metaheuristics. Future works can focus on other constraints in portfolio optimization model and facilitate its solving. As another work and based on this methodology other MADM methods and portfolio optimization models can be used.

References

Anagnostopoulos, K. P., & Mamanis, G. (2010). A portfolio optimization model with three objectives and discrete variables. *Expert Systems With Applications, 38*, 14208-14217. http://dx.doi.org/10.1016/j.cor.2009.09.009

Anagnostopoulos, K. P., & Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multi objective evolutionary algorithms. *Computers & Operations Research, 37*, 1285–1297. http://dx.doi.org/ 10.1016/j.eswa.2011.04.233

Basso, A., & Funari, S. (2001). A data envelopment analysis approach to measure the mutual fund performance. *European Journal of Operational Research, 135*, 477–492. http://dx.doi.org/10.1016/S0377-2217(00)0031-8

Benati, S., & Rizzi, R. (2007). A mixed integer linear programming formulation of optimal mean/Value-at-Risk portfolio problem. *European Journal of Operational Research, 176*, 423–434. http://dx.doi.org/10.1016/j.ejor.2005.07.020

Branke, J., Scheckenbach, B., Stein, M., Deb, K., & Schmeck, H. (2009). Portfolio optimization with an envelope-based multi-objective evolutionary algorithm. *European Journal of Operational Research, 199*, 684–693. http://dx.doi.org/10.1016/j.ejor.2008.01.054

Chang, C. H., Lin, J. J., Lin, J. H., & Chiang, M. C. (2010). Domestic open-end equity mutual fund performance evaluation using extended TOPSIS method with different distance approaches. *Expert Systems with Applications, 37*, 4642-4649. http://dx.doi.org/10.1016/j.eswa.2009.12.044

Chang, T. J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research, 27*, 1271–1302. http://dx.doi.org/10.1016/S0305-0548(99)00074-X

Chang, T. J., Yang, S. C., & Chang, K. J. (2009). Portfolio optimization problem in different risk measures using genetic algorithm. *Expert Systems With Applications, 36*, 10529-10537. http://dx.doi.org/10.1016/j.eswa.2009.02.062

Chen, L. H., & Huang, L. (2009). Portfolio optimization of equity mutual funds with fuzzy return rates and risks. *Expert Systems with Application, 36*, 3720-3727. http://dx.doi.org/10.1016/j.eswa.2008.02.027

Estrada, J. (2007). Mean-semivariance behavior: downside risk and capital asset pricing. *International Review of Economics and Finance, 16*, 169-185. http://dx.doi.org/ 10.1016/j.iref.2005.03.003

Golmakani, H. R., & Fazel, M. (2011). Constrained portfolio selection using Particle Swarm Optimization. *Expert*

Systems With Applications, 38, 8327-8335. http://dx.doi.org/10.1016/j.eswa.2011.01.020

Gupta, P., Mehlawat, M. K., & Saxena, A. (2008). Asset portfolio optimization using fuzzy mathematical programming. *Information Sciences, 178*, 1734-1755. http://dx.doi.org/10.1016/j.ins.2007.10.025

Jana, P., Roy, T. K., & Mazumder, S. K. (2009). Multi-objective possibilistic model for portfolio selection with transaction cost. *Journal of Computational and Applied Mathematics, 228*, 188-196. http://dx.doi.org/10.1016/j.cam.2008.09.008

Lin, C. C., & Liu, Y. T. (2008). Genetic algorithms for portfolio selection problems with minimum transaction lots. *European Journal of Operational Research, 185*, 393–404. http://dx.doi.org/10.1016/j.ejor.2006.12.024

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1959). Portfolio selection: efficient diversification of investments. New York: Wiley.

Mobius, M. (2007). Mutual funds: an introduction to the core concepts. Singapore: John Wiley & Sons (Asia) Pte Ltd.

Murthi, B. P. S., Choi, Y. K., & Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research, 98*, 408–418. http://dx.doi.org/10.1016/S0377-2217(96)00356-6

Opricovic, S. (1998). Multi-criteria optimization of civil engineering systems. Belgrade: Faculty of Civil Engineering.

Opricovic, S., & Tzeng, G. H. (2002). Multicriteria planning of post-earthquake sustainable reconstruction. *Computer-Aided Civil and Infrastructure Engineering, 17*(3), 211–220.

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research, 156*(2), 445–455. http://dx.doi.org/10.1016/S0377-2217(03)00020-1

Peng, Y., Wang, G., Kou, G., & Shi, Y. (2011). An empirical study of classification algorithm evaluation for financial risk prediction. *Applied Soft Computing, 11*, 2906-2915. http://dx.doi.org/10.1016/j.asoc.2010.11.028

San Cristobal, J. R. (2011). Multi-criteria decision-making in the selection of a renewable energy project in spain: the VIKOR method. *Renewable Energy, 36*, 498-502. http://dx.doi.org/10.1016/j.renene.2010.07.031

Sanayei, A., Farid-Mousavi, S., & Yazdankhah, A. (2010). Group decision making process for supplier selection with VIKOR under fuzzy environment. Expert Systems With Applications 37, 24-30. http://dx.doi.org/10.1016/j.eswa.2009.04.063

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.

Treynor, J. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.

Woodside-Oriakhi, M., Lucas, C., & Beasley, J. E., (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research, 213*, 538-550. http://dx.doi.org/10.1016/j.ejor.2011.03.030

Xia, Y., Wang, S., & Deng, X., (2001). A compromise solution to mutual funds portfolio selection with transaction costs. *European Journal of Operational Research, 134*, 564-581. http://dx.doi.org/10.1016/S0377-2217(00)00278-2

Yang, S. C., Lin, T. L., Chang, T. J., & Chang, K. J. (2011). A semi-variance portfolio selection model for military investment assets. *Expert Systems With Applications, 38*, 2292-2301. http://dx.doi.org/10.1016/j.eswa.2010.08.017

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Modeling of Correlated Two-Dimensional Non-Gaussian Noises

Vladimir Mikhailovich Artuschenko[1], Kim Leonidovich Samarov[2], Andrey Petrovich Golubev[1], Aleksey Yurievich Shchikanov[1] & Aleksey Sergeevich Kochetkov[1]

[1] The Federal State-Funded Educational Institution of Higher Professional Education Russian State University of Tourism and Service, Russia

[2] The State-Funded Educational Institution of Higher Professional Education Moscow region Finance and Technology Academy, 42, Gagarina street, Korolev city, Moscow region, Russia

Correspondence: Aleksey Sergeevich Kochetkov, The Federal State-Funded Educational Institution of Higher Professional Education Russian State University of Tourism And Service, 99, Glavnaya Street, Cherkizovo village, Pushkinskiy district, Moscow region, 141221, Russia.

Received: September 10, 2014 Accepted: September 25, 2014 Online Published: July 27, 2015

doi:10.5539/mas.v9n8p300 URL: http://dx.doi.org/10.5539/mas.v9n8p300

Abstract

The article describes and analyzes mathematical models of multiplicative and additive non-Gaussian noises affecting the useful signals. For synthesis and analysis, and, hence, the effective design of information systems and radio devices operating in conditions of intense perturbations it is necessary to choose not only the adequate mathematical model of the useful signals and information processes, but also the corresponding models of random effects, possessing in general non-Gaussian multiplicative and additive character. To describe the arbitrary non-Gaussian noises, which are quasiharmonic processes whose spectrum is close (or narrowband) to the band of the desired signal, the authors used a two-dimensional elliptic symmetric probability density function, including two extreme cases: Gaussian processes and a sinusoidal signal with random initial phase distributed uniformly in the interval [0, 2 π]. Model of correlated non-Gaussian narrowband noises of elliptically symmetric two-dimensional probability density function allows you to make a synthesis of information systems and devices based only on a priori knowledge of one-dimensional probability density function and the correlation function. Because knowing one-dimensional probability density function of the instantaneous values, we can determine the probability density function of the envelope; this makes it possible to use the elliptically symmetric probability density function to describe not only additive, but multiplicative (baseband) noises. To describe the real density of probability density function of the non-Gaussian process, the authors propose to approximate its a priori known one-dimensional probability density function and a specially designed transitional probability density function, and show the adequacy of this approximation of the real two-dimensional probability density function of correlated noises.

Keywords: information process, adequate mathematical models, additive and multiplicative noise, the probability density function, non-Gaussian processes

1. Introduction

For synthesis and analysis, and, consequently, for efficient design of radio systems and devices operating under intense perturbations, it is necessary to select not only adequate mathematical models of information processes $\lambda(t)$, but also accidental impacts, having, in general, the multiplicative $\eta(t)$ and additive $n(t)$ character (Artyushenko and Volovach, 2013; Artyushenko, 2013; Artyushenko and Samarov, 2013). Typically, the disturbance (noise) acting on radio systems and devices, are random processes with non-Gaussian probability density function (PDF) (stationary and non-stationary) (Artyushenko and Abbasova, 2011; Yong and Westerberg, 1971; Edward and Wegman, 1983; Bucy and Mallinckrodt, 1973). The most complete description of stochastic processes (sequences) is a method of multidimensional PDF. Several methods for describing and modeling of stochastic processes with multidimensional PDF. One of such methods is the mixing of random processes (Trofimov, 1986), based on the concept of a PDF random sequence $\{\lambda_h, h = \overline{1.H}\}$ by the sum of weighted PDF: $W(\lambda_1, \dots, \lambda_H) = \sum_{i=1}^{N} p_i W_i(\lambda_1, \dots, \lambda_H)$, where p_i is random weighting coefficients, provided that $\sum_{i=1}^{N} p_i = 1$; $W_i(\lambda_1, \dots, \lambda_H)$ – given H-dimensional distributions.

Sequence elements $\{\lambda_h, h = \overline{1.H}\}$ are interpreted as samples obtained by discretization of the corresponding

process $\lambda(t)$ at the time moment t_h, provided that, generally, $t_h - t_{h-1} = T_o = \text{const}$. The most widely used case is when as $W_i(\lambda_1, \dots, \lambda_H)$ there are used N-dimensional normal distributions.

Great opportunities for receiving multidimensional PDFs open Markov processes, allowing with the required accuracy to approximate the random process. In this article we consider the continuous-valued Markov processes.

2. Methodology

As is known, a common form of description of random Markov processes is the systems of statistical differential equations, as well as shaping filters. The complexity of the formation and the need to set a large number of a priori information, which is often difficult to obtain in practice (especially for non-Gaussian PDF), sometimes forced to abandon a full probabilistic description of random processes in favor of simplified. Most available information on any random process is one-dimensional correlation function and PDF. In these circumstances, to describe the real information and noises, there are widely used Markov models. Their high efficiency is widely known from the works about the Markov theory of nonlinear filtering (Yarlykov, 1980; Tikhonov and Mironov, 1977).

To describe arbitrary non-Gaussian noises, which are quasiharmonic processes whose spectrum is close (or narrowband) with a band of the desired signal, there can be used elliptical symmetric (ES) two-dimensional PDFs, including two extreme cases: Gaussian processes and a sinusoidal signal with random initial phase distributed uniformly in the interval $[0, 2\pi]$ (McGraw and Wagner, 1968; Middleton, 1973).

Elliptically symmetric two-dimensional PDFs $W_2(n_1, n_2)$ of the stationary process $n(t)$ depend on n_1 and n_2 ($n_1 = n(t)$, $n_2 = n(t + \tau)$) only in the combination $l = [n_1^2 + n_2^2 - 2r(\tau)n_1n_2]^{0.5}$, where $r(\tau) = B_n(\tau)/B_n(0)$ is a correlation coefficient of values n_1 and n_2, represents a form of ellipses.

Consequently, it is possible to write that

$$W_2(n_1, n_2) = Cf(R), \tag{1}$$

where C is a normalization constant; $R = l(1 - r^2)^{-0.5}$;

$$f(R) = [2\pi(1 - r^2)]^{-1} \int_0^\infty \Theta(\nu) J_0(R\nu)\nu d\nu \tag{2}$$

– function is the transformation zero-order Bessel-dimensional characteristic function $\Theta(\nu)$ of the random process being under the consideration.

As it can be seen from the relations (1), (2), $W_2(n_1, n_2)$ is completely determined by the one-dimensional PDF $W_1(n)$, related to the Fourier transformation with the characteristic function $\Theta(\nu)$ and a correlation coefficient $r(\tau)$ of the process being under the consideration. At the same time the one-dimensional PDF and the corresponding characteristic function are the even functions.

However, it should be noted that for the construction of ES-distribution (1), there may be used only such even functions $W_1(n)$, that lead to a non-negative and integrable function $W_2(n_1, n_2)$.

In this case fulfillment of the inequality $\int_0^R Rf(R)\, dR < \infty$, is a necessary and sufficient condition for the existence of ES two-dimensional distribution determined through $W_1(n)$ and $r(\tau)$ (Middleton, 1973).

Function

$$W(R) = 2\pi C(1 - r^2)^{0.5} Rf(R), \tag{3}$$

when describing the narrowband random process coincides with the PDF envelope (amplitude U) of the process. Therefore, the expression (3) can be written as:

$$W(U) = 2\pi C(1 - r^2)^{0.5} Rf(U).$$

It is a feature of ES distributions arising from their definition (Trofimov, 1986). As for the noises with a band spectrum the probability density function of amplitude (PDFA) is a sufficiently probable characteristic, so that it can be argued that the ES model of the corrected non-Gaussian process uniquely determines such noises.

In (Middleton, 1973; Artyushenko and Solenov, 1993) there are presented the main characteristics of random processes, which two-dimensional distributions possess elliptical symmetry.

Note that when adding arbitrary ES processes to the same correlation coefficients, the obtained process is also the ES process.

So, when you add a sinusoid to random initial phase uniformly distributed in the interval $[0, 2\pi]$, and narrowband Gaussian noise (at the equal correlation coefficients) taking into account the relations (1), (2) we get

the ES process, which PDFA is subject to the Rice law:

$$W_2(n_1,\ n_2) = [2\pi(1 - r^2)^{0.5}\sigma^2]^{-1}\exp\{(U^2 + R^2)(2\sigma^{-1})\}I_0(UR\sigma^{-2}) \quad (4)$$

where U is the amplitude of the sinusoidal component; σ^2 is the noise variance; $r(\tau) = \cos\omega_0\tau$.

You can come to the PDF described by (4) by using the relation (2) and the expression of the characteristic function for the total process:

$$\Theta_1(\nu) = J_0(U\nu)\exp\{- 0{,}5\sigma^2\nu^2\} \quad (5)$$

By integrating (2) we obtain (4).

Note that in the case of an arbitrary correlation $r(\tau) = r_0(\tau)\cos\omega_0\tau$, where $r_0(\tau)$ is a slowly decaying function, described by the expression (5), there can be no ES. In this case, the two-dimensional PDF (4) can be considered as a ES model provided that $\tau << \tau_{cor}$, where τ_{cor} is a correlation interval of the describing processes defined by the envelope $r(\tau)$ (Artyushenko and Solenov, 1993).

Let us consider as an example the case when the instantaneous values of additive noise described by generalized Gaussian distribution:

$$W(n) = [\nu\gamma(\sigma_n, \nu)/2\Gamma(\nu^{-1})]\exp\{- [\gamma(\sigma_n, \nu)|n|]^\nu\},$$
$$\text{where } \gamma(\sigma_n, \nu) = \sigma_n^{-1}[\Gamma(3/\nu)/\Gamma(1/\nu)]^{0.5}.$$

Considering that matching sampling time moments of quadrature components are uncorrelated, we obtain:

$$W_2(n_1, n_2) = [\nu\gamma_0^2(\sigma_n, \nu)/2\pi\Gamma(2/\nu)]\exp\{- [\gamma_0(\sigma_n, \nu)(n_1^2 + n_2^2)^{0.5}]^\nu\},$$
$$\text{where } \gamma_0(\sigma_n, \nu) = \sigma_n^{-1}[\Gamma(4/\nu)/2\Gamma(2/\nu)]^{0.5}; -\infty < n_1; n_2 < \infty.$$

At the same time the PDF noises:

$$W(U) = [\nu\gamma_0(\sigma_n, \nu)U/2\Gamma(2/\nu)]\exp\{- [\gamma_0(\sigma_n, \nu)U]^\nu\};\ 0 \le U < \infty.$$

If the narrow-band random process is stationary, then PDFA $W(U)$ and PDF of its instantaneous values are interrelated by (Artyushenko and Solenov, 1993):

$$W(U) = U\int_0^\infty \nu\Theta_n(\nu)J_0(U)d\nu, \quad (6)$$

where $\Theta_n(\nu) = \int_{-\infty}^{\infty} W(n)\exp\{jvn\}dn$ is a characteristic function of the process $n(t) = U(t)\cos F(t)$; $U(t)$ and $F(t)$ are the envelope and the total phase of a random process, respectively.

Making the necessary changes to (6), we obtain:

$$W(n) = \pi^{-1}\int_{|n|}^{\infty} W(U)(U^2 - n^2)^{-0.5}dU.$$

Thus, the description of correlated non-Gaussian narrowband noise of elliptically symmetric two-dimensional PDF allows to produce synthesis systems and devices based on a priori knowledge and one-dimensional PDF and correlation function. Knowing the one-dimensional PDF of instantaneous values $W(n)$, we can determine the PDF of the envelope $W(U_n)$, which makes it possible to use an elliptically symmetrical PDF for description of a multiplicative (modeling) noises.

Along with the described above methods for description of correlated non-Gaussian processes, let us consider the following.

The real PDF of non-Gaussian process, it is proposed to approximate by a priori known one-dimensional PDF $W(n_{h-1})$ and specially designed transitional PDF $W^A(n_h|n_{h-1})$. As a result the PDF of non-Gaussian process will be described as:

$$W^A(n_h,\ n_{h-1}) = W(n_h)W^A(n_h|n_{h-1}). \quad (7)$$

As a transitional PDF, we use the PDF of the following form:

$$W^A(n_h|n_{h-1}) = (2\pi G^2)^{-0.5}\exp\{- [n_h - M(n_{h-1})]^2/2G^2\}, \quad (8)$$

where G^2 characterizes the intensity of the random process $\{n_h\}$;

$M(n_{h-1}) = n_{h-1} - 0{,}5G^2\frac{d}{dn_{h-1}}\ln W(n_{h-1})$ is a function of a special form.

Note that in case when the process $\{n_h\}$ is described by a Gaussian PDF $W(n_{h-1}) = N(0, \sigma^2)$, equation (8) takes the form of:

$$W^A(n_h|n_{h-1}) = (2\pi G^2)^{-0.5}\exp\{- (n_h - rn_{h-1})^2/2G^2\}. \quad (9)$$

Considering $G^2 = \sigma^2(1 - r^2)$, come to the well-known expression for the conditional Gaussian PDF.

Representation of the two-dimensional PDF in the form (9) will be used further in the synthesis tasks, so it is necessary to substantiate the adequacy of the introduced approximation.

As a criterion we will use the information criterion:

$$\min\ I_K(W,\ W^A);\ n_h,\ n_{h-1} \in \Pi, \tag{10}$$

where I_K is the Kullback information characterizing the estimate of the average information, containing in the area Π of changes the component n_h and n_{h-1}, correlated random process at hypotheses testing

$$H_0: W(n_h|n_{h-1}) \text{ and } H_1: W^A(n_h|n_{h-1}).$$

There are two methods for the appraisal of the Kullback:

$$I_{12K}(W,\ W^A) = \iint_{-\infty}^{\infty} W(n_h, n_{h-1}) \ln \frac{W(n_h, n_{h-1})}{W^A(n_h, n_{h-1})} dn_h dn_{h-1}; \tag{11}$$

$$I_{21K}(W,\ W^A) = \iint_{-\infty}^{\infty} W^A(n_h, n_{h-1}) \ln \frac{W^A(n_h, n_{h-1})}{W(n_h, n_{h-1})} dn_h dn_{h-1}. \tag{12}$$

Criterion (10) and relations (11), (12) will be used at the testing phase when checking validity of PDF non-Gaussian processes PRB description made by the relations (7) and (8).

Let us consider a few examples of non-Gaussian processes for which $W(n_h, n_{h-1})$ is known, and then, turn to the description of non-Gaussian correlated processes for which are known only the one-dimensional PDFs.

As testing PDFs, we introduce the distribution of the form:

$$W(n_h) = \frac{\nu}{\Gamma(\nu^{-1})\sigma}\left[\frac{\Gamma(3/\nu)}{\Gamma(\nu^{-1})}\right]\exp\left\{-\left[\frac{\Gamma(3/\nu)}{\Gamma(\nu^{-1})}\right]^{\nu/2}\left[\frac{|n_h|^\nu}{\sigma^\nu}\right]\right\},\ \nu \geq 0{,}5; \tag{13}$$

$$W(n_h|n_{h-1}) = \frac{\nu}{2\Gamma(\nu^{-1})\sigma}\left[\frac{\Gamma(3/\nu)}{(1-r^2)\Gamma(\nu^{-1})}\right]^{0.5}\exp\left\{-\left[\frac{\Gamma(3/\nu)}{(1-r^2)\Gamma(\nu^{-1})}\right]^{\nu/2}\left[\frac{|n_h - r n_{h-1}|^\nu}{\sigma^\nu}\right]\right\}, \tag{14}$$

where $\Gamma(.)$ is a gamma function.

Note that as special cases of (13) and (14), there follows a Gaussian distribution (9) and PDF of Laplace, consequently at $\nu = 2$ and $\nu = 1$.

Following the procedure outlined above, as a transitional approximating $W^A(n_h|n_{h-1})$ for PDF of the considered form we will have:

$W^A(n_h|n_{h-1}) = (2\pi G^2)^{-0.5}\exp\{-[n_h - \mathrm{M}(n_{h-1})]^2/G^2\}$,

where $\mathrm{M}(n_{h-1}) = n_{h-1} - 0{,}5G^2 Z_A(n_{h-1})$;

$$Z_A(n_{h-1}) = -\frac{d}{dn_{h-1}}\ln W(n_{h-1}) = \frac{\nu}{2^{0.5\nu}\sigma^\nu}|n_{h-1}|^{\nu-1}\mathrm{sgn}(n_{h-1}). \tag{15}$$

Let us consider the example of the construction of two-dimensional PDFs of non-Gaussian process in accordance with the methodology set forth, if there is known only the true one-dimensional PDF $W(n_{h-1})$ of the form (13).

In accordance with (8) let us define the transitional PDF. Following (15), we write the function $\mathrm{M}(n_{h-1})$:

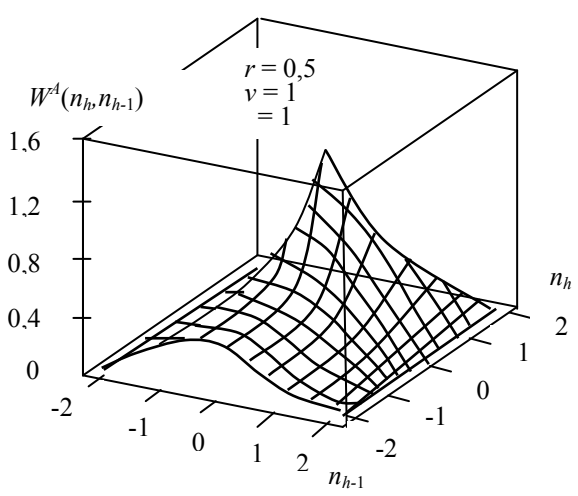
$$\mathrm{M}(n_{h-1}) = n_{h-1} - 0{,}5G^2\frac{d}{dn_{h-1}}\left\{-\left[\frac{\Gamma(3/\nu)}{\Gamma(\nu^{-1})}\right]^{\nu/2}\left[\frac{|n_{h-1}|^\nu}{\sigma^\nu}\right]\right\} = n_{h-1} - 0{,}5G^2\frac{\nu}{2^{0.5\nu}\sigma^\nu}|n_{h-1}|^{\nu-1}\mathrm{sgn}(n_{h-1}).$$

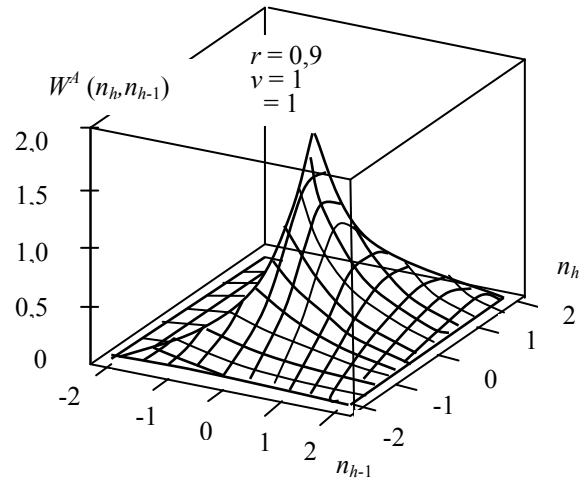
By introducing for the computational convenience the concept of equivalent correlation coefficient $r_э$, defined from the relation $G^2/\sigma^2 = 1 - r_э$, we finally write the expression of the transitional PDF:

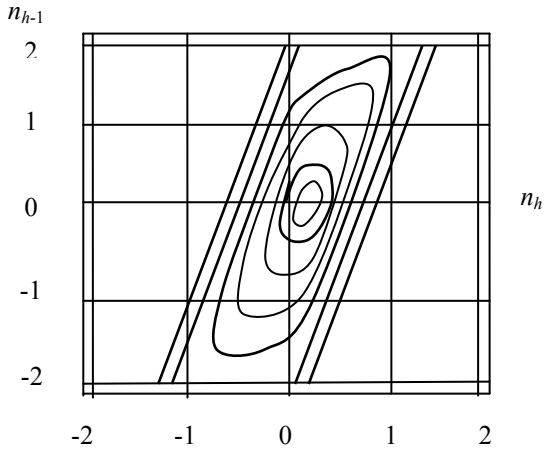
$$W^A(n_h|n_{h-1}) = [2\pi\sigma^2(1 - r_э)]^{-0.5}\exp\{-\frac{[n_h - n_{h-1} + 0.5\sigma^2(1 - r_э)z_э(n_{h-1})]^2}{\sigma^2(1 - r_э)}\}, \tag{16}$$

$$\text{where } z_э(n_{h-1}) = \frac{\nu}{2^{0.5\nu}\sigma^\nu}|n_{h-1}|^{\nu-1}\mathrm{sgn}(n_{h-1}).$$

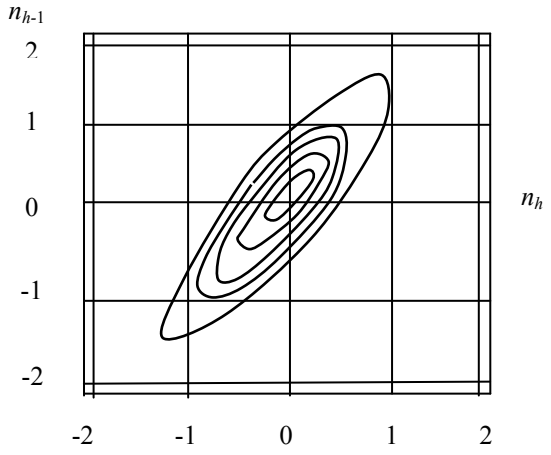
Results. The results of modeling the two-dimensional non-Gaussian PDFs in accordance with relations (7), (16) and (13) for different correlation coefficients $r_э$ and ν are shown in the Figure 1.


$W^A(n_h,n_{h-1})$
$r = 0,5$
$v = 1$
$= 1$
1,6
1,2
0,8
0,4
0
-2
-1
0
1
2
n_{h-1}
n_h

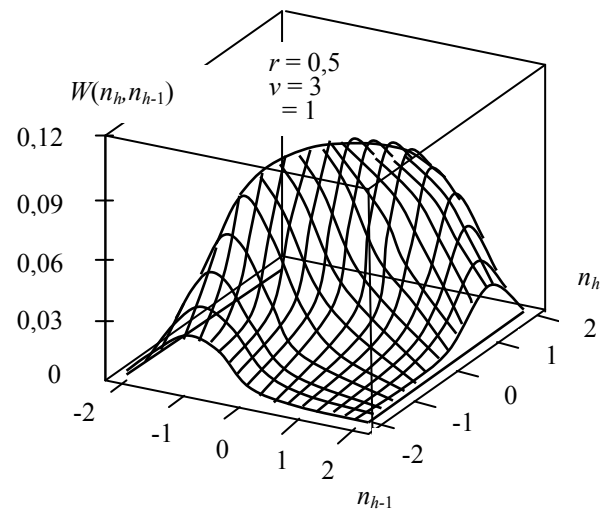

$W^A(n_h,n_{h-1})$
$r = 0,9$
$v = 1$
$= 1$
2,0
1,5
1,0
0,5
0
n_{h-1}
n_h

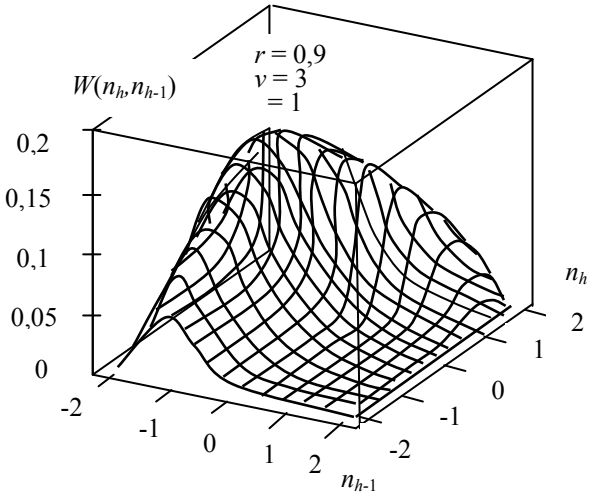

n_{h-1}
2
1
0
-1
-2
-2
-1
0
1
2
n_h

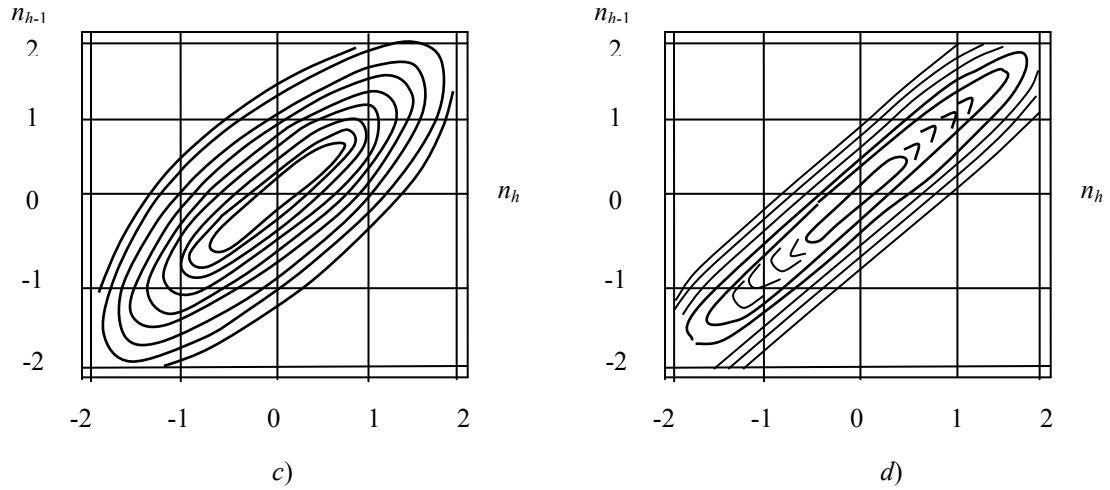
a)


n_{h-1}
2
1
0
-1
-2
-2
-1
0
1
2
n_h

b)


$W(n_h,n_{h-1})$
$r = 0,5$
$v = 3$
$= 1$
0,12
0,09
0,06
0,03
0
n_{h-1}
n_h


$W(n_h,n_{h-1})$
$r = 0,9$
$v = 3$
$= 1$
0,2
0,15
0,1
0,05
0
n_{h-1}
n_h

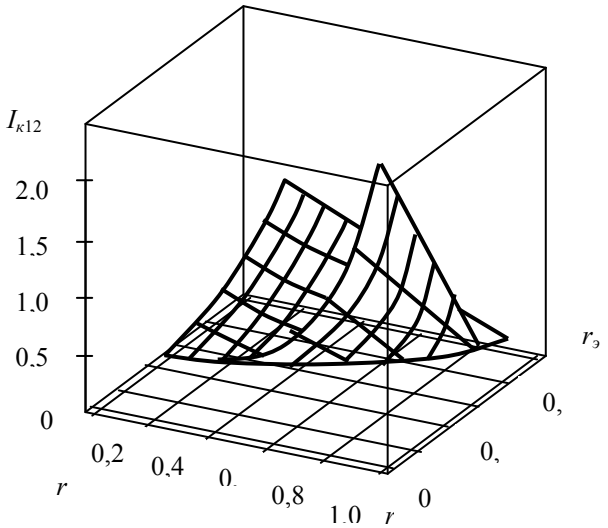


c) *d*)

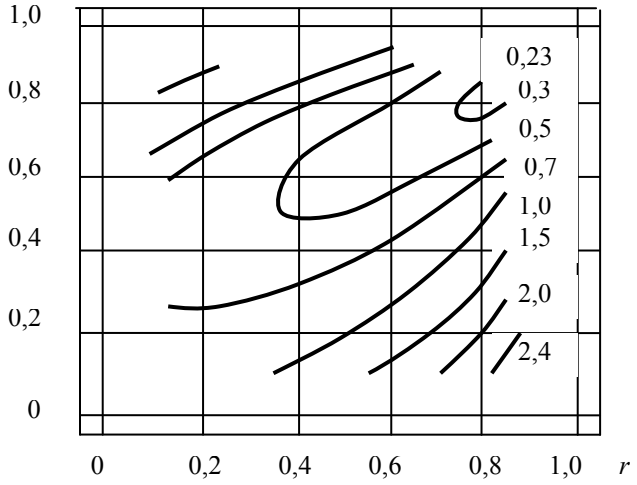
Figure 1. The results of modeling the two-dimensional non-Gaussian PDFs and isolines of their surfaces at: a – r = 0,5, ν = 1; b – r = 0,9, ν = 1; c – r = 0,5, ν = 3; d – r = 0,9, ν = 3

The analysis presented in PDF Fig. 1 shows that at large correlation coefficients the approximating $W^A(n_h, n_{h-1})$ and true $W(n_h, n_{h-1})$ PDFs are approaching each other. However, to accurately determine the degree of similarity of these distributions we use quantitative assessment of PDF similarity measure (10), (11), (1.28). Restrict ourselves to a particular case of PDF (14) – PDF of Laplace, which takes place at $\nu = 1$.

Graphs of dependancy I_{12K} and I_{21K} are shown in the Fig. 2*a* and Fig. 2*a,* respectively. Fig. 2*a* and Fig. 2*b* show the lines of equal level of displayed surfaces. The most informative surface, as seen from Fig. 2 and 3, is the surface $I_{21K}(n_h, n_{h-1})$, which illustrates that with increasing r and $r_э$ grows the degree of PDF closeness.

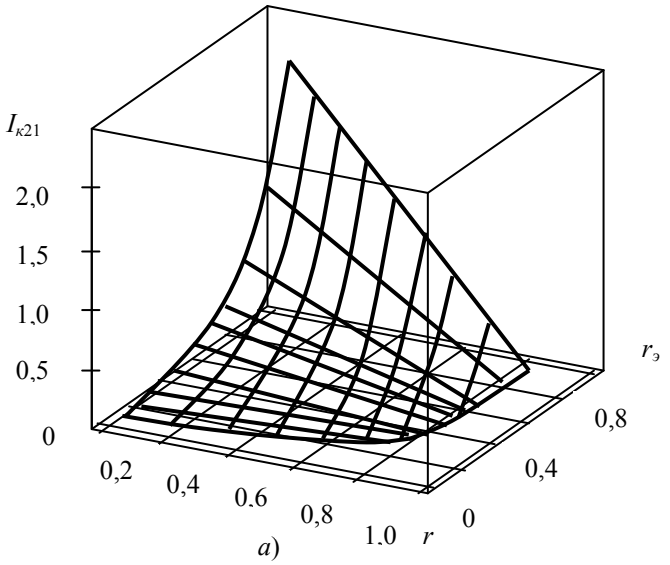


a)

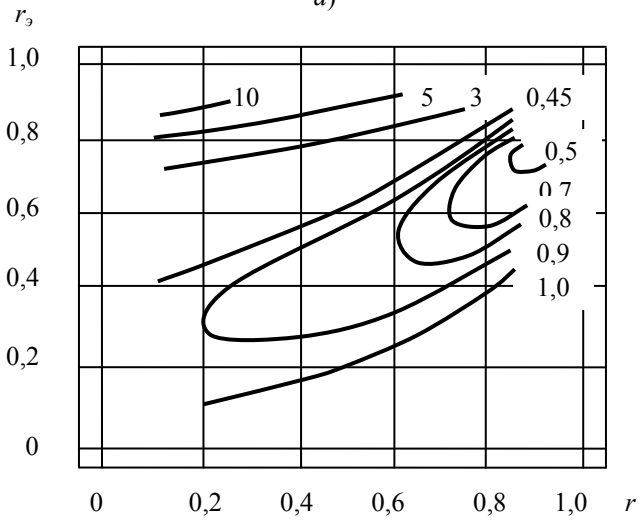


b)

Figure 2. Dependences Ik12 = f(r, rэ) and isolines of their surfaces



a)



b)

Figure 3. Dependencies Ik21 = f(r, rэ) and isolines of their surfaces

3. Discussion

Along with information on the Kullback, the wide practical importance, especially in problems of statistical synthesis of optimal algorithms for processing, received the information by Fisher, contained in a random process with PDFs $W(n)$ (Trofimov, 1986):

$$I_F = \left\{ m \left[\frac{dlnW(n)}{dn} \right]^2 \right\} = \int_{-\infty}^{\infty} \left[\frac{dlnW(n)}{dn} \right] W(n) dn.$$

It can be shown that in the case of one-dimensional PDFs, we have the identity: $m\left\{ \frac{d^2 lnW(n)}{dn^2} \right\} = -m\left\{ \left[\frac{dlnW(n)}{dn} \right]^2 \right\}$.

In case the random process is defined by the two-dimensional $W(n_1, n_2)$ or conditional $W(n_1|n_2)$ PDF by analogy with I_F, there is introduced the concept of the Fisher information matrix I_F with the elements:

$$I_{F.ij} = m\left\{ \frac{dlnW(n_1|n_2)}{dn_i} \frac{dlnW(n_1|n_2)}{dn_j} \right\} = \iint_{-\infty}^{\infty} \frac{d^2 lnW(n_1|n_2)}{dn_i dn_j} lnW(n_1, n_2) dn_1 dn_2 \; ; i, j = 1, 2.$$

It is supposed that the matrix $\|I_\Phi\|$ positively defined, i. e. $\det I_F \neq 0$.

In the particular case of a Gaussian random process $n(t)$, defined transition distribution, the Fisher information matrix obtains the form of:

$$\|I_{\Phi.\Gamma}\| = \begin{Vmatrix} I_{\Phi.11} & I_{\Phi.12} \\ I_{\Phi.21} & I_{\Phi.22} \end{Vmatrix} = [\sigma_n^2 (1 - r_n^2)]^{-1} \begin{Vmatrix} 1 & -r_n \\ -r_n & r_n^2 \end{Vmatrix},$$

where σ_n^2 is a variance, and r_n is a correlation coefficient of the random process $n(t)$.

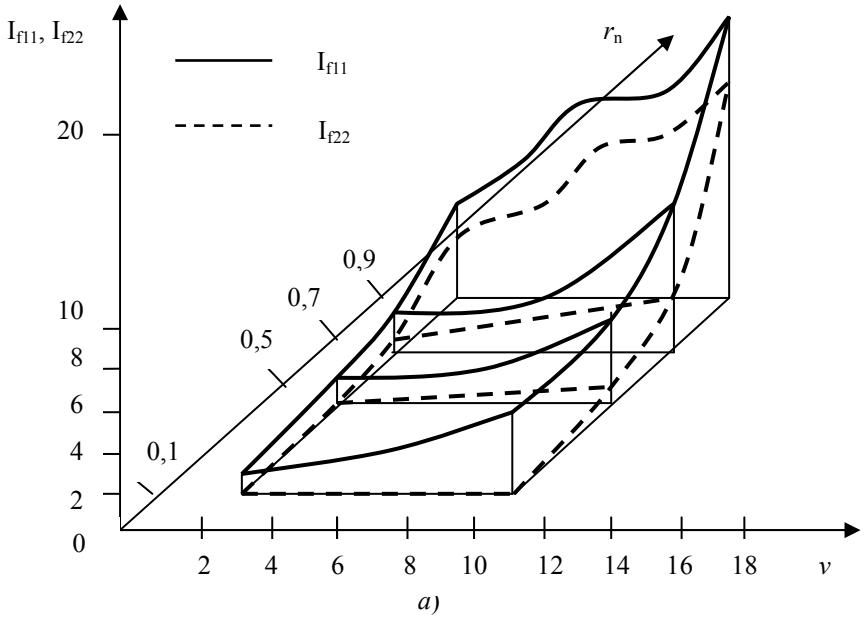
From a comparison of the member $I_{F.11}$ and I_F for PDF with independent values it follows that the factor at the right part $I_{F.11}$ at $r_n = 0$ coincides with I_F for the Gaussian PDF. While $0 \leq r_n^2 < 1$, it becomes clear that existence of a finite correlation between the values $n(t)$ results in an increase of Fischer information compared with the case of one-dimensional PDF (Artyushenko and Samarov, 2013).

In general, the definition of the information matrix for non-Gaussian correlation processes encounters considerable difficulties and, as a rule, cannot be obtained analytically. Only in some cases, the solution is possible to obtain analytically. For example, for the PDFs (14) information matrix takes the form:

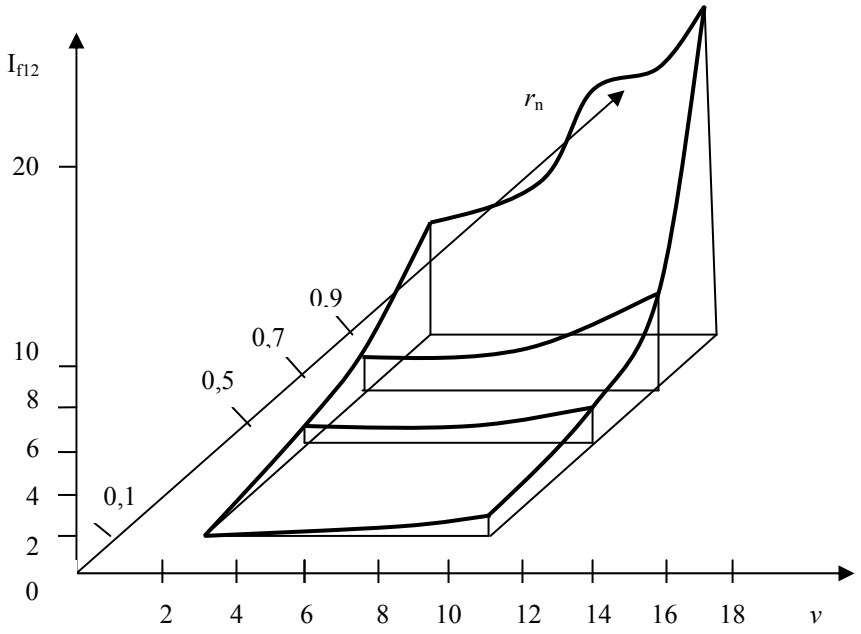
$$\|I_\Phi\| = \frac{\nu(\nu-1)\Gamma(3/\nu)\Gamma(1-\nu^{-1})}{\sigma_n^2 \Gamma^2(1/\nu)(1-r_n^2)} \begin{Vmatrix} 1 & -r_n \\ -r_n & r_n^2 \end{Vmatrix} = A(\nu) I_{F.\Gamma},$$

where $A(\nu) = \frac{\nu(\nu-1)\Gamma(3/\nu)\Gamma(1-\nu^{-1})}{\Gamma^2(1/\nu)}$; $\nu \geq 2$ – constant depending on the parameter of distribution.

Dependencies of matrix elements $I_{F.ij}$ on the PDF parameters are shown in Figure 4.



a)



b)

Figure 4. Dependence of matrix elements I_{fij} on the parameters of distribution: *a* – I_{f11}, I_{f22}; *b* – I_{f12}

Similary to the quasi-Shenon information there is introduced the concept of quasi-Fisher information:

$$\int_{-\infty}^{\infty} W_0(n)\left(\frac{W'(n)}{W(n)}\right)^2 dn = \int_{-\infty}^{\infty} W_0(n)\left(\frac{dlnW(n)}{dn}\right)^2 dn = m\left\{\left(\frac{W'(n)}{W(n)}\right)^2\right\} = I_\Phi(W, W_0).$$

The value $I_\Phi(W, W_0)$ is used in particular for the evaluating the effectiveness of the algorithms of asymptotically optimal reception.

Along with the introduced notions of Fisher information in assessing the characteristics of unilateral characteristic PDFs, for example, to describe the random amplitudes of narrowband radio signals, there is used Fisher dispersion information (dinformation) (Tsushkin, 1984; Artyushenko, Volivach and Samarov, 2014; Artyushenko and Volivach, 2013):

$$I_{F.D}(W) = m\left\{n^2\left(\frac{W'(n)}{W(n)}\right)^2\right\} = \int_{-\infty}^{\infty} n^2\left(\frac{dlnW(n)}{dn}\right)^2 dn.$$

Note that for the same PDFs the following inequality is performed: $I_{F.D}(W) > I_F(W)$.

The theoretical results presented in yielded work have been used by working out and creation of a multipurpose measuring instrument (MMI) of parameters of movement of extended objects, that have past successful tests in actual practice [2]. In its structure the block constructing a multiplicate hindrance has been used, allowed considerably to raise accuracy of work of a measuring instrument. For definition of efficiency again created MMI and existing measuring instruments the experimental researches providing check of their working capacity in a dynamic mode at measurement of speed of movement of extended objects with acceleration to 3 m/s2 in a mode of deep stand by regime of Doppler amplitude of a signal (to 80 % from duration of realisations) have been spent. As the examinee of the sample of a radar-tracking measuring instrument speeds (RTMI) was taken a serially produced measuring instrument RTMI-B3. On Figure 5 in uniform systems of co-ordinates schedules norm of errors of measuring instruments are presented $\delta_o = f(t_{зам}T_м^{-1}; F_д)$, where $t_{зам}T_м^{-1}$– relative time of a stand by regime of a processed signal (PS); $T_м = F_м^{-1}$, $F_м$ – frequency of the signal modulated on amplitude at 100 %-s' peak modulations, $F_д$ – frequency (PS) (Doppler frequency), for cases $F_м = 0{,}1F_д$.

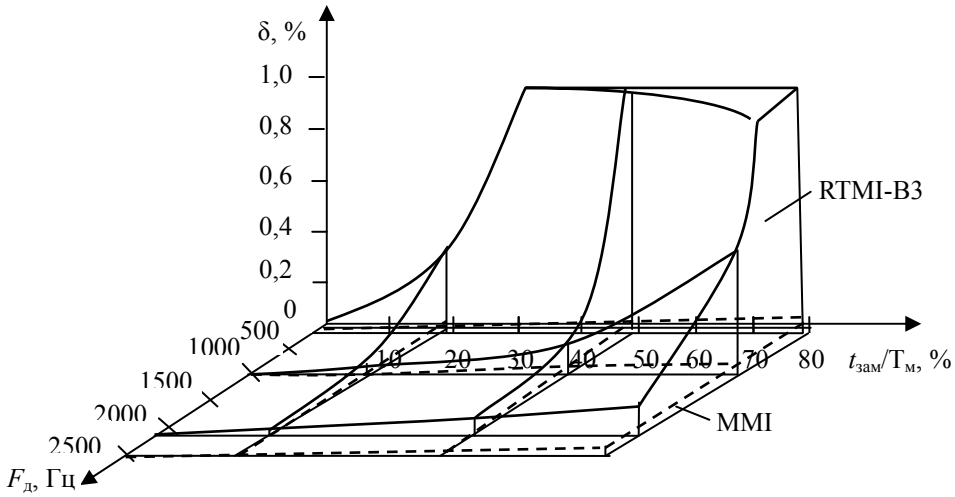


Figure 5. Comparative characteristics norm of errors of measuring instruments MMI and RTMI-B3

Errors are led to error size of RTMI-B3, where $\delta_o = \delta/\delta_{RTMI\text{-}B3}$, where $\delta_{RTMI\text{-}B3} = 100\%$; $\delta = [|F_{д.ф} - F_{д.и}|/F_{д.ф}]100\%$, $F_{д.ф}$и $F_{д.и}$ – the actual and measured frequency PS.

Apparently from the presented dependences, under identical conditions of measurement of PS, size of relative errors is ruther less, than at RTMI-B3. Fuller results of comparative researches of the created measuring instrument are stated in [2]. Calculation of an estimation of economic efficiency is resulted in (17).

4. Conclusion

Thus, there are considered and analyzed mathematical models of multiplicative and additive non-Gaussian noise affecting the useful signals. To perform the synthesis of radio systems and devices, there are introduced elliptical symmetric PDFs, allowing to describe not only narrowband correlated additive noise, but interference with multiplicative (modulating) character.

There is proposed the transitional PDF allowing to design two-dimensional PDFs of correlated non-Gaussian noises. There is shown the adequacy of designed with its help PDFs of two-dimensional real PDFs affecting correlated noise.

There are introduced the information characteristics of non-Gaussian additive and multiplicative noises.

It is shown that in the general case, the definition of information matrix for non-Gaussian correlation processes encounters considerable difficulties and, as a rule, cannot be obtained analytically.

In the future, the considerable interest arises to investigate the possibility of using the Fisher dispersion information to describe the statistical characteristics of multiplicative non-Gaussian noise, significantly affecting the quality of the information processed upon receipt.

Experimental researches created with use of the theoretical results presented in article, a multipurpose measuring instrument of parametres of movement of extended objects, show that an error of measurement at it much more lower than at similar serially produced measuring instruments.

References

Artushenko, V. M., & Volovach, V. I. (2012). Estimation of economic efficiency of work of a radar-tracking measuring instrument of speed. The bulletin of Volga region state university of service. *A series Economy, 6*(26), 182-192.

Artyushenko, V. M. (2013). Research and development of radar meter for motion parameters of extended objects. M., pp: 214.

Artyushenko, V. M., & Abbasova, T. S. (2011). Designing of multiservice systems under the influence of external electromagnetic noises. M., 264.

Artyushenko, V. M., & Samarov, K. L. (2013). Construction of two-dimensional correlated additive and multiplicative models of non-Gaussian noise. *Electrical and information complexes and systems, 4*(10), 83–92.

Artyushenko, V. M., & Solenov, V. I. (1993). Elliptically symmetric models of non-Gaussian noise. Kiev, pp. 4–8.

Artyushenko, V. M., & Volivach, V. I. (2013). Choice and grounding of the models of measured parametres of extended objects motion. School of the university science: *Development paradigm, 3*(10), 109–113.

Artyushenko, V. M., & Volovach, V. I. (2013). Statistical Characteristics of Envelope Outliers Duration of non-Gaussian Information Processes. Proceedings of IEEE East-West Design & Test Symposium (EWDTS'2013). Rostov-on-Don, Russia, September 27–30: 137–140.

Artyushenko, V. M., Volivach, V. I., & Samarov, K. L. (2014). Construction of two-dimensional correlated models of non-Gaussian noise. *Scientific and Technical Herald of the Volga region, 1*, 53–56.

Bucy, R. S., & Mallinckrodt, A. J. (1973). An Optimal Phase Demodulato. *Stochastics, 1.*

Edward, J., & Wegman, G. (1983). Tohic in Non-Gaussian Signal Processings. New-York: 235.

McGraw, D. K., & Wagner, J. F. (1968). Elliptially Symmetric Distributions. *IEEE Trans, 1*(IT-14), 76–84.

Middleton, D. (1973). Man-madenjise in urban environments and measuzements. *Trans. ITTT, 9*(AES-9), 371–376.

Tikhonov, V. I., & Mironov, M. A. (1977). Markov processes. MA., p. 488.

Trofimov, A. T. (1986). Estimation of nuisance parameters for the adaptive signal processing based on the usage of multi-Gaussian noise model. *Radio Engineering and Electronics, 11*(31), 2151–2159.

Tsushkin, Y. Z. (1984). Fundamentals of information theory of identification. M., p. 320.

Tzay, Y. Y., & Westerberg, R. A. (1971). Error bounds for stochastic estimation of signal parameters. IEEE Trans., IT-17, 549–557. http://dx.doi.org/10.1109/TIT.1971.1054696

Yarlykov, M. S. (1980). Application of Markov theory of nonlinear filtering in radio engineering.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Analysis of Rice Farming with System of Seeding Direct and Seeding Indirect: A Case Study in Buol Regency Indonesia

Max Nur Alam[1]

[1]Department of Agriculture Economics, Tadulako University, Indonesia

Corresponding author: Max Nur Alam, Department of Agriculture Economics, Faculty of Agriculture, Tadulako University, Indonesia.E-mail: nuralam_max@yahoo.co.id

Received: December16, 2014 Accepted: January 6, 2015 Online Published: July 27, 2015

doi:10.5539/mas.v9n8p311 URL: http://dx.doi.org/10.5539/mas.v9n8p311

Abstract

This research aimed to analyze production and income of rice farming with system of seeding direct and seeding indirect with a variety of input usage. This research was conducted with descriptive and comparative methods that described and compared rice farming with system of seeding direct and seeding indirect. The number of samples that were taken amounted 74 household heads. Samples on seeding direct system were taken by census, amounted 34 household heads. Seeding indirect system was determined by 40 household heads (20% of 197 farmers of seeding indirect system), sampling was done by simple random. The result of research showed that rice farming with seeding direct system could save labors when compared with seeding indirect system. The income average of rice farming with seeding direct system were bigger than seeding indirect system. Rice farming with seeding direct system had a prospect to be developed on irrigated rice land with labors that were rare and expensive. However, the application of this system needed to be balanced with the usage of fertilizers, herbicides and reapers, so that labors efficiency could occur.

Keywords: Production, Income, Rice, Seeding direct, Seeding indirect

1. Introduction

The Increasing of rice production in Indonesia was conducted through intensification and extensification program. Intensification is conducted by improving technology to increase land productivity. Extensification is intended to expand the production area (Soekartawi, 2003).

Areas that are not enable to do farming expansion, then one of business that is able to increase and improve the welfare is by increasing the carrying capacity of land. Technology of seeding direct system and seeding indirect system had been developing to support the carrying capacity of land (Adnyana, 1996). Seeding indirect system generally is conducted by farmers if the labors are many available and cheap. Areas with rare and expensive labors, while the machine price of seeding indirect system is not affordable by farmers, seeding direct system can be an alternative for farmers. The scarcity of labors can cause the cropping time delayed, so the rice production are low, then the seeding direct system is introduced (Rao et al. 2007). The usage objective of seeding direct system is to reduce the usage of labors that are concentrated at the same time such as land cultivation and cropping, as well as to avoid the creation and maintenance of seedbed. Labors efficiency can reduce the cost of labors which are expensive and pursue cropping season simultaneously with cost that are relatively low. Seeding direct system with a optimum management can give the higher crop yields (De Datta and Flinn, 1986; Washio, 1992).

The general objective of this research was to analyze the rice farming with system of seeding direct and seeding indirect. In particular, this research aimed to analyze production and income of rice farming with system of seeding direct and seeding indirect with a variety of input usage.

2. Research Methods

This research was conducted with descriptive and comparative methods that described and compared rice farming with system of seeding direct and seeding indirect in Tiloan Sub-regency Buol Regency, Indonesia. Tiloan Sub-regency was an area that had potential in increasing of rice farming production (intensification). There were 197 Household Heads (HH) that applied seeding indirect system and 34 HH applied seeding direct system. This research was conducted on September 2012 to November 2012.

The number of samples that were taken amounted 74 HH. Samples on seeding direct system were taken by census, amounted 34 HH. Seeding indirect system was determined by 40 HH (20% of 197 farmers of seeding indirect system), sampling was done by simple random. Descriptive method and independent t-test statistics were used to analyze production and income of rice farming with system of seeding direct and seeding indirect with a variety of input usage.

3. Results and Discussion

3.1 Input Usage on Seeding Direct System and Seeding Indirect System

3.1.1 The Usage of Seeds

Seeds are one of factors that determined the success in rice farming. The usage of seeds that are done proportionally with land area, and noticed the good quality would give good growth and high production. The result of observation and direct interview with farmers were known that the usage of seeds by respondents were very varies when seen by source, amount and term of seeds quality that used.

Rice seeds that often used were Cisantana seeds. Cisantana seeds were seeds that produced by cultivator farmers, which taken from their products. The number usage of rice seeds by farmers were varies respondent (Table 1).

Table 1. The usage average of seeds against rice production

Numberof seeds (kg)	Average ofrice production (kg)	
	Seeding direct system	Seeding indirect system
15	2,186.67	1,445.11
20	-	1,446.89
25	1,149.24	1,788.44
30	1,050.24	1,800.42
35	1,655.51	1,952.52
40	2,261.61	1,523.84
50	1,831.68	3,220.43
>50	2,888.51	-
Average/ha	1,826.28	1,748.21

The usage average of rice seeds on seeding direct system amounted 51.82 kg / ha and usage average of rice seeds on seeding indirect system amounted 27.17 kg / ha. This difference was due to the seeding indirect system used seeding. On average, seeding direct system produced higher rice production, especially in the usage of seeds were more than 50 kg / ha (Ameen et al. 2014). The density of seeds which are high can be considered as a weed management strategy (Juraimi et al. 2013; Ikeda et al. 2008). Besides the density of seeds, seeds types also affect the rice growth on seeding direct system (Islam et al. 2014).

3.1.2 The Usage of Fertilizers

Fertilizers are one of production factors that can increase the fertility and improve the growth of crops so that can increase crops production. Types of fertilizers that were used by rice farmers of seeding direct system and seeding indirect system were Urea, ZA, KCl, SP-36, and PPC (liquid complement fertilizer). The usage of fertilizers by farmers based on land area are shown in Table 2.

Table 2. The usage average of fertilizers in rice farming

System	The usage average of fertilizers				
	Urea (kg/ha)	ZA (kg/ha)	SP-36 (kg/ha)	KCl (kg/ha)	PPC (l/ha)
Seeding direct	128.97	36.41	53.10	49.25	4.95
Seeding indirect	125.00	28.06	54.85	49.11	2.14

The usage average of fertilizers for rice with system of seeding direct and seeding indirect were relatively same. The usage of Urea fertilizers were higher when compared to other types of fertilizers. Farming with seeding direct system, crops grew directly from seeds while with seeding indirect system, crops were from seeds that aged 21 days. Rice of seeding direct system were longer in rice land than rice of seeding indirect system, therefore Washio (1992) argued that rice fertilizers dosage of seeding direct system, especially N, should be

20-30% higher than rice fertilizers of seeding indirect system. The giving of N fertilizers were reduced in the half of first growth, but were increased in primordia stage and panicle establishment. According to Qi et al. (2012), the giving of N fertilizers have to be given on a regular basis, because it is related with the evaporation of ammonia that can inhibit the growth of rice crops.

3.1.3The Usage of Pesticides

Pests attack and diseases are one of factors that can cause the decreasing of production level on a farming, so that pests attack and diseases need to be controlled by using pesticides (Singh et al. 2008; Mahajan et al. 2009). Types of pesticides that were often used by rice farmers with seeding direct system and seeding indirect system namely DMA, Dharmabas, Dursban, and Clipper. Types and quantities of pesticides that used by farmers were very dependent on land area and state of rice crops against pests and crops diseases (Table 3).

Table 3.The usage average of pesticides in rice farming

System	The usage average of pesticides (l/ha)			
	DMA	Dharmabas	Dursban	Clipper
Seeding direct	0.88	1.37	0.56	0.35
seeding indirect	0.78	0.84	0.51	0.37

The usage of pesticides by farmers have to notice the problem of environmental sustainability and apply wisely. Overspray can cause pollution on environment, especially on water and turn off non-target organisms. The usage of various technologies of weed control that support each other and integrated are highly recommended. The integrated approach are recommended for control of weed sustainable om seeding direct system, such as the usage of certified seeds, cultivation of land that is proper, irrigation water management, crops rotation, seeding density that is higher, as well as the proper usage of pesticides are technologies that compatible each other and synergistic to support the weed control effectively (Juraimi et al. 2013; Khaliq and Matloob, 2011; Akbar et al. 2011; Khaliq et al. 2013).

3.1.4 The Usage of Labors

Labors are one of production factors that support the success of farming. The usage of labors that are effective as well as have abilities and sufficient skills tend to increase the production of farming. The usage of labors are very dependent on the types of work that contain in each farming. The labors that used by farmers were labors from within the family and outside the family. Types of farming activities in rice farming with seeding direct system included land cultivation, cropping, weeding, fertilizing, control of pests and crops diseases, harvest, transporting crops and drying while rice farming with seeding indirect system included seeding, land cultivation, cropping, weeding, fertilizing, HPT control, harvest, transporting crops, and drying. The usage average of labors in rice farming with seeding direct system were 75.22 HH / ha. The usage average of labors in rice farming with seeding indirect system were 83.74 HH / ha. It showed the usage of labors in seeding indirect system were higher when compared with seeding direct system (Bhushan et al. 2007).

Seeding direct system can be adopted by farmers, especially in production areas of rice with rare and expensive labors. Sowing seeds in the array can use a tool that is called "atabela" (direct seeding tool). With this system, the outpouring of labors to plant rice only 1-2 persons / ha. If rice farming with seeding direct system is developed, then the reapers need to be available on the farmers level. If the reapers are not available, the costs of harvest will still be expensive so the efficiency of production costs are not achieved.

3.2 Analysis of Rice Farming with System of Seeding Direct and Seeding Indirect

Analysis of farming is a way to calculate the amount of income in a farming. Income of rice farmers with system of seeding direct and seeding indirect could be known by calculating the difference between total revenue and total cost.

The income average of rice farming with system of seeding direct and seeding indirect are shown on Table 4.

Table 4. Income of rice farming

No.	Analysis	Rice Farming with Seeding Direct System / ha / season	Rice Farming with Seeding Indirect System / ha / season
1	Rice Production (kg)	1,826.28	1,748.22
2	Rice Costs / kg (IDR)	7,000.00	7,000.00
3	Revenue (IDR)	12,783,960.00	12,237,540.00
4	Costs:		
	a. Fixed Costs		
	Shrinkage (IDR)	79,627.91	85,084.75
	RentLand (IDR)	800,000.00	800,000.00
	Others (IDR)	288,953.49	268,813.56
5	Sub Total a (IDR)	1,168,581.40	1,153,898.31
	b.Variablecosts		
	Labors Costs (IDR)	2,632,531.98	2,933,474.58
	SeedsCosts (IDR)	206,511.63	107,796.61
	FertilizersCosts (IDR)	572,988.37	512,457.63
	Pesticides Costs (IDR)	312,593.02	270,406.78
6	Sub Total b (IDR)	3,724,625.00	3,824,135.59
7	Total Costs(5 + 6)	4,893,206.40	4,978,033.90
8	Income (3 - 7)	7,890,753.60	7,259,506.10

The statistical analysis result of two independent samples t test against income in rice farming with seeding direct system and rice farming with seeding indirect system showed the probability value 0.001 < 0.05 at 95% confidence level of two-tail test. It showed rejecting the null hypothesis that meant there was a significant difference between income of rice farming with seeding direct system and rice farming with seeding indirect system.

Table 4 shows that income of rice farming with seeding direct system are higher when compared with income of rice farming with seeding indirect system. The difference of income in this research was caused by the difference of rice production on seeding direct system and seeding indirect system, where rice production on seeding direct system were higher than rice production on seeding indirect system. The difference in this production tended was caused by the usage of production factors that had different amount, such as seeds, fertilizers and pesticides. It is accordance with the opinion of Supriadi and Malian (1993), that the seeding direct system technique by using the cropping tool do not need seedbed such as in seeding indirect system technique. Seeds were soaked directly for 24 hours then dried for 12-14 hours and planted directly in the array. The seeds that were used around 51.82 kg / ha while seeding indirect system only 27.17 kg / ha. The number of different seed then the population number of rice would also be different so it tended increasing the production of rice per hectare. Likewise, the usage of fertilizers more tended increasing rice production. The usage of pesticides tended suppressing the production loss caused by pests attack and diseases in rice crops with seeding direct system.

The income difference was also caused by the total costs of rice production with seeding direct system were lower when compared to the costs of rice production with seeding indirect system. The difference of total costs was caused by variable costs of rice production with seeding indirect system were bigger, where in rice crops with seeding indirect system used more labors. It was accordance with the opinion of Adnyana (1996), that when compared between the usage of seeding indirect system with with seeding direct system, then seeding direct system would give some advantages, among others:

1. Labor costs outside harvest were lower 25-30 percent
2. Production facilities costs were lower 5-10 percent.
3. Results per hectare were 10-25 percent higher and the grain costs as well as rice were higher (due to better quality).
4. The net income of farmers increased from IDR 1.2-1.5 million / ha / season became IDR 2.0-2.5 million / ha / season.

The income difference showed that rice farmers in areas which are rare and expensive labors recommended for planting in rice with seeding direct system because more profitable if compared with seeding indirect system.

Farming with seeding direct system then reapers need to be available on the farmers level. If the reapers are not available, the costs of harvest will still be expensive so the efficiency of production costs are not achieved.

4. Conclusion

Farming with seeding direct system could save labor when compared with farming with seeding indirect system. The income average of rice farming with seeding direct system were bigger than seeding indirect system. This difference was due to the rice production with seeding direct system were higher so that the total acceptance on seeding direct system were higher and total costs on seeding direct system were lower if compared with seeding indirect system. Rice farming with seeding direct system had a prospect to be developed in irrigated land with rare and expensive labor. However, the application of this system needed to be balanced with the usage of fertilizers, herbicides and reapers so the labor efficiency could occur.

Acknowledgements

The author would like to thank Prof. Dr. Ir. Muhammad Basir, SE., MS. and Prof. Dr. Ir. Made Antara, MP., who had guided and directed the author. The author also liked to thank the reviewers so this paper could be published.

References

Adnyana, M.O.(1996).ProspekdanKendalaPengembanganTehnikTanam Benih Langsung (Tabela) dalam SUTPA, dalam Made Oka Adnyana dkk (eds) monographSeriesNo. 19: PengkajianSUTPA: Konsep, Keragaan Empiris dan Prospek.

Akbar, N.,Ehsanullah, Jabran, K., &Ali, M. A. (2011). Weed management improves yield and quality of direct seeded rice.*Australian Journal of Crop Science,5(*6), 688 - 694.

Ameen, A.,Aslam, Z.,Uz Zaman,Q.,Ehsanullah,Zamir, S. I., Khan, I., &Subhani, M.J. (2014). Performance of Different Cultivars in Direct Seeded Rice (*Oryza sativa* L.) with Various Seeding Densities. *American Journal of Plant Sciences*,5, 3119-3128. http://dx.doi.org/10.4236/ajps.2014.521328

Bhushan, L., Ladha,Jagdish, K.,Gupta, Raj, K.,Singh, S.,Tirol-Padre, A.,. Saharawat, Y. S, Gathala, M., &Pathak, H. (2007). Saving of Water and Labor in a Rice-Wheat System with No-Tillage and Direct Seeding Technologies.*Agronomy Journal, 99*(5), 1288 - 1296.

De Datta, S.K., & Flinn, J.C. (1986). Technology and economics of weed control in broadcast seeded flooded tropical rice. *Asian Pacific Weed Science Society Conference, 10*, 51−74.

Ikeda, H., Kamoshita, A., Yamagishi, J., Ouk, M., & Lor, B. (2008). Assessment of management of direct seeded rice production under different water conditions in Cambodia. *Paddy Water Environ*, 6, 91 – 103. http://dx.doi.org/10.1007/s10333-007-0103-9.

Islam, M. K., Islam, Md. S., Biswa, J. K., Lee, Si. Y., Alam, I., & Huh, M. R. (2014). Screening of rice varieties for direct seeding method.*Australian Journal of Crop Science, 8*(4), 536 - 542.

Juraimi, A. S., Uddin, Md. K., Anwar, Md. P., Mohamed, M. T. M., Ismail, M. R., & Man, A. (2013). Sustainable weed management in direct seeded rice culture: A review. *Australian Journal of Crop Science,7*(7), 989-1002.

Khaliq, A., & Matloob, A. (2011). Weed-crop competition period in three fine rice Cultivars under direct-seeded rice culture. *Pak. J. Weed Sci. Res., 17*(3), 229 - 243.

Khaliq, A., Matloob, A., Ihsan, M. Z., Abbas, R. N., Aslam, Z., & Rasool, F. (2013). Supplementing Herbicides with Manual Weeding Improves Weed Control Efficiency, Growth and Yield of Direct Seeded Rice.*International Journal of Agriculture and Biology, 15*(2), 191 - 199.

Mahajan, G., Chauhan, B.S., & Johnson, D.E. (2009). Weed management in aerobic rice in Northwestern Indo-Gangetic Plains. *Journal of Crop Improvement, 23*, 366-382.

Qi, X, Wu, W., Peng, S., Shah, F., Huang, J., Cui, K., Liu, H., &Nie, L. (2012). Improvement of early seedling growth of dry direct-seeded rice by urease inhibitors application. *Australian Journal of Crop Science, 6*(3), 525 - 531.

Rao, A.N., Jhonson, D.E., Sivaprasad, B., Ladha, J.K., & Mortimer, A.M. (2007). Weed management in direct seeded rice. *Advances in Agronomy, 93*, 153 - 255.

Singh, S., Ladha, J.K., Gupta, R.K., Bushan, L., & Rao, A.N. (2008). Weed management in aerobic rice systems under varying establishment methods. *Crop Protection, 27*, 660 - 671.

Soekartawi(2003).*Teori Ekonomi Produksi*. PT. Raja Grafindo Persada. Jakarta.

Supriadi, H., & Malian, A.H. (1993). Kelayakan Agronomis Tekhnologi Budidaya SebarLangsungdiLahanSawahIrigasi.Dalam Muhyidin Syamdkk (eds),*ProsidingSimposium*III ;KinerjaPenelitianTanamanPangan. Buku 3, Puslitbang Tanaman Pangan Bogor.

Washio, O. (1992). Direct seeding rice culture in Japan: Its technical outlook. *Farming Japan, 26*(1), 11−19.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Statistical Data Processing in Rocket-Space Technology

Bulat-Batyr Saukhymovich Yesmagambetov[1], Zhambyl Talkhauly Ajmenov[1], Alexander Mikhailovich Inkov[1], Abdushukur Satybaldievich Saribayev[1] & Serik Umirbayevich Ismailov[1]

[1] M. Auezov South Kazakhstan State University, Kazakhstan

Correspondence: Bulat-Batyr Saukhymovich Yesmagambetov, M. Auezov South Kazakhstan State University, 160018, Shymkent, Tauke-Khana Street, 5, Kazakhstan. Tel: 7701-460-7116. E-mail: bulatbatyr@mail.ru

Received: December 3, 2014 Accepted: January 9, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p317 URL: http://dx.doi.org/10.5539/mas.v9n8p317

Abstract

The most measured data in the rocket-and-space technics is a broadband non-stationary random process. Using traditional methods of cyclic digitalization at processing of such data leads to heavy computational costs and heavy amounts of memory of on-board computing devices. As a rule, in processing of discontinuous processes, the recovery of original implementation on the receiving side is not required and processing involves calculation of probabilistic characteristics. In this case processing of data in the rocket-and-space technics has a number of features. First, random processes always represented by the sole implementation. Secondly, priori knowledge of measured random process probabilistic properties is not possible. Thirdly, there is need in on-line processing that predetermines use of rapid processing methods. The article considers application possibilities of nonparametric solution theory methods for evaluation of probabilistic properties of non-stationary broadband random processes in the rocket-and-space technics on-board systems.

Keywords: Rocket-and-space telemetry, random process, data compaction, non-parametric methods, order statistic, stationary hypothesis

1. Introduction

1.1 Introduce the Problem

At the rocket and space technology always has been a pressing problem transferring to Earth in real-time telemetry data rate (Mamchev, 2007). This problem buys a special urgency by transmission to the land of parameters of propulsion systems, such as temperature and combustion chamber pressure, a fuel rate. The knowledge of these parameters in real rate of time allows making operatively a solution at origin of emergencies in propulsion systems.

1.2 Explore Importance of the Problem

Physical parameters in propulsion systems (for example, temperature and combustion-chamber pressure), always look like broad-band no stationary random process with a breadth of a spectrum 2÷32 kilohertz, i.e. represent a non stationary trend from the broad-band random component superimposed on it. However, in the modern rocket-engineering systems for measurement of these parameters often use low-frequency transmitters which on an exit give not the random determined signal (trend) (Voronov, 2011). Thus, the random component of physical process is filtrated. Such mode of measurement and handling of parameters is applied for the purpose of abbreviation of computing costs and decrease of demands to radio engineering communication channels as frequency of an initial cyclical digitization Δt depends on a signal spectrum ΔF (Novikov, 2003). In practice, this relation is equal

$$\Delta t = \frac{1}{(5 \div 10)\Delta F}$$

Here (5÷10) means that the coefficient gets out in the range from 5 to 10 depending on the demanded restoration accuracy. From the ratio given above it is visible that measurement and processing of broadband signals requires increase in speed of onboard computing means and their memory sizes in tens times, and at data transmission it is necessary to increase by Earth in ten times the capacity of communication channels. Meanwhile, there is a practical need of reduction of characteristics of onboard computing means that it is connected with requirements of reduction of their mass-dimensional characteristics and energy consumption.

Representation of physical parameters in the form of a no stationary low-frequency trend is justified, if space vehicle flight passes in a regular condition. However, the history of assimilation of an outer space knows many facts of accidents of space vehicles. For example, accidents of the American orbiting space crafts of reusable use Space Shuttle, the Russian launchers «Proton» in the end of past century happened because of uninominal situations in propulsion systems. In addition, if in the second case business was limited to loss of satellites, in the first case numerous loss of human life took place. In our opinion, measurement of data on physical parameters in the combustion chamber of motive systems in full, i.e. measurement of a non-stationary broadband signal, and not just trend, has paramount importance. It is connected with that the casual component starts reacting to the coming nearer emergency situation much earlier, than a trend. Thus, the random component is as though a forerunner of an uninominal situation. In addition, its measurement and transmission to the Earth in real rate of time allows making operative solutions on the Earth, in particular to generate a pilot signal on a cut-off of drives and not to suppose accidents.

Thus, the problem of working out of accelerated processing methods of broadband random processes in onboard radio telemetry systems is extremely actual.

Importance of this problem strongly increases if necessary to treat huge data arrays in deadlines.

It gains a special urgency in connection with sharp magnification of a volume of the treated data called by expansion of a scale and research problems, and also rushing to more detailed research of measured processes.

Such demand, naturally, leads to essential thickening as telemetry systems, and radio engineering communication channels.

1.3 Solutions to the Problem

Solution of problems was possible at use of so-called methods of compression data (Ivanov, 2008; Salamon, 2004).

As a rule, in a spectrum of measured parameters the leading position is occupied with broad-band processes (Mala, 2007). Introduction of methods of quasireversible compression of data is effective only for slowly changing processes. Therefore, it does not solve a problem of cumulative reduction of volume of transmitted data and unloading of a communication channel.

Thus, the special urgency is gained by a problem of compression of broad-band signals which make all (10÷30) % from the general nomenclature of measured parameters, but load the transmission channel on (60÷80) %. (Ivanov, 2007; Molodchik, 2002)

For compression of broad-band signals it is possible to use methods of irreversible compression of data which consist in definition on a transmitting leg of is informational-measuring system estimations of probabilistic characteristics of measured random processes and their transmission over communication channels.

Difficulties in realization of methods of irreversible compression are significantly higher, in spite of the fact that their efficiency that is reductions of redundancy of data 10 times more. These results from the fact that the data processed onboard very often represent the only realization of non-stationary casual processes in the absence of aprioristic data on a type of function of distribution. The modern mathematical statistics does not arrange methods for an estimation of probabilistic characteristics of such processes.

2. Method

The assessment of probabilistic characteristics of non-stationary casual process is possible the basis on of adaptive division of temporary series of observations into stationarity sites with use of the nonparametric theory of decision-making. (Gamiz, 2011; Gibbons, 2010, Kvam, 2002; Richardson, 2005). Thus, the measured process structure can be described by additive-multiplicative model of the next form (Figure 1):

$$y(t) = X(t) + F(t) .$$

Here *y(t)* – the measured non-stationary casual process, *X(t)* – a stationary component of casual process, *F(t)* – a non-stationary trend.

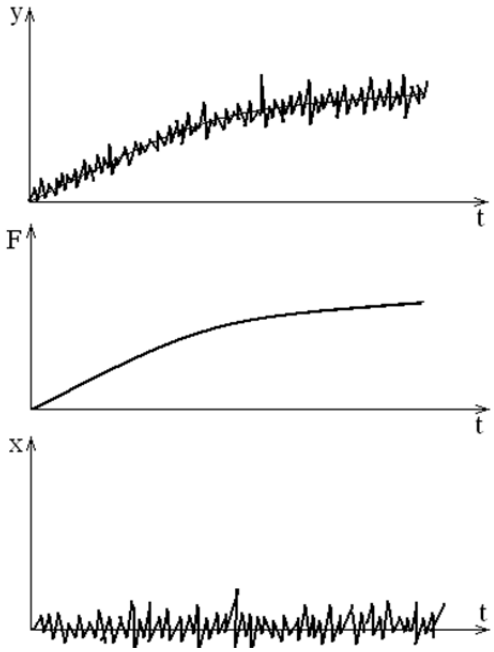


Figure 1. Additive-multiplicative model of the signal

The essence of methods consists in check of a statistical hypothesis of stationarity based on selective data of the measured process with use nonparametric statistics.

In this article the way of division of an interval of supervision into stationarity intervals according to the flowchart provided on figure 2 is offered.

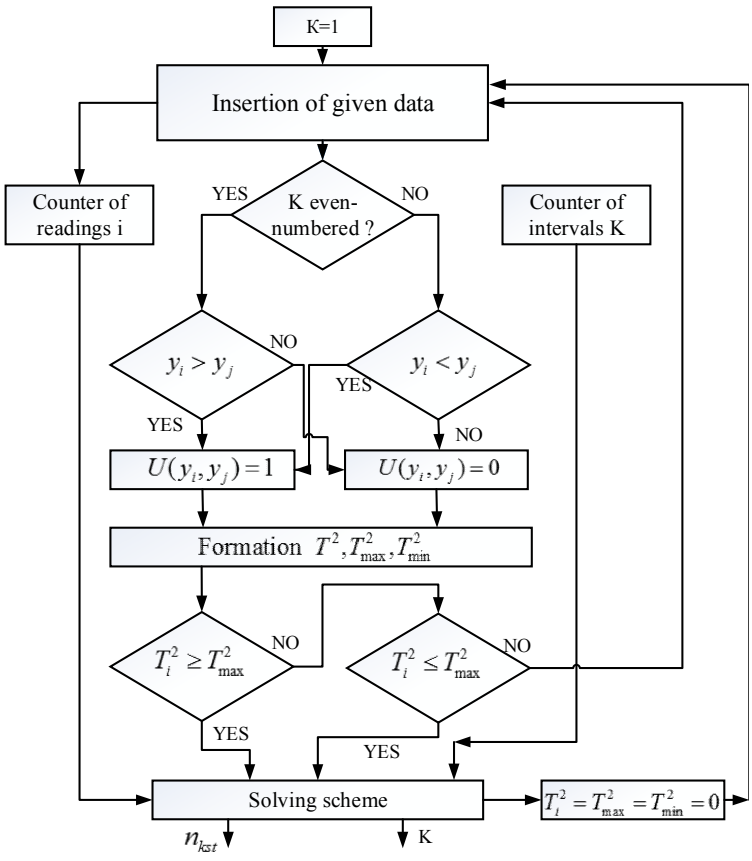


Figure 2. Algorithm of division of an interval of supervision into stationarity intervals

The flowchart in figure 2 is explained in section 2.1.

2.1 Division of an Interval of Observation into Intervals of Stationarity

The algorithm works as follows. Nonparametric statistics, known in literature as Kendal statistics (Tarasenko, 1976; Sheskin, 2003; Sprent, 2001) is formed by y_i, y_k sampling data of the measured process.

$$T^2 = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} u(y_i, y_k),$$

where:

$$u(y_i, y_k) = \begin{cases} 1, \ y_i \geq y_k \\ 0, \ y_i \leq y_k \end{cases}$$

Boundary $T^2_{\min}$ and $T^2_{\max}$ values are at the same time calculated. (Formulas for calculation of $T^2_{\min}$ and $T^2_{\max}$ aren't given as it is "know-how" of authors of article). Further we can see comparison of T_i^2 function with $T^2_{\max}$ and $T^2_{\min}$, i.e. the statistical hypothesis of stationarity is formed. Casual process with beforehand the set probability P = 1-α is considered stationary when performing the following condition: $T^2_{\min} \leq T_i^2 < T^2_{\max}$. If P it is probability of that the hypothesis of stationarity will be true, α is a probability of that the hypothesis of stationarity will be not true. α is called as a significance value. Calculation of T_i^2, $T^2_{\min}$ and $T^2_{\max}$ depends from α. If $T_i^2 < T^2_{\min}$ or $T_i^2 \geq T^2_{\max}$, a site of stationarity is considered certain. After this the $u(y_i, y_k)$ sign function "transfered" into the converse and T_i^2, $T^2_{\min}$ and $T^2_{\max}$ clearing into the null state, $T_i^2 = T^2_{\max} = T^2_{\min} = 0$ (Figure 3). Further, the evaluation process of $T_i^2, T^2_{\max}, T^2_{\min}$ and comparison procedure is repeated for a new stationarity interval (Yesmagambetov, 2007; Yesmagambetov, 2006). This means, the dynamic observational series is divided into any quantity of sections of different length on which a process with specified probability is considered stationary. The sequence of intervals of stationarity of different duration will be result of application of the specified procedure. On stationarity intervals average value of casual process with probability *P=1-α* is considered constant.

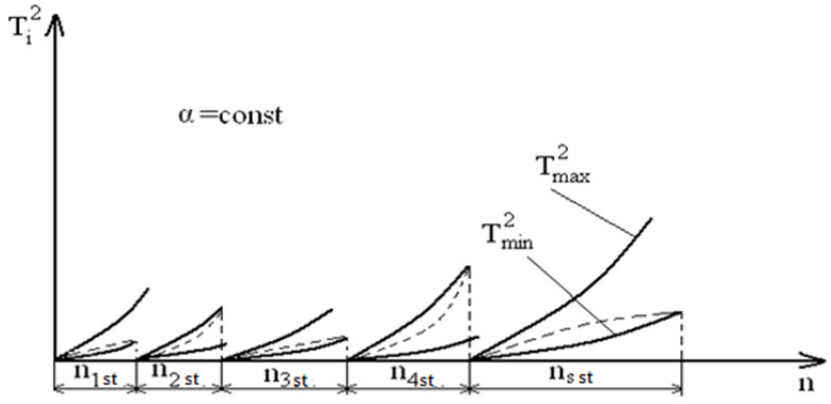


Figure 3. Separation of the observational data on the stationarity intervals

In figure 3 on a horizontal axis are shown number of counting on stationarity intervals. For example, the n_{3st} means that on the 3rd interval of stationarity there is a n_3 of counting.

2.2 Nonparametric Estimation of Mean Value Random Process

As known, that there are no universal estimations of probabilistic characteristics for the wide class of random processes (Golovnyh, 2005). For example, maximum likelihood estimates

$$(m_o = \frac{1}{N}\sum_{i=1}^{N} x_i, D_o = \frac{1}{N}\sum_{i=1}^{N}(x_i - m_o)^2)$$

are optimum for Gaussian distributed random numbers. However, they aren't effective for evenly distributed numbers, and also with correlation between counting of Gaussian casual process. Thus, the a priori knowledge of an aspect of a cumulative distribution function of measured random process is a necessary condition of correct sampling of estimations of statistical performances. It is especially important if to consider that at information handling, as a rule, it is necessary to deal with the only realization of nonstationary casual process. However, in practice a priori informations about measured process are often absent. That practically expels a feasibility of usual parametric modes for statistical handling.

This case, the various algorithms based on adaptation to absolute properties of implementation used: iterated methods of search of quasioptimum conditions of an estimation by volume samples, to length of implementation or a digitization pitch; methods of successive approximations, tests and errors (Karpenko, 2011; 2013).

However, such algorithms are labour-consuming enough and represent too large demands to microcomputer performances. Therefore, response of the microcomputer should make up 62 million simple 60 operations per second at handling of 32 signals with a breadth of a spectrum of 4 kilohertz, and core budget for storage of programs, input data and outcomes of handling should make 2 thousand 8-discharge words.

Thus, procedure of deriving of optimum estimations of probabilistic characteristics is tightly connected with problems of minimization of computing costs. Simplification of estimations of probabilistic characteristics was possibly at use of ordinal statisticians (OS) of ranked a series (RS).

$$x_{(1)} \le x_{(2)} \le \ldots \le x_{(R)} \le \ldots \le x_{(N)}$$

of random numbers gained on the interval of stationarity in sequence. Here $x_{(1)}$ - ordinal statistics of the ranged row on a stationarity interval. Ranging can be made both in ascending order, and in decreasing order.

In a number of studies carried out the errors of estimating the probability characteristics of order statistics. However, these activities were restricted to study of stationary random process while the particular interest represents deriving of estimations of probabilistic characteristics on samples of nonstationary random process.

Application of ordinal statisticians allows to use simple enough procedures for the calculation of average value of casual process, based on central ordinal statistics (COS) ranked beside (Tarasenko, 1975; Daivid, 1979).

$$\tilde{m}_{11} = x_{(c)}; \quad \tilde{m}_{12} = x_{(c+1)}; \quad \tilde{m}_{21} = \frac{1}{2}(x_{(c-1)} + x_{(c)});$$

$$\tilde{m}_{22} = \frac{1}{2}(x_{(c)} + x_{(c+1)}); \quad \tilde{m}_{2j} = \frac{1}{2}(x_{(c)} + x_{(c+j)});$$

$$\tilde{m}_{31} = \frac{1}{3}(x_{(c-1)} + x_{(c)} + x_{(c+1)}).$$

There are estimates based on the truncation ranked series

$$\tilde{m}_{41} = \frac{1}{N-2}\sum_{i=2}^{N-2} x_{(i)}; \quad \tilde{m}_{4j} = \frac{1}{N-j}\sum_{i=j}^{N-j} x_{(i)};$$

And also using extreme ordinal statistics

$$\tilde{m}_{51} = \frac{1}{2}(x_{(N)} + x_{(1)}); \quad \tilde{m}_{52} = \frac{1}{2}(x_{(N-1)} + x_{(2)});$$

$$\tilde{m}_{j} = \frac{1}{2}(x_{(N-j+1)} + x_{(j)}).$$

The estimations using various combinations of enumerated estimations can be synthesised:

$$\tilde{m}_{61} = \frac{1}{2}(x_{(K1)} + x_{(K2)}),$$

where $K1 = E[0.73N]$, $K2 = E[0.23N]$;

$$\tilde{m}_{62} = \frac{1}{2}(x_{(K1)} + x_{(K2)}),$$

where $K1 = E[0.75N]$, $K2 = E[0.25N]$;

$$\tilde{m}_{71} = \phi_1 \cdot x_{(K1)}; \quad \tilde{m}_{72} = \phi_1 \cdot x_{(K1)} + \phi_2 \cdot x_{(K2)};$$

$$\tilde{m}_{7j} = \phi_1 \cdot x_{(K1)} + \phi_2 \cdot x_{(K2)} + ...\phi_j \cdot x_{(Kj)}, j << N$$

$$\tilde{m}_{81} = \phi_{12} \cdot (x_{(K1)} + x_{(K2)});$$

$$\tilde{m}_{82} = \phi_{12} \cdot (x_{(K1)} + x_{(K2)}) + \phi_{34} \cdot (x_{(K3)} + x_{(K4)});$$

$$\tilde{m}_{8j} = \phi_{12} \cdot (x_{(K1)} + x_{(K2)}) + ...\phi_{ij} \cdot (x_{(Ki)} + x_{(Kj)}).$$

where φ – specially chosen multiplier which choice in this article isn't considered.

2.3 Nonparametric Estimation of a Random-Process Variance

When measuring the dispersion is advisable to use the same number of the ranked order statistics as in the estimation of the mean value. At the same time, it is best not to evaluate the dispersion of the process itself, and the ratio of the mean square deviation (MSD). To estimate MSD in nonparametric statistics using the simplest function of scope

$$W_1 = x_{(N)} - x_{(1)}$$

and function of "subrange"

$$W_j = x_{(n-j+1)} - x_{(j)},$$

using extreme order statistics ranked series:

$$\tilde{\sigma}_{11} = \nu(x_{(N)} - x_{(1)}),$$
$$\tilde{\sigma}_{12} = \nu(x_{(N-1)} - x_{(2)}).$$

It is possible to use estimations also:

$$\sigma_{3j} = \nu(x_{(K1)} - x_{(K2)})$$

where for σ31 $K1 = E[0.75N]$, $K2 = E[0.25N]$;
and for σ32 $K1 = E[0.73N]$ $K2 = E[0.25N]$;

$$\sigma_{3j} = \nu(x_{(K1)} - x_{(K2)})$$

$$\sigma_5 = \nu(x_{(K1)});$$

$$\sigma_{52} = v_{12}(x_{(K1)} + x_{(K2)}) + v_{34}(x_{(K3)} + x_{(K4)});$$

$$\sigma_{5j} = v_1 x_{(K1)} + v_2 x_{(K2)+..} + v_j x_{(Kj)+..}$$

As in the case of estimating the average, different combinations of central order statistics and extreme order

statistics (EOS):

$$\sigma_{4j} = \nu(x_{(c+j)} - x_{(c-j)})$$

$$\sigma_{2j} = \nu(x_{(N-j+1)} - x_{(c-j+1)})$$

The coefficient ν can be assigned from a wide range; however, the most effective factor values are as follows (Mamchev, 2007):

$$\nu = 1; ...1/2; ...1/3; ...1/4;$$

2.4 Nonparametric Estimation of Function of Distribution of Casual Process

The ranked number of order statistics can be used to estimate the function of distribution of casual process *F(x)* and function of density *f(x)*. In this case, it suffices to estimate one of them and indirectly to estimate the other, respectively, differentiating *F(x)* and integrating *f(x)*. With regard to the technique of transmission of telemetry data it is better assess the function of distribution of casual process *F(x)*. This connected with the greater complexity in the implementation of methods for estimating f(x) and better noise immunity transfer *F(x)* compared to *f(x),* because of the continuous increase in the ordinate *F(x)*. Therefore, consideration of methods of estimating function of distribution of casual process be paid more attention.

The classic definition of the distribution function, as the probability of the event *(x(t) <x)* allows us to write the following relation.

$$\tilde{F}_0(x) = \text{Prob}(x(t) < x) = \frac{N_x}{N} = \frac{1}{N}\sum c(x - x_i),$$

Where - *Prob(...)* means probability, *N* - sample size, N_x - number of samples of the process *x(t)*, not exceeding the value of *x*, $c(x - x_i)$ - the comparison function.

$$c(x - x_i) = \begin{cases} 1, x \geq x_i \\ 0, x < x_i \end{cases}$$

Statistical relationship between the sample value and its rank allows us to write the following approximate value:

$$\tilde{F}_1(x) = \tilde{F}_1(x_{(R)}) = \frac{R}{N+1.},$$

where *R* is a rank of counting of $X_{(R)}$ in the ranged row.

Modification of this method, based on fixation as quantile not order statistic *x(R)* of rank *R*, but a linear combination *Q* of order statistics

$$x_{(R)}^{Q} = \sum_{q=1}^{Q} A_q x_{(q)}$$

allow to generate the following estimates

$$\tilde{F}_2(\frac{1}{2}(x_{(R-1)} + x_{(R)})) = \frac{R}{N+1};$$
$$\tilde{F}_3(\frac{1}{2}(x_{(R)} + x_{(R+1)})) = \frac{R}{N+1};$$
$$\tilde{F}_4(\frac{1}{3}(x_{(R-1)} + x_{(R)} + x_{(R+1)})) = \frac{R}{N+1}.$$

At these estimations in the capacity of a quantile magnitude, average of two or three ordinal statisticians is fixed.

Other mode of the estimation of a cumulative distribution function is based on the evaluation of a nonparametric tolerant interval *(L2-L1)* where *L1* and *L2* name 100 β - percent independent of distribution *F(x)* tolerance limits at level γ and

$$\text{Prob}\left[(F_{(L2)} - F_{(L1)}) \geq \beta\right] = \gamma$$

If to suppose *L1=x(R)*, and *L2=x(S)*, where *R* <*S* the tolerant interval *[x(R), x(S)]* is equal to the sum of elementary shares from *R*-th to *S*-th, i.e.

$$\text{Prob}[(F_{x_{(R)}-x_{(S)}} \geq \beta] = \gamma = \frac{N!}{(S-R-1)!(N-S+R)!} \times \int_{\beta}^{1} z^{S-R-1}(1-z)^{n-S-R}dz =$$

$$= 1 - I_{\beta}(S-R, N-S+R+1) = \sum_{i=1}^{S-R-1} \binom{N}{i} \beta^{i}(1-\beta)^{N-i}.$$

Thus γ is a function of arguments *N*, *S-R* and β. There are some minimum value N_{min} to which in each specific case there matches quite certain combination *R* and *S*. It is possible to determine $\frac{1}{2}N(N-1)$ tolerant intervals with various level γ between which N/2 and N (N-1)/2 (depending on that even or odd N) will be symmetric. For security of symmetry of a rank should be connected a condition:

$$S = N - R + 1.$$

Then for an estimation of cumulative distribution function $F_5(x)$ in points *x(R)* and *x(S)* with a confidence coefficient γ is possible to accept the following magnitudes:

$$F_5(x_{(R)}) = \frac{1-\beta(R,S)}{2},$$

$$F_5(x_{(S)}) = \frac{1+\beta(R,S)}{2}.$$

Thus, changing value *R* from 1 to *N/2* and computing matching values S, it is possible to gain estimation $F_5(x)$ in N points.

Another mode of nonparametric estimation $F_6(x)$ can be generated from definition of a nonparametric confidence interval $[\, x_{(R)},\ x_{(R+K)}\,]$ for a quantile x_p level *p*. The Confidence level γ is determined from a relation:

$$\gamma = \text{Prob}(\tilde{F}_6(x_{(R)})) \leq p \leq F_6(x_{(R+K)}) = I_p(R, N-K+1) - I_p(R+K, N-R-K+1),$$

where $I_p(n,m)$ – Beta - Prison's function

$$I_p(n,m) = \frac{\Gamma(n+m)}{\Gamma(n)\cdot\Gamma(m)} \int_{O}^{P} x^{n-1}(1-x)^{m-1}dx.$$

In addition, the probability γ that the quantile x_p will appear between ordinal statistics $x_{(R)}$ and $x_{(R+K)}$ does not depend on an aspect of initial distribution *F(x)*.

3. Results

Comparative analysis of the following methods of estimation of the distribution function showed that in terms of the volume of the most satisfactory computational costs are estimates of the form F_1, F_2, F_3 and F_4. In addition, F_6 is much more complicated than an estimation of aspect F5 in implementation and on their use for this reason is undesirable.

Error analysis of the mean and mean-square deviation values in the random process had been carried out by statistical modeling method in Mathcad environment.

Random function *X(t)* with distribution function of $rnorm(N, \mu, \sigma)$ form having Gaussian (normal) distribution with mean μ and mean-square deviation σ (for example $\mu=1$, $\sigma=0$) have been used for the modeling.

Signal (trend) *F(t)* has been stipulated on the random function of the next form

$$F(t) = A \cdot (1 - \exp(-a_1 t)),$$

Where *A* and a_1 are varied in different limits for the modeling aims.

Trend of *F(t) = t* form has been stipulated in a number of cases.

In a result we have generated the non-stationary random process (Figure 4):

$$Y(t) = X(t) + F(t)$$

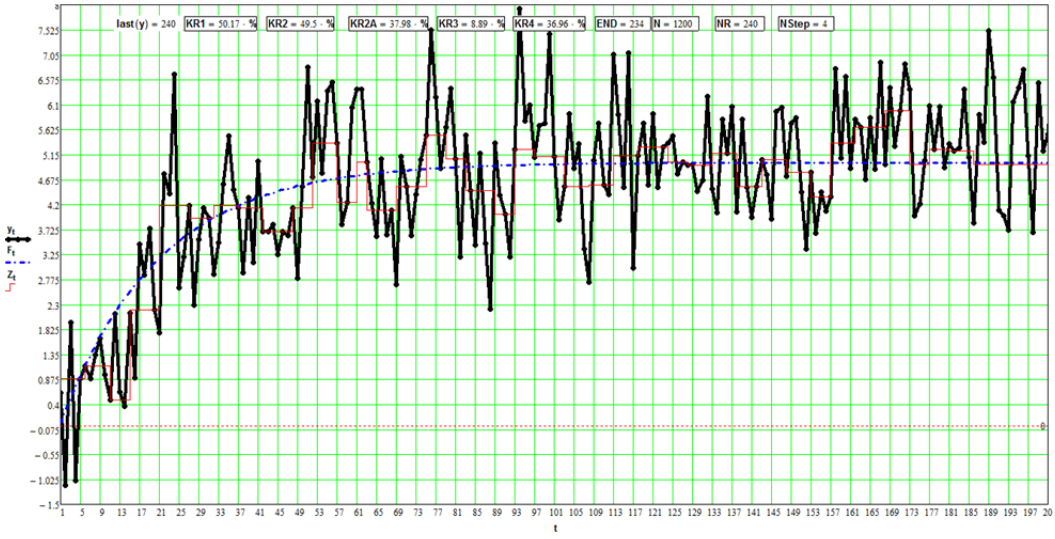


Figure 4. Generating example of the non-stationary random process

In figure 4 by black color non-stationary casual process *y(t)*, blue color – *F(t)* trend, red color – $\tilde{m}_{11} = X_{(c)}$ function is shown.

Valuation of the mean value by $\tilde{m}_{11} = X_{(c)}$ formula allows in point of fact approximate *F(t)* trend by graduated multinomial (multinomial of zero power) (Figure 5). On the figure 5 we have four sections of quasi-stationarity with the length in 9 readings, 35 readings, 57 and 59 readings respectively.

The modeling results showed the following.

Error of the estimated mean essentially depends on the trend form and on the proportion of random and nonrandom components (*noise/signal* relationship) (Karpenko, 2014; Mala, 2005).

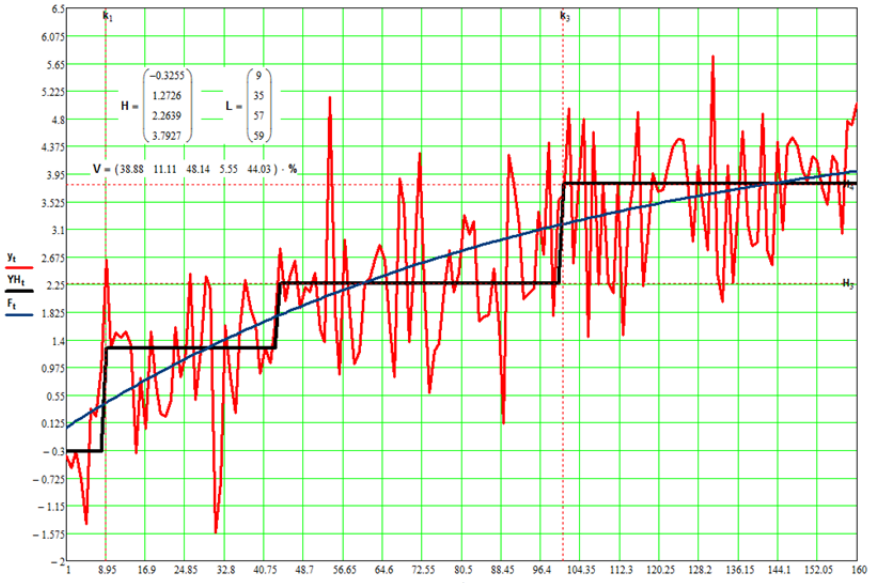


Figure 5. Separation of the non-stationary component (red color non - stationary casual process *y(t),* blue color – *F(t)* trend, black color – $\tilde{m}_{11} = X_{(c)}$)

The change of the relationship in the stationary component power to the non-stationary component amplitude

(*noise/signal* relationship) leads to the change in the calculation of average value of casual process. Thus, change in the *noise/signal* relationship within the range from 0.14 to 0.8 leads to the change in the calculation of average value of casual process by 3 times (Figure 6 and Figure 7).

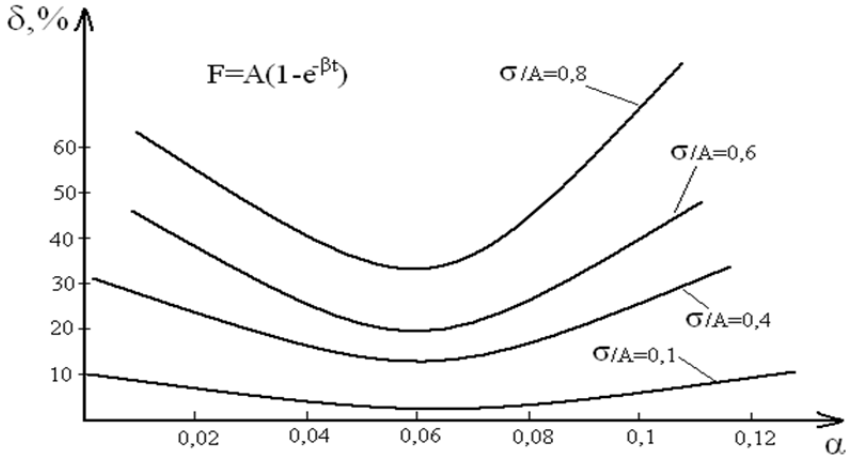


Figure 6. Dependence of an error of calculation of average value on a significance value

The fore cited dependences point on the presence of clear-cut error minimum of the non-stationary estimated mean which is in the range of significance value α = 0.05÷0.06l. Error increase at α value, which are out of this range explained by two properties. On the one hand, there is increase in the stationarity hypothesis region at α < 0.05 value that leads to the decrease in the non-stationary component separation accuracy. On the other hand, there is decrease in the quantity of readings for the stationarity participation at the decrease in the stationarity hypothesis region (α > 0.06) that leads to the increase in the estimated mean error. α=0.05÷0.06 minimum position does not depend on the trend character and the *noise/signal* relationship.

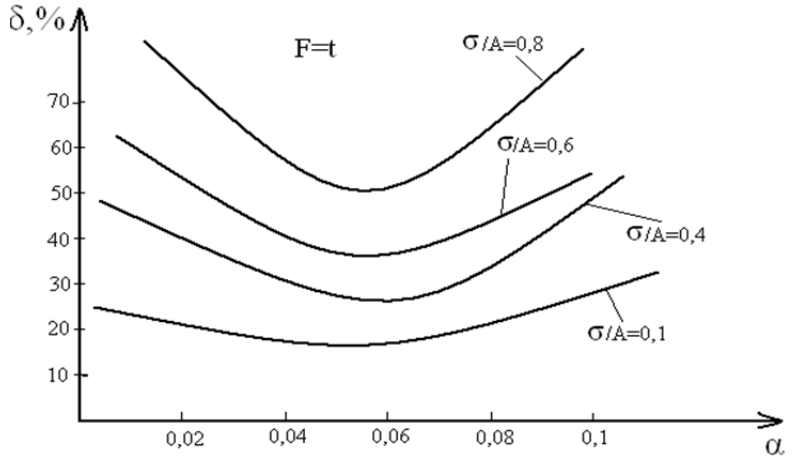


Figure 7. Dependence of an error of calculation of average value on a significance value

In figures 6 and 7 *σ/A* - *noise/signal* relationship, where σ – a mean square deviation of a casual component, *A* – trend amplitude.

It is obvious that the central order statistics are the easiest to implement. In Figure 8 shows a comparative analysis of different methods of computational cost estimation of the average. The minimum costs, apparently, have estimations of an aspect $\tilde{m}_{11}$ and $\tilde{m}_{12}$.

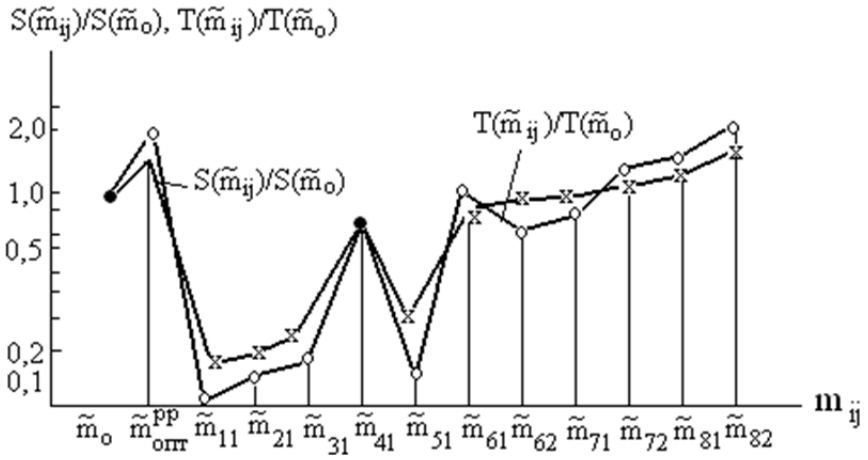


Figure 8. The comparative analysis of computing costs of an estimation

In figure 8 S and T – the memory size and average time of calculation at realization on the microcomputer of the elementary ways of estimation of average value.

Comparative analysis of the estimates shows the dependence of the mean error of the volume of uncorrelated samples N (Figure 9 and 10), and this dependence is similar to the corresponding dependence evaluation $\tilde{m}_0$, where it can be concluded about the consistency of the estimates based on the use of order statistics ranked series.

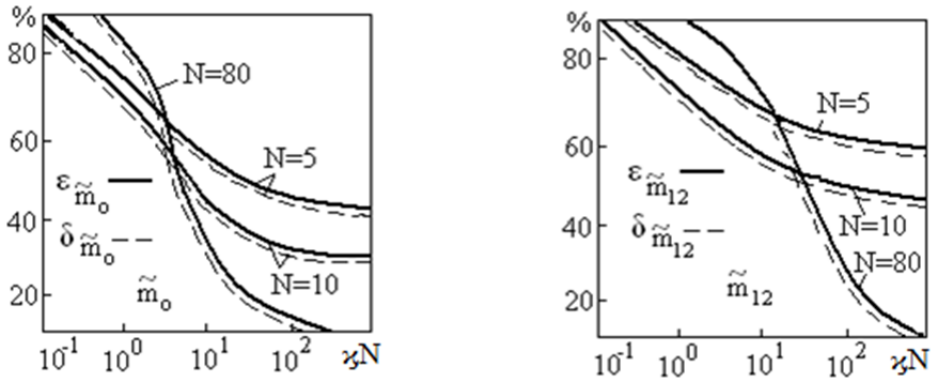


Figure 9. Dependences of an error of estimation of average value on the specified realization length

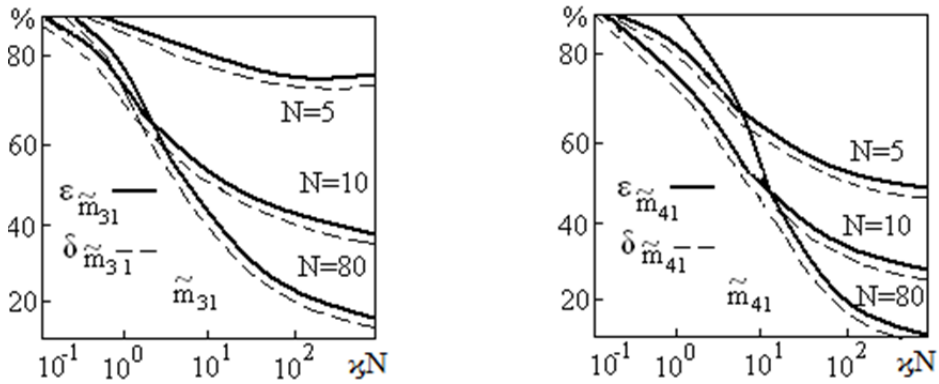


Figure 10. Dependences of an error of estimation of average value on the specified realization length

The following designations are given in figures 9 and 10:

ϰN – the specified realization length, where ϰ= $1/N_K$; N_K , - number of counting on a correlation interval; ε, δ – estimation errors.

The considered COS - count the best indicators of accuracy has a simple estimate of the form $\tilde{m}_{12}$, having besides the minimum computational cost. Uncertainties in estimates using extreme order statistics, for small N is not worse than the COS - estimates, but for large values of N, they are less effective than simple estimates of the form $\tilde{m}_{11}$ and $\tilde{m}_{12}$. In class evaluations, using truncation ranked series, the quality of estimates increases with decreasing j. This is especially noticeable in the region of small N. For large values of N (N> 50) decrease j virtually no effect on the quality of estimates. Estimates of species

$$\tilde{m}_{6j}, \tilde{m}_{7j}, \tilde{m}_{8j},$$

No preferential have areas of use in mind more complex computing operations to obtain them. These dependences for all modes of evaluation are at least attributable to the value of N is the sample size equal to 20÷30 counts. These dependencies are common to all types trender. Deterioration of the accuracy for N> 30 is due to non-stationarity of the process being measured, that for large sample sizes to have a greater impact on accuracy than the increase in sample size.

With a sample size of 20-30 counts, errors of all estimates does not exceed 30%.

From these dependences it is possible to conclude that the most effective, both in terms of accuracy of the estimations, and in the sense of minimizing the computational cost is a simple estimate of the form-DSP $\tilde{m}_{12} = x_{(c+1)}$

The analysis of the conducted researches at a variance estimation has shown that at the fixed value of factor ν there are no satisfactory estimations in a broad band of modification N. At values *N=const* the more correlative association between references, the more demanded value of a factor. On Figure 11 values of lapses for different estimations are reduced at various factors ν depending on a sample size for uncorrelated references. From the reduced associations it is visible that value of factor ν is necessary for appointing in the adaptive image, proceeding from a sample size.

Use of estimations of type σ_{2j} and σ_{4j} is unpromising, as any of these estimations does not give comprehensible accuracy $\delta\sigma < 30\%$ in all range of modification N at any values of factor ν. Aspect estimations σ_{3j} and σ_{5j} are less preferable in comparison with the estimations using extreme ordinal statisticians, in view of large computing costs.

Thus, the most comprehensible is estimations of an aspect $\tilde{\sigma}_{11}$ and $\tilde{\sigma}_{12}$, yielding satisfactory outcomes in a broad band of modification N. Optimum conditions of an estimation can be noted in a following aspect:

$$\tilde{\sigma}_{\text{opt}} = \tilde{\sigma}_{11} = \nu(x_{(N)} - x_{(1)}), \quad \begin{cases} \nu = \dfrac{1}{3}, & N < 15 \\ \nu = \dfrac{1}{4}, & N \geq 15 \end{cases}.$$

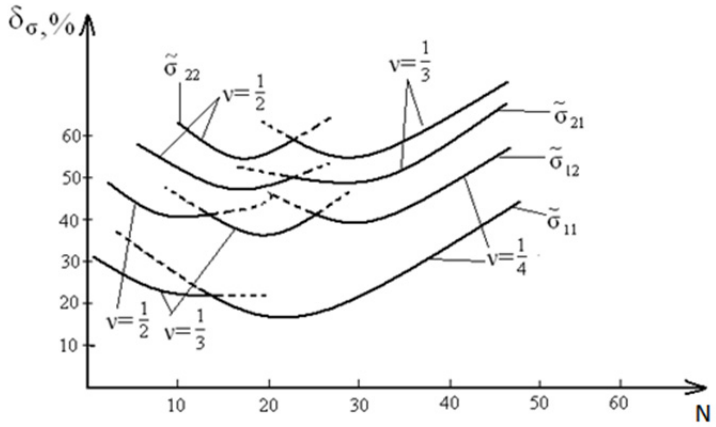


Figure 11. Dependence of an error of estimation of the mean square deviation of counting on a stationarity interval

Estimation application $\tilde{\sigma}_{11}$ leads to a lost significance in comparison with a maximum likelihood estimate $\tilde{\sigma}_{0}$ no more than 5 %. Almost satisfactory outcomes (δ <30 %) are gained at *5 < N < 40*. The decline of estimations at *N> 40* is connected, as well as in case of an average estimation, uncorrelated measured random process.

The study estimates the error $F_1(x)$ shows that the error in the ranked series is always lower than at the boundaries of the series, where the values $\tilde{F}_1(x)$ associated with rare events. Associations of lapses of an estimation on a sample size testifies that sample size magnification at a strong correlation between references does not lead to lapse decrease. Lapse decrease at the fixed sample size was possibly decrease of extent of correlation between references. In the absence of a correlation improving of quality of an estimation happens to sample size magnification. Comprehensible outcomes are gained at N> 10. However at *N > 40* the decline of estimations called by agency nonstationarity of process is observed. At uncorrelated references the estimation $\tilde{F}_1(x)$ is worse than the modifications of all on 2..3 % (Figure 12), and at magnification *N (N> 80*) this difference in lapses of estimations *F_1, F_2, F_3 and F_4* practically disappears. Therefore, considering that use of modified estimations gives an insignificant scoring on accuracy, it is possible to consider optimum more simple estimation

$$F_1(x) = \frac{R}{N+1}.$$

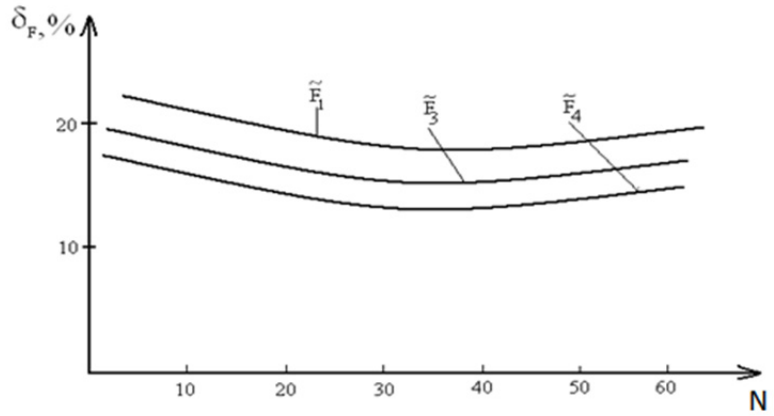


Figure 12. Dependence of an error of estimation of distribution function on number of counting on a stationarity interval

Thus, application of the best estimations of an aspect $\tilde{m}_{12}$, $\tilde{\sigma}_{11}$ and $\tilde{F}_1(x)$ allows to use the same ranked a series of ordinal statisticians for an estimation of such different probabilistic characteristics of random process, as an average, a variance and a cumulative distribution function. This fact is very important, as it allows first, substantially reduce the computational cost of the preparation of these estimates, and secondly, allows a single measurement to obtain almost full information about the measured process. The sample size, as appears from the analysis of lapses of estimations $\tilde{m}_{12}$, $\tilde{\sigma}_{11}$ also $\tilde{F}_1(x)$ should be appointed proceeding from an interval of 15÷30 references. Thus, lapses of estimations do not exceed 20%.

Unfortunately, at such approach it is not possible to gain correlation function estimations as deriving of estimations $\tilde{m}_{12}$, $\tilde{\sigma}_{11}$ and $\tilde{F}_1(x)$ is connected with the demand of statistical independence between references.

4. Discussion

Sampling of an optimum pitch of a digitization. In references there are various approaches to sampling of a pitch of a digitization Δt. The majority of recommendations assumes to make a digitization of random process, proceeding from the theorem of Katelnikov. In some activities it is supposed to determine a digitization pitch depending on accuracy of approximation of estimations of correlation function k_τ random process. However, too high demands to frequency of inquiry in the first case and a priori independence of a correlation function in the second does not allow to use these techniques at handling of broad-band random process.

In a number of activities it is recommended for the basis at sampling of a pitch of a digitization to accept absolute error of restoration of continuous implementation.

However, it is possible to consider as the most comprehensible criterion assigning of a pitch of a digitization proceeding from a statistical lapse of gained estimations. As analytical research of expressions for lapses is difficult for fulfilling because of bulkiness, the analysis of lapses was made by a modelling method on the computer.

As variances of estimations have a minimum sampling of a quasioptimum pitch of a digitization $\Delta t = K\tau_k$ K is possible, ensuring an estimation of probabilistic characteristics with a margin error, close to minimum at possible a smaller amount of references in implementation. In this case τ_k - a correlation interval for which at an average estimation to a variance accept a time of the first intersection normalized a level correlation function $\lambda = 0.30$. At an estimation of a cumulative distribution function, the sampling step can be appointed by the same technique, as qualitatively character of behaviour of a lapse of an estimation of a cumulative distribution function same, as at average estimations to a variance. Level λ in this case starts equal 0.7. Thus, accuracy of an

estimation worsens on 3÷5 %, and the sample size in comparison with an optimum estimation is divided out almost to an order.

However, as value of a correlation function a priori is not known, for a correlation interval τ_k it is better to adopt a value an integral interval of correlation

$$\tau_k = \int_0^\infty \rho(\tau) d\tau,$$

Where $\rho(\tau)$ - the normalized correlation function of the random process.

Singularity of application of an integral interval of correlation is that this magnitude connected with an effective breadth of a spectrum ΔF_{ef} simple association

$$\tau_k \Delta F_{ef} = \frac{1}{2}$$

and in this case it is enough for assigning of a pitch of a digitization a priori knowledge only a spectrum breadth.

Outcomes of researches have shown that at an average estimation an effective mode of assigning of a pitch of a digitization is sampling of its condition of independence of selective references as presence of a correlation worsens accuracy of an estimation (Figure 13). Therefore in this case

$$\Delta t = \tau_k = \frac{1}{2\Delta F_{ef}}.$$

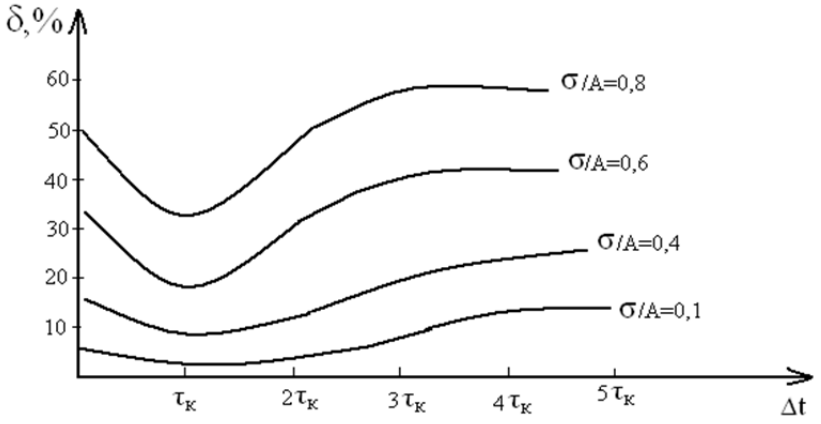


Figure 13. Dependence of an error of calculation of average value on a digitization interval

Agency of a correlation on accuracy is characteristic, though to a less degree, and for a variance and cumulative distribution function estimation (Figure 14 and 15). The conducted researches have shown that the pitch of a digitization for these cases should be appointed from a relation

$$\Delta t = \frac{1}{2}\tau_k = \frac{1}{4\Delta F_{ef}}.$$

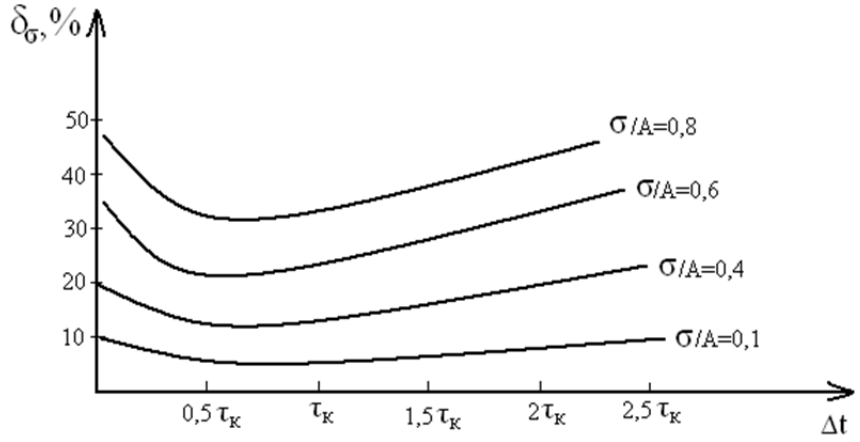


Figure 14. Dependence of an error of calculation of a mean square deviation on a digitization interval

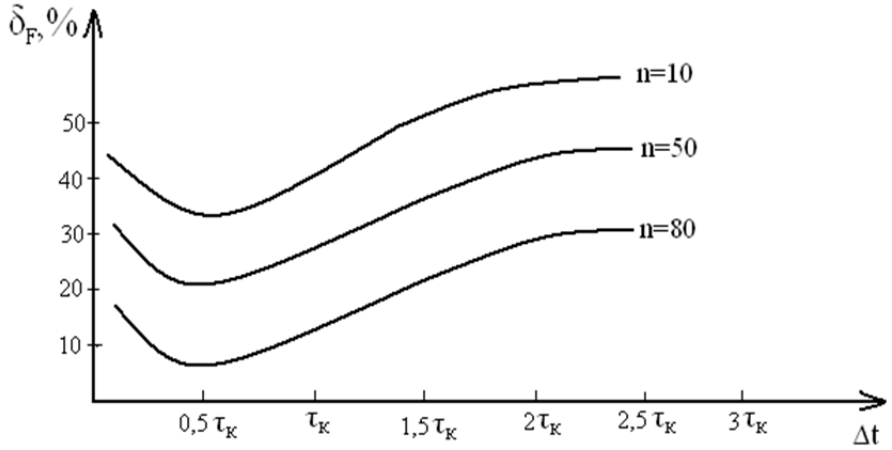


Figure 15. Dependence of an error of function evaluation of distribution on a digitization interval

Necessity of deriving of all estimations of probabilistic characteristics on the same sampled data does not allow gaining an optimality simultaneously for all estimations owing to different demands to a digitization pitch. In this case, as the most important is the information on mean value, it is expedient to choose a digitization pitch Δt equal to a correlation interval $\tau\kappa$ and to gain optimum estimations of a nonstationary average. Variance and cumulative distribution function estimations in this case though will not be optimum will be comprehensible enough.

5. Conclusion

Random processes in the rocket-and-space technics, as a rule represented by the sole implementation in conditions of expected uncertainty on the kind of distribution function. Processing of such processes on-line is possible using nonparametric solution theory methods. The present work considers technique for the dynamic observational series division on stationarity intervals with the following valuation of mean values and dispersion.

In the future, we are planning to perform researches on valuation of the distribution function and correlation function of the non-stationary broadband random process.

References

Daivid, G. (1979). Ordinal statisticians-M: the Science, 336.

Gamiz, M. L. (2011). Applied Nonparametric Statistics in Reliability. Springer, p. 230. http://dx.doi.org/10.1007/978-0-85729-118-9

Gibbons, J. D. (2010). Subhabrata Chakraborti. Nonparametric Statistical Inference (5th ed.). Chapman and Hall/CRC.

Golovnyh, A. (2003). Digital inhabitancy. CHIP. Computers and communications. K.: The publishing dwelling

The Software the Press, 1, 68-70.

Ivanov, V. G., Lomonosov, U. B., & Lyubarsky, M. G. (2008). Analiz and classification of methods of compression of the information/.//Bulletin NTU HPI. Thematic issue: Information science and modelling. *Kharkov: NTU KhPI, 49*, 78-86.

Ivanov, V. G., Lyubarsky, M. G, & Lomonosov, U. V. (2007). Abbreviation of substantial redundancy of images on the basis of classification of plants and a hum. *Control and Information Science Problems, 3*, 93-102.

Karpenko, A. P. (2013). Hybrid population algorithms of parametric optimization of design solutions. *Information Technology Application, 12*, 6-15.

Karpenko, A. P., & Babichenko, A. B. (2014). Nejrosetevoe forecasting of a fuel rate by the aircraft. *Aerospace instrument making, 4*, 47-56.

Karpenko, A. P., Kulesh, D. S., & Fedin, V. A. (2011). Construction of domain boundary of reachability of dynamic system by a combination of methods multi finish and approximation of a field of vectors. *Science and Education: Electronic Scientifically-Engineering Issuing, 5*. Retrieved from http://technomag.edu.ru/doc/185335.html

Kvam, P. H., & Vidakovic, B. (2007). Nonparametric Statistics with Applications to Science and Engineering. *John Wiley AND Sons, Inc.* http://dx.doi.org/10.1002/9780470168707

Mala, S. (2005). Wavelets in handling of signals.-M: Mir, 27-30.

Mamchev, G. V. (2007). The radio communication and television bases. The manual for high schools.-m: the Hot line-a Telecom, 416.

Mionov, S. (2005). Electronic archives for the industry. *Open Systems, 2*, 56-60.

Molodchik, P. A. (2002). video compression: the present and future//the Computer review. *The Publishing Dwelling ITC, 33*, 49-51.

Novikov, S. A. (2005). Video observation Data transfer on IP-webs. *Open Systems, 9*, 57-59.

Richardson, J. (2005). Video code. H. 264 and MPEG-4 Standards of new generation. *The Technosphere,* p. 368.

Salamon, D. (2004). Compression of data, images and a sound. *Technosphere,* 368.

Sheskin, D. J. (2003). Handbook of Parametric and Nonparametric Statistical Procedures. *Chapman AND Hall/CRC,* p. 1184. http://dx.doi.org/10.1201/9781420036268

Sprent, P., & Smeeton, N. C. (2001). Applied Nonparametric Statistical Methods. ChapmanandHall/CRC, p. 462.

Tarasenko, F. P. (1976). Nonparametric statistics-Tomsk: Publishing house of Tomsk university, p. 293.

Voronov, E. M., Karpenko, A. P., & Fedin, V. A. (2011). Parallel construction of assemblage of reachability of the highly-maneuverable aircraft by a "multifinish" method. *Bulletin NNSU, 3*, 194-200.

Yesmagambetov, B. (2007). Using of not parametrical criteria atirreversible compression of the data. *Turk Dunyasi Arastirmalari, 171*, 1-6.

Yesmagambetov, B. B. S. (2006). Optimization of parameters of the associative device of compression of data. *The bulletin of the Kazakh National Engineering University of K.Satpaev, 3*(53), 126-130.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Development of Fire-Resistant Multilayer Materials for Working Clothes of Welders

Zaure Dauletbekovna Moldagazhiyeva[1] & Raushan Orazovna Zhilisbayeva[1]

[1] Almaty Technological University, Kazakhstan

Correspondence: Zaure Dauletbekovna Moldagazhiyeva, Almaty Technological University, 100, Tole bi Street, 050012, Almaty, Kazakhstan.

Received: October 5, 2014 Accepted: October 15, 2014 Online Published: July 31, 2015

doi:10.5539/mas.v9n8p334 URL: http://dx.doi.org/10.5539/mas.v9n8p334

Abstract

In the paper several types of multilayer materials are proposed, which consist of fire-resistant woven and nonwoven materials. The developed non-woven material is a fabric made from woolen and m-aramide fibers connected in a combined way and which are components of a new multilayer material for welder's uniform. The proposed new woven materials due to their specific properties are suggested for an application as internal layers of protective clothing. Several types of fire-resistant fabrics with improved service properties are studied.

Keywords: Non-woven material, multilayer materials, working clothes, wool, aramid fibers

1. Introduction

Light industry development program in the Republic of Kazakhstan for 2010-2014 years is a step in an implementation of activities for a development of competitive consumer goods of high quality and wide variety (Program, 2010)

Development of scientific and technical progress, transformation in a field of production at modern industrial enterprises prioritized a task of an optimal solution of industrial safety problem, including a development of working clothes for protection from harmful effects of industrial environment

For a long time in a design of working clothes a major approach was functional-economic approach, in which the maximum unification of structure, its details, parts and elements minimized expenses of industrial enterprises related with provision (free of charge distribution) of working clothes to employees, working in various branches of the economy (Mokeeva, Picinskaya, 2013) .

At a pre-project analysis phase a problem was indicated – lack of working clothes for welders on market, which would fully meet consumers requirements. Currently, there are several companies in Kazakhstan, which produce working clothes: "KazCentre", " Master-Spec ", "WorkClothers", "Amirsana", "Modni rabotnik" (Stylish worker), "Altyn Zhuldiz", etc. The majority of these enterprises uses canvas made from half-linen fabric with fire-protective treatment or special cotton fabrics soaked with a special fire-retardant treatment, at the same time those materials are not effective from the point of view of their heat-resistance: they deteriorate in 4 to 6 months instead of required 12 months, also their manufacturing process is labor-intensive.

In spite of the presence on the market of a diversity of fabrics and materials with special finishing for protection from extreme temperatures, enterprises, producing costumes for welding, do not fully use new materials in because of high cost in case high performance properties and, backwards, use low performance materials with an affordable cost.

That's why a supply of working clothing to Kazakhstan market is economically profitable and improvement of its quality is an important scientific and industrial problem.

On of the promising direction in a production of fire-protective nonwoven materials are materials, produced from chemical and synthetic heat-resistant fibers.

1.1 Review of Fire-Protective Materials

Nowadays there is a big number of fire-protective materials. Methods of their production vary from soaking a fabric with special compositions to a creation of new materials from non-inflammable fibers and threads. Nowadays special chemical fibers had become widely spread, thus transfering scientific advances to industry.

Properties of the fibers such as high strength, thermal stability, chemical stability, fire resistance, non-inflammability, electrical conductivity, shock resistance and light weight make them and products based on them necessary and irreplaceable in various branches of industry, including textile industry (Zhilisbayeva, & Kozhabergenova, 2012).

Many foreign and domestic companies are developing fire-protective materials, which have standard mechanical and thermal and physical properties. However, it's worth mentioning, that high cost is a significant disadvantage of imported materials, and materials, produced by domestic companies, do not alway meet requirements of complex protection from harmful industrial factors, in particular, they are not fire-resistant.

Variety of fire-protective materials on international market allows to select a fabric for welder's working clothing, which have a certain set of characteristics, which will allow to use it in different companies. One of the leading Russian companies producing fabric for working clothing and uniform is "Chaikovski textil'" INC., which produces "Phenix" (Phoenix) brand of fabric with fire-resistant treatments TEFLON and PYROVATEX, which are produces by CIBA company (Fomchenkova,2002). Russian company "Kadoteh" produces fire-resistant fabrics from Nomex thread (produced by DuPont) and working clothes from it (Gushina, 2004). Westex concern produces INDURA and INDURA UltraSoft brands of fabric, which use PROBAN treatment. INDURA fabric is produces from cotton fibers with 240-472 g/m^2 density. In a case of INDURA UltraSoft production, small amount (12%) of polyamide (nylon) fibers with 190-405 g/m^2 density are added to cotton fibers, which constitute the major part (88%). An addition of nylon increases service life of clothing up to 50% and increases protection properties of the fabric. Fabrics protect from electric arc, open fire, splashes of melted metal and are designated for oil and gas industry workers, power engineering specialists, welders, metallurgist and other applications (Ognezashitnie materialy Westex (USA), 2005).

Weldersafe 100% cotton fabric (420 g/m^2) provides an absolute protection from burning injury of 3 level during the test with open fire and also protects from splashes of meltd metal and sparkles. The fabric is recommended for an application in welder working clothing (Rabochaya odezhda, 2004). «Klopman» company produces fire-resistant fabrics for working clothing and uniforms, which protect from fire and melted metal (cotton, mixed (Fabrics Klopman, 2005). Russian company «Baltiiski textil'» is on the biggest produces of fabrics for working clothing, uniform and camouflage on Russian market. XM FIRELINE fabrics are produced from 100% high-quality cotton (Gefest-420) for manufacturing of working clothing for welders and metallurgist (Colombo-350), workers in oil and gas industry and power engineering speialists (Etna-350, Electra-320, 420, Madeira-320) («Baltiiski textil'», 2014). Also, textile materials with treatment based on greycoat cloth (Zhilisbayeva, 2011), which have several disadvantages such is high surface density (weight) and that a treatment can be washed off during use of working clothing.

Studying of working conditions in aforementioned industries, many foreign and domestic companies develop fire-protective materials, which have standard mechanical and thermal and physical properties. However, it's worth mentioning, that high cost is a significant disadvantage of imported material and materials, produced by domestic companies, do not alway meet requirements of complex protection from harmful industrial factors.

Until recently in Kazakhstan, as a rule, the most necessary and primitive textile materials and protection methods were used. In particular, in the case of working clothing production, a development was directed at use of inexpensive fabrics of natural and synthetic origin, which protective properties were determined by a fabric's thickness, weight and type of treatment. At the same time, new synthetic and artificial protective materials with improved protection, technological and consumer properties had become accepted in the world.

Nonwoven industry is the one of the most dynamically developing industries in the world. The most important point in the rapid development and commercialization of nonwoven materials is an ability to produce materials with special properties in the shortest time and with an affordable price (Scott, 2005).

The considerable part of all textile materials is nonwoven linens, which successfully compete with fabrics and substitute them, because they posses significant advantages (high productivity, an ability to reduce a production cycle, expenses of money and labor, maximum possible replacement of natural fibers with artificial ones).

2. Methodology

After complex analysis of the most acceptable textile material used as a protective layer for welder's working clothing, it was established, that the most heat-resistant and economically feasible are nonwoven materials, produced from wool and aramid fibers.

Technical effect of set objective consists of widening raw materials range for manufacturing of nonwoven and fire-resistant material with an implementation of low-grade wool fibers of sheep, bred in Kazakhstan, and, in the

same time, of utilization of textile industry waste materials. From the point of view of flammability, wool is incombustible in comparison with regular fibers and it has low flame temperature (from 600 to 1000 ***EC***) (Scott, 2005).

Mixing of meta-aramid fibers with wool fibers and their binding by needle punching provides an increase of thermal stability, watertightness and a decrease of thermal conductivity evenly along whole volume of produced material, which increase a protective capability of produced material.

Needle-punching is the most widely spread method of nonwoven materials production in textile industry in the world (Kamath et al., 2004).

That kind of nonwoven materials are produced using needle-punching equipment. Binding of fibers in nonwoven linen is conducted by mechanical mixing with repeated punching of nonwoven linen with notched needles. When they pass through linen, needles take bundles of fibers with notches and pull them through the thickness of linen. In results position an orientation of fibers in linen are changed. In the positions of punches bundles of fibers perpendicular to a linen's plate are formed and with a help of that bundles binding of linen's elements is conducted, which increase strength of the material.

An implementation of wool fibers in the discussed method allows to traditionally bind the material by fulling, which increase thermal stability, watertightness and decrease thermal conductivity (Suharev, 1973).

In order to obtain the best combination of physical, chemical and consumer properties of a nonwoven material in a corresponding segment, it is necessary to select a structure of material and method of its formation. In the same time, a structure of fibrous layer (linen), as a rule, is characterized by thickness of web, number of web's additions, unevenness of the web, ratio of various fibers in a mix, composition of a fibers, position of fibers (coefficient of straightness and extent, orientation angle), length of fiber, diameter of fiber, fiber distribution along thickness, presence or absence of frame layer, its structure, density of punches, depth of punches, presence of binder and adhesion bonds.

On the basis of the conducted studies optimal methods of new nonwoven materials production are selected, which includes such operations as preparation and mixing of fibers, greasing, combing, formation of fibrous layer, needle punching, calendering, heat shrinkage, and can be characterized by effectiveness of raw materials loosening, effectiveness of scutching, composition of emulsion and binder, coefficient of fiber distribution's unevenness, speed of web, linear speed of a linen's movement, temperature of thermal treatment, etc. Those specimens were produced using fulling, needle-punching and combined methods.

Mixed production of nonwoven linens consists of mixing of wool and aramid fibers on an initial stage of linen's production,

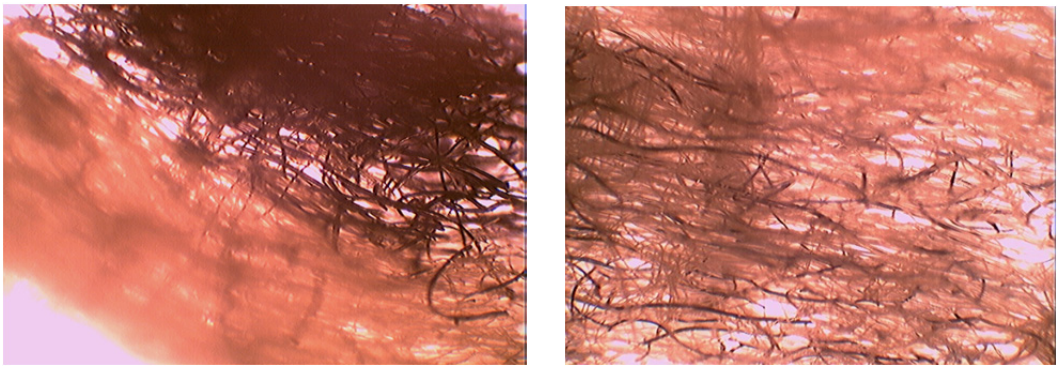
Combined technology of nonwoven linen's production is based on a combiantion of mechanical and physicochemical methods of binding in form of layer of wool and aramid fibers. Combinations of methods can vary: for example, initial linen binding by needle punching or jet-stream and consequent binding with binder, broaching of frame by fleecy threads and their binding by means of binder reagents, etc. Also, jet-stream processing of linen, containing fusible fibers, fibrids or biocompnent fibers by hot air or water, can be attributed to a combined method. At that, not only linen fibers' entaglement takes place, but also their thermal binding (Sheromova, 2006). In the table 1 examples of testing specimens of nonwoven material, which differ in number of layers, weight and production method are presented.

Table 1. Examples of obtained specimens of nonwoven fire-protective linens

No	Composition of layers	Temperature inside the fabric after high temperature treatment	Ratio of componenets, mass. %	Method of production	Thickness of specimens, mm
1	M-aramide	53	50	combined	20
	Wool		50		
2	M-aramide	52	60	mixed	15
	Wool		40		
3	M-aramide	50	70	mixed	25
	Wool		30		
4	M-aramide		70	combined	25
	Wool	45	30		

	M-aramide				
5	M-aramide		60	combined	18
	Wool	47	40		
	M-aramide				
6	M-aramide		60	combined	30
	Wool	49	40		
	M-aramide				
	Wool				

Study of the micro structure of nonwoven was conducted with an implementation of «MC-300TX» trinocular microscope. The view of testing specimens is presented on figure 1.



a) ×4000 b) ×4000

Figure 1. Microstructure of testing specimens of nonwoven fire-resistant linens: a) Specimen produced with combined method; b) Specimen produced with mixed method.

Figure 1a shows a prototype produced in a combined way, in the form of layers of woolen and meta-aramid fibers. Figure 1b shows a sample of the resulting mixed manner in the form of entanglement of wool and meta-aramid fibers.

Note the high degree of interlinkage fibers together, which has a positive effect on improving the physical and mechanical properties of fire-resistant nonwoven fabric as a whole.Thickness of fibers is: m-aramid - 12 microns, wool - 25 microns.

3. Results

According to the results of investigation on a dependence between weight and number of layers, as well as various ways of nonwoven materials production, it was established, that the optimal way of new nonwoven materials' production is combined method, because of optimal parameters of a worker's protection from high temperature, in particular, acceptable temperature specified by standards, on an inner side of linens.

Working clothing from new generation materials is of course 3-4 times more expensive than regular one, but higher then traditional canvas (3-4 months) welder's working clothing actual service life, as well as essentially new level of protective, mechanical, aesthetical and performance parameters of new materials must perceive customers to use working clothing from that type of fabrics (Nonwovens & Technical Textiles, 2007).

During a design of welder's working clothing it is proposed to use known fire-resistant materials as additional protective pieces and outer layer, which characteristics are presented in table 2.

Table 2. Main properties of analogue materials

Properties and characteristics, units of measurement	Values of properties and characterisitcs of the meaterial			
Name of the material	Premier FR 350A	FlameFort W280 Protect	Brezent OP	FlameFort 210A
Code number	10202 AM	50402 Kl4	11255	60405 a-M
Composition of fibers	100% cotton+ antistatic thread	100% aramide	60% flax+40% cotton	100% aramide + antistatic thread
Surface density, g/m2	340	430	495	220
Weave	Satinlike5/2	Satinlike	Reps	Serge 2/1
Treatment	PyrovatexTo+HMBO	Kl4	OP	HMBO
Breaking load (base/weft)	1200 N / 700 N	1200 N/ 1200 N	1400 N / 700 N	1000 N / 800 N

For a characterization of fire-protective fabrics from functinal protection point of view, it is necessary to conduct a series of studies, which will be proving their suitability for an application as a protection from high temperatures and melted metal. The first step in a selection of working clothing is a determination of danger, evaluation of the impact's potential and level of protection (Golubaev, 2008).

Resistance to splashes of melted metal and fire-resistance are main indicators of welder's protective clothing's quality.

Considering that topography of melted metal splashes during welding varies and at the same time sparkles and splashes hit at various angles, it was proposed to study materials resistance to inflammation.

3.1 Inflammability Tests

Combustion is oxidation chemical reaction occurring at fast rate, which is accompanied by light and heat emission. An initial moment of combustion, which occurred due to open flame action, is called inflammation (Kel'bert, 1971).

According to that, studies were conducted on «OVT» equipment using GOST 50810 (GOST50810-95) standards. The standard specifies a method of determination of textile materials (fabrics, nonwoven linens) capability to resist inflammation, steady combustion, as well as evaluation of their fire-resistant properties. For tests, specimens of 220×170 mm sizes were fabricated, 8 in a direction of warp (along the length) and 8 in a direction of weft (along the thikness).

Before testing, specimens were conditioned with temperature of (20±2) °C and relative humidity of (65±2) % during 24 hours. During tests from a surface, a burner is positioned in a horizontal direction 40 mm higher than lower edge of a specimen and moved to a specimen on 17 mm distance. During tests from an edge, a burner is positioned in 60°, in order to make flame touch lower edge of a specimen. Time of exposure for a new specimen is 5 sec. In a case of testing steady combustion, time of exposure to flame is increase to 15 sec.

Tests showed that the materials belong to flame-resistant fabrics.

Results of inflammability tests of analogue materials are presented in table 3.

Table 3. Results of the classification trials of materials (on warp)

	Name of the material			
Type of tests	Premier FR 350A	FlameFort W280 Protect	Brezent OP	FlameFort 210A
Time of ignition from a surface, sec	15	15	11	15
Time of ignition from an edge, sec	15	15	15	15
Time of independent combustion, sec	0	0	0	0
Burnout to an edge	No	No	No	No
Inflammation of cotton wool	No	No	No	No
Length of charred area, mm	19	12	35	22

Table 4. Results of tthe classification trials of materials (on weft)

Type of tests	Name of the material			
	Premier FR 350A	FlameFort W280 Protect	Brezent OP	FlameFort 210A
Time of ignition from the surface, sec	15	15	11	15
Time of ignition from an edge, sec	15	15	15	15
Time of independent combustion, sec	0	0	0	0
Burnout to an edge	No	No	No	No
Inflammation of cotton wool	No	No	No	No
Length of charred area, mm	25	19	42	27

Diagram of the test results is presented on Figure 2.

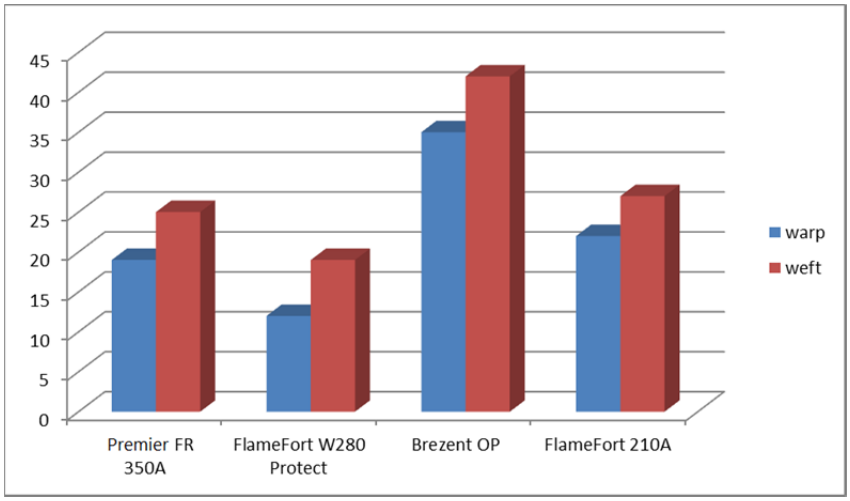


Figure 2. Resistance to burning along length of charred area

Results of the studies allowed to establish, that Brezent OP material has low resistance to burning, Premier FR 350A and FlameFort 210A didn't demonstrates signs of combustion.

After numerous experimental studies, an optimal compostion of layers of multilayer material, which are resistant to high temperatures, was determined.

Table 5.

No	Composition of layers of multilayer material	Surface density, g/m2	Fire-resistance (Oxygen index), %
1	FlameFortW280 Protect	430	
	Premier FR 350A	340	37.4
	Nonwoven material M-Sh (20)	200	
2	FlameFortW280 Protect	430	
	Premier FR 350A	340	38.0
	Nonwoven material M-Sh-M (25)	150	
3	FlameFort W280 Protect	430	
	Premier FR 350A	340	38.2
	Nonwoven material M-Sh-M (1.8)	140	
4	FlameFortW280 Protect	430	
	Premier FR 350A	340	38.1
	Nonwoven material M-Sh-M-Sh (30)	305	

Thus, optimal layers composition of multilayer fire-resistant materials and nonwoven linens are compositions #2, #3.

3.2 Melted Metal Splashes Testing

Resistance to splashes of melted metal and fire-resistance are main indicators of welder's and metallurgist's protective clothing's quality.

In textile materials and working clothing testing laboratories of personal protection equipment center in Dongguk University (South Korea) a number of studies of textile fire-resistant materials resistance to high temperature was conducted.

Impact of sparkles and melted metal splashes are of special interest. Testing method for tests of materials in a case of melted metal is specified in GOST12.4.237-2007 (ISO 9150:1988) (GOST12.4.237-2007). The method consists of counting number of splashes of melted metal.

10 specimens were fabricated with 120×20 mm size. Tested specimen was conditioned for at least for 24 hours with relative humidity of (65±2) % and (20±2) °C temperature.

In a case of testing according to ISO 9150 each material or multilayer material, used in working clothing, must sustain at least:

- not less than 15 splashes of melted metal with an increase of temperature of sensor behind a tested specimen on 40°C for 1 class;

- not less than 25 splashes of melted metal with an increase of temperature of sensor behind a tested specimen on 40°C for 2 class;

Obtained values of number of splashes of melted metal and temperature inside a multilayer material are presented in table 6. Weight of splashes varied from 10.1 to 12.8 g.

Table 6. Values of temperatures during an influence of melted meal on a multilayer material

No of multilayer material	Composition of layers of multilayer material	Temperature of an inner layer of multilayer material, °C	Temperature of inside of multilayer material, °C	Number of melted splashes
1	FlameFortW280 Protect Premier FR 350A Nonwoven material - M-Sh (20)	36.8	32	25
2	FlameFortW280 Protect Premier FR 350A Nonwoven material M-Sh-M (25)	36.8	29	25
3	FlameFort W280 Protect Premier FR 350A Nonwoven material M-Sh-M (1.8)	36.8	30	25
4	FlameFortW280 Protect Premier FR 350A Nonwoven material M-Sh-M-Sh (30)	36.8	28	25
5	FlameFort W280 Protect Premier FR 350A Nonwoven material M-Sh mixed (25)	36.8	31	25
6	FlameFortW280 Protect Premier FR 350A Nonwoven material M-Sh mixed (15)	36.8	36	25

Results of tests of multilayer materials are presented in Figure 3.

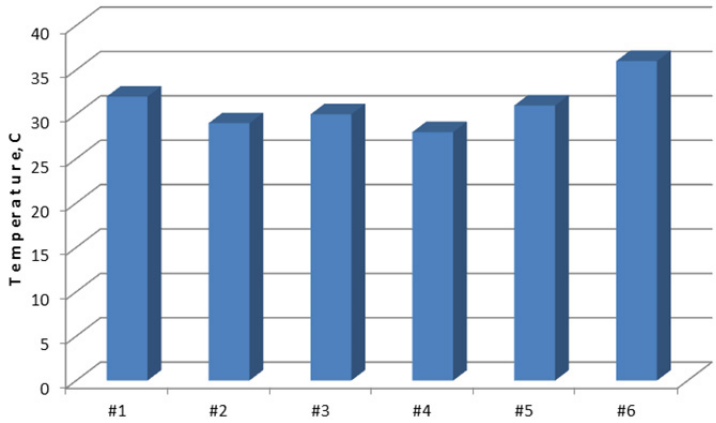


Figure 3. Temperatures of inside of specimens of multilayer materials during an influence of melted metal splashes

Analysis of results (figure 3) shows, that the specimen #4 possesses higher resistance to high temperatures.

4. Discussion

On the basis of the results of conducted studies new compositions of fire-resistant multilayer materials with improved performance are obtained.

General characteristic of selected composition of multilayer fire-resistant material is presented in Table 7.

Table 7. Characteristics of the optimal composition of multilayer material for working clothing

Properties and characteristics, units of measurement s	Values of properties and characterisitcs of the meaterial		
Name of the material	Premier FR 350A	FlameFortW280	Nonwoven material stitched with moleskine
Code number	10202 AM	50402 Kl4	-
Composition of fibers	100% cotton+ antistatic thread	100% aramide	(meta-aramide+ wool)
Surface density, g/m2	340	430	150
Weave	Satinlike5/2	Satinlike	Multilayer linen (4 layers)+1
Treatment	PyrovatexTo+HMBO	Kl4	-
Breaking load (base/weft)	1200 N / 700 N	1200 N / 1200 N	346 N / 205 N
Application	Upper layer	Piece	Inner stitched detachable layer with lining

In the designed working clothing as a lining material detachable, stitched layers of nonwoven material with moleskin are used.

Thus, it is proposed to design new multilayer material for a manufacturing of product using Premier FR 350 A as warp, nonwoven material stitched with moleskin as lining with additional protection from high temperatures and FlameFortW280 as additional protective fabric in a specified places (on front side of jumper, front side of pants, lower parts of sleeves).

For designing of working clothing, protecting from high temperatures and splashes of melted metal, new multilayer materials are proposed, which consists of fire-resistant clothing and newly designed nonwoven linens, produced from second grade wool and aramide fibers, which posses improved heat-resistance properties.

5. Conclusion

Nowadays, a lot of companies in Kazakhstan show a demand for more reliable types of working clothing, which

protect from sparkles and splashes of melted metal and meet other requirements of a customer.

That's why a supply of working clothing to Kazakhstan market is economically profitable and improvement of its quality is an important scientific and industrial problem.

At a pre-project analysis phase a problem was indicated - the lack of working clothes for welders on market, which would fully meet consumers requirements. Despite of variety of fabrics and materials with special fnishing for a protection from high temperatures on market, companies, producing working clothes for welders, do not completely use new materials because of high cost in a case of high performance, and, backwards, use fabrics with lower performance and lower cost.

Thus, the proposed multilayer material consists of: 3 layer nonwoven linen made from wool and aramide fibers and working as an additional fire-resistant lining, for the top part - protective material Premier FR350A, which is whole costume made from, for an additional protection FlameFortW280 Protect was used as protective pieces.

In Kazakhstan there are a number of researches in a field of new composite materials, which require further investigation.

Essential difference of the proposed solution from existing ones in that field consists of an implementation of domestic, natural wool fibers with m-aramid fibers, which allows to significantly decrease the product's cost and in the same time improve fire-protection and performance. In a case of a new implementation of wool fibers in a composition with m-aramid, it is possible to solve a problem of imports substitution in that field.

In the further research it is planned to increase a range of studies in a filed of testing and design of an optimal composition of multilayer materials for working clothing. Because the proposed solution is very simple and allows any interested company to organize mini-production of protective fire-resistant multilayer materials, it is planned to organize manufacturing of such materials.

References

Dahiya, A., Kamath, M., & Hegde, R. (2014). *Introduction to Nonwovens.* Textiles Introduction. Retrieved July 25, 2014, from http://web.utk.edu

Fabrics, K. (2005). Advanced protective materials for clothing, (2005 November 18). *Working clothing, 4*(31). Retrieved September 20, 2014, from http://www.termolin.com/catalogue.htm

Fomchenkova, L. (2002). Modern materials for working clothing and uniforms. *Textile industry, 7*, 15-17.

Golubaev, M. (2008). Problems of welder's working clothing design. *Working clothing, 4*, 43.

GOST. (2007). *12.4.237-2007 (ISO 9150:1988)-Occupational safety standards system. Protective clothing. Methods of testing the material on impact of splashes of molten metal.*

GOST. (2007). 50810-95-Fire hazard of textiles. Decorative textiles. Flammability test method and classification.

Gushina, K. (2004). *Working clothing.* Moscow: Consumer goods industry and food industry.

Kamath, M., Dahiya, A., & Hegde, R. (2004). *Needle Punched Nonwovens*. Textiles Introduction. Retrieved July 25, 2014, from http://web.utk.edu

Kel'bert, D. (1971). *Labour protection in textile industry.* Moscow: Consumer goods industry.

Mokeeva, N., & Pitzinskaya, O. (2013). *Features of welder's working clothing design.*

Nonwovens & Technical, T.(2007). *Indian textile.* Retrieved July 25, 2014, from indiantextilejournal.com.

Program of the development of consumer products industry in the Republic of Kazakhstan for 2010-2014. (n./d.)

Scott, R. (2005). *Textiles for Protection*. Woodhead Publishing Limited.

Sheromova, I. (2006). *Textile materials. Production, structure, properties: Textbook.* Vladivostok.

Suharev, M. (1973). *Material science.* Moscow: Consumer products industry.

Westex (USA) fire-protective materials. (2005). *Working clothing, 4*(31).

Working clothing and personal protection equipment. (2004). *3*, 23-28.

XM FIRELINE. (2014). Baltiski textile presents fire-resistant fabrics *Working clothing, 1*(6).

Zhilisbaeva, R., & Kozhabergenova, K. (2012). *Reliability problems in a design of working clothing for metallurgists.* Almaty: Almaty Korkem.

Zhilisbayeva, R. (2011). Peculiarities of Kinetics of New Fire Proof Fabrics Thermal Breakdown. *International Conference on Textile Engineering and Materials, China.*

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Comparative Evaluation of Fattening, Slaughter and Meat Qualities of Purebred and Hybrid Swine

Zhanna Aleksandrovna Perevoyko[1]

[1] Perm State Agricultural Academy named after academician D.N. Pryanishnikov, Russia

Correspondence: Zhanna Aleksandrovna Perevoyko, Perm State Agricultural Academy named after academician D.N. Pryanishnikov, 23Petropavlovskaia st., 614990 Perm, Russia

Received: October 29, 2014 Accepted: November 12, 2014 Online Published: July 31, 2015

doi:10.5539/mas.v9n8p344 URL: http://dx.doi.org/10.5539/mas.v9n8p344

Abstract

The creation of more productive breeds, lines, types, and their rational use in purebred breeding and hybridization contributes to the efficiency of swine industry. Comparative study on meat quality of pigs is essential for an objective assessment of the bred species, lines, and hybrids. We carried out a study on the meat quality of purebred and hybrid pigs fattening up to 100 kg and 125 kg of live weight under the conditions of an industrial complex. It was established, that of higher nutritional value, as well as of good fattening qualities, was the meat of two-breed hybrids LW × DL and three-breed combinations (LW × DL) × DIb both fattening to100 kg and fattening to125 kg of live weight.

Keywords: breed, breeding, hybridization, crossing, fattening qualities, meat quality, water binding ability

1. Introduction

1.1 Introduce the Problem

Shortage of animal protein in the diet of the population in the Russian Federation remains an important problem of food supply in the country. The main source of protein is meat. One of the most difficult and important problems to be addressed the agriculture of Russia is increase in production of meat, and first of all, pork as the most precocious livestock industry. Digestibility and bioavailability of swine meat in human's organism is 95%. The uniqueness of the pig meat is its high energy content, balanced amino acid composition of proteins, the presence of biologically active substances and its high digestibility, all of which ensures the normal physical and mental human's activities (Babailova et al., 2010).

The main raw material for the production of high-quality meat products supply in most countries of the world remains the pork. Muscle tissue of pigs in comparison with beef contains three times bigger amount of polyunsaturated fatty acids, eight times more vitamin B1, has a soft texture and pleasant aroma and taste, so the food value of the pork is very high (Birta, 2008; Kortz & Czarnecki, 1999).

Quality of pork is determined by the set of properties that provide nutritional value of products, taste and nutritional properties. Pork quality management is a very complex, laborious and multifactorial process. In this regard, focused selection remains the most effective way (Burmistrov & Pustovit, 2005, Okolyshev, 2008).

1.2 Explore Importance of the Problem

Priority direction in modern swine production is breeding pigs to increase the muscle mass. The main factor of improving the fattening and meat qualities of pigs is usage of the best selection options of specialized lines and breeds to produce hybrids.

1.3 State Hypotheses and Their Correspondence to Research Design

In young swine obtained as a result of two-and three-breed hybridization, the heterosis effect appears constantly. Two-and three- breed young animals have a higher precocity, adaptation flexibility and are better adapted to breeding at swine production complexes than purebred peers (DuninI et al., 2009; Kabanov & Titov, 2013; Ovchinnikov & Zatsarinin, 2013).

The difficulty of breeding to increase the quantity and improve the quality of the meat is that these rates negatively correlate with each other. Therefore, breeding pigs on meat must be accompanied by complex assessment of the quality of meat after slaughter of animals (Pogodaev & Peshkov, 2011, Rybalko et al., 2011).

Currently, requirements for quantity and quality of meat are growing. Therefore, methods of increasing the muscle mass in swine are of particular importance.

On this basis, the aim of our study was to examine the quality of the meat of pigs, produced by purebred and breeding of three-breed hybridization in fattening to 100 kg and 125 kg of live weight under conditions of an industrial complex.

2. Methods

All studies in the period of experiment were conducted using conventional methods.

2.1 Participant (Subject) Characteristics

To conduct the investigation under the principle of analogue groups dependent on the origin we formed 10 groups of yelts – 15 animals in each group on the following scheme: the first group – control group– LW × LW; 2^{nd}–10^{th} – experimental groups: 2^{nd} – LW × D, the 3^{rd} group – LW × DL, the 4^{th} group – LW × FL, the 5^{th} – LW × B b/w, the 6^{th} group (LW × DL) × D, the 7^{th} group (LW × DL) × Dib, the 8^{th} group (LW × DL) × MG, the 9^{th} group LW × (D × FL), the 10^{th} group LW × (FL × D), where LW – Large White, D – Duroc, DL – Dutch Landrace, FL – Finnish Landrace, Bb/w – Belarusian black-and-white, DIb – Duroc of the Irish breeding, MG – synthetic line of the Irish breeding.

The experience was conducted in the OJSC "Perm pig farm" in Permskii krai.

2.2 Sampling Procedures

During the studies all the animals were held in the same conditions of feeding and keeping. All animals in the groups were fed with specialized combined feed for animals.

2.3.1 Sample Size, Power, and Precision

The evaluation of slaughter qualities was conducted on the results of control slaughter of swine at reaching 100 and 125 kg of live weight with the relevant studies of carcasses. In each group of slaughtered animals a full deboning of half carcass was conducted to obtain fair presentation of meat quality of pigs.

2.3.2 Measures and Covariates

For the qualitative study of muscle tissue at slaughter of pigs we took samples of longissimusdorsi on a site between 9-12 scapula vertebraes, 400g. Chemical analysis of meat was conducted by conventional methods of zoo-technical analysis. In addition, we identified pH by potentiometric method, the water-binding ability – by Grau-Gram press method in the VNIIMP modification. The results obtained were processed by means of variation statistics method by E.K. Merkurieva in Microsoft Excel programme.

3. Results

Scientific and economic experience was held in the timeframe 2013-2014 with two-and three-breed hybrids young. Basic study on feeding, meat and meat qualities were held during the period of the fattening, within 100 days, and when the young pigs were send to a meat-processing plant.

3.1 Recruitment

Creation of highly productive hybrids or crosses involves, above all, improving the genetic potential of the source breeds, types, lines, consolidations of their basic productivity qualities, breeding for the compatibility. (Bekenev, 2012).

3.2 Statistics and Data Analysis

It is known that in swine of different breeds, the heritability of features such as average daily gain and feed consumption is quite high (average daily gain of 0.31 -0.77, feed consumption per 1 kg of gain – 0.2-0.5), fattening qualities have the average value of heritability, and meat qualities – high. It is therefore considered that the increase in the intensity of growth must be based on the relevant conditions of feeding and keeping, taking into account the formation of strong type of animals. (Zhebrovskii, 2002; Lesley, 1982; Bazhov, 2006).

The indicators that determine the intensity of swine husbandry are feed consumption per unit of gain, earliness and growth speed. Among two-breed crossing options, yelts LW × DL reached 100 kg of live weight in 182.1 days; that was the minimum age. They surpassed by precocity the purebred animals on 9.1 days, analogues of other two-breed combinations on 2.1-10.2 days (table 1).

Table 1. Fattening qualities of purebred and mixed young

Group	Hybridization variant	Indicator		
		Age of gaining live weight 100kg, days	Average gain a day, g	Feed consumption per 1 kg of gain, feed unit
I	LW × LW	191.2±1.48	674.7±5.94	4.22±0.11
II	LW × D	186.6±2.20	712.8±4.32***	3.87±0.03***
III	LW × DL	182.1±1.84**	735.8±6.21***	3.89±0.03***
IV	LW × FL	184.2±1.34**	724.3±4.44**	3.92±0.02***
V	LW × Bb/w	192.3±2.17	667.3±4.15	4.31±0.12
VI	(LW × DL) × D	172.4±1.18***	761.2±6.06***	3.73±0.14**
VII	(LW×DL) × DIb	168.5±2.11***	818.7±6.42***	3.68±0.13***
VIII	(LW×DL) × MG	167.8±1.92***	821.9±4.71***	3.62±0.15**
IX	LW ×(D×FL)	172.9±1.63***	764.6±5.32***	3.71±0.12***
X	LW × (FL×D)	173.4±2.08***	760.5±5.64***	3.77±0.13**

Note: - *P<0.05; ** P<0.01;*** P<0.001.

We established cross-group differences in the values of average daily weight gain. The mixed LW × DL had the highest intensity of growth from the beginning of the fattening until the weight of 100 kg among two-breed animals. The daily gains of body weight amounted to 735.8 g. Two-breed LW × DL exceeded purebred pigs on this indicator on 61.1 g or 9.1% (p > 0.999), two-breed peers – on 11.5-68.5 g or 1.5-10.3%.

The main indicator of efficiency of fattening pig is feed consumption per 1 kg of gain. In the cost structure of pork feed constitutes the major share (about 70%). It was found that two-breed animals LW × D, LW × DL and LW × FL differed by less feed consumption per 1 kg of gain compared to purebred animals of Large White breed. The difference was 0.35, 0.33 and 0.30 feeding units or 9.0; 8.5 and 7.7 % (P> 0.999), respectively. The two-breed LW × D hybrids were characterized by the minimum feed consumption – 3.87 feeding units.

The advantage of three-breed crossing compared to purebred breeding is distinguished through less feed consumption per gain in fattening offspring on 0.45-0.60 feeding units (P>0,999).

Fattening qualities of swine under experiment were also estimated by the quality and age of reaching 125 kg of live weight, the average daily weight gains and feed consumption per 1 kg of body mass gains (table 2).

The analysis of the data shows that the best qualities of feeding among two-breed crossed options were in mixed animals LW × DL. They reached the 125 kg of live weight 11.8 days earlier than the pure-bred peers, surpassed them on average daily gain on 57.8 g two-breed LW × DL and LW × FL had better pay for feed – 4.35 feed units.

The advantage of three-breed animals (LW × DL) × DIb over the purebred animals of Large White breed in age of reaching 125 kg of live weight constituted 16.7 and 17.7 days, in average daily gain – 85.0 and 90.8 g, in feed consumption –0.60 and 0.67 feeding units, respectively.

Main indicators of quality meat carcasses are: length, thickness and lard smoothness, muscle eye size and mass of the rear third of half carcass. These indicators are used for evaluating meat and slaughter qualities both in Russia and abroad.

Table 2. Fattening qualities of offspring reaching 125 kg of live weight

Group	Hybridization variant	Indicator		
		Age of reaching live weight 125 kg, days	Daily average gain, g	Food consumption per 1 kg of gain, feed unit
I	LW × LW	235.2±1.84	654.3±7.12	4.76±0.13
II	LW × D	229.6±2.11	680.5±6.43***	4.37±0.18
III	LW × DL	223.4±1.14***	712.1±7.12***	4.35±0.19
IV	LW × FL	226.3±1.18***	696.9±7.40***	4.35±0.11*
V	LW × Bb/w	236.5±1.23	648.5±8.12	4.87±0.17
VI	(LW × DL) × D	223.8±1.82***	710.0±8.82***	4.18±0.16**
VII	(LW×DL) × DIb	218.5±1.95***	739.3±8.92***	4.16±0.10**

VIII	(LW×DL) × MG	217.5±2.14***	745.1±7.14***	4.09±0.13**
IX	LW ×(D×FL)	224.5±1.80***	706.3±7.94***	4.19±0.20*
X	LW × (FL×D)	225.2±1.75***	701.1±8.01***	4.18±0.18*

Note: - *P<0.05; ** P<0.01;*** P<0.001

Meat qualities in swine are determined by a number of factors: conditions of feeding and housing, fatness, age and weight of an animal at slaughter, gender, breed, individual features (genotype), growth pace, what can be explained by the change of the organs and tissue growth character and exchange processes during the growth and development of an animal. In addition, meat qualities of swine in many aspects depend on breed, crossing variants of different genotypes, stress sensitivity and a number of other factors (Perevoiko, 2005; Matoušk et al., 1997; Mucami & Jamada, 1992).

Crossing sows of Large White breed with boars of meat breeds (Landrace and Duroc) and mixed boars is one of the ways to improve the meat qualities of the fattening young animals in conditions of industrial swine production. The degree of influence of meat breeds on the formation of the muscle mass in mixed livestock was confirmed by the results of slaughter and deboning of carcasses of swine when they reach 100 and 125 kg of live weight.

It is established that at slaughter in 100 kg, half carcasses of two-breed piglets were longer than of purebred (table 3).

Table 3. Meat qualities of offspring at slaughter in 100 kg

Group	Hybridization variant	Indicator			
		Half carcass length, cm	Fat thickness, cm	Muscle eye area, cm^2	Mass ham, kg
I	LW × LW	95.2±1.33	2.33±0.24	39.5±1.23	8.0±0.31
II	LW × D	94.9±2.14	2.27±0.15	44.7±1.12*	10.5±0.23***
III	LW × DL	103.9±1.68**	2.12±0.19	46.2±1.73*	10.9±0.23***
IV	LW × FL	98.2±1.86	2.25±0.22	45.3±1.73*	10.8±0.38***
V	LW × Bb/w	96.5±0.91	2.76±0.22	38.6±0.94	8.7±0.16
VI	(LW × DL) × D	98.9±1.69	2.18±0.17	60.0±4.36**	10.0±0.54*
VII	(LW×DL) × DIb	102.4±1.29**	1.96±0.25	63.5±4.63***	12.4±0.24***
VIII	(LW×DL) × MG	101.6±1.28*	2.10±0.08	61.4±2.88***	12.1±0.32***
IX	LW ×(D×FL)	99.4±1.50	2.24±0.11	59.7±2.73***	10.0 ±0.35*
X	LW × (FL×D)	100.4±2.50	1.96±0.19	58.5±2.21***	10.3±0.72*

Note: - *P<0.05; ** P<0.01;*** P<0.001.

The length of the half carcass of piglets in these groups ranged from 94.9 cm to 103.9 cm. Half carcasses of piglets LW × DL were the longest, they exceeded the purebred offspring LW × LW on the investigated indicator on 8.7 cm (P>0.99).

Therefore, the greatest increase in the length of the half carcass in the options of two-breed mating was observed when using landrace breeds of Dutch selection.

The analysis of the data shows that half carcasses of swine LW × DL had the smallest thickness of lard over 6-7 pectoral vertebrae of two-breed animals at slaughter in 100 kg, and amounted to 2.12 cm, that is lower than the figures of LW × DL on 0.21 cm.

One of the main indicators characterizing meat quality of pigs is muscle eye area. It was found that mixed animals LW × DL at slaughter in 100 kg exceeded the two-breed analogues and Large White breed on 0.9-7.6 and 6.7 cm^2 or on 1.98 -19.6 and 16.9%, respectively.

The back third of half carcass is the most valuable part in term of food. Therefore, the quality of carcass depends on the mass of the back ham and its morphological content.

Based on the obtained data, the substantial differences between the animals of the test and control groups can be noted. The greater weight of the back leg of two-breed crossing options were hybrids LW × DL –10.9 kg, what is

on 2.9 kg or 36.3% more than that of purebred animals (P > 0.999) and 0.1 -2.2 or 0.9 -25.3% than in two-breed analogues.

The evaluation of the half carcass of two-breed hybrids and purebred animals has shown that the most noticeable meat qualities were in two-breed animals LW × DL, LW × FL and LW × D with thin lard, large area of the muscle eye and ham mass.

In three-breed hybrids compared to purebred peers of Large White breed at slaughter in 100 kg, the half carcass is on 3.7-7.2 cm longer or on 3.9-7.6 % longer, muscle eye area – on 19.0-24.0 cm^2 or on 51.1-60.8 %, ham muss – on 2.0-4.4 kg or on 25.0-55.0 %, and lard is thinner on 0.09-0.37 mm or on 4-37.0 %.

At slaughter of fattening swine in 125 kg of live weight among two-breed crossing options, the carcasses of offspring LW × DL were characterized by thinner lard (2.39 cm), larger muscle eye area (48.3 cm^2); their parameters exceeded analogue indicators at slaughter in 100 kg on 12.7 and 4.5 %, respectively (table 4).

Table 4. Meat qualities of offspring at slaughter in 125 kg

Group	Hybridization variant	Indicator			
		Length of half carcass, cm	Lard thickness, cm	Muscle eye area, cm^2	Mass ham, kg
I	LW × LW	97.0±1.32	2.70±0.13	41.7±1.12	8.3±0.22
II	LW × D	97.3±2.14	2.56±0.14	46.8±1.32*	10.8±0.35**
III	LW × DL	106.6±1.24**	2.39±0.39	48.3±1.62*	11.2±0.34***
IV	LW × FL	101.0±1.52	2.40±0.16	47.7±1.71*	11.1±0.26***
V	LW × Bb/w	98.8±1.78	3.02±0.12	40.6±0.83	9.0±0.96
VI	(LW × DL) × D	102.0±1.58	2.00±0.15*	62.0±1.68***	10.0±0.83
VII	(LW×DL) × DIb	104.8±1.18**	2.23±0.18	65.5±1.62***	12.7±0.34***
VIII	(LW×DL) × MG	103.6±1.17*	2.35±0.21	63.4±1.58***	12.2±0.41***
IX	LW ×(D×FL)	98.0±1.61	2.40±0.21	61.7±1.63***	10.4 ±0.44**
X	LW × (FL×D)	105.0±1.49**	2.20±0.28	60.7±1.51***	10.6±0.62**

Note: - *P<0.05; ** P<0.01;*** P<0.001

Two-breed and three-breed hybrids at slaughter in 125 kg exceeded the analogues at slaughter in 100 kg on the length of half carcass, muscle eye area and ham mass; however, they were lower on lard thickness in average 0.23 cm or 9.7% (two-breed hybrids) and 0.15 cm on 7.2% (three-breed hybrids).

The most noticeable meat qualities at slaughter in 125 kg were three-breed animals (LW × DL) × DIb that had the longest half carcass – 104.8 cm, the largest area of the muscle eye – 65.5 cm^2 with the smallest thickness of back fat over 6-7 breast vertebra – 2.23 cm.

The slaughter output is a summarized indicator of swine slaughter qualities which decreases from 73-84 % to 62-71% in the first three months after birth, and in subsequent age periods increases to 78-81% (Kabanov V. et al,Titov I., 2013).

At slaughter of two-breed hybrids of 100 kg of live weight it was established that the mass of fresh carcass in hybrids LW × FL is higher in comparison to the analogues of the control group LW × LW on 0.9 kg or on 1.4% (table 5).

Table 5. Slaughter qualities of offspring at slaughter in 100 kg

Group	Hybridization variant	Weight at slaughter, kg	Weight of hot carcass, kg	Slaughter output without skin, %	Morphological content of half carcass, %		
					meat	fat	bones
I	LW × LW	101.2±2.36	64.2±1.41	62.8±2.23	56.2±3.79	32.5±1.50	11.3±0.21
II	LW × D	100.6±1.58	64.3±2.34	64.0±3.17	58.6±2.78	29.1±1.21	12.3±0.69
III	LW × DL	101.5±2.17	64.8±2.81	63.8±1.74	61.9±2.51	26.2±2.18*	11.9±0.25
IV	LW × FL	100.9±3.92	65.1±1.32	64.5±1.58	58.9±4.31	28.9±0.36*	12.2±0.35
V	LW × Bb/w	102.4±1.23	64.1±1.83	62.6±2.80	55.9±2.20	32.0±0.75	12.1±0.43
VI	(LW × DL) × D	100.3±7.66	69.1±5.53	68.9±0.95*	60.9±2.32	28.4±2.93	10.7±0.52

VII	(LW×DL) × DIb	102.0±0.49	70.5±0.90**	69.1±0.92**	61.4±2.64	27.5±1.72	11.1±0.56
VIII	(LW×DL) × MG	101.6±0.75	69.8±0.70**	68.7±0.23**	60.3±2.03	28.8±1.93	10.9±0.38
IX	LW ×(D×FL)	101.3±4.34	69.3±2.95	68.4±0.88*	60.6±0.65	28.4±0.44	11.0±0.30
X	LW × (FL×D)	100.4±4.76	68.4±4.41	68.1±1.32*	59.3±0.75	29.6±1.14	11.1±0.30

Note: - *P<0.05; ** P<0.01;*** P<0.001.

The largest slaughter output in two-breed combinations was obtained from crosses of the genotype – 64.5%. The two-breed hybrids LW × FL exceeded by its value purebred animals on 1.7%, peers LW × D, LW × DL and LW × B b/w on 0.5%, 0.7% and 1.9%, respectively.

Morphological composition of swine half carcass of investigated genotypes significantly differed in muscle, fat and bone tissues.

Two-breed hybrids LW × DL at slaughter in 100 kg were characterized by the highest output of meat in the carcass – 61.9%. Animals of Large White breed yield to them in the content of meat on 5.7%. By the highest content of body fat in carcass among two-breed hybrids, the animals LW × B b/w were distinguished –32.0%. The least amount of fat the carcass LW × DL – 26.2% contained, what was on 6.3% less than in purebred animals (P > 0.95).

Analyzing the morphological composition of carcasses, it can be concluded that two-breed animals differed from their purebred analogues by high content of meat and less fat.

Three-breed yelts at slaughter in 100 kg were superior to purebred peers of Large White breed by mass of fresh carcass on 4.8-6.9 kg or on 7.5-10.8%, by slaughter output – on 5.3-6.3%. The yelts (LW × DL) × DIb were characterized by the best slaughter qualities; their slaughter output constituted 69.1 % (P>0.99), the fat content in the carcass – 27.5 %. Then follow the animals (LW × DL) × D, (LW × DL) × MG. The comparative assessment of the compatibility of the breeds on the slaughter qualities indicates that the superiority on slaughter output, content of meat in the carcass and the least amount of fat in three-breed animals (LW × DL) × MG and (LW × DL) × DIb is determined by the genotype of the investigated breeds.

Three-breed hybrids exceeded the purebred animals of Large White breed on meat output on 4.3% and two-breed animals – on 1.7 %. At slaughter of three-breed and two-breed animals the least amount of crude fat was obtained (28.5 % and 29.1%, respectively), mating sows of Large White breed with meat boars leads to significant reduction of the fat content in meat.

There were no significant differences between the groups on the bone output at slaughter in 100kg. However, it should be marked, that in two-breed variants, the yelts LW × D and LW × FL were distinguished through the highest content of bones – 12.3 и 12.2 %, what is on 1.0 and 0.9 % more than the indicators of LW × LW.

The obtained data and its analysis prove that three-breed animals had the highest output at slaughter in 125 kg (LW × DL) × DIb – 72.2 %, then (LW × DL) × D – 71.1 %, then follow (LW × DL) × MG –70.7 % against 64.7 % in control (Table 6).

Based on the results of monitoring of control slaughter we established the superiority of three-breed hybrids (LW × DL) × DIb in mass of fresh carcasses at slaughter in 125 kg to 8 kg or 9.8% compared with the control. The highest content of meat in half carcass at slaughter in 125 kg was in hybrids (LW × DL) × DIb (59.5 %) what is higher than in the control group on 6.2%.

Table 6. Slaughter qualities of young pigs at slaughter in 125 kg

Group	Hybridization variant	Weight at slaughter, kg	Weight of half carcass, kg	Slaughter output without skin, %	Morphological content of half carcass, %		
					meat	fat	bones
I	LW × LW	125.3±1.28	81.3±2.41	64.7±1.23	53.3±2.67	35.6±1.47	11.1±0.35
II	LW × D	125.0±1.60	82.0±1.34	65.6±1.26	55.0±2.51	33.3±1.18	11.7±0.52
III	LW × DL	125.0±1.89	82.4±1.71	65.9±1.63	57.7±2.71	31.1±1.18	11.2±0.28
IV	LW × FL	125.3±1.82	81.7±1.32	65.2±1.47	57.0±2.24	31.2±1.47	11.8±0.48
V	LW × Bb/w	125.1±1.27	81.2±1.73	64.9±1.78	50.9±2.25	37.3±2.64	11.8±0.26
VI	(LW × DL) × D	125.3±2.56	89.1±2.42*	71.1±1.85*	59.1±2.68	30.4±2.82*	10.5±0.43
VII	(LW×DL) × DIb	125.4±1.37	89.3±2.19*	72.2±1.32**	59.5±2.54	29.6±2.62	10.9±0.67

VIII	(LW×DL) × MG	125.6±1.75	88.8±2.17	70.7±1.43**	58.7±2.81	30.6±4.53	10.7±0.32
IX	LW ×(D×FL)	125.0±2.24	86.9±2.95	69.5±1.78*	58.0±2.51	31.2±2.62	10.8±0.84
X	LW × (FL×D)	125.0±1.56	87.0±2.31	69.6±1.33*	57.5±2.75	31.6±2.25	10.9±0.57

Note: - *P<0.05; ** P<0.01;*** P<0.001

On the content of fat in carcass, the animals of Large White breed exceeded three-breed analogues on 4-6 % in average.

For the comparative assessment of carcasses at slaughter an illustrative criterion is the ratio of the tissue: meat/bones – fleshing index and the meat/fat – leanness index (table 7).

Table 7. Carcass quality evaluation on correlation of tissue at slaughter in 100 and 125 kg

Group	Hybridization variant	Index			
		Fleshing (meat/bones)		leanness (meat/crude lard)	
		at slaughter			
		in 100 kg	in 125 kg	in100 kg	in125 kg
I	LW × LW	4.97	4.8	1.73	1.49
II	LW × D	4.75	4.70	2.01	1.65
III	LW × DL	5.20	5.15	2.36	1.85
IV	LW × FL	4.83	4.91	2.04	1.85
V	LW × Bb/w	4.62	4.34	1.74	1.37
VI	(LW × DL) × D	5.69	5.63	2.14	1.94
VII	(LW×DL) × DIb	5.53	5.46	2.23	2.01
VIII	(LW×DL) × MG	5.53	5.48	2.09	1.92
IX	LW ×(D×FL)	5.51	5.37	2.13	1.86
X	LW × (FL×D)	5.34	5.28	2.00	1.82

Note: - *P<0.05; ** P<0.01;*** P<0.001

The findings prove that at slaughter in 100 kg on the fleshing index and on the leanness index the leaders were two-breed animals LW × DL and three-breed (LW ×DL) × D and (LW ×DL) × DIb. It is significant that with increased live weight at slaughter leanness and muscle mass indices reduce; it means that the animal carcasses at fattening swine up to high grade become lardy.

A completed study of the chemical composition of the longissimus dorsi in young pigs revealed that indicators such as moisture, fat, protein content depend on the genotype of the pigs and slaughter terms.

An important component of the meat is water. On the basis of the received data, in the sample of meat animals with increasing weight at slaughter we observed the reduction of total moisture content.

The analysis of physical and chemical properties of meat revealed that in young pigs of all experimental groups in the sample of meat at slaughter of 125 kg, moisture content was less than in young pigs with the weight before slaughter 100 kg of similar groups, on 1.3; 1.2; 1.1; 1.0; 0.8; 1.1; 1.1; 1.2; 1.3 and 1.0%, respectively (tables 8, 9).

Table 8. Chemical composition of meat in two- and three-breed crossing at slaughter in 100 kg

Group	Hybridization variant	Indicator					
		total moisture, %	dry matter, %	fat, %	protein, %	waterbinding ability, %	pH
I	LW × LW	72.1±0.54	27.9±0.54	4.90±0.18	22.4±0.26	61.0±0.80	5.65±0.11
II	LW × D	72.6±0.33	27.4±0.33	3.70±0.13***	23.1±0.22	65.7±0.60**	5.65±0.05
III	LW × DL	72.9±0.61	27.1±0.61	3.61±0.18***	23.1±0.54	65.3±0.91**	5.58±0.02
IV	LW × FL	72.7±0.42	27.3±0.42	3.72±0.12***	23.5±0.81	65.7±0.33***	5.61±0.01
V	LW × Bb/w	70.6±0.18	29.4±0.18	4,17±0.22*	22.0±0.41	62.7±0.43	5.63±0.08
VI	(LW × DL) × D	72.8±0.62	27.2±0.62	2.59±0.12***	24.1±0.38*	66.0±0.56***	5.69±0.22

VII	(LW×DL)× DIb	73.2±0.61	26.8±0.61	2.50±0.18***	24.8±0.54*	67.3±0.91**	5.71±0.02
VIII	(LW×DL)× MG	73.6±0.33	26.4±0.33	2.55±0.13***	24.5±0.22**	67.7±0.60**	5.75±0.05
IX	LW ×(D×FL)	73.3±0.76	26.7±0.76	2.57±0.13***	24.1±0.59	66.7±0.88**	5.83±0.20
X	LW × (FL×D)	73.4±0.66	26.6±0.66	2.58±0.22***	24.4±0.50	66.3±0.28*	5.79±0.03

Note: - *P<0.05; ** P<0.01;*** P<0.001

Table 9. Chemical composition of meat in two- and three-breed crossing at slaughter in 125 kg

Group	Hybridization variant	Indicator					
		total moisture, %	drymatter, %	fat, %	protein, %	waterbinding ability, %	pH
I	LW× LW	70.8±0.14	29.2±0.14	5.73±0.18	22.6±0.12	59.8±0.78	5.45±0.12
II	LW × D	71.4±0.19	28.6±0.19	4.33±0.21***	23.5±0.10	64.4±0.53**	5.48±0.07
III	LW × DL	71.8±0.20	28.2±0,20	4.26±0.19***	23.6±0.16	63.9±0.85**	5.48±0.11
IV	LW × FL	71.7±0.15	28.3±0.15	4.39±0.12***	23.9±0.17	64.7±0.52***	5.42±0.05
V	LW × Bb/w	69.8±0.25	30.2±0.25	4.97±0.18**	22.4±0.18	61.4±0.62	5.44±0.12
VI	(LW × DL) × D	71.7±0.17	28.3±0.17	3.03±0.18***	24.5±0.15	64.7±0.61**	5.49±0.09
VII	(LW×DL)×DIb	72.1±0.24	27.9±0.24	2.93±0.23***	25.3±0.18	65.9±0.74***	5.46±0.07
VIII	(LW×DL)×MG	72.4±0.29	27.6±0.29	3.01±0.14***	24.9±0.22	66.3±0.58***	5.44±0.08
IX	LW ×(D×FL)	72.0±0.30	28.0±0.30	3.03±0.19***	24.6±0.19	65.4±0.68***	5.48±0.11
X	LW× (FL × D)	72.4±0.20	27.6±0.20	3.02±0.18***	24.9±0.20	64.9±0.46***	5.44±0.06

Note: - *P<0.05; ** P<0.01;*** P<0.001

The muscle tissue of three-breed hybrids (LW×DL) × MG had the highest amount of moisture in all the studied periods. Their superiority over their peers in the control group by size of the studied indicator was at slaughter at 100 kg – 1.5%, at slaughter at 125 kg –1.6%.

Fat content has a significant impact on the nutritional value of meat, because it has a high energetic value and gives the pleasant taste to meat products. The fat content of 3.8-4.0% is considered as normal, and allows the meat to save its technological properties and high qualities.

The lowest fat content found in two-breed hybrids LW × DL – 3.61 and 4.26% and three-breed hybrids (LW×DL) × DIb – 2.50 and 2.93% at slaughter at 100 and 125 kg respectively, what is 1.29 and 1.47% less compared to purebred animals at slaughter in 100 and 125 kg and 2.4 and 2.8% for three-breed hybrids. Therefore, meat of the hybrids LW × DL and (LW × DL) × DIb is quite leanness.

The protein content in the meat of young animals of all groups increased slightly with age. The three-breed hybrids (LW × DL) × DIb were characterized by the highest amount of protein and fat. At slaughter at 100 kg of live weight, two-breed hybrids LW × D, LW × DL and LW × FL were superior to their peers of Large White breed on the content of protein in the average of 0.83%, at slaughter in 125 kg to 1.07% and yield by quantity of fat on 1.22 and 1.40%, respectively.

One of the main indicators of the meat is water-binding ability to illustrate the ability of muscle proteins to hydration. High content of bound water in meat testifies to its juiciness and best technological properties. It is known that water-binding ability of pork is divided into 3 categories: high (67% and more), normal (53-66 %) and low (52% and less) (Rybalko V. et al, 2005).

It was established that the content of bound water in meat of purebred and hybrid young pigs was in the norm – 61.0-67.7 %.Therefore, the meat of these animals meets the requirements for normal qualities of pork. The highest water-binding ability of meat was observed in hybrids (LW× DL) × MG – 67.7% at slaughter in 100 kg and 66.3% at slaughter in 125 kg, which was significantly higher than in the animals in the control group on 6.7% (P > 0.95) and 6.5%.

One of the main indicators of the quality of meat is pH. pH indicates the glycolysis degree in muscle tissue, that is suitable for the storage of meat and cooking. Hydrogen-ion concentration in meat depends on the content of

glycogen and lactic acid in the muscles at the time of slaughter, and as a result, is a derivative of the physiological condition of the animal before slaughter, and also displays the flow of after-slaughter processes in the carcass(Sheiko I.P. et al., 2008, Sevčílková S. et al, 2002).

In meat of healthy animals in 12 hours after their slaughter, pH ranges from 5.6 to 7.2.The normal value of pH is from 5.4 to 6.3. In our research, active acidity of meat of young pigs in all experimental groups was within the rates and was at slaughter in 100 kg 5.65 -5.83 units; at slaughter in 125 kg -5.42 -5.48 units.

4. Discussion

Comprehensive assessment of fattening qualities of pigs with different genotypes indicates the efficiency of hybridization of sows of Large White breed with boars of Dutch Landrace and Finnish Landrace breeds and also sows LW × DL with Irish Duroc boars and with synthetic MG boars. It allows increasing production of pork with high consume qualities.

Summarizing of obtained data on three-breed hybridization with hybrid boars and sows using shows that advantage when using of hybrid sows in comparison with hybrid boars reveals in increasing their daily average growth while fattening by 57.7 g in average in group. Offspring of hybrid sows has an advantage over their contemporaries from hybrid boars in early ripeness and feed consumption per 1 kg of growth in average in group 5 days and 0.09 feeding units, respectively.

Breeding of LW × DL pigs were notable for good early ripening (182.1 days), high daily average growth (735.8 g), lower feed consumption per 1 kg of growth among two-breed variants. Hybrids (LW × DL) x MG reached 100 kg of live weight at younger age and had maximum daily average growth – 167.8 days and 821.9 g among three-breed variants.

Three-breed hybrids excel purebred Large White pigs in meat receiving by 4.3 % and two-breed hybrids – by 1.7 %. Three-breed and two-breed hybrids gave the lowest amount of fat: 28.5 % и 29.1 %, respectively.

The two-breed LW × DL and three-breed (LW × DL) x D and (LW × DL) x DIb animals led by fleshing and leanness indices at slaughter in the 100 and 125 kg.

Therefore, obtained results of research are evidence of that two-breed hybrids LW × DL and three-breed hybrids (LW × DL) × DIb had higher meat productivity and good fattening and slaughtering qualities both as at fattening up to 100 kg, as to 125 kg of live weight.

5. Conclusions

1. The animals LW× DL had the highest precocity among two-breed hybrids; reaching 100 kg they surpassed their peers from other experienced groups on 2.1 – 10.2 days. Among three-breed hybrids, animals (LW × DL) × MG had the highest precocity, they surpassed their peers on 0.7 – 5.6 days. A similar trend was observed when animals reached 125 kgoflive weight.

2. The highest slaughter output was obtained from two-breed hybrids LW×FL and amounted to 64.5% at slaughter in 100 kg, which is higher than from two-breed hybrids of other groups on 0.7-1.9%.Themass of hot carcass in two-breed hybrids LW×FLwas higher compared with their peers in other test groups on 0.3-1.0 kg. In three-breed hybrid gilts (LW × DL) × Dib were characterized by the best slaughter qualities, their slaughter output accounted for 69.1%, and mass of hot carcasses – 70.5 kg.

3. The use of two-breed and three-breed hybridization contributed improvement of meat qualities: the muscle eye area in gilts LW × DL amounted to 46.2 cm^2, what is on 0.9-7.6 cm^2 higher than in their peers, and the hock mass amounted to 10,9 kg, what is in compare to their peers 0.1 – 2.2 kg, respectively. In gilts (LW × DL) × DIb, the muscle eye area amounted to 63.5 cm^2, what is on 2.1-5.0 cm^2 higher than in their peers, and hock mass amounted to 12.4 kg, what is on0.3 – 2.4 kg higher that in their peers.

References

Babailova, G. P., Zhdanov, S. L., & Dubinin, A. A. (2010). Chemical composition of meat of large white pigs breed. Materials of the international scientific-practical conference dedicated to the 80th anniversary of the Vyatka State Academy, 11- 14.

Bazhov, G. M. (2006). Breeding swine production. *Saint Petersburg,* p. 384.

Bekenev, V. A. (2012) Technologies of breeding and keeping swine. *Saint Petersburg,* p. 416.

Birta, G. (2008) Meat and lard quality of pigs of different breeds. *Swine production, 5*, 11-12.

Burmistrov, V., & Pustovit, I. (2005). Physical and chemical composition of muscle and adipose tissue in pigs of

different genotypes. *Swine production, 2*, 14-16.

DuninI, M., Garai, V. V., & Pavlova, S. V. (2009). State, development strategy of pedigree base and hybridization system in pig-breeding in Russia. *Zootechny, 1*, 4-8.

Kabanov, V., & Titov, I. (2013). Yopkshire, Landrace, Duroc and hybrids. Russian livestock. Special issue, p. 19.

Kortz, J., & Czarnecki, R. (1999). Effects of genetic line on the quality of meat produced by purebred Polish large white (PLW) swine and by PLW crosses with Czech genetic lines. *Adv. Agr. Sci., 6*(2), 93-98.

Lesley, J. F. (1982) Genetic base of agriculture animal selection. *Publishing House Kolos,* p. 391.

Matoušk, V., Kernerová, N., & Václavovsý, J. (1997). Analýzakvalitymasa u hybridnípopuaceprasat. *Zivoc. Zyroba, 42*(11), 511-515.

Mucami, H., & Jamada, J. (1992). Economical efficiencies of crossbreeding systems in pigs. *Bulletin of National Institute of Animal Industry, 38*, 5–19.

Okolyshev, S. (2008). Meat-lard qualities of swine. *Animal husbandry of Russia, 4*, 43- 45.

Ovchinnikov, A. V., & Zatsarinin, A. A. (2013). Fattening and slaughtering qualities of pigs of different genotypes when fattening up to high weight standards. *Zootechny, 2*, 18-20.

Perevoiko, Zh. A. (2005). Fattening and meat quality of purebred and crossbred swine fattening up to 100 kg. Proceedings of Russian scientific conference of scientists and specialists. pp. 169-172.

Pogodaev, V. A., & Peshkov, A. D. (2011). Quality of muscle and adipose tissue in purebred and hybrid pigs. *Swine Production, 4*, 24-26.

Rybalko, V. P., Bankovska, I. B., & Getia, A. A. (2005). Meat quality management under the intensive breeding of swine. *Agricultural Journal,* 4-5, 28-29).

Rybalko, V., Birta, G., & Burgu, I. U. (2011). The quality of pork meat in different genotypes. Swine production.No. 6. pp.23-25.

Sevčílkoví, S., Koucký, M., & Laštovková, J. (2002). Meat performance and meat quality in different genotypes of F1 generation gilts.*Czech J. Anim. Sci., 47*(9), 395-400.

Sheiko, I. P., Fedorenkova, L. A., & Chramchenko, N. M. (2008). Comparative assessment of quality of pork in young animals of different genotypes. *Swine Farm, 3*, 10-13.

Zhebrovskii, L. S. (2002). Animal selection. SPb.: Publishing house Lan, p. 256. Saint-Petersburg.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

A New Method for Detecting Cerebral Tissues Abnormality in Magnetic Resonance Images

Mohammed Sabbih Hamoud Al-Tamimi[1,2] & Ghazali Sulong[1]

[1] UTM-IRDA Digital Media Centre (MaGIC-X), Faculty of Computing, Universiti Teknologi Malaysia, Malaysia

[2] Department of Higher Studies, University of Baghdad, Al-Jaderia, Baghdad, Iraq

Correspondence: Ghazali Sulong, UTM-IRDA Digital Media Centre (MaGIC-X), Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia. Tel: 60-177-467-128. E-mail: ghazali@utmspace.edu.my/m_altamimi75@yahoo.com

Received: December 9, 2014 Accepted: January 14, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p354 URL: http://dx.doi.org/10.5539/mas.v9n8p354

Abstract

We propose a new method for detecting the abnormality in cerebral tissues present within Magnetic Resonance Images (MRI). Present classifier is comprised of cerebral tissue extraction, image division into angular and distance span vectors, acquirement of four features for each portion and classification to ascertain the abnormality location. The threshold value and region of interest are discerned using operator input and Otsu algorithm. Novel brain slices image division is introduced via angular and distance span vectors of sizes 24˚ with 15 pixels. Rotation invariance of the angular span vector is determined. An automatic image categorization into normal and abnormal brain tissues is performed using Support Vector Machine (SVM). Standard Deviation, Mean, Energy and Entropy are extorted using the histogram approach for each merger space. These features are found to be higher in occurrence in the tumor region than the non-tumor one. MRI scans of the five brains with 60 slices from each are utilized for testing the proposed method's authenticity. These brain images (230 slices as normal and 70 abnormal) are accessed from the Internet Brain Segmentation Repository (IBSR) dataset. 60% images for training and 40% for testing phase are used. Average classification accuracy as much as 98.02% (training) and 98.19% (testing) are achieved.

Keywords: abnormal brain tissues, brain slices images, cerebral tissues, image classification, magnetic resonance imaging, statistical features.

1. Introduction

Lately, MRI is employed in diverse medical fields, including heart diseases, cancer research and brain diseases (El-Dahshan et al., 2009). It creates high-quality two or three dimensional images of an object to accurately visualize and detect the brain tumors.

Furthermore, the images provide valuable and detailed information regarding normal and abnormal tissues. Brain tumor being the most common and deadly brain diseases can affect and devastate many lives. According to International Agency for Research on Cancer (IARC) report, over 126000 people having brain tumor are diagnosed worldwide per year with more than 97000 mortality rate (Al-Tamimi & Sulong, 2014b; Ferlay et al., 2010). At present, MRI is the most common test for diagnosing and confirming the presence of brain tumor. It identifies the tumor location for recommended specialist treatment options (Horská & Barker 2010; Al-Tamimi & Sulong, 2014b; Al-Tamimi & Sulong, 2014a).

The diagnostic flaws needs considerable reduction through precise diagnosis and detection of the brain abnormality related to the type, size and effected area to aid appropriate treatment planning. Computer Aided Diagnosis (CAD) systems are introduced to enhance the diagnostic accuracies (Roy et al., 2013; Al-Tamimi & Sulong, 2014b; Al-Tamimi & Sulong, 2014a). In fact, provides CAD output of the computer system second opinion to help the radiologists image interpretations by considerably reducing the image reading time. These computer based auto detection systems often help in improve the accuracy of radiation analyses (Selvanayaki & Karnan, 2010).

In the medical field, image processing and analysis is of great significance, particularly in noninvasive treatment and clinical study (Spisz & Bankman, 2000). Medical image processing emerged as one of important and

reliable methods in identifying and diagnosing different tissues disorders and abnormalities. Imaging diagnostics assist the doctors/radiologists to visualize, analyze and understand the nature of abnormalities in internal organ structures for specific diagnosis and treatment (Veloz et al., 2011). MRI possessing notable advantages of non-invasiveness and soft tissue contrast remains one of the widely-used techniques for brain imaging (Gang et al., 2013; Sikka et al., 2009).

Undoubtedly, the brain being the command centre is the key organ responsible for controlling diverse functions related to body movement, blood flow, cardiac activities, temperature of the body, emotional behaviour, learning and memorization, to cite a few. Despite intensive research, the classification of MRI data in terms of normal and abnormal brain tissues for performing fast and accurate diagnostics remains challenging. Segmentation of brain for abnormality detection in slice image is the most a daunting task due to its complex anatomy and problems inherent to the nature of the image. The heterogeneous and diffuse manifestation of pathology in medical images often prohibits the employment of computational methods. Furthermore, all imaging modalities carry limitations and artifacts which must be addressed and solved by segmentation methods. In this study the brain images are selected for the image reference as the injuries or abnormalities which can affect the large part of an organ.

Formerly, different methods used to classify the MRI data are divided into two categories. First one includes the supervised learning techniques such as SVM (Salankar & Bora, 2014), Artificial Neural Network (ANN) (Joshi et al., 2010; Ibrahim et al., 2013) and K-Nearest Neighbor (KNN) (Al-Badarneh et al., 2012; Rajini & Bhavani, 2011; Badran et al., 2010). The other one containing unsupervised learning methods as Self-Organizing Map (SOM) (Goswami & Bhaiya, 2013), k-means Clustering (Deepak et al., 2013), Statistical Images and Analysis (SIA) (Technologies, 2014) are mainly used for data clustering.

We use SVM to classify MR images based on normal and abnormal brain tissues. The proposed new method combines four steps such as extraction of the cerebral tissues, their classification into angular and distance spans, determination of four features for each portion and identification of the abnormality location. A Brain slice image division using angular and distance span is introduced by moving one portion of the vector entries clockwise. Rotation invariance of the angular span vectors of the sizes 24˚ and 15 pixels is applied. The results are analysed, compared and understood.

2. Brain Anatomy

The human brain is the command centre that controls all body parts and allows in adapting varying environmental conditions. The brain as displayed in Figure 1 is composed of two types of tissues called Gray Matter (GM) and White Matter (WM). GM consists of neuronal and glial cells often termed as neuroglia or glia which controls brain activity. The basal nuclei called GM nuclei are located deep within the WM. The basal nuclei contain caudate nucleus, putamen, pallidum and claustrum. WM fibers consist of elinated axons which connect the cerebral cortex to other parts of the brain. The left and the right hemispheres of the brain are connected by the corpus callosum, which is a thick band of WM fibres (Noback et al., 2005; Al-Tamimi & Sulong, 2014b).

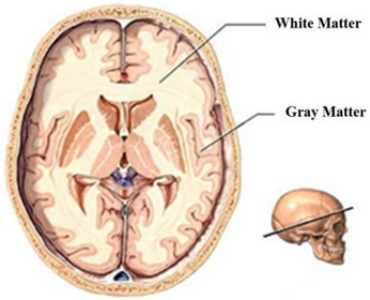


Figure 1. Grey Matter and White Matter Tissues (Nolte, 2013)

The Cerebrospinal Fluid (CSF) in the brain as shown in Figure 2 is comprised of enzymes, glucose, salts and white blood cells. Cerebrospinal fluid circulates through ventricles around the brain and the spinal cord to protect from injuries. Another type of tissues called meninges that present in form of membrane to cover the brain and spinal cord (Noback et al., 2005; Al-Tamimi & Sulong, 2014b).

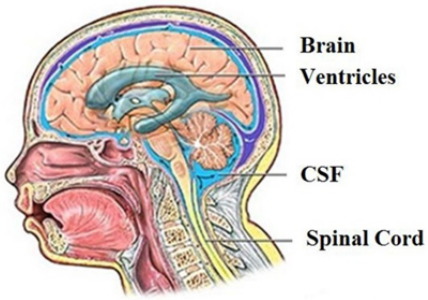


Figure 2. Normal circulation of CSF in the brain (Nolte 2013)

Figure 3 illustrates the brain anatomy which consists of brain stem and cerebrum. The cerebrum inhibits the largest part of the brain and responsible for the movement, conscious thoughts and sensations. The cerebrum is divided into left and the right hemispheres. Each cerebrum hemisphere controls the opposite side of the body and each half is further partitioned into four parts including frontal, temporal, parietal and occipital lobes. The cerebellum being the second largest structure of the brain is connected to the controlling motor functions of the body, such as balance, walking, posture and the general motor coordination. It is located at the rear side of the brain and attached to the brain stems. Both, cerebrum and cerebellum possess ultra-thin outer cortex of gray matter, internal white matter and small but deeply situated masses of the gray matter. The spinal cord is joined to the brainstem and positioned towards the brain bottom. The brainstem controls various fundamental functions such as motor, sensory pathways, cardiac, repository and reflexes. It is divided into the midbrain, pons and medulla oblongata membrane (Noback et al., 2005; Al-Tamimi & Sulong, 2014b).

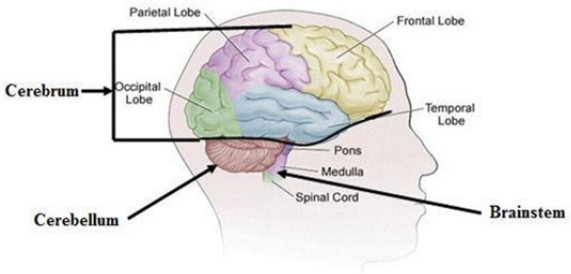


Figure 3. Major subdivisions of the human brain (Nolte 2013).

3. Related Work

Radiologists analyze the MR images by visual inspection to detect and determine if a tumor or abnormal tissue. The huge number of such images makes this visual interpretation process expensive and often erroneous. Furthermore, the sensitivity of the human eye and brain to elucidate such images reduces with the increase of number of cases, especially when only a small number of slices contain information of the affected area. Therefore, an automated system for analysis and classification MR images is essential.

Practically, MR images include both normal and defective slices. Firstly, these defective or abnormal slices are detected and separated from the normal slices. Secondly, these abnormal slices are further examined to identify the exact nature and location of tumors. Following slice segmentation with fuzzy c-means (FCM) algorithm, Clarke et al., (Clark et al., 1994) identified the abnormality in MR images. Antonie (Antonie, 2008) introduced a method for automated segmentation and categorization of brain MR images using SVM classifier. The normal and abnormal slices are identified based on statistical features. However, recent investigations in data classification reveal that Least Squares Support Vector Machines (LS-SVMs) approach is highly capable of providing higher sorting accuracy compared to existing algorithms (Cawley & Talbot, 2004; Improved et al., 2005).

The likelihood of detecting a premature dementia without using rigid registration of MRI is established (Klein et al., 2010). Based on the dissimilarity matrix, a k-nearest neighbors (k-NN) classifier is developed. The efficiency and performance of the classifier are tested in a leave-one-out experiment on 58 images. This method achieves

an efficiency of 81%. Hybrid techniques consisting of three steps including feature extraction via Discrete Wavelet Transform (DWT), reduce the dimensions size by Principal Component Analysis (PCA) and classification of the outputs using two classifiers are proposed (El-Dahshan et al., 2010). The classifiers are based on ANN and k-NN. The dataset comprised of T2-weighted, axial dimension, 256×256 pixels, image size 70 (with 10 normal and 60 abnormal) are employed. Remarkably, the number of extracted features is reduced from 1024 to seven using PCA. Accuracy as much as 97% and 98% are achieved from DWT+PCA+ANN and DWT+PCA+k-NN, respectively.

In the past, MR brain images are classified using ANN and SVM method (Chaplot et al., 2006). The pre-processing phase involving DWT is used as input for Neural Network (NN) and SVM. The dataset consisting of T2-weighted, axial, 256×256 pixels MRI, images size 52 with 46 for abnormal (marked by Alzheimer's disease) and 6 for normal are applied, where 4761 features are extracted. The achieved accuracy of the classifier DWT+SOM is 94%, DWT+SVM with linear kernel is 96.15%, DWT+SVM with polynomial kernel is 98.00% and DWT+SVM with radial basis function based kernel is 98.00%. An automatic classification of MR images for normal or abnormal tissues is proposed (Ariffanan & Basri, 2008). This classifier follows two steps such as feature extraction by PCA and classification by the neuro fuzzy. Using an input dataset of size 35 (with 20 as training set and 15 as testing set) the accuracy of 93.33% is achieved.

4. Methodology

Our proposed method comprised of four steps, including the extraction of cerebral tissues, division of cerebral tissues into angular and distance span, attainment of four features for each portion and finally the use of classification to detect the abnormality location. Figure 4 represents the flowchart of the proposed method. The slice image of MR representations is used to extract the brain tissues from non-brain one by skull stripping. Threshold value is automatically selected via Otsu's algorithm (Otsu, 1979). Following angular and distance span, new brain slice image division is introduced. The divided brain MR images are employed as input in the feature extraction algorithm. Four statistical features are calculated for each merger space. The features pattern vectors obtained from the divided images are utilized as input of SVM classifier. Five brain MR images with each image containing 60 slices are used for testing and validation. Total of 230 slices are considered as normal brain images and rest 70 slices are taken as abnormal brain images. Sixty percent MR brain slice images are utilized for the training phase and remaining 40% are used for the testing phase.

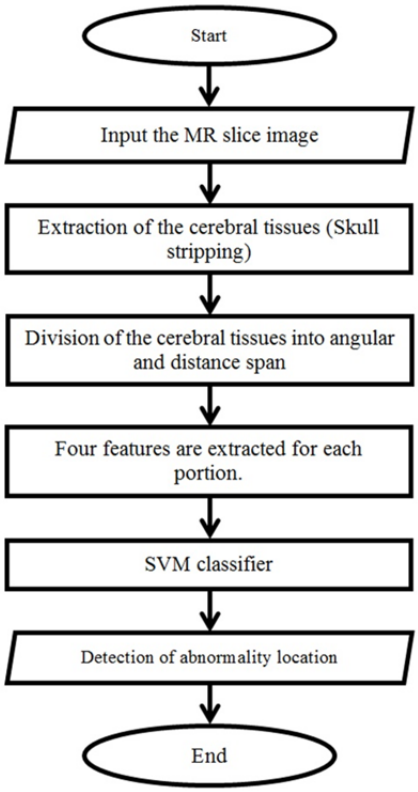


Figure 4. Flowchart of the proposed method.

4.1 Skull Stripping

Skull stripping is one of the significant phases in neurology, where brain tissues are extracted from non-brain tissues within MR images. Several skull stripping methods are developed in the past (Zhang et al., 2011; Sadananthan et al., 2010; Ségonne et al., 2004). Skull stripping algorithms are generally categorized into four groups based on morphological, deformable surface, atlas and hybrid nature (Ségonne et al., 2004). In usual practice the skull stripping is performed using threshold intensity (Ségonne et al., 2004; Sadananthan et al., 2010).

The proposed skull stripping method consisting of three steps is shown in Figure 5. Firstly, the image binarization is performed using threshold value via Otsu's automatic selection algorithm. Secondly, the largest connected portion of the binarized image is chosen, assuming skull as one of the major parts surrounding the head. Thirdly, the skull stripped brain image is obtained after extracting the cerebral tissues.

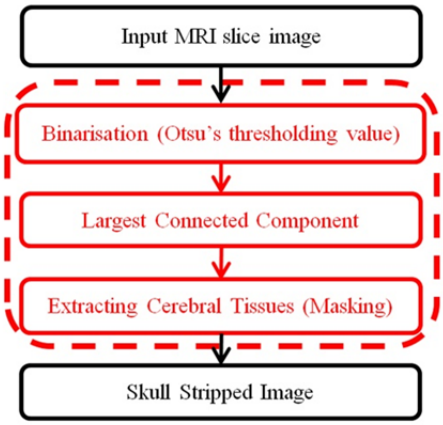


Figure 5. Steps of skull stripping.

4.1.1 Binarization

Image binarzation is the simplest method for image segmentation that transforms an image gray level into two values only. Selection of optimum threshold value "K" ensures the separation of MR image into the background (formed by the very low intensity pixels as part of CSF) and foreground (includes the GM and WM of the brain) tissues as shown in Figure 6.

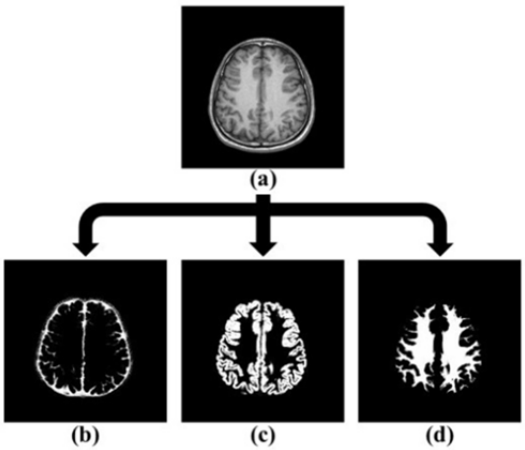


Figure 6. Components of the brain. (a) MRI Slice, (b) segmented CSF, (c) segmented GM and (d) segmented WM (Chuang et al., 2012).

Otsu's method (Otsu, 1979) is employed to obtain cluster-based image threshold value by transforming a gray level image into binary image. It is further assumed that the image comprises of two types of pixels, or bi-model histogram such as the background and the foreground. The optimum threshold values for these two types of pixels are calculated separately to minimize their combined spread (intra-class variance). Multi-level threshold is achieved by extending Otsu's algorithm. The threshold value is carefully selected by reducing the intra-class

variance (Chen, 2004) which is defined as the weighted sum of variances of two classes (Otsu 1979):

$$\sigma_{\omega}^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (1)$$

where ω_1 and ω_2 are the probabilities of the two classes separated by a threshold t and σ_1^2 and σ_2^2 are their variances.

Otsu's demonstrated that the process of intra-class variance minimization is equivalent to inter-class variance maximization (Otsu, 1979):

$$\sigma_b^2(t) = \sigma^2 - \sigma_{\omega}^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \quad (2)$$

Where ω_i class probabilities and μ_i are their Mean values. The class probability $\omega_1(t)$ is evaluated from the histogram using the expression,

$$\omega_1(t) = \sum_0^t p(i) \quad (3)$$

The class Mean $\mu_1(t)$ is calculated from,

$$\mu_1(t) = \frac{\sum_0^t p(i)x(i)}{\omega_1} \quad (4)$$

Where $x(I)$ is the value of the centre i of the histogram bin and $\omega_2(t)$ and μ_2 on the right-hand side of the histogram can compute for bins larger than t. The class probabilities and Means are calculated iteratively using the algorithm.

4.1.2 Largest Connected Component

Binarization on brain MR images classifies it into the background and foreground with the foreground into a number of connected components. The component with large area is considered as displayed in Figure 7 with the assumption that the skull is the principal connected structure surrounding the head.

4.1.3 Extracting Cerebral Tissues (Masking)

The earlier steps contribute to obtain the binary mask for head MRI scan. The cerebral tissues are extracted by performing bitwise operations between the original head MRI scans with the binary mask Figure 7. In this step, the background noise and other non-brain artifacts are removed.

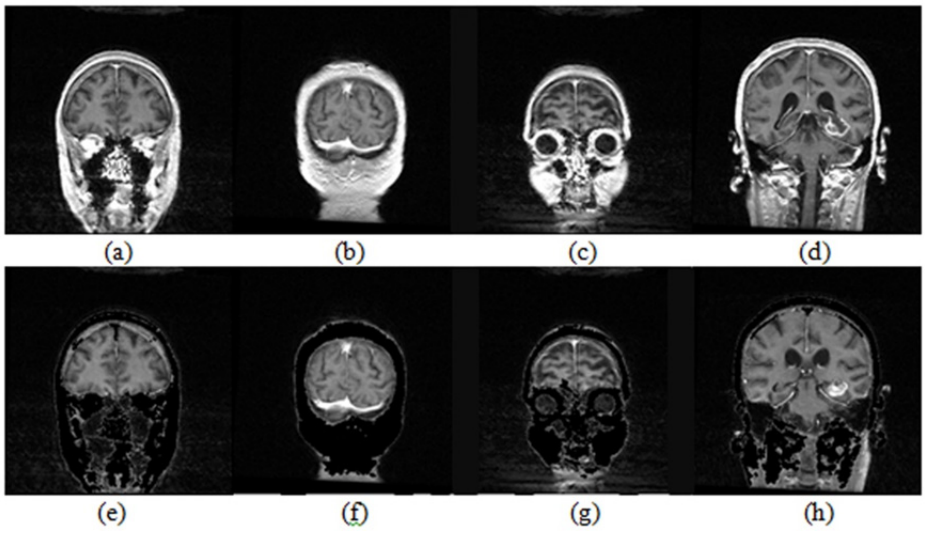


Figure 7. Extraction phase with (a), (b), (c) and (d) as the original MR images. Here (e), (f), (g) and (h) are the corresponding cerebral tissues of MR brain images achieved using the proposed method.

4.2 Gravity Center Point

The first step in extracting the features of input brain MR slice images is the calculation of the gravity center point on image enabling the proposed method as translational invariant. The Center Of Gravity (COG) method is employed to determine the image features. The coordinates for COG of an MR image (x_c, y_c) is determined using,

$$(x_c, y_c) = \left(\frac{\sum_{j=1}^{m} \sum_{i=1}^{n} iI[i,j]}{\sum_{j=1}^{m} \sum_{i=1}^{n} I[i,j]}, \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} jI[i,j]}{\sum_{j=1}^{m} \sum_{i=1}^{n} I[i,j]} \right) \quad (5)$$

Where i is a brain MRI skull stripped slice image of dimension $m \times n$ and x_c and y_c are the x and y coordinates of the brain COG.

4.3 MRI Slice Image Division

First two features Figure 7 being very useful for tumor recognition is used to prune the search space during MR image slice determination through the distance and angular span merging division. Subsequently, the rotation invariance of the angular span vector is determined by moving a part of vector entries clockwise. The angular and distance span vectors of sizes 24˚ with 15 pixels are used. Figure 8 illustrates the angular and distance span vectors.

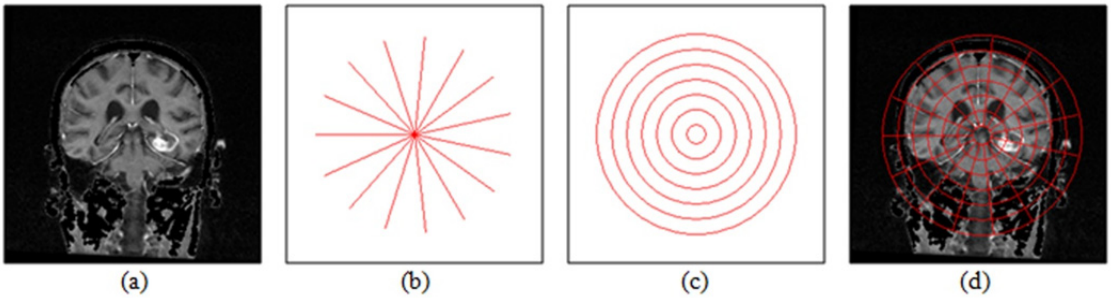


Figure 8. (a) Cerebral tissues of the brain, (b) Angular span, (c) Distance span and (d) Division of brain MR slice image using these spans.

4.4 Feature Extraction

Comparisons are made using different authentic feature extraction methods. They are operated on each division following angular and distance span. The merger between them is applied on the slice image. For each merger space, four statistical features are calculated. Four texture measures are determined from the distributed skull stripped brain slice image assuming that there is no correlation with the neighboring pixel. The texture analysis is performed using intensity value concentrations in the whole or a part of an image represented by a histogram. Histogram approach is followed to extract features including Standard Deviation, Mean, average Energy and Entropy (Jahne, 2005). The histogram intensity levels are simply the summary of the statistical information of the image and particular pixel determining the gray-level. Therefore, the histogram carries the first-order statistical data related to the division space for images. These features are mathematically defined below.

4.4.1 Mean

It is an average value which measures the general brightness of an image,

$$\text{Mean } (\mu) = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} I_i(x,y)}{M \times N} \quad (6)$$

4.4.2 Standard Deviation

It is the measure of the amount of variation or dispersion from the average value. A low value indicates that the data points tend to be very close to the Mean, and a high one implies their spread out over a large range of values. It is expressed as,

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum_{x=1}^{M} \sum_{y=1}^{N} (I(x,y) - \mu)^2}{M \times N}} \quad (7)$$

4.4.3 Energy

It returns the sum of squared elements of gray level values of all pixels in the image. It is equal to 1 for constant image and given by,

$$\text{Energy } (e) = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} I^2(x,y) \quad (8)$$

4.4.4 Entropy

It is the measure of non-uniformity in the image based on the probability of the gray level values of all pixels and is defined as,

$$\text{Entropy} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} I\,(x,y)\big(-\text{In}I(x,y)\big) \tag{9}$$

5. Abnormality Detection and Classification Via SVM

SVM algorithm was introduced by Vapnik and Lerner (Vapnik, 1963) in 1963 and later on extended as Generalized Portrait algorithm by Vapnik and Chervonenkis (Vapnik & Chervonenkis, 1964). This algorithm is based on the framework of statistical learning theory and termed as VC theory. It increases the generalization capability of learning machines to unseen data (Smola & Schölkopf, 2004; Vapnik, 2000). In the past, SVM have provided excellent performance in different real-world applications including hand written digit recognition (Sch et al., 1997), object recognition (Pontil & Verri, 1998), speaker identification (Wan et al., 2000), face detection in images (Osuna et al., 1997) and text categorization (Thorsten, 1999), to cite a few. This classification scheme is based on kernel methods (Bernhard & Smola 2002; Hastie et al., 2009; Cristianini & Shawe-Taylor 2000; Zhang et al., 2009).

In contrast to linear classification techniques, the kernel methods draw the original parameter vectors to a higher (possibly infinite) dimensional feature space via nonlinear kernel function. Here, the dot-products can efficiently be determined in higher dimensional space without computing the nonlinear mapping explicitly. The striking features that make SVM very attractive are the nonlinearly separable classes in which the original space can be linearly separated in the high dimensional feature space. Thus, SVM is able to solve complex nonlinear classification problems. Important characteristics of SVM lie in its ability to resolve classification problems using convex Quadratic Programming (QP) and the sparseness results from them. The learning is based on the principle of structural risk minimization. SVM tends to reduce the bound on the generalization error (error caused by learning machine on the test data which is not used for training) rather than minimizing the objective function which is based on the training samples (such as mean square error). Consequently, SVM performs well for the data outside the training set and enables the advantage of focusing on the training examples that is most difficult to categorize. These "borderline" training examples are called Support Vectors (SV).

Suykens has introduced Least Squares SVM (LS-SVM) (Suykens & Vandewalle, 1999; Van Gestel et al., 2002) using the notion of Vapnik's SVM algorithm where the least squares term is inserted in the cost function. The variant circumvents is required to solve the complex QP problems consisting of a set of linear equations. This significantly reduces the problem complexity and makes the computation process efficient.

The hyper-planes in SVM technique define the decision boundaries by separating the data points of different types of pixels. SVM not only performs simple and linear classification, but also complex problems including non-linear one. Separable and non-separable problems are tackled according to their linear and nonlinear nature. The key concept of SVM is to plot the original data points from the input space to a higher or even infinite dimensional feature space to simplify the classification problem.

A training dataset $(x_i, y_i)^N$ with $x_i \in R^4$, where x_i Is the four input feature vector and $y_i \in \{-1,1\}$ represents the class labels of normal space and tumor brain space is considered. SVM's mapping from the four-dimensional input vector (x) space to the four-dimensional feature space is performed. The non-linear function is defined as $\varphi: R^4 \rightarrow R^4$. Thus, the separated hyper-plane in the feature space is expressed as $w^T\varphi(x) + b = 0$, where $b \in R$ and w is an unknown vector having the same dimension as $\varphi(x)$. The data point x is allocated to the first class if $f(x) = \text{sign}\,(w^T\varphi(x) + b = 0)$ equals to +1 or to the second class if f(x) equals −1.

The hyper-plane can be defined in different ways for linearly separable data. However, SVM depends on the maximum margin principle where the hyper-plane can be constructed with utmost distance between the two different types of pixels. Classification in SVM is initiated following the relations,

$$w^T\varphi(x_1) + b \geq +1 \quad \text{for} \quad y_i = +1 \tag{10}$$

$$w^T\varphi(x_1) + b \leq -1 \quad \text{for} \quad y_i = -1 \tag{11}$$

This is equivalent to,

$$y_{i(w^t\varphi(x_i)+b)\geq +1} \tag{12}$$

Therefore, the classifier can be written as,

$$f(X) = \text{sign}\,(w^T \varphi(x) + b) \quad (13)$$

Figure 9 illustrates the SVM classification using a hyper-plane, where the separating margin between the two types of pixels is minimized using data points denoted by X and O. Here, the support vectors belong to the training set and reside on the hyper-planes boundary separating two classes.

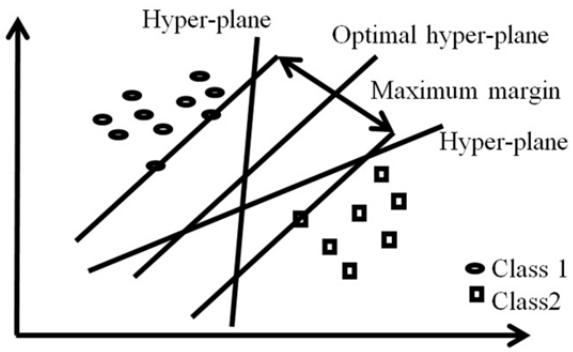


Figure 9. A separating hyper-plane (solid line) with maximum margin determined by linear SVM (Selvaraj et al., 2007)

6. Results and Discussion

6.1 Dataset

The proposed method is tested for the standard dataset called Internet Brain Segmentation Repository (IBSR) which is accessible at the Center for Morphometric Analysis, Massachusetts General Hospital (USA) as shown in (Figure 10) (NITRC, 2011). This being the most popular MRI segmentation database is extensively used in several studies (Gang et al., 2013; Ji et al., 2012; Balafar 2012; Eggert et al., 2012; Yousefi et al., 2012; Zhang et al., 2012; Iglesias et al., 2011; Hwang et al., 2011; Balafar et al., 2011; Rousseau et al., 2011; Somasundaram & Kalaiselvi 2011; Alia et al., 2011; Ortiz et al., 2011; Tian et al., 2011; Bourouis & Hamrouni, 2010; Sikka et al., 2009). The dataset (created in 1999) consists of 300 sliced MR images with 8-bit gray scale images of (256×256) pixels. It contains multiple scans of a patient's tumor images and outlines which is acquired roughly in every six month intervals over a period of two and half years. Each series is stored in 60 slice image after conducting manual diagnosis and execution by well-trained experts (NITRC, 2011).

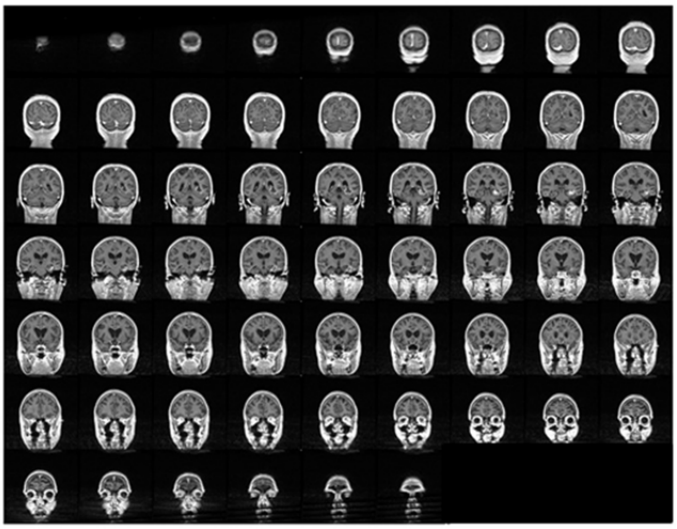


Figure 10. Three dimensional view of sixty slice images of the patient's living brain obtained from IBSR dataset.

6.2 Feature Extraction

SVM technique determines the abnormality in the training step, where the selection of feature is regarded as an important step (Abe, 2010). There exist 364 (91 portion×4 features) values of extracted features for each slice

image. Therefore, for every 3D MR image a total of 21840 (364 features × 60 slice) features value are classified using SVM method. The classification stage composed of training and testing phase. The feature vectors are comprised of approximate coefficients which are extracted from MR brain images and used as inputs for SVM. Besides, the distribution of these features is analyzed using three dimensional plots in which z-axis represents the normalized feature values while x and y-axis characterizes the angular and distance span, respectively.

The distribution of the Mean of the divided slice images as displayed in Figure 11 exhibits the spreading of average values over the whole image with comparatively higher Mean in the tumor region than others. The results provide a clear indication concerning the Mean value of the non-tumor region. The value of Mean for tumor is observed to be about 160.

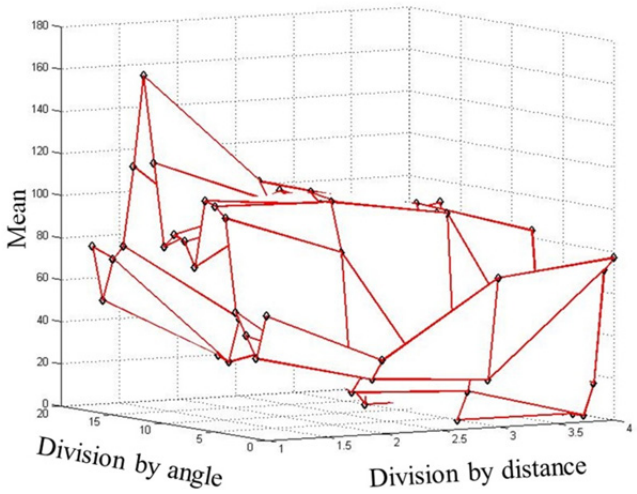


Figure 11. Distribution of Mean in MR slice image.

The Standard Deviation as shown in Figure 12 reveal comparatively higher values in the tumor region than the non-tumor zone. Clearly, the values of Standard Deviation are not distributed over the entire image as in the case of Mean. The approximate value of the Standard Deviation in the tumor space is found to be greater than 70.

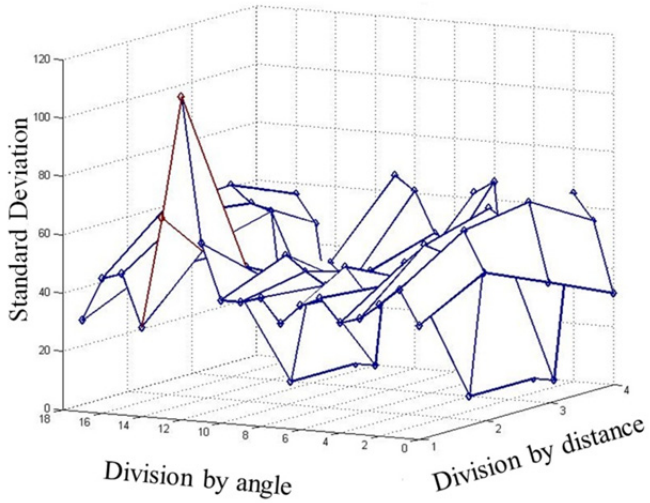


Figure 12. Distribution of Standard Deviation in MR slice image.

Figure 13 depicts the brain slice image Entropy distribution. They are distributed over the entire image with relatively higher occurrence in the tumor space than the non-tumor one. The Entropy value of the tumour space is discerned to be more than 500.

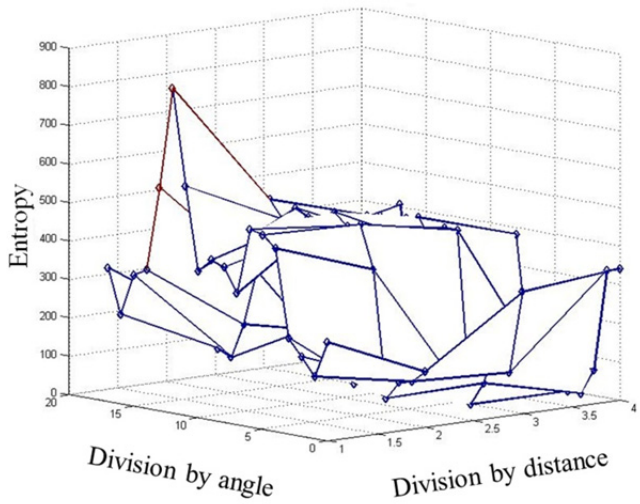


Figure 13. Distribution of Entropy in MR slice image.

Figure 14 showing the Energy distribution in the tumor space is found to be comparatively higher than the non-tumor one. Energy values are determined to be lower than 10.000 for the non-tumor space.

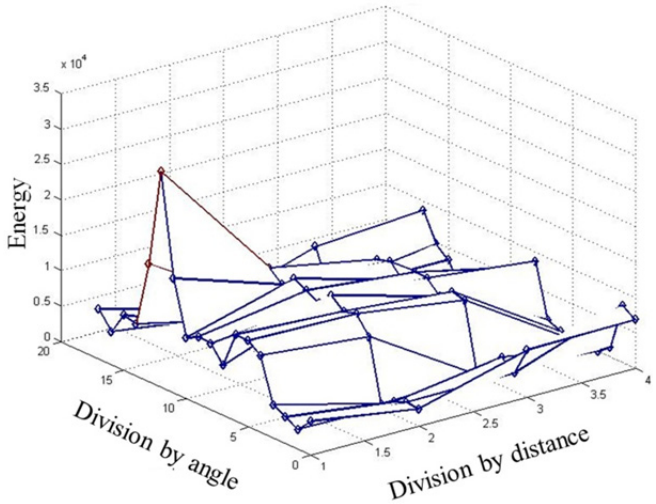


Figure 14. Distribution of Energy in MR slice image.

The generalized representations of the results are shown in Figure 15. The analysis is performed from the manual detection of tumor in MR slice image based on the relation between the Mean, Energy and Entropy. The brown color signifies the tumor space and blue one the non-tumor zone.

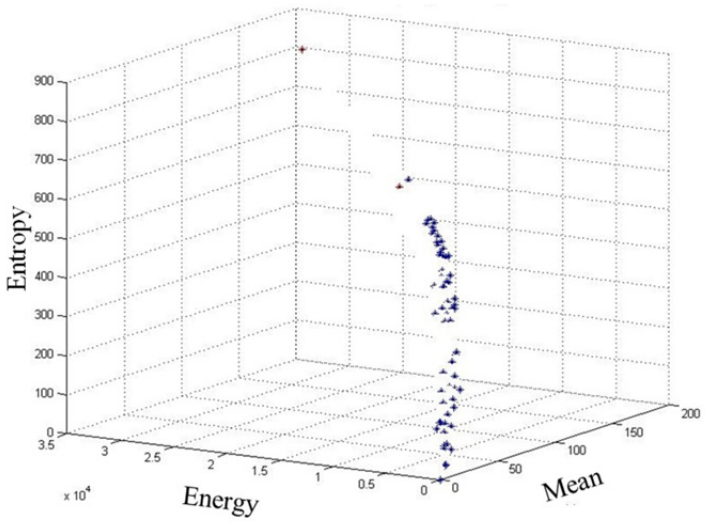


Figure 15. Relation among Mean, Energy and Entropy in MR slice image.

Figure 16 displays the overall comparison of Mean, Energy and Standard Deviation in the slice MR image. The values of Standard Deviation are observed to be distributed over the whole image. Furthermore, the Energy and the Mean values are found to reduce with the decrease of Standard Deviation for both tumor and non-tumor regions. The brown color symbolizes the tumor space and blue color for the non-tumor one.

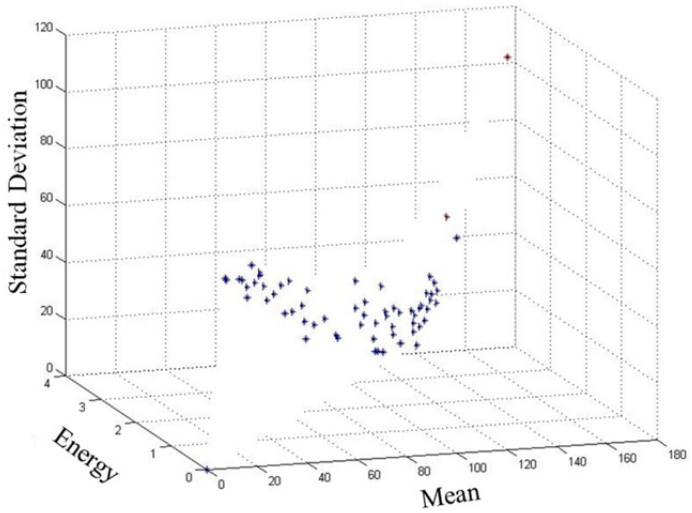


Figure 16. Relation between Mean, Energy and Standard Deviation in MR slice image.

Practically, 180 slice images are used for training, and rest 120 for testing phase, from all the 300 slice images, 70 are abnormal and 230 is the normal one.

6.3 Classification Results

IBM SPSS Modeler Clementine 12.0 is used for classification (Khabaza & Shearer, 1995; Ansheng, 2011). This being one of the influential, adaptable data and text analytical work-bench renders accurate predictive models quickly and intuitively, without programming. The data mining process can easily be visualized using this modeler's intuitive graphical interface. Firstly, the program is designed to model the classifier problem. Figure 17 displays the creation of model stream for SVM classification.

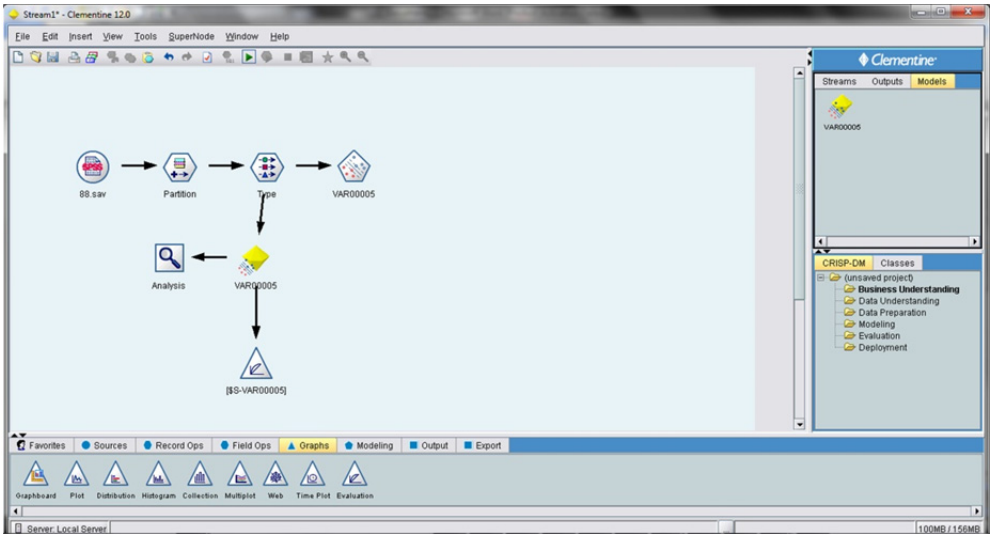


Figure 17. Model stream for SVM classification.

Experiments are executed on MR images to verify the efficiency and robustness of the present classifier with the utilization of training and test data sets. The simulations are performed on a DELL PC Pentium with 3.4 GHz processor and 16 GB RAM. The acquired features from MR images formed the input of SVM classifier. The formulation is carried out to ensure the adaptability and generalization of the classifiers. The first training set is found to be more biased towards the abnormal class. However, the second set having equal numbers of samples from both classes is used to achieve unbiased classifier training. Classification results obtained from with three MR images for training and two MR images for testing phases are listed in Table 1. We achieved an excellent average classification accuracy of 98.02% for training and 98.19% for testing.

Table 1. SVM classifier performance.

IBSR Image Number	Training Accuracy	Testing Accuracy
536_32	98.9%	99.05%
536_45	98.2%	98.49%
536_47	97%	97.67%
536_68	98.45%	98.26%
536_88	97.57%	97.48%

7. Conclusion

A computer based technique using SVM classifiers is developed for automatic sorting of MR image slices corresponding to normal or abnormal cerebral tissues. The performances of the classifiers are analyzed in terms of statistical measures such as sensitivity, specificity and classification accuracy. The method consists of four steps such as cerebral tissue extraction, division, features categorization and classification. Threshold value and region of interest are determined using operator input and Otsu's algorithm. New brain slices image division through angular and distance span vectors are introduced. Following a histogram approach, the values of Standard Deviation, Mean, average Energy and Entropy are calculated for each merger space and compared with existing methods. Sixty slice images of the brain are from IBSR dataset are used. The proposed method achieves a value of average classification accuracy as much as 98.02% for training and 98.19% for testing. It is demonstrated that the new approach provides relatively higher performance than existing methods. Excellent features of the results suggest that the proposed technique can potentially be applied for medical image classification. Furthermore, it may contribute towards the development of computer aided intelligent health care systems. This automated analyses system can further be extended for the image classification with different pathological condition, types and disease status.

Acknowledgements

The author(s) would like to thank the Ministry of Higher Education & Scientific Research-Iraq and Universiti

Teknologi Malaysia for providing the financial support and facilities for this research.

References

Abe, S. (2010). Support Vector Machines for Pattern Classification. Springer.

Al-Badarneh, A., Najadat, H., & Alraziqi, A. M. (2012). A Classifier to Detect Tumor Disease in MRI Brain Images. In Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference. IEEE. IEEE, pp. 784–787. http://dx.doi.org/10.1109/ASONAM.2012.142

Alia, O. M., Mandava, R., & Aziz, M. E. (2011). A Hybrid Harmony Search Algorithm for MRI Brain Segmentation. *Evolutionary Intelligence, 4*(1), 31–49. http://dx.doi.org/10.1007/s12065-011-0048-1

Al-Tamimi, M. S. H., & Sulong, G. (2014a). A Review of Snake Models in Medical MR Image Segmentation. *Jurnal Teknologi, 2*(1), 101–106.

Al-Tamimi, M. S. H., & Sulong, G. (2014b). Tumor Brain Detection Through MR Images: A Review of Literature. *Journal of Theoretical and Applied Information Technology (JATIT), 62*(2), 387–403.

Ansheng, L. (2011). Application Analysis of Clementine-Based Data Mining Algorithm. In 2011 International Conference on Intelligence Science and Information Engineering. *IEEE.,* pp. 479–482. http://dx.doi.org/10.1109/ISIE.2011.76.

Antonie, L. (2008). Automated Segmentation and Classification of Brain Magnetic Resonance Imaging. C615 Project.

Ariffanan, M., & Basri, M. (2008). Medical Image Classification and Symptoms Detection Using Neuro Fuzzy. PhD diss., Universiti Teknologi Malaysia, Faculty of Electrical Engineering.

Badran, E. F., Mahmoud, E. G., & Hamdy, N. (2010). An Algorithm for Detecting Brain Tumors in MRI Images. In The 2010 International Conference on Computer Engineering & Systems. IEEE, pp. 368–373. http://dx.doi.org/10.1109/ICCES.2010.5674887.

Balafar, M. A. (2012). Gaussian Mixture Model Based Segmentation Methods for Brain MRI Images. *Artificial Intelligence Review,* 429–439. http://dx.doi.org/10.1007/s10462-012-9317-3

Balafar, M. A., Ramli, A., & Mashohor, S. (2011). Brain Magnetic Resonance Image Segmentation Using Novel Improvement for Expectation Maximizing. *Neurosciences, 16*(3), pp.242–247.

Bernhard, S., & Smola, A. J. (2002). Learning With Kernels. MIT Pres, Cambridge, pp.73–76.

Bourouis, S., & Hamrouni, K. (2010). 3D Ssegmentation of MRI Brain Using Level Set and Unsupervised Classification. *International Journal of Image and Graphics, 10*(1), 135–154.

Cawley, G. C., & Talbot, N. L. C. (2004). Fast Exact Leave-One-Out Cross-Validation of Sparse Least-Squares Support Vector Machines. *Neural Networks : The official journal of the International Neural Network Society, 17*(10), 1467–1475. http://dx.doi.org/10.1016/j.neunet.2004.07.002

Chaplot, S., Patnaik, L.M., & Jagannathan, N. R. (2006). Classification of Magnetic Resonance Brain Images Using Wavelets as Input to Support Vector Machine and Neural Network. *Biomedical Signal Processing and Control, 1*(1), 86–92.

Chen, S. (2004). Chaotic Spread Spectrum Watermarking for Remote Sensing Images. *Journal of Electronic Imaging, 13*(1), 146–165. http://dx.doi.org/10.1117/1.1631316

Chuang, C. C. et al. (2012). Patient-Oriented Simulation Based on Monte Carlo Algorithm by Using MRI Data. Biomedical engineering online, 11(1), p.21. http://dx.doi.org/10.1186/1475-925X-11-21

Clark, M.C. et al. (1994). MRI Seamentation Usina Fuzzy Clustering TechnGues. *Engineering in Medicine and Biology Magazine, IEEE, 13*(5), pp.730–742.

Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods First., United Kingdom, Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511801389

Deepak, K. S. et al. (2013). An Efficient Approach To Predict Tumor In 2D Brain Image Using Classification Techniques. In Information Communication and Embedded Systems (ICICES), 2013 International Conference . IEEE. pp. 559–564.

Eggert, L. D. et al. (2012). Accuracy and Reliability of Automated Gray Matter Segmentation Pathways on Real and Simulated Structural Magnetic Resonance Images of the Human Brain Y. Fan, ed. Real and

simulated structural magnetic resonance images of the human brain. *PloS One, 7*(9), 1–9. http://dx.doi.org/10.1371/journal.pone.0045081

El-Dahshan, E. A., Salem, A. B. M., & Younis, T. H. (2009). A Hybrid Technique for Automatic MRI Brain Images Classification. Studia Univ. Babes-Bolyai, *Informatica, 54*(1), 55–67.

El-Dahshan, E. S. A., Hosny, T., & Salem, A. B. M. (2010). Hybrid Intelligent Techniques for MRI Brain Images Classification. *Digital Signal Processing, 20*(2), 433–441. http://dx.doi.org/10.1016/j.dsp.2009.07.002

Ferlay, J. et al. (2010). GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase. *Lyon, France: International Agency for Research on Cancer, 29*(10).

Gang, Z. et al. (2013). An Unsupervised Method for Brain MRI Segmentation. *International Journal of Emerging Technology and Advanced Engineering, 3*(10), 8–13.

Goswami, S., & Bhaiya, L. K. P. (2013). Brain Tumour Detection Using Unsupervised Learning Based Neural Network. In Communication Systems and Network Technologies (CSNT), 2013 International Conference. IEEE. IEEE., 573–577. http://dx.doi.org/10.1109/CSNT.2013.123

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning Second Edi., New York - USA: Springer.

Horská, A., & Barker, P. B. (2010). Imaging of Brain Tumors: MR Spectroscopy and Metabolic Imaging. *Neuroimaging clinics of North America, 20*(3), 293–310. http://dx.doi.org/10.1016/j.nic.2010.04.003

Hwang, J., Han, Y. & Park, H. (2011). Skull-Stripping Method for Brain MRI Using a 3D Level Set with a Speedup Operator. *Journal of Magnetic Resonance Imaging, 34*(2), 445–456. http://dx.doi.org/10.1002/jmri.22661

Ibrahim, W. H., Osman, A. A. A., & Mohamed, Y. I. (2013). MRI Brain Image Classification Using Neural Networks. In Computing, Electrical and Electronics Engineering (ICCEEE), 2013 International Conference. IEEE. pp. 253–258. http://dx.doi.org/10.1109/ICCEEE.2013.6633943

Iglesias, J. E. et al. (2011). Robust Brain Extraction Across Datasets and Comparison With Publicly Available Methods. *Medical Imaging, IEEE Transactions, 30*(9), 1617–1634.

Improved, A., Gradient, C., & Svm, L. S. (2005). An Improved Conjugate Gradient Scheme To The Slution of Least Squares SVM. *Neural Networks, IEEE Transactions, 16*(2), 498–501.

Jahne, B. (2005). Digital Image Processing 6th Editio., Springer-Verlag Berlin, germany.

Ji, Z. et al. (2012). Fuzzy Local Gaussian Mixture Model for Brain MR Image Segmentation. *Information Technology in Biomedicine, IEEE Transactions, 16*(3), 339–347.

Joshi, D. M., Rana, N. K., & Misra, V. M. (2010). Classification of Brain Cancer Using Artificial Neural Network. In Electronic Computer Technology (ICECT), 2010 International Conference. IEEE. 112–116. http://dx.doi.org/10.1109/ICECTECH.2010.5479975

Khabaza, T., & Shearer, C. (1995). Data Mining With Clementine. The digest of the IEEE Colloquium on Knowldge Discovery in Data-Base, Digest 1995 / 021 (B) - London, pp.1–5.

Klein, S. et al. (2010). Early Diagnosis of Dementia Based OnIntersubject Whole-Brain Dissimilarities. In Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium. IEEE. 249–252. http://dx.doi.org/10.1109/ISBI.2010.5490366

NITRC. (2011). NITRC: IBSR: Tool/Resource Info. Retrieved from http://www.nitrc.org/projects/ibsr

Noback, C. R. et al. (2005). The Human Nervous System Structure And Function Sixth Edit., New Jersey: Humana Press.

Nolte, J. (2013). The Human Brain in Photographs and Diagrams Fourth Edi., Elsevier Health Sciences.

Ortiz, A. et al. (2011). MR Brain Image Segmentation By Growing Hierarchical SOM and Probability Clustering. *Electronics Letters, 47*(10), 585–586. http://dx.doi.org/10.1049/el.2011.0322

Osuna, E., Freund, R., & Girosit, F. (1997). Training Support Vector Machines: An Application to Face Detection. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. *IEEE Comput. Soc.,* 130–136. http://dx.doi.org/10.1109/CVPR.1997.609310

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems. *Man and Cybernetics, 9*(1), 62–66.

Pontil, M., & Verri, A. (1998). Support Vector Machines for 3D Object Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(6), 637–646. http://dx.doi.org/10.1109/34.683777

Rajini, N. H., & Bhavani, R. (2011). Classification of MRI Brain Images Using k-Nearest Neighbor and Artificial Neural Network. In Recent Trends in Information Technology (ICRTIT), 2011 International Conference. IEEE. 563–568. http://dx.doi.org/10.1109/ICRTIT.2011.5972341

Rousseau, F., Habas, P. A., & Studholme, C. (2011). Human Brain Labeling Using Image Similarities. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference. 1081–1088.

Roy, S. et al. (2013). A Review on Automated Brain Tumor Detection and Segmentation from MRI of Brain. *arXiv preprint arXiv, 1*(1), 1–41.

Sadananthan et al. (2010). Skull Stripping Using Graph Cuts. *NeuroImage, 49*(1), 225–239.

Salankar, S. S., & Bora, V. R. (2014). MRI Brain Cancer Classification Using Support Vector Machine. In Electrical, Electronics and Computer Science (SCEECS), 2014 IEEE Students' Conference. IEEE. 1–6. http://dx.doi.org/10.1109/SCEECS.2014.6804439

Sch, B. et al. (1997). Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers. Signal Processing. *IEEE Transactions, 45*(11), 2758–2765.

Ségonne et al. (2004). A Hybrid Approach to the Skull Stripping Problem in MRI. *NeuroImage, 22*(3), 1060–1075.

Selvanayaki, K., & Karnan, M. (2010). CAD System for Automatic Detection of Brain Tumor Through Magnetic Resonance Image-A Review. *International Journal of Engineering Science and Technology, 2*(10), 5890–5901.

Selvaraj, H. et al. (2007). Brain MRI Slices Classification Using Least Squares Support Vector Machine. *International Journal of Intelligent Computing in Medical Sciences & Image Processing, 1*(1), 21–33.

Sikka, K. et al. (2009). A Fully Automated Algorithm Under Modified FCM Framework for Improved Brain MR Image Segmentation. *Magnetic Resonance Imaging, 27*(7), 994–1004. http://dx.doi.org/10.1016/j.mri.2009.01.024

Smola, A. J., & Schölkopf, B. (2004). A Tutorial On Support Vector Regression. *Statistics and computing, 14*(3), 199–222.

Somasundaram, K., & Kalaiselvi, T. (2011). Automatic Brain Extraction Methods for T1 Magnetic Resonance Images Using Region Labeling and Morphological Operations. *Computers in biology and medicine, 41*(8), 716–725. http://dx.doi.org/10.1016/j.compbiomed.2011.06.008

Spisz, T. S., & Bankman, I. N. (2000). Medical Image Processing and Analysis Software, Academic Press, Inc.

Suykens, J. A., & Vandewalle, J. (1999). Least Squares Support Vector Machine Classifiers. *Neural Processing Letters, 9*(3), 293–300.

Technologies, C. (2014). Histogram Equalization for Image Enhancement Using MRI Brain Images. Computing and Communication Technologies (WCCCT), 2014 World Congress, IEEE, 80–83. http://dx.doi.org/10.1109/WCCCT.2014.45

Thorsten, J. (1999). Transductive Inference for Text Classification Using Support Vector Machines. ICML, 99, 200–209.

Tian, G. et al. (2011). Hybrid Genetic and Variational Expectation-Maximization Algorithm for Gaussian-Mixture-Model-Based Brain MR Image Segmentation. Information Technology in Biomedicine. *IEEE Transactions, 15*(3), 373–380.

Van Gestel, T. et al. (2002). Least Squares Support Vector Machines, Singapore: World Scientific.

Vapnik, V. (1963). Pattern Recognition Using Generalized Portrait Method. *Automation and Remote Control, 24,* 774–780.

Vapnik, V. (2000). The Nature of Statistical Learning Theory Second Edi. M. Jordan, ed., Springer-Verlag New York - USA.

Vapnik, V., & Chervonenkis, A. (1964). A Note On Class of Perceptron. Automation and Remote Control.

Veloz, A., Orellana, A., & Vielma, J. (2011). Brain Tumors : How Can Images and Segmentation Techniques Help ? Diagnostic Techniques and Surgical Management of Brain Tumors, 67–92.

Wan, V. et al. (2000). Support Vector Machines for Speaker Verification and Identification. *Neural Networks Signal Process PROC IEEE, 2,* 775–784.

Yousefi, S., Kehtarnavaz, N., & Gholipour, A. (2012). Improved Labeling of Subcortical Brain Structures in Atlas-Based Segmentation of Magnetic. Biomedical Engineering. *IEEE Transactions, 59*(7), 1808–1817.

Zhang, H. et al. (2011). An Automated and Simple Method for Brain MR Image Extraction. *Biomedical engineering online, 10*(1), 81. http://dx.doi.org/10.1186/1475-925X-10-81

Zhang, N. et al. (2009). Multi-Kernel SVM Based Clasification For Brain Tumor Segmentation Of MRI Multi-Sequence. In Image Processing (ICIP), 2009 16th IEEE International Conference. IEEE. 3373–3376.

Zhang, T., Xia, Y., & Feng, D. D. (2012). Clonal Selection Algorithm for Gaussian Mixture Model Based Segmentation of 3D Brain MR Images. Intelligent Science and Intelligent Data Engineering. pringer Berlin Heidelberg., 295–302.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Vertical-Flow Constructed Wetlands in Cooperating with Oxidation Ponds for High Concentrated COD and BOD Pig-Slaughterhouse Wastewater Treatment System at Suphanburi-Provincial Municipality

Piyaporn Pitaktunsakul[1], Kasem Chunkao[1], Narouchit Dampin[1] & Satreethai Poommai[1]

[1] Department of Environmental Science, Faculty of Environment, Kasetart University, Bangkok, Thailand

Correspondence: Piyaporn Pitaktunsakul, Department of Environmental Science, Faculty of Environment, Kasetart University, Bangkok, Thailand. Tel: 66-81-684-3284.

Received: December 10, 2014 Accepted: November 25, 2014 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p371 URL: http://dx.doi.org/10.5539/mas.v9n8p371

Abstract

Oxidation Pond (OP) as engineering tool is generally used for treating the pig slaughterhouse wastewater which normally contains high concentration of COD and BOD in effluent. Unfortunately, it cannot reduce the organic substance (blood, hairs, grease, meats, solid dunks and some contaminants) from pig slaughtering areas under the 2-consecutive oxidation ponds by producing the minimum values of COD 151.92 mg/L, BOD 79.14 mg/L, coliform bacteria 2.6 x 10^{-5} MPN/100mL, and fecal bacteria 1.5 x 10^{-5} MPN/100mL but all of them above the standard values. After treating the effluent by VFCW-Typha from the 2-consecutive oxidation ponds, the results found COD 90.92 mg/L, BOD 31.67 mg/L, coliform bacteria 1.5 x 10^{-4} MPN/100mL and fecal bacteria 2.0 x 10^{-3} MPN/100 mL which were almost above the standard values. It is noted that the modification of 2-m consecutive ponds to 4-m consecutive ponds in cooperating the prolongation of VFCW-Typha length instead of 30 meters to 40-50 meters, width 3-5 meters, and still keeping 1-m depth would be enough to support the pig-slaughterhouse wastewater treatment system.

Summarily speaking, the experimental results have been brought to say that the Oxidation Pond as the engineering tool could not be applicable in slaughterhouse wastewater treatment that containing high concentration COD and BOD from slaughtering and dissecting activities.

Keywords: vertical-flow constructed wetland, oxidation pond, wastewater treatment, pig-slaughterhouse

1. Introduction

Wastewater treatment in Thailand are normally engineering devices which are seemingly designed for on-site establishment exactly at the point sources. The Royal Thai Government (RTG) on behalf of DOPC-MONRE (2010) has paid the annual budget more or less 50,000 M Baht to construct the community wastewater treatment system at every municipal and local administration, and being accomplished not later than 2015. Additionally, the government issued laws to obligate all industrial factories, housing estates, high buildings, hotels and shopping centers as to treat their wastewater before releasing to the public water sources (MOPH 1998, DOPC-MONRE (2010), and DIW-MOI 1996). In general, the engineering devices are used for wastewater treatment system in case of Filtration, Floatation, Skimming, Sedimentation, Asa Grit Chamber, Precipitation, Chemical Coagulation, Neutralization, Oxidation Pond (OP), Activated Sludge (AS), Stabilization Pond (SP), Aerated Pond (AP), Aerated Lagoon (AL), Bio Ditch Filter (BDF),Trickling Filter (TF), Upflow Anaerobic Sludge Blanket (UASB), Rotating Bio Contractor (RBC), Phytoremediation/Phytoextraction, Carbon Adsorption (CA), Activated Carbon Absorption (ACA), Reverse Osmosis (RO), Ion Exchange (IE) and some modified equipments to decrease the specific pollutants. (Metcalf and Eddy 1979, Ramsey et al 2013, Borja et al 1998, Metcalfe et al 1979). Unfortunately, the failure of implementing wastewater treatment has been found all the country (Faerge et al 2001, Berkun 2005, Faulkner et al 2000, Padgett 1975 ,Streeter and Phelps 1925, Tyagi et al 1999, Hosetti and Frost 1995), especially the settlement of factories and cities along the riverbanks as the serious point sources for causing stream pollution (Streeter and Phelps 1925, Tyagi et al 1999, Faulkner et al 2000, Berkun 2005, Pattamapitoon 2013). Even though the RTG still spent a lot if budget but the stream pollution has been spread out in every river. It cannot be put into words why wastewater pollution in worse condition almost every river. The basic mistakes might come from too high expensive engineering wastewater

treatment tools, costly electricity, and intensive care for effective function during mechanical operation, This is the reason why the engineering tools for wastewater treatment from any sources cannot meet the requirement of the standards in all parts of the country.

Moreover, the slaughterhouse is among those point sources of effluent as the same as industrial factory, communities, households, municipals, transportation, poultry farms, and cultivating areas, in which the public water sources have been intruded by their treated wastewater from those ineffective engineering tools.

In Thailand, there are many reports concerning with pork consumption, approximately 30 kg/person/year (each pig weight more or less 120 kg) which brought to slaughter about 12 million hogs/year (approximately 33,000 hogs/day) for Thai people. Consequently, the water use for washing and cleaning in the slaughtering processes inside the slaughterhouses about 3.3 MCM/day around the country. In the same manner, each pig slaughterhouse produces organic waste in terms of Biochemical Oxygen Demand (BOD) up to 1,200 mg/L to be treated before draining into the public water sources (such as streams, canals, rivers), in order to keep water sources avoiding the stream pollution after receiving its treated wastewater. The Royal Thai Government has declared the Water Quality Standard Laws for slaughterhouse effluent since 1960 and having one key issue on establishing the Oxidation Pond which is the engineering tool for wastewater treatment. Unfortunately, this technology seems blur to apply in Thai style for pig slaughtering techniques which are needed not only pork meet but also blood and another organs as well. Keeping hogs in pigpen before killing would be another reason why slaughterhouse wastewater was comprised of concentrated Chemical Oxygen Demand (COD), Biochemical Oxygen Demand (BOD), Total Dissolved Solid (TDS), and Suspended Solid (SS) than they were found in the developed country. This characteristics of slaughterhouse wastewater have been shown the evidence that the engineering-tool Oxidation Pond could not meet the target to have effluent under standard value (Tritt and Schuchardt 1992, Borja 1995, Masse and Masse 2000 a and b, Cao and Mehrvar 2011, Othman et al 2013, and Palatri et al 2011).

It can make the conclusion here that the Oxidation Pond for wastewater treatment tool might not be applicable for such high concentration of slaughterhouse wastewater for Thai-slaughtering style. In general, the anaerobic digestion is usually introduced in order to decrease the high concentration down to the suitable concentration level before transferring through the Oxidation Pond to obtain the effluent equal or less than standard value (for example, 120 mg/L for Chemical Oxygen Demand (COD), 20 mg/L for Biochemical Oxygen Demand (BOD), 100 mg/L for Total Kjeldahl Nitrogen (TKN). Unfortunately, the device for producing an effect to decrease high concentration of pig slaughterhouse wastewater is more expensive and complicated to install along with the existing Oxidation Pond as seen in every municipal. For the tropical climate like Thailand, the vertical-flow constructed wetlands (VFCW) are proposed to retreat the OP effluent as directly drained out from the ready-installed Oxidation Pond, the previous research results were satisfactory and applicable to all parts of the country (LERD 1999).

In principles, wetland is the swamp unit area which works together among soils, aquatic plants and organic wastewater for assimilating balance of inorganic-nutrient storage, and nutrient absorption for aquatic plant growth as well as the remediation of toxic chemicals. In another points, the bacterial organic digestion is naturally in both sub-soils and wastewater over soil surface by obtaining the oxygen as energy supplying through the processes of thermo-osmosis, and thermo-siphon from the atmosphere and also from phytoplankton and algae photosynthesis. (Deubigh and Raumann 1952, Grosse 1989, Mirmov and Belyakava 1982, Grofse and Bauch 1991, Bearman 1957,Ameth and Stchimair 2001).The products of bacterial organic digesting process as mentioned are exactly identified as plant nutrients and some toxic chemicals as accumulated in soils for growing submerged aquatic plants (such as typha, cyperus) and dispersing in treated wastewater for growing phytoplankton (Suchkov et al 2010, Luangsoonton 2010, Stottmeister et al 2003, Keddy 2010, Maine 2006, Juwarkar et al 1995, Khan et al 2009, Boyd 1970, Jenssen et al 1993, Hammer 1989, Hammer and Bastian 1989, Ahn and Mitsch 2002, Kayser and Hunst 2005, Cui et al 2010). The balancing condition is really needed to harvest the submerged aquatic plants by clear cutting of the zero-growth rate of the oldest plants with leaving the stem about 30-cm height, and to consume the herbivore fishestogetherwith zero-growth rate of biggest size (LERD 1999 and 2012, Sklarz et al 2009, Molle et al 2008). The said concept has been shown in success of organic wastewater treatment under the biological processing principles through both the vertical and horizontal flow constructed wetlands (VFCW and HFCW) as evidenced from the previous researches (Soroko 2007, Cauillard et al 1989, Kayser and Hunst 2005, Ahn and Mitsch 2002, Hammer and Bastian 1989, Hammer 1989, Boyd 1970, Khan et al 2009, Juwarkar et al 1995, Stottmeister et al 2003, Cui et al 2010). The wastewater treatment efficiencies were varied from 85 % to 95 % for COD, BOD, TDS, SS, DOC, DOP, DON, NH3-N, color, coliform and fecal bacteria; 80 % to 90 % for sulfate, TKN, EC, alkalinity, acidity, and organic compounds (Boyd 1970, Khan et al 2009, Juwarkar et al 1995, Main et al 2006, Nopparatanaporn 1992, Molle et al 2008,

Sklarz et al 2009). The previous researches were also indicated that the most effective efficiency of small vertical flow constructed wetlands could be applicable on the BOD concentration more or less 200 mg/L which found in effluent values of Oxidation Ponds. Such indicated effluent values cannot be released into the public water sources without any exempts. In order to sovereign the problem, the coopering Oxidation Ponds and VFCW is proposed to treat high concentrated COD and BOD of pig slaughterhouse wastewater by following the previous researches such as LERD 1999, Soroko 2007, Molle et al 2008 and Sklarz et al 2009).

2. Methods and Procedures

2.1 Location of Project Site

Project site for applying the VFCW units for pretreatment to high concentration of organic matters in slaughterhouse wastewater is localized inside the solid waste management land area at AmphoeMuang municipal of Suphanburi province (population about 60,000 persons) in tropical climate (annual rainfall 1,200 - 1,400 mm, average temperature 32 ^{o}C with ranging been 25 - 39 ^{o}C) on the upper part of perennially Suphanburi river (about 720 km. length of S-shaped watercourse), as called only the upstream of Thachin river, in westerly Bangkok about 110-km distance as shown in Figure 1. In reality, Suphanburi river flows from Chainat headwater and connecting by watergate with Chainat diversion-dam reservoir which can supply water flowing through Suphanburi rivet all year round without even long-time water shortage. This situation can make Suphanburi province all-time green with paddy rice fields and vegetables growing, and also follow-up livestock farming, particularly pig, poultry, and cow.

Those activities have been induced more plat nutrients in stream water as the products of bacterial organic digesting from releasing the organic wastes into the stream water. This is why the glory morning vegetables growing can be done on the Thachin river water surface from Suphanburi to Nakhon Pathom provinces, appraising more than 500 million Baht per annum. Besides, there have been a lot of households settling along on both sides of Suphanburi river as the same as the food industrial factories, paddy rice, cropping areas and sightseeing tours would be another point sources. Moreover, thousands of acres are in use for growing rice and another economic crops but the chemical fertilizers and pesticides have been applied for increasing the agricultural products. Due to high fertility of Thachin river water as obtained from bacterial organic digestion, there are a lot of water surface areas on about on both riversides that have been used for growing morning glory vegetables all years round and producing yields more than 100 tons per annum. Such cropping activity could enhance to decrease a lot of nitrogen and phosphorus content in Thachin river water and another elements.

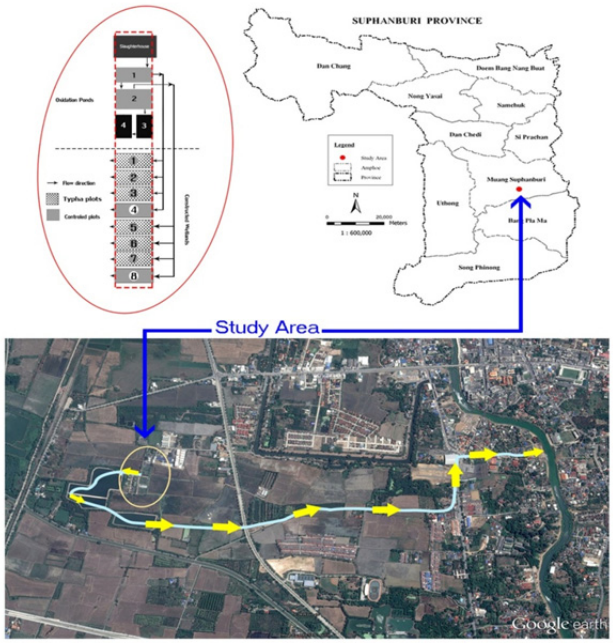


Figure1. Location of slaughterhouse the Suphanburi municipal AmphurMuang, Suphanburi province, including office building, pigpen, slaughtering and cleaning areas, wastewater drainage system, and four-sectional ponds of Oxidation Disc as engineering technology for slaughterhouse wastewater treatment as well as eight vertical flow constructed wetlands.

2.2 Suphanburi Slaughterhouse Establishment

The slaughterhouse is located inside the occupied area (1.5 sq.km) for managing solid waste and MuangSuphanburi municipal wastewater treatment system as indicated in Figures 1 and 2. In fact, the study site is composed of Oxidation Disc (OD) and eight vertical flow constructed wetlands (VFCW) which are surrounded by the sanitary landfills and neighboring fermented-stool boxes and lagoons (oxidation ponds) for treating Suphanburi municipal wastewater about 120 cu.m./day as obtained from population of about 60,000 persons' activities. Anyway, the wastewater from pig slaughterhouse would be expected to transfer directly or indirectly to Thachin river one way or another.

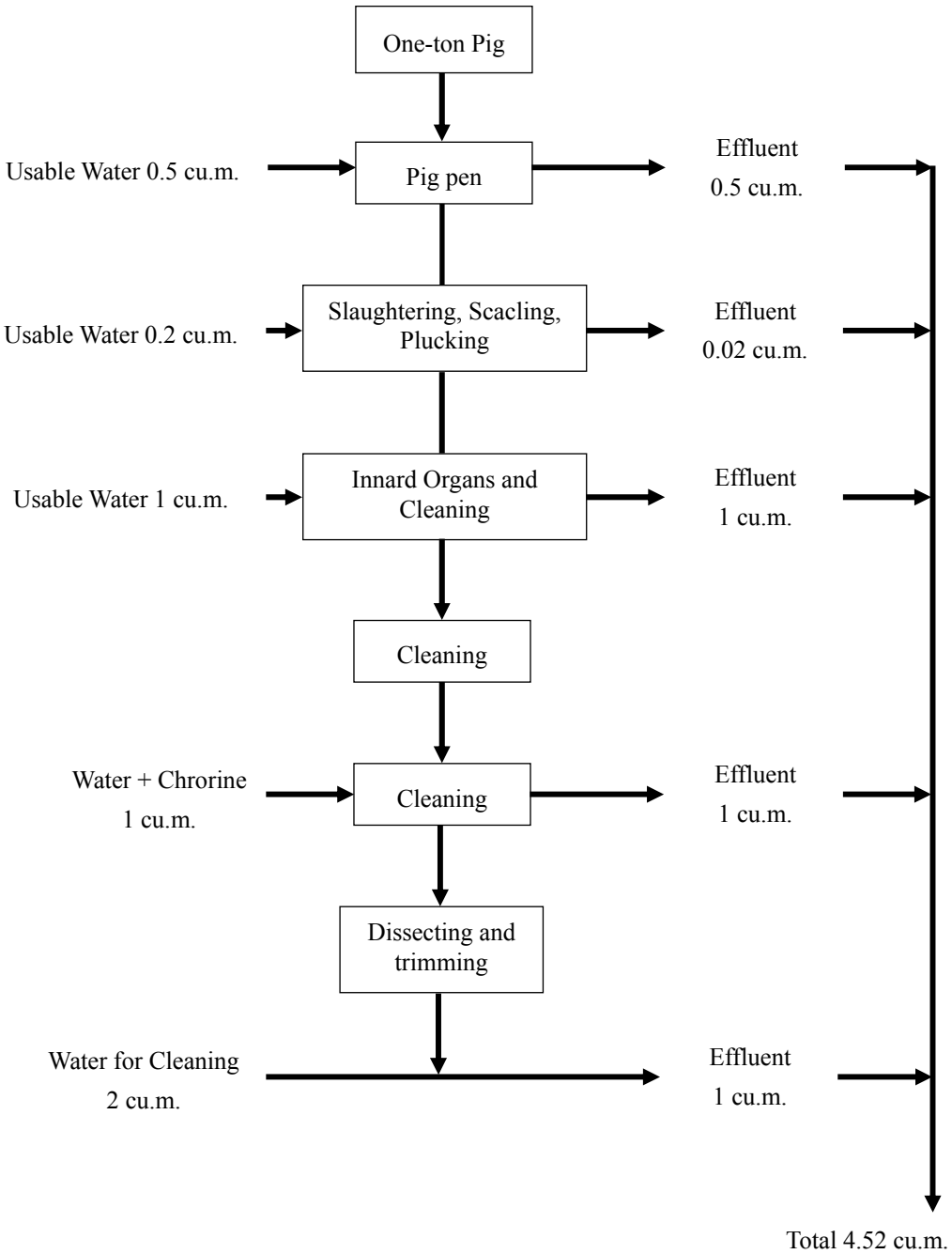


Figure 2. Layout of water use for Hog slaughtering and dissecting of Suphanburi slaughterhouse at AmphoeMuang municipal, Suphanburi province

2.3 Establishment of Vertical Flow Constructed Wetland (VFCW)

Eight VFCW units (30-m long, 3-m wide, and 0.75-m deep) with typha planting (25x25-cm spacing) were constructed on line of the 8-consecutive sectional ponds for second treatment to the effluent of Oxidation Pond (Figure 3). Usually, the slaughterhouse wastewater is received the slaughterhouse wastewater with high concentrated organic matters approximately between 6 - 10 cubic meters per days for 24 - 32 hogs (4.52 cu.m.water for one tone of hog) as shown in Figure 2. Consequently, this amount of slaughterhouse wastewater will be drained into the first of the OD four-sectional sectional ponds to the second, third and fourth between 2:00 - 6:00 am for the 6-day slaughtering of a week (except Buddhist holyday). However, the past records of water quality measurement from all 4-sectional ponds were found higher standard values of BOD (200 mg/L) and COD (120 mg/L) in which they was selected for taking the slaughtering wastewater for 8 vertical flow constructed wetlands as second treatment in order to meet the effluent standard requirement before releasing to the public water sources.

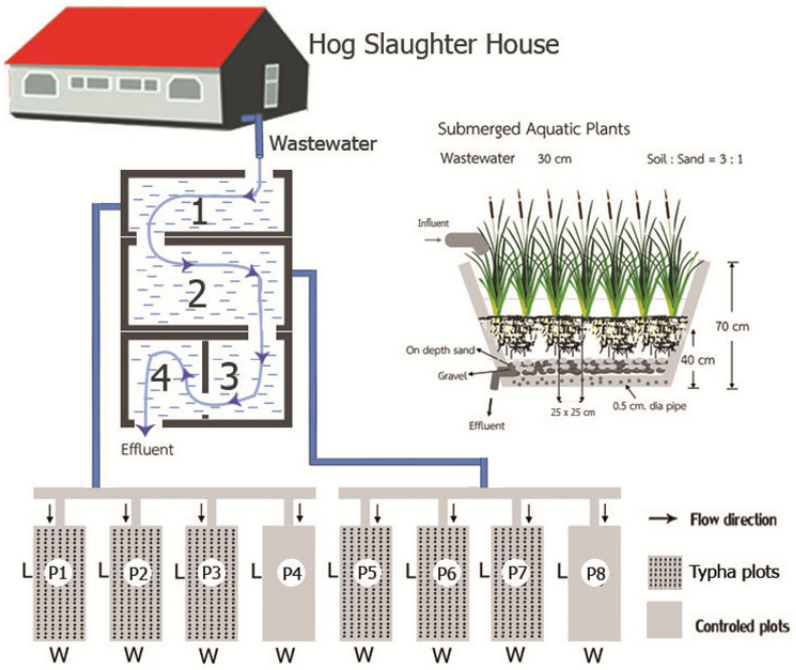


Figure 3. Hypothetical characteristics of Oxidation Pond (OP) together with X-section and longitudinal highlighting of small vertical flow constructed wetlands(VFCW) as grown typha (25-spacing) to retreat the effluent from the higher concentration sectional OP pond 2 for producing an effective slaughterhouse wastewater in Suphanburi province

2.4 Contaminants of Slaughterhouse Wastewater

Living hogs are transported from the swine farms to put in pigpen inside the slaughterhouse for a few days before killing them. Surely, the solid waste from pigpen floor is necessary to wash off the pig dunks into the sectional OP pond 1. A great number of blood and some organic materials as well as grease has to wash off by 4.52 cu.m./one-ton pig to became the slaughterhouse wastewater into the sectional OP pond 1 as shown in Figure 2. Actually, the slaughtering pigs are prohibited on important days in Buddhism, such a Buddhist holy days, beginning and ending Buddhist Lent, Songkran, (Thai new year), new year (international), Magha Bucha Days, VisaghaBucha Day, and the middle of the lunar month (DIW-MOI 1996).Totally, they come together to 60 days with no organic wastewater from slaughtering houses. As stated in former times, the water use for pig slaughtering can be calculated more or less 120 cu.m./d, as become the slaughterhouse wastewater for about 100 cu.m. forfulling OP treatment system (DOPC-MONRE 1998).

2.5 Techniques for Encouraging Organic Digestion Processes

The slaughterhouse wastewater was pumped from sectional OP pond 2 to distribute for all VFCW units together

with alternating between stagnating storage for 5 days (to provide the occurrence of bacterial organic digestion processes and soil filtration by vertical flowing) and releasing the treated wastewater for the next 2 days (to drain out treated wastewater) until the the cutting period of growing aquatic plants with 25-cm spacing in which the visible light will be induced to water surface as used for photosynthesis while rapid evapotranspiration occurs for oxygen diffusion through the water surface by thermo-siphon processing. Nevertheless, the air from the atmosphere will be brought through the thermo-osmosis process by the oxygen-producing photosynthesis in spongy cells (aerenchyma cells) of very young leaf that belonging to aquatic plants. The growing aquatic plants (such as typha, cyperus,etc.) plays vital role in absorbing plant nutrients and toxic chemicals under the phytoremediation process for keeping nutrients balancing in soils.

2.6 Wastewater Sample Collection and Water Quality Analysis

As mentioned beforehand, the VFCW treatment efficiency for slaughterhouse wastewater of OD ponds1 and 2 was intentionally proposed to take part in economizing operation cost in relation to the toxicant effluent elimination after releasing to the public water sources. In order to meet the target, the treated wastewater from OP ponds1 and 2 were applied for second treatment by two-consecutive sets of 3 VFCW (including 1 control in each set). At the same time, the typha aquatic plants were grown in all six VFCW units (except 2 control units) along with 25 x 25 - cm spacing. Flowing slaughtering wastewater at depth of 30 cm above VFCW beds for 5-day stagnation and 2-day releasing until the planting typha met the matured stage, about the twelfth weeks and/or flowering period. The matured typha plants were cut for determining biomass before analyzing some chemical composition.

Wastewater samples were collected before transferring into the small VFCW units and on the seventh days during releasing for analyzing COD, BOD, TDS, SS, EC, DO,Temperature, pH, coliform bacteria in order to determine the existing slaughterhouse wastewater and treatment efficiency of VFCW units for the second treatment of the effluent from OD effluent which are shown the BOD and COD concentration normally higher than standard values. Practically, there are six VFCW-Typha units (plus another two control units each) in size of 3-m width, 30-m length and 1-m depth at AmphoeMuang municipal in Suphanburi province as located in the western part of Thailand about 110 km from Bangkok. The hog slaughterhouse wastewater was pumped from the sectional ponds1 and 2 into two sets of one for control and three for vertical flow constructed wetlands at the level of 30-cm depth from the bottom; the sectional pond1 for the first set and the sectional pond2 for the second. Consecutively, the hog slaughterhouse wastewater was alternatively treated with the stagnated storage for 5 days and releasing for 2 days to the public watercourse. The said operation was continually taken in time after time until typha growing up to the harvesting stage at the age about 12-week period. During planting and maturing age, the effluent as obtained from the vertical flow through soil, sand, and gravel was sampled at the ended-pipe outlet every 7 days in order to analyze COD and BOD as well as the concerned factors,I.e., DO, TDS, TSS, TKN, total and Ortho phosphorus, EC, alkalinity, pH and coliform bacteria. At the same manner, the hight growth of typha was measured every 7 days on one spot at the middle and another four spots at the conners of the vertical flow constructed wetlands for determining the maximum height-growth age (mature age) in which the aquatic typha should be functioned with the most effectiveness for wastewater treatment through the VFCW units. Immediately after harvesting the typha leaves had to sample randomly for determining its biomass and also the accumulated nutrients of soils at 15-cm soils to extract the chemicals of both the toxicants and plant nutrients as well. The decreases of COD and BOD concentration as well as their related quality indexes such as DO, TDS, TKN, Total P, Ortho P, EC, Temperature, and pH were analyzed under the standard methods as recommended by APHA,AWWA,WEF (2005), and also determining soil nutrients by Black et al (1965)

3. Results and Discussion

3.1 Nature of Pig Slaughterhouse Wastewater

As mentioned before, the Oxidation Pond System with 4-sectional ponds is well known among environmental engineers and technologists as well as the the environmental administrators for pig slaughterhouse wastewater treatment (Palatri et al 2011, Othman et al 2013, Cao and Mehrvar 2011). Regrettably, it was found less efficiency for decreasing high concentrated COD and BOD of slaughterhouse wastewater. Accordance with such results, the primarily experiment was conducted by taking the role of 24, 27, and 32 pigs (DIW-MOI 1996) from slaughtering and dissecting in relation to water-use quantity and wastewater quality. according with experimental results, the OP effluent quality indicators COD and BOD were over standard values while the others are opposite as illustrated in Table1.

Table 1. Quantity and quality of wastewater (effluent) from hog slaughtering and dissecting activities of Suphanburi municipal slaughterhouse through the 4-sectional ponds of Oxidation Pond

No.	Water Quality Indicators	Unit	Number of Hogs				Standard Values
			24	27	32	Average	
1	COD	mg/l	1,450.9	1,424.9	1,403.8	1,426.3	120
2	BOD	mg/l	34.7	36.7	30.6	34.03	20
3	TDS	mg/l	254.7	315.3	215.0	261.7	3,000
4	EC	µs/l	506.3	631.3	453.3	530.3	-
5	DO	mg/l	1.9	3.5	3.6	3.0	4.0
6	Alkalinity	mg/l	286.0	392.7	297.3	325.3	-
7	pH		7.8	7.9	7.9	7.9	5.5
8	Temperature	^{o}C	31.8	37.4	35.5	34.6	40.0
9	Wastewater Quantity	cu.m.	5.95	7.3	9.8	-	-

It is surprising to point out that the COD values were high up to 1,450.9 mg/L for 24 slaughtering pigs, 1,424.9 mg/L for 27 pigs, 1,403.8 mg/L for 32 pigs and averaged 1,426.30 mg/L DOPC-MONRE (2010)because the higher ions could be obtained from antibiotics and growth stimulants in blood and liquid-solid wastes during and after pig slaughtering. Consequently,such rapid organic digestion rate has theoretically paid significant role in lowering the effluent quality as indicated in BOD values down to 34.7 mg/L for 24 pigs killing, 34.7 mg/L for 27 pigs, 30.6 mg/L for 32 pigs, and averaged 34.3 mg/L. In other words, the rapid decrease of BOD can make increasing COD during transferring the slaughterhouse wastewater through not only killing points but also flowing from the pond1 through ponds 2, 3, and 4 in form of effluent of Oxidation Pond (Masse and Masse 2000 a and b, Borja 1995, Tritt and Schuchard 1992, Borja et al 1998). Also, killing diseases after pig slaughtering activities with chlorine solution as obligated by Public Health lawsn (MOPH 1998) declaring would be another cause of drastic increasing COD in OD effluent (Couillaed et al 1989).

Table 2. Averaged values of slaughterhouse wastewater quality from OD sectional ponds1 and 2 before and after draining into VFCW-Typha and VFCW-Control units at pig slaughterhouse in Suphanburi province

No.	Item/Unit	Averaged Values of Wastewater Quality													
		COD	BOD	TDS	TKN	NH_3-N	Total P	Ortho P	DO	Alkalinity	EC	pH	Temp.	Coliform Bacteria	Fecal Bacteria
		mg/l	mg/l	mg/l	mg/l	mg/l	mg/l	mg/l	mg/l	mg/l	µs/cm		^{o}C	MPN/ 100 ml	MPN/ 100 ml
After pond 1	1.VFCW-Typha														
	Before	328.80	181.98	441.98	56.03	32.55	12.38	11.01	0.50	431.00	846.00	6.88	31.40	$2.1x10^6$	$1.1x10^6$
	After	129.19	57.67	310.13	29.71	19.42	10.78	9.47	1.52	396.67	554.73	7.55	27.92	$2.9x10^4$	$2.0x10^4$
	Efficiency (%)	59	70	30	47	40	13	13	-	8	36	-	11	96	98
	2.VFCW-Control														
	Before	328.80	181.98	441.98	56.03	32.55	12.38	11.01	0.50	431.00	846.00	6.88	31.40	$2.6x10^6$	$1.9x10^6$
	After	174.68	51.60	212.33	26.98	18.29	11.15	10.45	4.00	379.20	495.20	7.73	29.32	$4.0x10^4$	$2.0x10^4$
	Efficiency (%)	44	66	52	52	43	10	15	-	12	42	-	6	95	99
After pond 2	1.VFCW-Typha														
	Before	151.92	79.14	271.86	26.62	18.64	12.73	11.98	1.72	369.12	616.80	7.50	29.30	$2.6x10^5$	$1.5x10^5$
	After	90.92	31.67	217.07	5.71	2.05	10.73	10.21	1.52	331.73	474.33	7.37	27.74	$1.3x10^4$	$2.9x10^3$
	Efficiency (%)	40	61	20	79	89	15	13	-	10	23	-	5	95	98
	2.VFCW-Control														
	Before	151.92	79.14	271.86	26.62	18.46	12.73	11.98	1.72	369.12	616.80	7.50	29.30	$1.3x10^5$	$1.4x10^3$
	After	95.92	28.20	266.53	4.00	2.10	11.10	10.21	1.50	293.67	441.60	8.50	28.98	$6.0x10^3$	$1.1x10^3$
	Efficiency (%)	36	66	2	85	89	15	15	-	20	33	-	1	97	99
	Standard	120	20	3,000	100	25	-	-	4.00	-	-	5.5	40	5,000	1,000

3.2 Pig-Slaughterhouse Wastewater Treatment and VFCW

According to pig slaughterhouse wastewater is normally contained a lot of organic content in form of chemical oxygen demand (COD) and biological oxygen demand (BOD) which found from the previous researches. In order that statement, the primary experiment was focused on COD concentration in pig slaughterhouse wastewater but BOD excluded because it could be estimated in the range between 40 % to 60 % of COD. Anyway, the DOPC-MONRE (2010)studied on the measurement of COD which contaminated on the ground surface of pig slaughtering areas found in averaged value of 1,426.3 mg/L, higher than standard 120 mg/L (Table 1). After it was travelled through OP pond 1 with HRT about 2 days, the effluent contained with COD 328.8 mg/L that turning to be wastewater treatment efficiency of OP pond1 about 77 %. In consequence, it became OP-pond 2 effluent approximately 151.92 mg/L (Table 2) and in turn to OP pond 2 efficiency 84 %. In case of Oxidation Pond for pig slaughterhouse wastewater treatment was supposed to have only 2 OP ponds, (pond 1+2) the treatment efficiency could be calculated about 96 % as obtained from COD 1,426.3 mg/L before draining into pond1 and becoming effluent COD 328.8 mg/L, and finally pond2 effluent COD 151.92 mg/L. (COD standard 120 mg/L). Summarily speaking, the treatment efficiency of Oxidation Pond alone cannot be able to decrease the high concentrated COD of pig slaughterhouse wastewater but it really needs another supporting wastewater treatment technology to eliminate the above-standard COD remainders down to satisfactory level.

Besides COD concentration 1,426.30 mg/L as found on the ground surface of pig slaughtering areas, another water quality indicators were used for making sure before selecting the wastewater treatment to support the OP technology. Anyway, the most popular indicators were used for evaluating the better way on how to take them identifying the accuracy of treatment efficiency of pig slaughterhouse wastewater treatment as informed by Cauillard et al (1989) and Hammer (1989) and Hammer and Bastian (1989), Paltri et al (2011), Pattamapitoon (2013),Hadad et al (2006) andTritt and Schuchardt (1992). For response the previous words, the study will pinpoint on BOD, coliform bacteria, and fecal bacteria as the added indicators for precise hypothesizing the supporting technology for pig slaughterhouse wastewater treatment.

For response previous statement, the second study was paid attention on the quantity and quality of pig slaughterhouse wastewater from the ground surface after passing through OD Pond 1 as called the effluent of OP pond 1, which was found the status of COD 328.80 MPN/100mL, BOD 181.98 mg/L, coliform bacteria 2.1 x 10^{-6} MPN/100mL, and fecal bacteria 1.1 x 10^{-6}, while OD pond 2 effluent indicated COD 151.92 mg/L, BOD 79.14 mg/L, coliform bacteria 2.6x10^{-5}MPN/100mL, and fecal bacteria 1.5 x 10^{-5} MPN/100mL (see Table2). The existing measured values of four indicators were still higher than standards that existed more contaminants of pig meat, grease, hairs, urine, dunk, and blood in which some of them might be the bacteria inhibitors as described by Cauillard et al (1989) and Hammer (1989) and Hammer and Bastian (1989) Hammer (1989) Pattamapitoon (2013) Hadad et al (2006) and Tritt and Schuchardt (1992). In the same way, Hammer (1989) Pattamapitoon (2013) Hadad et al (2006) Palatsi et al (2011) and Tritt and Schuchardt (1992) concluded that there were uncertain condition to complete the bacterial organic digestion processes, it depends on solar radiation, air temperature, atmospheric moisture, washing water quality, organic properties, bacteria types, and topography. Fortunately, the vertical flow constructed wetlands (VFCW) can eliminate all stated pig organs as found by LERD (1999), Kayser and Hunst (2005), Cui et al (2010), Soroko (2007), Molle et al (2008), Sklarz et al (2009), Tripathi and Shukla (1991), and Ye et al (2001). In principles, wastewater is transferred from point sources on the water surface in VFCW units, then wastewater as treated by bacterial organic digestion processes is sunk to the bottom before penetrating through in order of soils, sand, gravel, and flowing in pipe holes, finally flowing out at ended-pipe part as effluent of VFCW units. For clear understanding, the special was aimed at the relation between COD, BOD, coliform and fecal bacteria in treated wastewater moving out from OD ponds (1 and 2) and VFCW (Typha and Control) units, alternately to COD, BOD, coliform and consideration fecal bacteria in effluent moving out from both VFCW units as shown in Table 2.

3.3 Role of VFCW for Pig Slaughterhouse Wastewater Treatment

Special consideration was aimed at the relation between COD, BOD, coliform and fecal bacteria in treated wastewater moving out from OD ponds (1 and 2) and VFCW (Typha and Control) units, alternately to COD, BOD, coliform and fecal bacteria in effluent moving out from both VFCW units as shown in Table2 and described as follows:-

Case1: VFCW-Typha Capacity

Due to the fact that VFCW-Typha efficiency for retreating the high concentrated COD and BOD effluent from was paid attention on influent from OP ponds1 and 2 which were needed to convince the failure of this engineering tool being not suited for the pig slaughtering in Thailand. Also, it would be the basic understanding

to look for the supported technology decreasing COD and BOD content under the standard concentration. The details will be focused on the minimum concentration of COD, BOD, coliform and fecal bacteria rather than efficiency of wastewater treatment capacity.

1) OP Pond 1 Influent Treatment Capacity

The effluent quality from VFCW-Typha units as obtained treated wastewater from OD pond1 found COD 129.80 mg/L, BOD 57.67 mg/L, coliform bacteria 2.9 x 10^{-4} MPN/100mL, and fecal bacteria 2.0 x 10^{-4} MPN/100mL, in turn of treatment efficiency, COD 59 %, BOD 70 %, coliform bacteria 96 %, and fecal bacteria 98 %. (see Table2).

2) OP Pond2 Influent Treatment Capacity

The effluent quality from VFCW-Typha units as obtained treated wastewater from OD pond2 found COD 90.92 mg/L, BOD 31.67 mg/L, coliform bacteria 1.5x10^{-4} MPN/100mL, and fecal bacteria 2.9 x 10^{-3} MPN/100mL, in turn of treatment efficiency, COD 40 %, BOD 61 %, coliform bacteria 95 %, and fecal coliform 98 % (see Table 2).

The findings were satisfied in COD concentration after VFCW-Typha for pig slaughterhouse wastewater treatment in terms of minimum COD 90.92 mg/L (120 mg/L as standard value) from OP pond2 but opposite from OP pond 1. In addition, the concentration of BOD, coliform and fecal bacteria were not quite acceptable under standard values. However, the said concentrated COD, BOD, coliform and fecal bacteria effluent from OP ponds 1 and 2 were shown in higher than standard values and the same trend from research results as found by Juwarkar et al (1995), Maine et al (2006), Khan et al (2009), Ahn and Mitsch (2002), Cui et al 2010), Couillard et al (1989), Hosetti and Frost (1995), Soroko (2007), Molle et al (2008), Sklarz et al (2009), Luangsoonton (2010), and Pattamapitoon (2013). So, it is necessary to conclude that not only VFCW-Typha is able to be the supported technology but also OP ponds 1 and 2 have to more than 2 ponds and also enlarging unit size (longer 30 m. and/or wider 3 m.) in order to make sure for lowering COD, BOD, coliform and fecal bacteria in VFCW-Typha effluent down to under standard values.

Case2: VFCW-Control Treatment Capacity

The effluent quality from VFCW-Control units as obtained treated wastewater from OD pond1 found COD 174.68 mg/L, BOD 51.60 mg/L, coliform bacteria 4.0 x 10^{-4} MPN/100mL, and fecal bacteria 2.0 x 10^{-4} MPN/100mL, in turn of treatment efficiency , COD 44 %, BOD 66 %, coliform bacteria 95 %, and fecal bacteria 99 %.

The effluent quality from VFCW-Control units as obtained treated wastewater from OD sectional pond 2 found COD 95.92 mg/L, BOD 28.20 mg/L, coliform bacteria 6.0 x 10^{-3} MPN/100mL, and fecal bacteria 1.1 x 10^{-3} MPN/mL, in turn of treatment efficiency, COD 36 %, BOD 66 %, coliform bacteria 97 %, and fecal bacteria 99 %.

Expectedly, the concentration of COD, BOD, coliform and fecal bacteria in VFCW-Control and VFCW-Typha effluent were the same trend as found from previous research results. Anyway, it is quite obvious that the treatment capacity of VFCW-Control cannot competed against VFCW-Typha because of an increasing oxygen supply as energy for bacterial organic digestion processes to convert organic substances in forms of COD and BOD becoming inorganic materials, but in opposite direction for coliform and fecal bacteria accordance with the digesting bacterial death from solar radiation, especially ultraviolet (Pattamapitoon 2013, Cao and Mehrvar 2011, Kayser 2005, Grosse1989, Mirmov and Belyakava 1982, Nopparatanaporn 1992, Palatri et al 2011).

Case3: VFCW-Typha and VFCW-Control Relations

Above results are informed that treatment efficiencies between OD ponds 1 and 2 between sectional ponds 1-2 and VFCW (Typha-Control units) have shown in low efficiencies, except coliform and fecal bacteria seemed high but the effluent still higher than standards, as illustrated in Table 2. There is somewhat higher efficient COD and BOD in effluent from VFCW-Typha units than VFCW-Control units but the another water quality indicators were opposite, especially coliform and fecal bacteria (see Table 2). High bacterial organic digestion processing rates in VFCW-Typha were expected to convince such mentioned statement but direct solar energy from sky to the open space of VFCW-Control was surely affected to kill the coliform and fecal bacteria one way or another.

Close observation on the Tables 1 and 2 found that theanother effluent quality indicators,i.e. TDS, EC, alkalinity, and pH were under standard values except DO was opposite. When the chemical analysis of slaughterhouse wastewater was taken in consideration, this experiment did not extract ammonia-nitrogen, total P, Ortho P, coliform and fecal bacteria which plays vital role in contamination from pig slaughtering and dissecting activities

and also for phytoremediation by Typha growth in vertical flow constructed wetlands. Therefore, they must be included the in-depth study on VFCW application for second treatment of OD effluent which were identified as too high concentrated COD and BOD effluent.

3.4 VFCW-Typha for Supporting OD Effluent Treatment

In principles, the VFCW units would be functioned like "value added technology" for second treatment of OD effluent for feasibility study by using only the OD ponds 1and 2. In operational point of view, the daily amount of slaughterhouse wastewater is not only small quantity but also discontinuity flow from OD ponds 1 to 2,3, and 4. This would be the reason of having a small part of inflow than OD storage capacity with stagnation conditioning as well as high evaporation rate all four sectional ponds on daytime, especially in wet season of the year. However, the previous scientific observation indicated that the decreasing organic wastes by bacterial organic digestion was relied on the oxygen supplying from both the photosynthesis of algae and thermo-siphon processing as caused by OD water surface evaporation process (Mirmov and Belyakava 1982, Deubigh and Raumann 1952, Poommai et al 2013, Grofse and Bauch 1991,Grosse 1989, Ameth and Stichimair 2001 ,Bearman 1957). Moreover, there was no existing submerged aquatic plants and less phytoplankton in turbid OD wastewater to produce the oxygen as energy supplying for bacterial organic digestion process. This would be another reason why the OD effluent found COD, BOD, and DO out of standards. Inevitably, the wastewater quality indicators of ammonia-nitrogen, total P, and Ortho P have to add up for gaining in-depth knowledge (Hammer 1989, Hammer and Bustian 1989, Jenssenet al 1993, Maine et al 2006, Mackey and Smail 1995)

3.5 Typha Useful Life for Pig Slaughterhouse Wastewater Treatment

Typha (Typhaangustifolia Linn.), cattail as common name, is classified as emergent aquatic plant which has aerenchyma cells (spongy cells) in leaf, especially very young leaf which allows the photosynthesizing oxygen going through its soft membrane to form rhizosphere around rhizomes before diffusing to bacteria in soils and wastewater as energy supplying for bacterial organic digestion processing as seen phenomena in Table 2. Accordance with previous saying, why the treatment efficiency of VFCW-Typha units looked better than VFCW-Control because of nature-by-nature processes through the thermo-osmosis techniques to supply oxygen from rhizosphere to supply energy for bacterial organic digestion until coming an end of typha useful life.

Naturally, wetland is supposed to be the balancing unit area among soils and its properties, organic wastes in wastewater, microorganisms, oxygen as energy supplying for bacterial organic digestion, aquatic plants and up-taking nutrient capacity as well as its toxic-chemical accumulative organs (Tripathi and Shukla 1991, Reddy et al 1991, Reddy et al 1990, Hadad et al 2006, Xia and Ma 2006). Therefore, an applicable VFCW-Typha for pig slaughterhouse wastewater is subjected to pay attention to convert the organic matters to become inorganic materials as plenty nutrients for typha growing, such as P 236 to 417 ppm and N 1.033 to 1,500 ppm after flooding OP pond1; and P 243 to 412 ppm and N 833 to 1,367 ppm after flooding OP pond2 (Table3). At the same time, the growing Typha are able to up-take toxic-chemicals to accumulate in all parts of its organs in order to maintain the constructed wetland in balancing stage. According to treat wastewater along with continuous flowing, Typha species as the aquatic plants have to be trimmed at useful age (Figure4) above submerged soil surface at full-water level in order to provide the new rhizome budding for supplying oxygen by photosynthesis and thermo-osmosis processes; and also to keep nutrient balance of constructed wetlands (Reddy et al 1990, Ye et al 2001, Xia and Ma 2006, Hadad et al 2006, Trpathi and Shukla 1991).

Table 3. Plant nutrient accumulation in soils and Typha leaves above ground and also including biomass and average height of Typha for pig slaughterhouse wastewater treatment

No.	Items	Nutrients		weeks	Average Height (cm.)	
		Sect.Pond1	Sect.Pond2		Sect.Pond1	Sect.Pond2
1	Biomass, kg/m^2	0.44	0.28	1	111	109
2	P in soils, ppm			2	147	147
	1.VFCW-Typha			3	182	169
	1) Before	236	243	4	199	188
	2) After	417	412	5	223	211
	2.VFCW-Control			6	241	224
	1) Before	348	312	7	262	244
	2) After	415	407	8	271	252

3	N in Soils, ppm			9	283	261
	1.VFCW-Typha			10	287	270
	1) Before	1033	967	11	294	279
	2) After	1500	1367	12	294	283
	2.VFCW-Control					
	1) Before	1100	833			
	2) After	1400	1300			

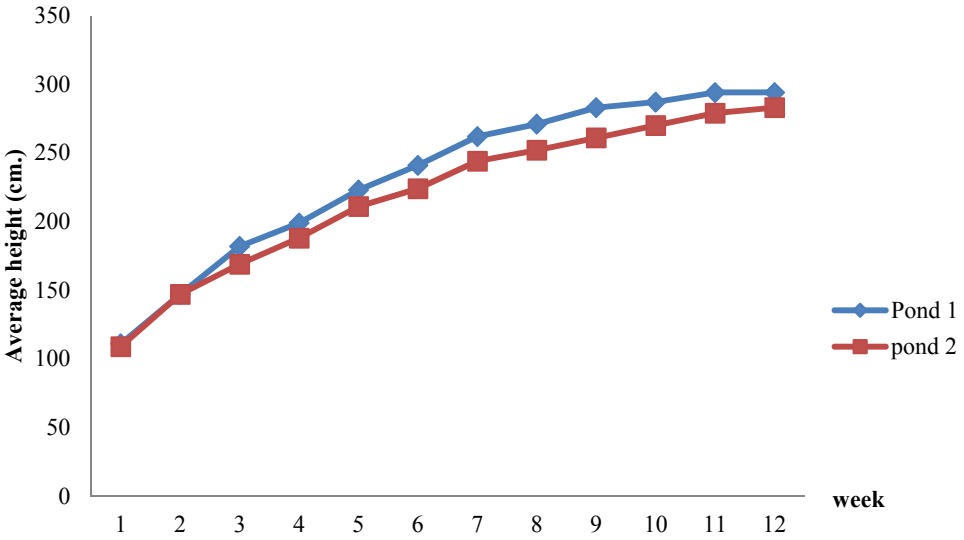


Figure 4. Average height for useful age determination as the maximum treatment of pig slaughterhouse wastewater in relation to flowering of Typha growing in vertical flow constructed wetlands

3.6 VFCW-Typha Treatment Modification

Previous findings learnt that the high concentration of organic matters as contaminated in pig slaughterhouse wastewater from activities of pig slaughtering and dissecting cannot be well treated good enough by OP engineering tool (ponds 1 and 2) in pond size of 3-m width, 30-m length and 5-m depth. Principally, the completeness of bacterial organic digestion process is really needed the longer HRT (hydraulic retention time) in order to have more time to convert the organic forms into the inorganic materials as indicated by LERD (1999) research on HRT 21 days with 5-consecutive oxidation ponds that would be enough to complete the bacterial digestion processing. Therefore,It is impossible to take the pig slaughterhouse 4-consecutive Oxidation Ponds with only 12-day HRT to complete such sophisticated processing. For future study, it is recommended to use the 4-consecutive Oxidation Ponds not to replace the old ones but it should evacuate the sludge in order to encourage the bacterial organic digestion process for effective reduction of COD, BOD, P, N, etc.

At the same procedure, the VFCW-Typha unit size of 3-m width, 30-m length, and 1-m depth is not enough size to obtain the bacterial organic digestion completeness for alternating 5-day stagnating and another 2-day releasing to out sources. Looking at the research results from effluent quality after passing through both the VFCW-Typha and VFCW-Control found that the COD and BOD values were lower than findings from Oxidation Ponds but all of them were still over standards. It is foreseen under research results of LERD (1999) that if the VFCW length were prolonged the HRT instead of 30 meters to 40 meters, the concentrated COD, BOD and another indicators would become under standard values. If the experimental areas are not enough to prolong up to 40-m length, the prolongation of 3-m width to become 4-m or 5-m width would be enough to solve the ineffectiveness of pig slaughterhouse wastewater treatment by 4-consecutive Oxidation Ponds in cooperating with vertical-flow constructed wetlands with the size of 4-m wide, 40-m long, and 1-m deep vertical-flow constructed wetlands which is integrated for pig slaughterhouse wastewater treatment from previous papers by Stottmister et al (2003), Borja et al (1998), Grosse (1989), Grofse and Bauch (1991), Deubigh and Raumann (1952), Jenssen et al (1993), Hammer (1989), Ahn and Mitsch (2002), Cui et al (2010), Boyd (1970), Khan et al (2009), Juwarkar et al 1995), Borkar and Mahatma (2011), Maine et al (2006), Keddy (2010), Ramsey et al

(2013), Hosetti and Frost (1995), Noparatanaporn (1992), Mackey and Smail (1995), and LERD (1999).

4. Conclusion

Due to Oxidation Pond (OP) as engineering tool is generally used for pig slaughterhouse wastewater treatment that containing high concentrated COD and BOD in effluent but it becomes point sources of stream pollution in concerned areas. For problem solving, the vertical-flow constructed wetland (VFCW) was proposed for increasing effectiveness in effluent retreating on excess organic pollutants. The experiment was focused on high concentrated COD and BOD effluent from OP ponds 1 and 2 as the influent of 2 sets of VFCW units, each set containing 3-units of VFCW-Typha plus 1-unit of VFCW-Control. The high COD and BOD content effluent as obtained from OP ponds1 and 2 were taken in the VFCW-Typha and VFCW-Control units for conditioning the alternative 5-day stagnation and 2-day before releasing to the public water sources. The results found as follows:

In primary study, the pig slaughterhouse wastewater samples on the ground floor were taken before draining into Oxidation Pond as the point sources and found COD 1,426.3 mg/L and still above standard after treating by the 2-consecutive oxidation ponds. Moreover, the secondary experimental study found COD 151.92 mg/L, BOD 79.14 mg/L,and coliform bacteria 2.6 x 10^{-5} MPN/100mL. They were above the standard values that needed the VFCW-Typha units for decreasing all excess organic forms. After treating the effluent by VFCW-Typha from the 2-consecutive oxidation ponds, the results found COD 90.92 mg/L, BOD 31.67 mg/L and coliform bacteria 1.5 x 10^{-4} MPN/100mL which were almost above the standard values. Furthermore, the VFCW-Control units cannot compete against VFCW-Typha.

References

Ahn, C., & Mitsch, W. J. (2002). Evaluating the use of recycled coal combustion products in constructed wetland: an ecologic-economic modeling approach. *Ecological Modeling, 150*, 117-140. http://dx.doi.org/10.1016/s0304-3800(01)00477-x

Ameth, S., & Stichimair, J. (2001). Characteristics of thermo-siphon rebillers. *International Journal of Thermal Sciences, 40*, 385-391. http://dx.doi.org/10.1016/s1290-0729(01)01231-5

APHA, AWWA, WEF. (2005). Standard Methods for the Examination of Water and Wastewater. second edition, American Public Health Association, Washington, USA, 523 p.

Bearman, R. J. (1957). The thermo-osmosis of rare grass through a rubber membrane. *Journal of Physical Chemistry, 61*, 708-714.

Berkun, M. (2005). Effects of Ni, Cr, Hg, Cu, Zn, Al on the dissolved oxygen balance of streams. *Chemosphere, 59*, 207-215. http://dx.doi.org/10.1016/j.chemosphere.2004.11.086

Black, C. A., Evans, D. D., Ensminger, J. L., & Clark, F. F. (1965). Method of Soil Analysis. American Society of Agronomy, Inc., New York, 521 p.

Borja, R. (1995). Effect of organic loading rate on anaerobic treatment of slaughterhouse wastewater in a fluidized-bed reactor. Bioresource, *Technology, 52*, 157-162. http://dx.doi.org/10.1016/0960-8524(95)00017-9

Borja, R., Banks, C. J., Wang, Z., & Mancha, A. (1998). Anaerobic digestion of slaughterhouse wastewater usinga combination sludge blanket and filter arrangement in a single reactor. *Bioresource Technology, 68*, 125-133. http://dx.doi.org/10.1016/s0960-8524(98)00004-2

Borkar, R. P., & Mahatma, P. S. (2011). Wastewater treatment with vertical flow constructed wetland. *International Journal of Environmental Science, 2*, 590-603.

Boyd, C. E. (1970). Vascular aquatic plants for mineral nutrient removal from polluted water. *Economic Botany, 2444,* 94-103.

Cao, W., & Mehrvar, M. (2011). Slaughterhouse wastewater treatment by combined anaerobic baffled reactor and UV/H2O processes. Chemical Engineering Research and design. *Chemical Engineering Research and Design, 89*, 1139-1143. http://dx.doi.org/10.1016/j.cherd.2010.12.001

Cauillard, D., Gariepy, S., & Tran, F. T. (1989). Slaughterhouse effluent treatment by thermophilic aerobic process. *Water Research, 23*, 573-579. http://dx.doi.org/10.1016/0043-1354(89)90024-9

Cui, L. H., Ouyang, Y. Q., Lou, F., Yang, L., Chen, Y., Zhu, W. L., & Luo, S. M. (2010). Removal of nutrients from wastewater with Canna indica L. under different vertical flow constructed wetland conditions. *Ecological Engineering, 36*, 1083-1088. http://dx.doi.org/10.1016/j.ecoleng.2010.04.026

Deubigh, K. G., & Raumann, G. (1952). The thermo-osmosis of gas through a membrane.II. *Experimental Proceeding of Royal Science, 21*(A), 518-533.

DIW-MOI. (1996). Handbook of Environmental Management for Slaughterhouse. Department of Industrial Works, Ministry of Industry, p. 190.

DOPC-MONRE. (2010). Annual Report on Existing Pollution in Thailand. Department of Pollution Control, Ministry of Natural Resources and Environment, Bangkok Thailand, p. 158.

Faerge, J., Magid, J., & de Vries, F. T. P. (2001). Urban nutrient balance for Bangkok. *Ecological Modeling, 139*, 63-74. http://dx.doi.org/10.1016/s0304-3800(01)00233-2

Faulkner, H., Edmonds-Brown, V., & Green, (2000). Problems of quality designation in diffusely polluted urban streams--the case of Pymme's Brook, north London. *Environmental Pollution, 109*, 91-107. http://dx.doi.org/10.1016/s0269-7491(99)00227-4

Grofse, W., & Bauch, B. (1991). Gas transfer in floating-leaved plants. *Plant Physiology, 97,* 185-192. http://dx.doi.org/10.1007/BF00035391

Grosse, W. (1989). Thermo-osmosis air transport in aquatic plants affecting growth activities and oxygen diffusion to wetland soils: In Constructed Wetland for Wastewater Treatment. Edited by D.A. Hammer, Lewis Publishers, Inc., Chelsea, p.469-478.

Hadad, H. R., Maine, M. A., & Bonetto, A. (2006). Macrophyte growth in a plot-scale constructed wetland for industrial wastewater treatment. *Chemosphere, 63*, 1744-1753. http://dx.doi.org/10.1016/j.chemosphere.2005.09.014

Hammer, D. A. (1989). Constructed Wetlands for Wastewater Treatment -- Municipal, Industrial, Agricultural. Lewis Publishers, Inc., Chelsea, Michigan, 831 p.

Hammer, D. A., & Bastian, R. K. (1989). Wetland ecosystem: Natural Water Purifiers: In Constructed Wetland for Wastewater Treatment. Lewis Publishers, Michigan, 485 p.

Hosetti, B. B., & Frost, S. (1995). A review of the sustainable value of effluents and sludges from wastewater stabilization ponds. *Ecological Engineering, 5*, 421-431. http://dx.doi.org/10.1016/0925-8574(95)00005-4

Jenssen, P., Maehlum, T., & Krogstad, T. (1993). Potential use of constructed wetlands for wastewater treatment in northern Argentina environments. *Water Science and Technology, 28*, 149-157.

Juwarkar, A. S., Oke, B., Juwarkar, A., & Patnaik, S. M. (1995). Domestic wastewater treatment through constructed wetland in India. *Water Science and Technology, 32*, 291-294. http://dx.doi.org/10.1016/0237-1223(95)00637-0

Kayser, K., & Hunst, S. (2005). Processes in vertical flow reed beds: nitrification, oxygen transfer and soil clogging. *Water Science and Technology, 51*, 177-184.

Keddy, P. A. (2010). Wetland Ecology: Principles and Conservation. Second Edition, Cambridge University Press, Cambridge, UK, 497 p.

Khan, S., Ahmed, I., Shah, M. T., Rehman, S., & Khaliq, A. (2009). Use of constructed wetland for the removal of heavy metals from industrial wastewater. *Journal of Environmental Management, 90*, 3451-3457. http://dx.doi.org/10.1016/j.jenvman.2009.05.026

LERD. (1999). Economized Technology for Community Garbage Disposal and Wastewater Treatment by Aquatic Plants. The King's ainitiative Laem Phak Bia Environmental Studies, Research and Development Project at Laem Phak Bia Sub-District, Ban Laem District, Petchaburi Province. *Thailand,* 420 p.

LERD. (2000). Science for Garbage Disposal and Wastewater Treatment towards the King's Initiative Project. Technical Papers for 10-Year LERD Seminar, Organized by ChaipatanaFoundation, during 24-25 August 2000 at Kasersart University. *Bangkok Thailand,* 560 p.

LERD. (2012). Annual Report on The King's Royally Initiated LaemPhak Bia sub-district, Ban Laem district, Petchaburi province Thailand. (serialnumbers of 2005 to 2011), 86 p.

Luangsoonton, S. (2010). Growth and wastewater treatment in Suphanburi municipal slaughterhouse wastewater.M.S. Thesis, College of Environment, Kasetsart University. *Bangkok Thailand,* 82 p.

Mackey, A. P., & Smail, G. (1995). Spatial and temporal variation in litter fall of Avicennia marina (Forssk) Vierh.in the Brisbane river, Queensland, Australia. *Aquatic Botany,* 133-142. http://dx.doi.org/10.1016/0304-3770(95)00490-Q

Maine, M. A., Sune, N., Hadad, H., Sanchez, G., & Bonetto, C. (2006). Nutrient and metal removal in a constructed wetland for wastewater treatment from a metallurgic industry. *Ecological Engineering, 26*, 341-347. http://dx.doi.org/10.1016/j.ecoleng.2005.12.004

Masse, D. I., & Masse, L. (2000a). Treatment of slaughterhouse wastewater in aerobic sequencing batch reactors. *Canadian AgriculturalEngineering, 42*, 131-137.

Masse, D. I., & Masse, L. (2000b). Characterization of wastewater from hog slaughterhouse in Eastern Canada and evaluation of their in-plant wastewater treatment systems. *Canadian Agricultural Engineering, 42*, 139-146.

Metcalf & Eddy (1979). Wastewater Engineering: treatment, disposal, reuse. Second Edition, McGraw-Hill, New York, 542 p.

Mirmov, N. I., & Beyakava, I. G. (1982). Heat liberation during vapor condensation in a thermo-siphon. *Journal of Engineering Physics, 43*, 970-974.

Molle, P., Prost-Boucle, S., & Lienard, A. (2008). Potential for total nitrogen removal by combining vertical flow and horizontal flow constructed wetlands: an full-scale experiment study. *Ecological Engineering, 34*, 23-29. http://dx.doi.org/10.1016/j.ecoleng.2008.05.016

MOPH. (1998). Technical Guidelines for Environmental Management of slaughterhouse in 1998. *Ministry of Public Health*, 210 p.

Nopparatanaporn, N. (1992). Microorganisms in Wastewater.Department of Microbiology, Faculty of Science, Kasetsart University, *Bangkok Thailand,* 412 p.

Othman, I., Amar, A. N., Rosman, Z. N. H., Harun, H., & Chelliapan, S. (2013). Livestock wastewater treatment using aerobic granular sludge. *Bioresource Technology, 133*, 630-634. http://dx.doi.org/10.1016/j.biortech.2013.01.149

Padgett, W. J. (1975). A stochastic model for stream pollution. *Mathematical Biosciences, 25*, 309-317. http://dx.doi.org/10.1016/0025-5564(75)90008-5

Palatri, J., Vinas, M., Guivernau, M., Fermandez, B., & Flotsi, X. (2011). Anaerobic digestion of slaughterhouse waste: main process limitation and microbial community interactions. *Bioresource Technology, 102,* 2219-2227. http://dx.doi.org/10.1016/j.biortech.2010.09.121

Pattamapitoon, T. (2013). Natural phenomena of solar radiation to reduce coliform and pathogenic bacteria of oxidation ponds for wastewater treatment system.Ph.D. Thesis, Graduate School, Kasetsart University, *Bangkok Thailand,* 75 p.

Poommai, S., Chunkao, K., Dumpin, N., Boonmang, S., & Nimpee, C. (2013). Determining the in-pipe anaerobic processing distance before draining to oxidation pond of municipal wastewater treatment. *International Journal of Environmental Science and Development, 4*, 157-162. http://dx.doi.org/10.7763/IJESD.2013.V4.326

Ramsey, D., Soldevila-Lafon, V., & Viladomiv, L. (2013). Environmental regulations in the hog farming section: A comparison of Catalonia, Spain and Manitoba, Canada. *Land Use Policy, 32*, 239-249. http://dx.doi.org/10.1016/j.landusepol.2012.10.020

Reddy, K. R., Agami, M., & Tucker, J. C. (1990). Influence of phosphorus on growth and nutrient storage by water hyacinth (*Eichhornia crassipes (Mart.)Solms*) plants. *AquaticBotany, 37*, 355-365. http://dx.doi.org/10.1016/0304-3770(90)90021-c

Reddy, K. R., Agami, M., & Tucker, J. C. (1991). Influence of potassium supply on growth and nutrient storage by water hyacinth. *Bioresource Technology, 37*, 79-81. http://dx.doi.org/10.1016/0960-8524

Sklarz, M. Y., Gross, A., Yakirevich, A., & Soasres, M. I. M. (2009). A recirculating vertical flow constructed wetland for the treatment of domestic wastewater. *Desalination, 246*, 617-624. http://dx.doi.org/10.1016/j.desal.2008.09.002

Soroko, M. (2007). Treatment of wastewater from small slaughterhouse in hybrid constructed wetland systems. *Eco-Hydrology and Hydrology, 7*, 339-343. http://dx.doi.org/10.1016/s1642-3593(07)70117-9

Stottmeister, U., Wiebner, A., Kuschk, P., & Kappelmeyer, U. (2003). Effects of plants and microorganisms in constructed wetlands for wastewater treatment. *Biotechnology Advances, 22*, 93-117. http://dx.doi.org/10.1016/j.biotechadv.2003.08.010

Streeter, H. W., & Phelps, E. B. (1925). A study of the pollution and natural purification of the Ohio River. *US Public Health Service. Bulletin Number, 146*, 96.

Suchkov, N., Darakas, E., & Ganoulis, J. (2010). Phytoremediation as a prospective method for rehabilitation of areas contaminated by long-term sewage sludge storage: A Ukrainian-Greek case study. *Ecological Engineering, 36*, 373-378. http://dx.doi.org/10.1016/j.ecoleng.2009.11.002

Tripathi, B. D., & Shukla, S. C. (1991). Biological treatment of wastewater by selected aquatic plants. *Environmental Pollution, 69*, 69-78. http://dx.doi.org/10.1016/0269-7491(91)90164-R

Tritt, W. P., & Schuchard, F. (1992). Materials flow and possibilities of treating liquid and solid wastes from slaughterhouses in Germany. *Bioresources Technology, 41*, 235-245. http://dx.doi.org/10.1016/0960-8524(92)90008-L

Tyagi, B., Gakkhar, S., & Bhargava, D. S. (1999). Mathematical modelling of stream DO-BOD accounting for settleable BOD and periodically varying BOD sources. *Environmental Modelling and Software, 14*, 461-471. http://dx.doi.org/10.1016/S1364-8152(98)00091-7

Xia, H., & Ma, X. (2006). Phytoremediation of ethion by water hyacinth fromwater. *Bioresource Technology, 97*, 1050-1054. http://dx.doi.org/10.1016/j.biortech.2005.04.039

Ye, J., Wang, L., Li, D., Han, W., & Ye, C. (2012). Vertical oxygen distribution trend and oxygen source analysis for vertical flow constructed wetlands treating domestic wastewater. *Ecological Engineering, 41*, 8-12. http://dx.doi.org/10.1016/j.ecoleng.2011.12.015

Ye, Y., Nora, F., Tam, Y., & Wong, Y. S. (2001). Livestock wastewater treatment by amangrove pot-cultivation system and the effect of salinity on the nutrient removal efficiency. *Marine Pollution Bulletin, 42*, 513-521.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Defining Thermophysical Parameters of Electric Devices Based on Solution of Inverse Heat Transfer Problem

Yuriy Alekseevich Bachvalov[1], Nikolai Ivanovich Gorbatenko[1] & Valeriy Viktorovich Grechikhin[1]

[1] Platov South-Russian State Polytechnic University (NPI), Russia

Correspondence: Valeriy Viktorovich Grechikhin, Platov South-Russian State Polytechnic University (NPI), 346428, Novotcherkassk, Prosveshenija, 132, Russia.

Received: December 18, 2014 Accepted: January 2, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p386 URL: http://dx.doi.org/10.5539/mas.v9n8p386

Abstract

This paper describes application of the study methods based on the solution of inverse problems of mathematical physics to define thermophysical parameters of electric devices. The mathematical model of the device is developed based on equations of non-stationary heat conductivity. The algorithm to define thermohysical parameters is developed; this algorithm uses the finite element method to solve a direct heat transfer problem and the gradient method to minimize the objective function. Examples of the algorithm application are given. The problem to define an equivalent heat transfer coefficient of the solenoid area covered with heavy winding and a heat emission from its inside surface is considered. In the second example thermophysical parameters of electromagnetic valve actuator of an ICE (internal combustion engine) gas distribution mechanism are defined. The obtained results show that thermophysical parameters and temperature distribution in non-stationary and steady-state operating conditions of electrical devices may be evaluated with adequate efficiency based on the solution of inverse heat transfer problems.

Keywords: inverse problems, heat transfer, non-stationary heat conductivity, electrical device

1. Introduction

The need to save energy and resources leads to creation of technical units characterized by marginal operation, high thermal, electro-magnetic and mechanical loads on their materials and structures. All of these require a reliable identification of the study subjects i.e. parameters and characteristics of materials and structures used in mathematical modeling shall be defined with an adequate accuracy. At the same time some of the parameters and characteristics can not be directly defined. In such cases the only way to obtain necessary information is to use a study approach based on solving inverse problems of mathematical physics (Alifanov, 1994), (Bachvalov *et al*., 2013), (Beck *et al*., 1985), (Bui, 1994), (Grechikhin and Grecova, 2011), (Korovkin *et al*., 2006), (Ozisik and Orlande, 2000), (Tikhonov and Arsenin, 1977), (Vatulyan, 2007). It is one of the main trends in studying physical processes and optimizing operating parameters of technical units and operating procedures.

Functioning of a number of technical units, including electric devices, is accompanied with heat transfer processes, which in their turn have an impact on the unit technical characteristics. Tougher tolerances are being established for temperature ranges of parts and devices, requirements to reliability in maintaining these ranges and reduction of material intensity of structures are getting tougher. Therefore thermophysical study of electric devices and calculation of their thermal rates becomes very important. Effectiveness of decisions made depends a lot on completeness and accuracy of heat transfer study. This justifies a necessity to conduct full-scale and modeling tests of devices (Grechikhin and Grecova, 2011). During the tests temperature $T_*(M_i,t)$ is measured in points M_i of the device on a bounded time interval $[0,t_e]$. Then, a thermal mathematical model of the device is adjusted by variation of a system of n thermophysical parameters $x_1,x_2,...,x_n$ in such a way, that the values $T_*(M_i,t)$ of the device agree with $T(M_i,t)$ of the model on the interval $[0,t_e]$ with an adequate accuracy. Then, using the model, desired time to achieve the steady-state condition t_s and temperature distribution of the device in this condition are defined.

Heat transfer and heat conductivity coefficients which values either unknown or known but with low accuracy can be used as adjustable thermophysical parameters. It is known, that temperature measurement data remain the main source of inaccuracy in solving applied problems to define these parameters (Beck *et al*., 1985). This is caused by performance features of temperature sensors. Therefore it is practical to identify heat transfer and heat

conductivity coefficients based on the solution of inverse heat transfer problems (Alifanov, 1994), (Beck *et al*., 1985), (Ozisik and Orlande, 2000). This paper describes a mathematical model and an algorithm to define thermophysical parameters of electric devices using such an approach.

Two problems are solved. The first one determines an equivalent heat transfer coefficient for a solenoid heavy winding of a measuring system designed to define magnetization curves (B-H curves) and hysteresis loops for ferromagnetic materials as well as a heat transfer coefficient between the solenoid inside surface and the ambient air. The second problem defines thermophysical parameters of an electromagnetic valve actuator in an ICE gas distribution mechanism.

This paper shows the results of using thermal testing methods for the stated devices based on the solution of an inverse heat transfer problem resulting in essential decrease of testing time and power consumption.

2. Methodology

2.1 Statement of the Problem

The studied electric device together with the ambient environment are represented as a multiply connected domain V (Fig. 1), with subdomains V_i , and heat sources with volume density q_v.

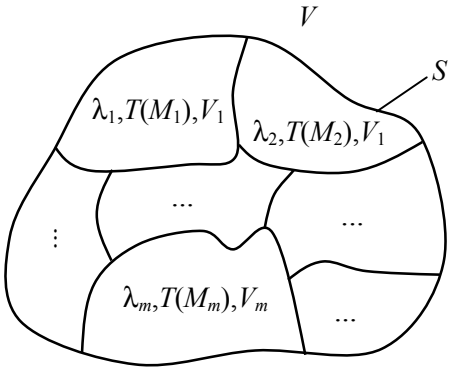


Figure 1. Multiply connected domain

Non-stationary temperature distribution in the domain V is described with a system of equations

$$\rho_i c_i \frac{\partial T_i(M_i)}{\partial t} = \mathrm{div}\left(\lambda_i \mathrm{grad} T_i(M_i)\right) + q_{v_i}\ ;\quad i = 1,2,...,m\ , \tag{1}$$

where $T(M_i)$ – temperature in the point $M_i \in V_i$; λ_i – heat transfer coefficient of the medium in V_i; q_{v_i} – volume density of the heat sources in a subdomain V_i; ρ_i – density of the medium in V_i; c_i – specific heat capacity of the medium in V_i; m – number of bodies in the studied domain V.

Boundary conditions are added to the system of equations (1):

$$\lambda \frac{\partial T}{\partial n} = -\alpha\left(T - T_{\mathrm{amb}}\right) \ \text{ on } \ S; \tag{2}$$

$$T(M_i) = T(M_{i+1}),\ \ \lambda_i \frac{\partial T}{\partial n}(M_i) = \lambda_{i+1} \frac{\partial T}{\partial n}(M_{i+1}) \tag{3}$$

at interfaces of media with different λ_i . Here, α – heat emission coefficient, T_{amb} – ambient temperature.

Initial temperature distribution in the domain V at a point of time $t = 0$ is considered to be known:

$$T(M_i,0) = T(M_i), i = 1,2,...,m\ . \tag{4}$$

Problem (1) – (4) describes heat transfer in linear and non-linear media. (1) – (4) forms a direct problem to find the function $T(M_i,t)$. Analytical (Polyanin *et al*., 2005) and numerical (Samarskii, 2001), (Zienkiewicz and Taylor, 2000) methods of solving this direct heat transfer problem are well known.

Let us consider an inverse problem where, for instance, in addition to $T(M_i,t)$, a heat transfer coefficient $\lambda(T)$ and a heat emission coefficient $\alpha(T)$ are unknown. The unknown coefficients shall be restored by solving the system (1) with conditions (2) − (4) and additional information

$$T_*\left(M^*,t\right)=\widetilde{T}_M\left(t\right),\ \ T_*\left(N^*,t\right)=\widetilde{T}_N\left(t\right),\ \ M^*,N^*\in V,\ \ t\in\left[0,t_{\text{и}}\right]. \tag{5}$$

where M^*, N^* – fixed points, where temperature is measured with an error ΔT_*.

The formed problem belongs to inverse heat transfer problems of a mixed type (coefficient and boundary). The studies (Alifanov *et al*., 1995), (Borukhov *et al*., 2005) prove the existence and uniqueness of the solutions to such problems. Solution stability is ensured by selecting them in a class of functions with a bounded norm (Tikhonov's stability) (Tikhonov and Arsenin, 1977).

2.2 Computational Algorithm

To define thermophysical parameters of electric devices based on the solution of an inverse heat transfer problem we shall use a conjugate gradient method (Alifanov *et al*., 1995), (Dinh and Reinhardt, 1998), (Rumyantsev, 1985). It is an iterative process of minimizing the objective function

$$J(\lambda,\alpha)=\int_0^{t_{\text{и}}}\left[T\left(N^*,t\right)-T_*\left(N^*,t\right)\right]^2dt+\int_0^{t_{\text{и}}}\left[T\left(M^*,t\right)-T_*\left(M^*,t\right)\right]^2dt. \tag{6}$$

Iterations, which define minimizing sequence of the function (5), are imposed with recursion

$$\lambda^{(n+1)}=\lambda^{(n)}-\rho\frac{\partial J^{(n)}}{\partial\lambda},\ \ n=1,2,\ldots; \tag{7}$$

$$\alpha^{(n+1)}=\alpha^{(n)}-\rho\frac{\partial J^{(n)}}{\partial\alpha},\ \ n=1,2,\ldots, \tag{8}$$

where $\lambda^{(n)}$, $\lambda^{(n+1)}$ – n and $(n+1)$ approximation for $\lambda(T)$; $\alpha^{(n)}$, $\alpha^{(n+1)}$ – n and $(n+1)$ approximation for $\alpha(T)$; $\frac{\partial J^{(n)}}{\partial\lambda},\frac{\partial J^{(n)}}{\partial\alpha}$ – partial derivatives of the function (6).

The iterative process is terminated when the equation

$$\int_0^{t_{\text{и}}}\left[T^{(n)}\left(N^*,t\right)-T_*\left(N^*,t\right)\right]^2dt+\int_0^{t_{\text{и}}}\left[T^{(n)}\left(M^*,t\right)-T_*\left(M^*,t\right)\right]^2dt\le\varepsilon\Delta T_* \tag{9}$$

is true,

where ε – given empirical parameter, defined using the criteria described in (Samarskii, 2001), (Tikhonov *et al*., 1995), $T^{(n)}\left(N^*,t\right),T^{(n)}\left(M^*,t\right)$ – problem solution (2) – (4) when $\lambda(T)=\lambda^{(n)}(T)$ and $\alpha(T)=\alpha^{(n)}(T)$.

Considering (6) – (9) we shall present the sequence for solving the inverse problem (2) – (5). We shall select initial approximation $\lambda(T)=\lambda^{(0)}(T)$, $\alpha(T)=\alpha^{(0)}(T)$ and increments $\Delta\lambda$, $\Delta\alpha$, ρ and Δt.

Iteration loop of the algorithm for each $n=0,1,2,\ldots$ consists of the steps as follows:

1. We solve a direct problem (2) – (4) where $\lambda(T)=\lambda^{(n)}(T)$, $\alpha(T)=\alpha^{(n)}(T)$ and define a temperature field including the values $T\left(N^*,t_j\right)$ and $T\left(M^*,t_j\right)$, $j=1,2,\ldots,n$, at the time points t_j.

2. We find a value of the function (6).

3. We check the condition for termination of the calculations (9). If the condition (9) is satisfied, values $\lambda(T)=\lambda^{(n)}(T)$, $\alpha(T)=\alpha^{(n)}(T)$ and $T^{(n)}\left(N^*,t\right),T^{(n)}\left(M^*,t\right)$ are considered to be the solution of the problem.

If the condition (4) is not satisfied, then go to 4.

4. We find the next values of the desired variables using a gradient method and the equations (7) and (8).

5. Go back to 1.

As a rule numerical implementation of this algorithm is based on the application of a finite element method (FEM) or a finite difference method.

3. Results and Discussion

3.1 Thermophysical Parameters of the Measuring System Solenoid

The applicable standards (IEC Standard 60404) for measuring static magnetic characteristics of soft magnetic materials are not adopted for testing under changing temperature and mechanical stresses. For instance, during operation the temperature of automobile starters and generators may reach 200 °C, their stress – 200 MPa. In cases like that it is practical to take measurements on an open magnetic circuit, using a solenoid to magnetize the material (Hall *et al*., 2009).

Let us consider a problem of finding thermophysical parameters of a solenoid of a measuring system designed to define magnetization curves (B-H curves) and hysteresis loops for soft magnetic materials. They include an equivalent heat transfer coefficient λ_2 of area *2* of the solenoid, covered with heavy winding, and a heat emission coefficient α_1 between the solenoid inside surface with a radius r_1 and the ambient air (Fig. 2).

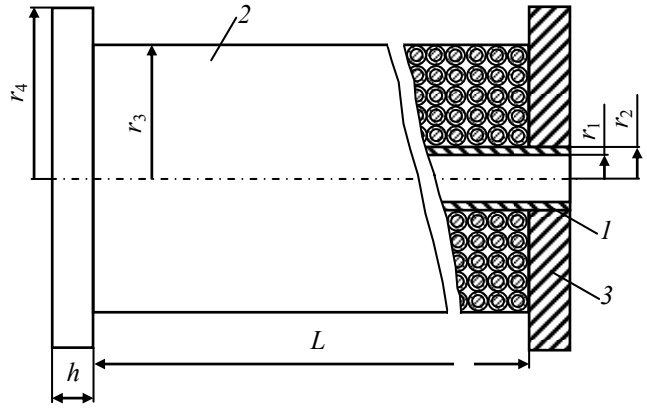


Figure 1. Sketch of a solenoid

Smallness of the radius r_1 makes it harder to measure thermophysical parameters inside the solenoid and allows for only one temperature sensor to be placed inside that area.

Considering axial symmetry of the solenoid, we will use a cylindrical coordinate system $r0z$.

In this case a system of equations (1) is

$$\rho_i c_i \frac{\partial T_i}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(\lambda_i r \frac{\partial T_i}{\partial r}\right) + \frac{\partial}{\partial z}\left(\lambda_i \frac{\partial T_i}{\partial z}\right) + q_{V_i},\ i = 1,2,...,m\ . \tag{10}$$

The initial temperature distribution in the solenoid at the point of time $t = 0$ $T(N_i,0) = T(N_i),\ i = 1,2,3$; and the ambient temperature T_{amb} are considered to be known.

Boundary conditions are:

- On the inside surface of solenoid area *1*

$$\lambda_1 \left.\frac{\partial T_1}{\partial n}\right|_{r=r_1} = -\alpha_1 [T_1(r_1,t) - T_{\text{amb}}];$$

- On the outside surface of solenoid area *2*

$$\lambda_2 \left.\frac{\partial T_2}{\partial n}\right|_{r=r_3} = -\alpha_2 [T_2(r_3,t) - T_{\text{amb}}];$$

– On the surfaces of area *3*, which are in contact with the ambient environment:

$$\frac{\partial T_3}{\partial n} = 0;$$

– At the media interfaces

when $r = r_2$

$$T_1(r_2,t) = T_2(r_2,t);\ \ \lambda_1 \frac{\partial T_1}{\partial n}\bigg|_{r=r_2} = \lambda_2 \frac{\partial T_2}{\partial n}\bigg|_{r=r_2};$$

– at the interfaces of areas *1* and *3*

$$T_1(N,t) = T_3(N,t);\ \ \lambda_1 \frac{\partial T_1}{\partial n} = \lambda_3 \frac{\partial T_3}{\partial n};$$

– at the interfaces of areas *2* and *3*

$$T_2(N,t) = T_3(N,t);\ \ \lambda_2 \frac{\partial T_2}{\partial n} = \lambda_3 \frac{\partial T_3}{\partial n},$$

where α_1, α_2 – heat emission coefficients between the inside and outside surfaces of the solenoid and the ambient air, respectively.

Computational domain is given in Fig. 3.

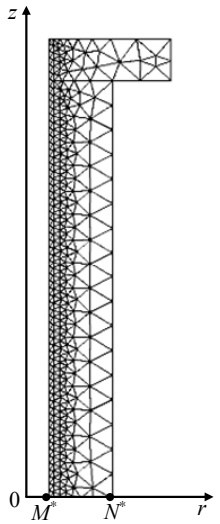


Figure 3. Computational domain with FEM mesh (709 elements)

M^*, N^* – points where temperature sensors are placed

Unknown values are λ_2, α_1 and $T_i = T(N_i,t), i = 1,2,3$.

We shall formulate a problem: it is required to define the equivalent heat transfer coefficient λ_2 of the area covered with the solenoid winding, the heat emission coefficient α_1 and the functions $T_i = T(N_i,t), i = 1,2,3$, which satisfy the system of equations (10) and the above given initial and boundary conditions; there are additional known data– functions $T_*(N^*,t)$ and $T_*(M^*,t)$, obtained by taking measurements in N^* and M^* (Fig. 3) with error ΔT_* on the time interval $[0,t_e]$.

Let us convert the solution of this problem to the solution of a sequence of direct problems – a system of equations (10) with the above given initial and boundary conditions using the FEM.

We shall select initial approximation $\lambda_2^{(0)}$, $\alpha_1^{(0)}$ and increments $\Delta\lambda_2$, $\Delta\alpha_1$, ρ_1, ρ_2 and Δt.

Then, according to the algorithm, we shall solve a direct problem using the FEM when $\lambda_2 = \lambda_2^{(n)}$, $\alpha_1 = \alpha_1^{(n)}$ and define the temperature field, including $T\left(N^*, t_j\right)$ and $T\left(M^*, t_j\right)$, $j = 1, 2, ..., n$, at the points of time t_j.

As in our case the functions $T_{\text{и}}\left(N^*, t\right)$ and $T_{\text{и}}\left(M^*, t\right)$ are represented as table data, instead of (6) we will use a function type

$$J^{(n)}\left(\lambda_2^{(n)}, \alpha_1^{(n)}\right) = \beta_1 \sum_{j=1}^{p} \left[T\left(N^*, t_j\right) - T_*\left(N^*, t_j\right)\right]^2 + \beta_2 \sum_{j=1}^{p} \left[T\left(M^*, t_j\right) - T_*\left(M^*, t_j\right)\right]^2 ,$$

where p – number of temperature measurements on the interval $[0, t_e]$; β_1, β_2 – weight factors, $\beta_1 = 1/3$; $\beta_2 = 2/3$.

According to (9), we shall check the condition for calculation termination. If the condition (9) is not satisfied we start calculating, similar to (7) and (8), the next values of the desired variables:

$$\lambda_2^{(n+1)} = \lambda_2^{(n)} - \rho_2 \frac{\partial J^{(n)}}{\partial \lambda_2};$$

$$\alpha_1^{(n+1)} = \alpha_1^{(n)} - \rho_1 \frac{\partial J^{(n)}}{\partial \alpha_1},$$

and solve the direct problem again using the FEM when $\lambda_2 = \lambda_2^{(n+1)}$, $\alpha_1 = \alpha_1^{(n+1)}$.

Let us consider using the algorithm to define the heat emission coefficient α_1 and the equivalent heat transfer coefficient λ_2 of the solenoid area, covered with heavy winding of a copper wire with the following parameters: $r_1 = 12 \cdot 10^{-3}$ m; $r_2 = 15 \cdot 10^{-3}$ m; $r_3 = 50 \cdot 10^{-3}$ m; $r_4 = 85 \cdot 10^{-3}$ m; $h = 25 \cdot 10^{-3}$ m; $L = 505 \cdot 10^{-3}$ m; $\lambda_1 = \lambda_3 = 0.3$ W/(m·K) (material – textolite); specific heat capacity of copper $c_{Cu} = 385$ J/(kg·K); density of copper $\rho_{Cu} = 8.9 \cdot 10^3$ kg/m^3; number of turns $w = 2514$; wire diameter (copper only) $2.1 \cdot 10^{-3}$ m; insulation thickness $0.2 \cdot 10^{-3}$ m; volume of copper $V_{Cu} = 1.02 \cdot 10^{-3}$ m^3; we shall consider that specific heat capacity and density values of other materials equal to zero; test time $t_e = 30$ min.

Table 1 gives measurements of the temperature $T_*\left(N^*, t\right)$ and $T_*\left(M^*, t\right)$ in the points N^* and M^*, the heat flux density q_*, from the surface of area *2*, the ambient temperature T_{amb} measured with ITP–MG4.03/3(1) Potok heat flux density and temperature measuring device with the solenoid winding powered from d.c. power source with voltage $U = 18.75$ V, current $I = 7.53$ A. The instrument error: ±6 % relative error for heat flux density, and ±0.2 °C absolute error for temperature.

Table 1. Experimental data

t, min	0	5	10	15	20	25	30
$T_*\left(N^*, t\right)$, °C	22	25.5	29.2	32.0	34.6	37.7	40.1
$T_*\left(M^*, t\right)$, °C	22	24.5	29.3	33.8	38.6	43.2	47.8
q_*, W/m^2	0	43.1	77.2	106.1	137.1	165.6	188.7

T_{amb}, °C	22	22	22	22	22	22	22

Using data from Table 1, we shall define the mean value of the heat emission coefficient from the solenoid surface

$$\alpha_2 = \frac{q_*}{T_*(N^*,t) - T_{amb}} = 11{,}1 \quad \text{W/(m}^2\cdot\text{K)}.$$

Now we shall find other solenoid parameters, reduced to the volume of area *2* (solenoid winding) V_2:

$$V_2 = \pi(r_3^2 - r_2^2)L = 3{,}1416\left[(50\cdot10^{-3})^2 - (15\cdot10^{-3})^2\right]\cdot 505\cdot10^{-3} = 3.61\cdot10^{-3} \quad \text{m}^3;$$

$$q_{V_2} = \frac{UI}{V_2} = \frac{18.75\cdot 7.53}{3.61\cdot10^{-3}} = 39122 \quad \text{W/m}^3;$$

$$\rho_2 c_2 = \rho_{Cu} c_{Cu} \frac{V_{Cu}}{V_2} = 8.9\cdot10^3\cdot 385\cdot\frac{1.02\cdot10^{-3}}{3.61\cdot10^{-3}} = 1.64\cdot10^6 \quad \text{J/(kg·K)}.$$

We consider that $\lambda_2^{(0)} = 0.01$ W/(m·K), $\Delta\lambda = 0.003$ W/(m·K), $\varepsilon = 10^{-4}$, $\alpha_1^{(0)} = 5$ W/(m^2·K), $\Delta\alpha_1 = 0.5$ W/(m^2·K), $\Delta t = 5$ min.

Computational domain (Fig. 2) is covered with the FEM mesh which consists of 709 triangles. Using the above algorithm, the fifth iteration gives us $\lambda_2 = 0.027$ W/(m·K), $\alpha_1 = 6.5$ W/(m^2·K).

Relative error of λ_2 and α_1 is not more than 3 % for the method used.

The standard deviation $T(M^*,t)$ of $T_*(M^*,t)$ with the found values λ_2 and α_1 on the interval [0, 30 min] was 1.3 °C (Fig. 3), which is acceptable. The same deviation is in the point N^*.

The developed mathematical model and the algorithm allowed defining temperature distribution in the solenoid in the steady-state condition (Fig. 4), which was reached within 12 h, as well as dependences $T(M^*,t)$, $T(N^*,t)$ (Fig. 5) and the steady-state temperatures $T_{st}(M^*) = 173.4$ °C, $T_{st}(N^*) = 80.9$ °C.

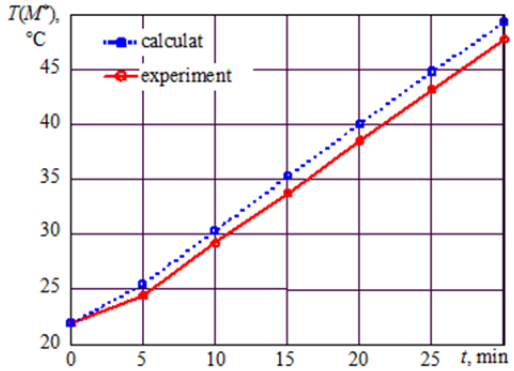


Figure 4. Results of measurements and temperature calculation in the point M^*

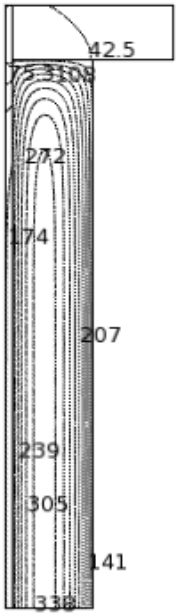


Figure 5. Temperature distribution in the solenoid cross-section in the steady-state condition

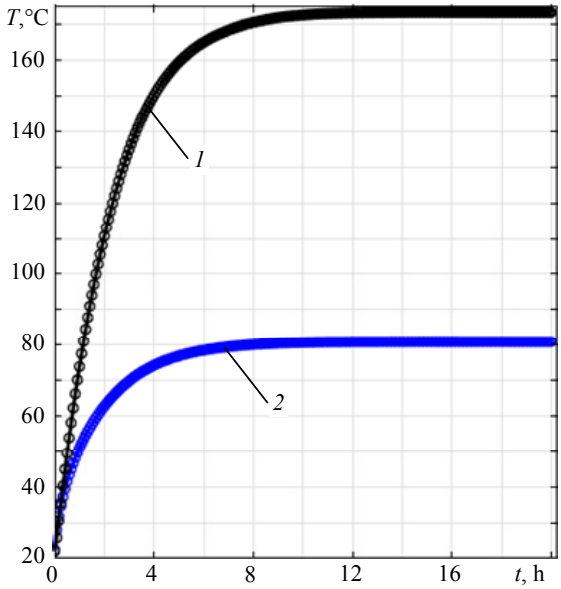


Figure 6. Temperature-time curve in the points M^* and N^*: *1* – $T(M^*, t)$; *2* – $T(N^*, t)$

Let us check if the first law of thermodynamics is satisfied in the reached steady-state (temperature T_{st}) – the capacity supplied to the solenoid winding shall be equal to the sum of heat energy radiated from the solenoid surface to the ambient environment:

$$UI = \alpha_1 S_{int}\left[T_{st}\left(M^*\right) - T_{amb}\right] + \alpha_2 S_{ext}\left[T_{st}\left(N^*\right) - T_{amb}\right], \tag{11}$$

Where $S_{int} = \pi \cdot 2r_1 L = 3.14 \cdot 24 \cdot 10^{-3} \cdot 505 \cdot 10^{-3} = 0.0384$ m^3, $S_{ext} = \pi \cdot 2r_3 L = 3.14 \cdot 0.1 \cdot 505 \cdot 10^{-3} = 0.1586$ m^3.

Based on (11) we have:

$18.75 \cdot 7.53 = 6.5 \cdot 0.0381 \cdot (173.4 - 22) + 11.1 \cdot 0.1586 \cdot (80.9 - 22)$ or $141.2 \approx 141.17$.

Consequently, heat transfer and heat emission coefficients of the test unit and temperature distribution in it can

be evaluated with an adequate accuracy based on the solution of the inverse heat transfer problems.

3.2 Thermophysical parameters of an Electromagnetic Valve Actuator of an ICE Gas Distribution Mechanism

Quick-action electromagnetic actuators are used as basic parts in many fuel delivery and air/gas mixing assemblies of ICEs. They are used for gas distribution mechanisms with individual valve actuators, blow off valves, gas recirculation systems and other devices, which improve energy, economic and environmental performance indicators of motors (Bolshenko, (2013), (Dresner and Barkan, (1989). The actuators work under conditions of excessive power loads, which makes it necessary to study heat parameters of the equipment operation.

A mathematic model represented by a system of equations (10) with boundary conditions on the outside surface of the actuator (2) is used to analyze heat transfer processes in the valve actuator of the ICE gas distribution mechanism (Figure 7).

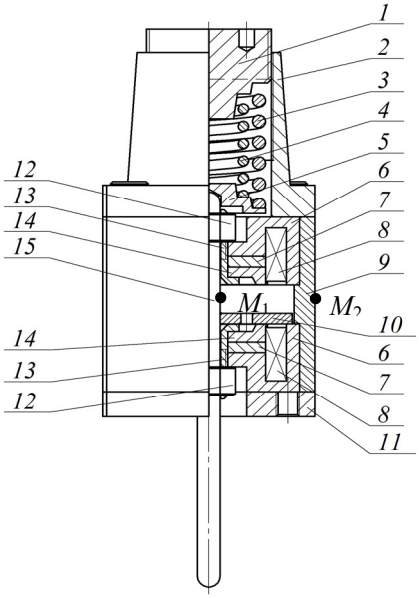


Figure 7. Electromagnetic valve actuator of a gas distribution mechanism

1 –adjusting screw; 2 – upper flange; 3 – external valve spring; 4 – internal valve spring; 5 – stop; 6 – solenoid mounting plate; 7 – constant magnet; 8 – coil; 9 – housing; 10 – anchor; 11 – lower flange; 12 – nut; 13 – bush; 14 – hub; 15 – rod

Unknown parameters in our case are equivalent coefficients of heat transfer λ_{eq}, specific heat capacity c_{eq} and equivalent density ρ_{eq} of the coils *8,* as well as a heat emission coefficient α of the housing *9*. Other parameters are known.

To solve an inverse heat transfer problem on a trial actuator and the time interval $[0,t_e]$ functions $T_*(M_1,t)$ in the point M_1 with the relative error $\delta(T_*(M_1))$ and heat flux density $q_*(M_2,t)$ in the M_2 with the relative error $\delta(q_*(M_2))$ are measured (Figure 7).

The inverse problem shall be solved using the following algorithm. We shall select initial values $\lambda_{eq}^{(0)}$, $c_{eq}^{(0)}$, $\rho_{eq}^{(0)}$. Then, we solve the system (10), (12) and find the function to be minimized:

$$J^{(n)} = \sum_{j=1}^{p} \beta_1 \left(\frac{T_j^{(n)}(M_1) - T_{j*}(M_1)}{T_{j*}(M_1)} \right)^2 + \beta_2 \left(\frac{q_j^{(n)}(M_2) - q_{j*}(M_2)}{q_{j*}(M_2)} \right)^2, \quad (12)$$

where β_1, β_2 – weight factors, $\beta_1 = 1/3$; $\beta_2 = 2/3$; $T_j^{(n)}(M_1) = T^{(n)}(M_1, t_j)$; $q_j^{(n)}(M_2) = q^{(n)}(M_2, t_j)$; $T_{j*}(M_1) = T_*(M_1, t_j)$; $q_{j*}(M_2) = q_*(M_2, t_j)$; t_j – point of time on the interval $[0, t_i]$: p – number of j^{th} points on the interval $[0, t_e]$.

Then, we shall check the condition

$$J^{(n)} \le \delta_{\max}^2, \tag{13}$$

where $\delta_{\max} = \max[\delta(T_*(M_1)), \delta(q_*(M_2))]$.

If the condition (13) is satisfied, the problem is solved. If not, then new values of the desired variables shall be found using a gradient method of minimizing the function (12), and we shall go to the start of the algorithm.

Let us consider using the described algorithm to solve an inverse heat transfer problem for the given actuator, (Figure 8): $D_{\text{ext}} = 56$ mm; $h_{\text{em}} = 49$ mm; h = 65 mm; H = 108.5 mm.

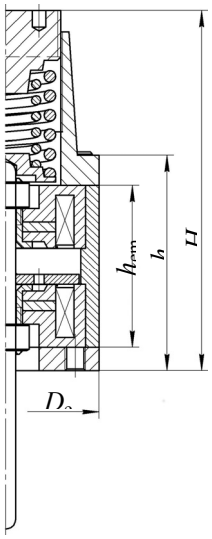


Figure 8. Main dimensions of the actuator

Design of the actuator, its dimensions, properties of the accessories (ρ, c, λ – density, specific heat capacity, heat transfer coefficient), except for the same parameters of the coils *8*, relative errors of the temperature measurement $\delta(T_*(M_1))$ and the heat flux density $\delta(q_*(M_2))$ are known.

The actuator parts: *1 – 6*, *10 – 12*, *14*, *15* are made from steel (grade 1010); *7* – NdFe35; *8* – copper; 9 – aluminum; *13* – brass .

Testo 922 is used to measure temperature, relative error of the temperature measurement $\delta(T_*(M_1)) = 1$ %, ITP–MG4.03/3(1) Potok is used to measure heat flux density, relative error $\delta(q_*(M_2)) = 6$ %.

As the result of the solution of the inverse heat transfer problem the following parameters are defined: $\lambda_{eq} = 0.03$ W/(m·K); $c_{eq} = 164$ J/(kg·K); $\rho_{eq} = 6800$ kg/(m^3); $\alpha = 18$ W/(m^2·K). Figures 8 and 9 show the functions on the time interval $[0.3 \times 10^3 \text{s}]$: *1* – experimental data; *2* – data received from the calculation using

λ_{eq}, c_{eq}, ρ_{eq} and α. The standard deviation $T_*(M_1,t)$ of experimental data is 2.7 %. The standard deviation $q_*(M_2,t)$ of experimental data is 5.7 %.

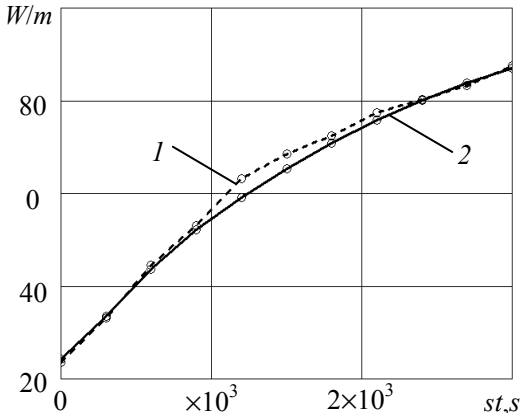


Figure 9. Results of temperature measurement and calculation in the point M_1

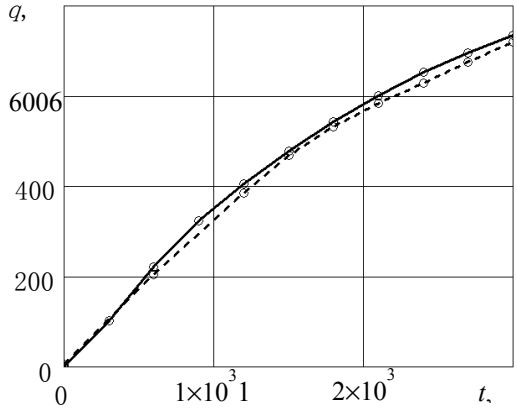


Figure 10. Results of heat flux density measurement and calculation in the point M_2

The undertaken studies of the heat transfer processes in the actuator using the developed model showed that at the ambient temperature of 150 °C, actuating pulses with the duration of 0.5 ms, the pulse period of 9.5 ms and the maximum current amplitude of 116 A, the temperature of the actuator winding in the steady-state is 196 °C and the temperature limit for such insulation class is 200 °C.

4. Conclusions

The results of the conducted study show that thermophysical parameters and temperature distribution in non-stationary and steady-state operating conditions of electrical devices may be evaluated with adequate efficiency based on the solution of inverse heat transfer problems.

The developed model and algorithm allow defining maximum temperature of a device as well decreasing time on thermal testing and electric power consumption.

It is planned that further studies will consider internal convection in a heat transfer model, and will help optimize the structure and the operating parameters of electric devices based on the obtained results.

Acknowledgements

The study results are obtained with the support of the project #1.2690.2014/K *Methods to solve inverse problems of complex system diagnostics (in engineering and healthcare industry) based on full-scale and modeling testing*, implemented within the scope of a designing stage for the state-given job. We would like to thank Grichenkova V.P and Bolshenko I.A. for their assistance in conducting the tests.

References

Alifanov, O. M. (1994). *Inverse Heat Transfer Problems*. New York: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-76436-3

Alifanov, O. M., Artyukhin, E. A., & Rumyantsev, S. V. (1995). *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*. New York: Begell House. ISBN:156700038X.

Bachvalov, Y. A., Gorbatenko, N. I., Grechikhin, V. V., & Grecova, A. N. (2013). Application inverse problems theory magnetic fields in the design of energy-saving electromechanical devices. *Izvestiya Vysshikh Uchebnykh Zavedeniy. Elektromekhanika, 5*, 28-32. Retrieved from http://electromeh.npi-tu.ru/assets/files/archive_5_2013.pdf

Beck, J. V., Blackwell, B., & Clair, C. R. jr. (1985). *Inverse Heat Conduction: Ill-Posed Problems*. New York: John Wiley and Sons. ISBN:0471083194. http://dx.doi.org/10.1002/zamm.19870670331

Bolshenko, I. A. (2013). Review and analysis of designs of electromechanical valve actuators for internal combustion engine. *Izvestiya vuzov. Severo-Kavkazskii region. Technical Sciences, 6*, 66-73. Retrieved from http://technauka.npi-tu.ru/assets/files/2013-6_annotations.pdf

Borukhov, V. T., Zayats, G. M., Tsurko, V. A., Timoshpol'skii, V. I., & Andrianov, D. N. (2005). Structural properties of dynamic systems and inverse problems of mathematical physics. *Journal of Engineering Physics and Thermophysics, 2*(78), 201-215. http://dx.doi.org/10.1007/s10891-005-0050-5

Bui, H. D. (1994). *Inverse problems in the Mechanic of Materials: An Introduction*. CRC Press. ISBN:0849384710.

Dinh, N. H., & Reinhardt, H. J. (1998). Gradient Methods for Inverse Heat Conduction problems. *Inverse Problems in Engineering, 3*(6), 177-211. http://dx.doi.org/10.1080/174159798088027675

Dresner, T., & Barkan, P. (1989). A Review and Classification of Variable Valve Timing Mechanisms. SAE Technical Paper, 890674. http://dx.doi.org/10.4271/890674

Grechikhin, V. V., & Grecova, A. N. (2011). Definition of parameters of mathematical models of potential fields by the full-scale modeling method. *Izvestiya Vysshikh Uchebnykh Zavedeniy. Elektromekhanika, 1*, 18-21. Retrieved from http://electromeh.npi-tu.ru/assets/files/archive_1_2011.pdf

Hall, M., Harmon, S., Patel, H., & Thomas, O. (2009). Obtaining the d.c. properties of soft magnetic materials using an open circuit measurement technique. *Pezeglad Elektrotechniczny, 1*(85), 28-30. Retrieved from http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.baztech-article-BPOC-0050-0007

IEC Standard 60404 part 4, Methods of measurement of d.c. magnetic properties of iron and steel.

Korovkin, N. V., Chechurin, V. L., & Hayakawa, M. (2006). *Inverse Problems in Electric Circuits and Electromagnetics*. Springer. http://dx.doi.org/10.1007/978-0-387-46047-5

Ozisik, M. N., & Orlande, H. R. B. (2000). *Inverse Heat Transfer: Fundamentals and Applications*. New York: Taylor and Francis. ISBN:156032838X.

Polyanin, A. D., Zaitsev, V. F., & Zhurov, A. I. (2005). *Methods for the Solution of Nonlinear Equations of Mathematical Physics and Mechanics*. Moscow: Fizmatlit. ISBN:5922105396

Rumyantsev, S. V. (1985). Ways of allowing for a priori information in regularizing gradient algorithms. *Journal of engineering physics, 6*(49), 1418-1421. http://dx.doi.org/10.1007/bf00871290

Samarskii, A. A. (2001). *The theory of difference themes*. New York – Basel: Marcel Dekker, Inc. http://dx.doi.org/10.1201/9780203908518

Tikhonov, A. N., & Arsenin, V. Y. (1977). *Solution of Ill-Posed Problems*. Washington, DC: Winston & Sons. ISBN:0470991240.

Tikhonov, A. N., Goncharsky, A. V., Stepanov, V. V., & Yagola, A. G. (1995). *Numerical Methods for the Solution of Ill-Posed Problems*. Dordrecht: Kluwer Academic Publishers. http://dx.doi.org/10.1007/978-94-015-8480-7

Vatulyan, A. O. (2007). Inverse Problems in Solid Mechanics. Moscow: Fizmatlit. ISBN:5922108352.

Zienkiewicz, O. C., & Taylor, R. L. (2000). *The finite element method* (1st ed.). Vol. 1: The Basis. Oxford: Butterworth-Heinemann. ISBN:0340759844.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Local Community and Tourism Development: A Study of Rural Mountainous Destinations

Mastura Jaafar[1], Norjanah Mohd Bakri[1] & S. Mostafa Rasoolimanesh[1]

[1] School of Housing, Building, and Planning, Universiti Sains Malaysia, Penang, Malaysia

Correspondence: S. Mostafa Rasoolimanesh, Post- Doctoral Fellow, School of Housing, Building, and Planning, Universiti Sains Malaysia, Penang, Malaysia. Tel: 604-653-5278. E-mail: mostafa@usm.my

Received: December 24, 2014 Accepted: January 14, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p399 URL: http://dx.doi.org/10.5539/mas.v9n8p399

The research is financed by the Ministry of Higher Education, Malaysia under the Long-Term Research Grant Scheme 2011 [LRGS Grant No. JPT.S (BPKI)2000/09/01/015Jld4(67)].

Abstract

Malaysia is internationally regarded as a popular rural destination because of its natural heritage. Rural tourism is increasingly viewed as a panacea for increasing the economic viability of marginalized areas, stimulating social regeneration, and improving the living conditions of rural communities. This study explores local community involvement in a rural tourism development in Kinabalu National Park, Sabah. We explore how the local community perceives their involvement in a local rural tourism development and look to identify the benefit of tourism destination development for this community. To address these objectives, we employed quantitative research methodologies and a sample of 378 respondents drawn from villages surrounding Kinabalu National Park. Sampled residents indicated having positive perceptions of tourism development in the area. Local communities enjoy being involved in the tourism sector because it improves their key income resources and quality of life.

Keywords: rural tourism, local community, community involvement, mountainous destinations, kinabalu national park

1. Introduction

Rural tourism has attracted renewed interest from researchers. In defining rural tourism, Lane (1994) notes several defining characteristics, such as a destination located in a rural area, functionally rural, small-scale, traditional, and largely locally controlled. Top rural tourism destinations, particularly in developing countries, normally consist of national parks, wilderness areas, mountains, lakes, and cultural sites. Okech et al. (2012) observes that these sites are important features of the rural economy.

Chaudhry and Gupta (2010) report that 75% of the world's poor live in rural areas and that rural tourism is a tool for rural revitalization. Rural tourism benefits local communities in terms of stimulating economic growth, valuing social cultural heritage, triggering the growth of service industries, and raising the standard of living; these benefits in turn encouraging positive attitudes and behaviors among these communities toward regard to tourism development (Jaafar et al., 2013; Nunkoo & Gursoy, 2012). Hall (1994) indicates that tourism is a viable means of promoting economic activity in developed and developing countries alike, and that rural tourism often enjoys substantial encouragement from both the public and private sectors (Fleischer & Pizam, 1997)

Malaysia has great potential in terms of nature tourism and ecotourism (Backhaus, 2003). Khalifah and Tahir (1997) indicate that Malaysia's tropical rainforests are among the oldest and most diverse ecosystems in the world. Covering an area of 753.7 km^2, Kinabalu National Park, in Malaysia, first became a national park in 1964. In 2000, the United Nations Educational, Scientific, and Cultural Organization (UNESCO) designated Kinabalu National Park a World Heritage Site (WHS). National parks function for the preservation of natural habitats and wilderness areas; however, they also facilitate educational, recreational, and tourism access. Consequently, the state government of Sabah has undertaken a series of initiatives to position Kinabalu Park as an international attraction and to encourage the local community to participate in its development. These initiatives, Liu (2006) suggests, demonstrate the government's use of rural tourism a mechanism for racial and spatial economic

restructuring to reduce regional disparities and to increase the economic control of the indigenous Bumiputra people.

Studies of rural tourism remain in their infancy, an opinion shared by Frederick (1993). While Suh and Gartner (2004) argues that tourism in peripheral areas has been extensively investigated, a review of the rural tourism literature reveals that disproportionately less attention has been afforded to rural tourism as compared to other forms of tourism; furthermore, tourism would appear to have been largely been ignored by rural economists. The rural tourism studies that do appear in the literature are diverse, indicative of a lack of focus in the area. For instance, Liu (2006) analyzes the socio-economic effect of tourism based on local participation in the tourism industry. Scheyvens (2002) investigates the participation of local communities in the decision-making processes, whereas Lacher and Nepal (2010) examine the participation of minority groups in the tourism industry. However, few studies have investigated the economic distribution of the benefits of tourism across a local community (Harrison & Schipani, 2007).

Researchers have rarely examined the economic advantages of mountain tourism in Malaysia. Most Malaysian highland studies have focused on the physical and social effects of mountain tourism, such as the effect on the diversity of small mammals on Mount Kinabalu (Shukor, 2001), community land and forest management systems (Horowitz, 1998), or the ecology (Maryati et al., 1996). Only two studies (Hasegawa et al., 2006; Liu, 2006) have correlated rural tourism in Malaysia with economic development. In view of this knowledge gap, the following research questions are posed:

1. To what extent is the local community involved in the development of rural tourism in Kinabalu Park?
2. How does the local community benefit from tourism development?

To answer these research questions, we undertake an extensive review of the literature concerning rural tourism and local community involvement in the tourism sector. In describing our methodologies, we give an account of our sampling procedures. In the results section, we describe the characteristics of our sample, their level of involvement in the tourism sector, and what benefits community members receive from tourism. Before concluding this paper, we undertake a discussion of the findings of this study in relation to previous studies.

2. Literature Review

Various efforts have been made to explain the key elements of rural tourism. Keane (1992) and Sharpley (2000) characterize rural tourism in terms of the destination's attractiveness to a type of tourist who enjoy unique or themed experiences exhibiting qualities of peace, relaxation, inspiration, recreation, local culture, and entertainment. Sharpley and Sharpley (1997) suggest that the sense of space, peace and tranquility, and escape from modern pressures are qualities which are intrinsic to rural tourism. In this century, however, researchers have come to view rural tourism from a wider perspective. For example, Macdonald and Jolliffe (2003) refer to rural tourism as a tool for the development of rural and isolated areas and as facilitating the growth of traditional industries. Eruera (2008) propose that rural tourism encompasses a wide range of activities, natural or man-made attractions, amenities and facilities, transportation, and marketing and information systems. While Ghaderi and Handerson (2012) associates low population density and minimal land area usage—providing tourists with an impression of space—with rural tourism.

The International Year of Mountains, in 2002, was a unified response to increased global awareness of mountain and tourism issues (Nepal & Chipeniuk, 2005). Mountain regions have high levels of both ecological and cultural diversity (Stepp, 2000), and Lama and Sattar (2002) argue that tourism is vital for the conservation and development of these highland regions. Highlands are rich in natural resources, including water, timber, minerals, biodiversity, and cultural heritage, making them attractive tourism destinations (Beedie & Hudson, 2003). Consequently, highland tourism has developed rapidly worldwide in recent decades (Moss & Godde, 2000), constituting 15–20% of worldwide tourism or US$70–90 billion per year, and has become a significant influence on the economies of many countries (Lama & Sattar, 2002).

2.1 Importance of Rural Tourism

Tourism is a means to stimulating local economic development (Gurung & DeCoursey, 2000). In Europe and elsewhere, rural tourism provides economic and social benefits to rural destination communities (Iorio & Corsale, 2010). Sharpley (2000) observes that rural tourism can act as a catalyst for socio-economic development and regeneration. Furthermore, rural tourism can supplement the incomes of impoverished agricultural cooperative settlements (Fleischer & Pizam, 1997), and provide new sources of income for families living in remote rural areas (Gale, 2006; Su, 2011).

Tourism contributes toward rural development because it provides another avenue for employment and income

generation, expands the market for local products, and revitalizes traditional economies (Azman et al., 2011; Ghaderi & Henderson, 2012). Hall (2004) notes that rural tourism benefits local communities by providing a supplementary income to the farming, craft, and service sectors. Hall (2004) also observes that rural tourism allows the economic value local food products to be reassessed, especially higher quality products which might otherwise escape special attention, and that rural tourism provides the opportunity to re-evaluate local heritage and its symbols, environment, and identity.

2.2 Community Participation and Rural Tourism

Thongma et al. (2011) suggest that the involvement of local communities is instrumental to the success of rural tourism development because these communities build more personable relationships wih visiting tourists and impress visitors with local cultural activities. Consequently, having had an enjoyable experience during their visit, the visitors leaves satisfied and more likely to revisit the same destination (Lo et al., 2013). Furthermore, local community participation provides locals with opportunities to enjoy the benefits of the development activities and empowers them to mobilize their capabilities through small business ventures (May-Ling et al., 2014). Therefore, the participation of the local community in rural tourism is a positive force for change and a catalyst for development (Claiborne, 2010).

Community participation need not necessarily be direct, as noted by Telfer and Sharpley (2007). For example, community participation can often focus on the decision-making process and non-economic benefits of tourism development (Tosun, 2000). Only when local communities are involved in decision making can their benefits be ensured, and their traditional lifestyles and values respected (Sheldon & Abenoja, 2001).

In the context of tourism planning, community participation can be defined thusly: "A process of involving all [stakeholders] (local government officials, local citizens, architects, developers, business people, and planners) in such way that decision-making is shared" (Okazaki 2008, p. 511). Scheyvens (1999) argues that the host community should be involved in tourism planning because they (a) have a historical understanding of how the region adapts to change, (b) are the ones most closely affected by tourism, and (c) are expected to become an integral part of the tourism product. Consequently, the literature would seem to strongly advocate the position that the sustainability of a tourism development might contingent upon the direct and indirect involvement of the local community (Hall, 2008; Mowforth & Munt, 2008).

3. Methodology

This study investigates the contribution of community involvement to rural tourism development. We used a survey questionnaire to explore the involvement of the local community in rural tourism development and to identify the what tourism development benefits the local Kinabalu National Park community in Sabah were recipients of. Stratified random sampling was used to distribute the questionnaire among local communities around Kinabalu National Park; these communities being divided according to districts and villages. The sampling frame (see Table 1) consisted of local residents who lived near the predetermined sampling areas within the vicinity of Kinabalu National Park. Respondents from villages that contributed the most to tourism activities were selected for participation through the recommendation of local authorities, particularly those authorities responsible for Sabah's park management. The respondents included mountain guides, porters, service staff, and hospitality-related workers.

Table 1. Sampling frame

Population	Distributed Questionnaires	Returned Questionnaires	Valid Responses
3822	450	401 (89.1%)	378 (84%)

Descriptive methods of statistical analysis, including frequency and mean, were used to analyze collected data per the research objectives.

4. Analysis and Findings

Of the 378 respondents, 38.6% were male, and 61.4% were female. In descending order of frequency, respondents were aged 21–30 (31.7%), 31–40 (24.3%), 41–50 (20.9%), and 51 and above (16.9%). In terms of their educational background, most of the respondents had some formal education (93.7%), whereas only a handful had no prior education (6.3%). Half of the respondents (54.5%) attained only a secondary education.

Involvement in the tourism sector among the local communities is shown in Table 2. Data analysis shows that

most of the respondents were indigenous to the area (74.9%). Most of the respondents had previous experience in the tourism sector (68.5%), although most (52.1%) had less than 10 years involvement in the industry. Most of the respondents were either self-employed (38.4%), or employed full-time (24.6%). Before working in the tourism sector, most of the respondents were employed in other occupations, mostly mechanics or students (82%). The survey also revealed that 32% of the local community working in the tourism sector earned less than MYR 1080 (approx. US$ 313 at the time of this writing) per month, 20.9% earned MYR 1081–2080, and 8.7% earned MYR 2081–3080. Before being involved in the tourism sector, 42.1% earned less than MYR 1080 per month, and only 9.3% earned MYR 4081–5080.

Table 2. Involvement in tourism sector

Questions	Scale	Frequency	Percentage%)
Indigenous Population	Yes	283	74.9
	No	95	25.1
Previous Involvement in Tourism Sector	Yes	258	68.5
	No	120	31.5
Years of Involvement	1–10	197	52.1
	11–20	37	9.8
	21–30	20	5.3
	31–40	4	1.1
Previous **Occupation**	Farmer	71	18.8
	Businessman/Self-employed	80	21.2
	Government employee	25	6.6
	Other	82	21.7
Previous Monthly **Income**	Below1080*	159	42.1
	1081–2080	50	13.2
	2081–3080	8	2.1
	3081–4080	4	1.1
	4081–5080	35	9.3
	Above 5081	2	0.6
Current Monthly **Income**	Below 1080	121	32.0
	1081–2080	79	20.9
	2081–3080	33	8.7
	3081–4080	5	1.3
	4081–5080	6	1.6
	Above 5081	14	3.0
Employment status	Self-employment	145	38.4
	Employees/Staff	87	23.0
	Part-time worker	14	3.7
	Government worker	11	2.9
	Others	1	0.3

Note: *MYR 3.45 = USD 1.00

The local communities' sectoral involvement in tourism development is depicted in Table 3. Five tourism sectors; namely, transport, services, food and beverage, handicraft shops, and services and support, were selected. Table 3 categorizes the six tourism-related sectors that involving the local community: transportation, accommodations, food and beverage, handicraft, services, and others. As a service industry, tourism needs support from other industries through a complex chain of supply providing goods and services. Respondents were divided across the food and beverage (20.6%), accommodation (18.3%), in handicraft (11.1%), and other service (10.3%) industries. The "others" sector, comprising only 6.3% of the sample, included fish farmers, government servants, lifeguards, flower and vegetable stall vendors.

Table 3. Sectoral involvement

Involvement sector		Frequency	Percentage
Transportation	Car/van rental	5	1.3
	Taxi driver	2	0.5
Accommodations	Hotel/budget hotel	18	4.8
	Apartment/condominium	2	0.5
	Guest room/chalet/homestay	49	13.0
Food and beverage	Bar	17	4.5
	Food stall	51	13.5
	Restaurant	10	2.6
Handicraft shops		42	11.1
	Information center	3	0.8
	Grocery store	7	1.9
	Laundry soap	1	0.3
	Cyber café	1	0.3
Services	Tourist guide	7	1.9
	Site guide	9	2.4
	Mountain guide	7	1.9
	Porter	1	0.3
	Others	2	0.5
Others	Fish farm	2	0.5
	Government servant	10	2.6
	Lifeguard	2	0.5
	Flower stall	3	0.8
	Vegetable stall	7	1.9

Respondents' perspectives on their involvement in rural tourism is presented in Table 4. Seven questions were asked of the respondents to evaluate their perspectives about community involvement in tourism development. These questions were answered along a 5-point Likert scale, from 1 (*strongly disagree*) to 5 (*strongly agree*). The highest mean score belonged to, "The local community should be given opportunity in decision making" ($\bar{x}$ = 3.94). At the opposite end of the spectrum, the lowest score belonged to, "The progress of the local tourism sector is more significantly dominated by outsiders than by locals" ($\bar{x}$ = 3.36). These mean scores suggest that the local community is interested in engaging throughout the tourism development.

Table 4. Community involvement in rural tourism development

Items	Mean	Std. Deviation
The local community should be given opportunity in decision making.	3.94	0.859
The local community has been given an opportunity in giving an opinion.	3.65	0.975
There is good cooperation between the local community and important persons (stakeholders).	3.54	0.958
Support for tourism development depends on the level of awareness and knowledge of the local community.	3.82	0.846
The government encourages local community involvement in tourism activity.	3.92	0.820
The local community should have the authority to control tourism development.	3.83	0.901
The progress of the local tourism sector is more significantly dominated by outsiders than by locals.	3.36	1.294

Table 5 describes the advantages of tourism development from the perspective of the respondents. The highest mean score was for, "Tourism development can create new business opportunities and can increase household income" ($\bar{x}$ = 4.12). As shown in Table 5, the mean value of all the responses were high (>3.64), the mean scores for these items demonstrating that tourism positively affects the life of the local community. The local

community agree that tourism development gives them a chance to improve their lives, and in turn, they support tourism through entrepreneurial activities.

Table 5. Advantages of tourism development from perspective of community

Items	Mean	Std. Deviation
Tourism helps to improve the economic conditions of the local community.	4.02	0.741
Tourism creates jobs and investment appropriate for the local community.	3.91	0.829
An increase in the number of tourists enhances the economy of local communities.	4.02	0.784
The local community generates greater income from tourism.	3.64	0.931
The local community earns money from selling local products.	3.95	0.788
Tourism development can create new business opportunities.	4.12	0.751
Tourism can increase household income.	4.12	0.723
Local communities should be given priority in the field of tourism entrepreneurship.	4.10	0.807
The benefits of tourism to the local community are appropriate in relation to its costs.	3.72	0.841
Standard of living has increased because of tourist spending in the local community.	3.84	0.840
Tourism development generates more benefits than losses for the local community.	3.80	0.892
Tourism development can improve infrastructure, such as roads, electricity, and schools.	4.05	0.877

5. Discussion

This paper explored the involvement of local communities in the development of a rural mountain tourism industry and the advantages of that tourism development for the local community in Kinabalu National Park, Malaysia. Situated in the heart of Sabah, almost to the north of Borneo Island, Kinabalu National Park was awarded WHS by UNESCO in 2000 for its outstanding universal value, being a rich source of biodiversity. As a nature-based attraction, it is this biodiversity, more than its local communities, that has captured the attention of researchers. This study sought to identify the perceptions of the local community regarding the tourism development of the surrounding the area.

The data revealed that mountain tourism development was well received by the local community. The majority of the local community, those aged 21–30, had begun their careers in various tourism-related sectors. In light of the general paucity of formal education among the local populace, the tourism sector provided an ideal career option for many. Many of the local youth we surveyed had become involved in the tourism industry 1–10 years ago, within a few years of Kinabalu National Park having been recognized as a WHS. Thirteen years since having been recognized, tourism development in and around Kinabalu National Park has dramatically altered the pattern of community employment, from farming to self-employment, thereby explaining the increased salary of residents as shown in Table 2. This finding supports those of previous studies (Gale, 2006; Gurung & DeCoursey, 2000; Iorio & Corsale, 2010; Liu, 2006; Petric, 2003). Ghaderi and Henderson (2012) observed that the benefits of rural tourism include the generation of income and employment, the creation of a market for local products, and the vitalization of usually stagnate traditional economies.

Our analysis of the community's perceptions of tourism development in the area revealed the positive effects that development had had—locals felt that tourism development provided unparallelled opportunities for them to enter new economic sectors. Rural tourism provides a new market to small businesses and encouraged the development of such businesses directly and indirectly. The spinoff of this effect was a more general increase in economic multipliers (Campbell, 1999; Wild et al., 1994). The food and beverage, accommodation, and handicraft industries were among the more popular sectors supporting the development. Many traditional industries can be developed and improved in response to the increased demand following tourism development (Macdonald & Jolliffe, 2003).

Consequently rural tourism serves a channel for socio-economic development and revival. As a rising sector in the overall tourism market (Sharpley, 2000) and a significant source of income and employment for rural

economies (Hummelbrunner & Miglbauer, 1994), rural tourism amplifies the quality of life of local communities (Simpson et al., 1998) and stimulates flagging rural economies (Din, 1993). Respondents were positive toward tourism development, most being willing to support tourism development so long as they are given the opportunity be involved in tourism-based activities. This desire to exert some control over the tourism development, and to be more aware of tourism, helps the local community to influence the course of development in the area.

Furthermore, the respondents recognized other positive effects of tourism development, including the creation of new business opportunities, promoting entrepreneurship, increasing household income, better roads, electricity, and other forms of infrastructure. Most of the respondents indicated that rural tourism positively affected the development of the local area.

This study ascertained the relationship between tourism development and local community participation. Thongma et al. (2011) reported that the participation of local communities was instrumental to the success of rural tourism development because locals build strong relationships with tourists and impress visitors with local cultural activities. Community participation also gives locals the opportunity to participate in development activities, empowers them to mobilize their capabilities in managing their resources, enables them to make decisions, and to exert some control over activities that affect their lives.

To that end, our findings demonstrated the interest of the local community to engage in decision making, planning, and tourism development. Inskeep (1991) proposed that the local community should be involved in tourism activities and that the greater their involvement the greater their socio-economic benefits are from tourism. Through the involvement of local communities in decision-making processes, traditional lifestyles and values are respected (Sheldon & Abenoja, 2001; Tosun, 2000). With the establishment of any kind tourism there exists the possibility to supplement traditional sources of income.

However, in our study, it became apparent that tourism benefits, and even employment opportunities, were narrowly distributed across the locality. Tourism can be an effective means of development; however, historically tourism has not always been well integrated into the development strategies of rural communities. As a result, tourism often becomes an unfulfilled promise in terms of the creation of indigenous employment and a supplementary source of household income. Benefits to rural residents, both in economic and social terms, become minimal where there is a lack of local involvement in tourism development, or where locals are unable to or prevented from responding to the employment opportunities arising through tourism. Mowforth and Munt (2008), therefore, advocate for community participation in the context of sustainable tourism. Hall (2008) adds that successful tourism planning entails the involvement and participation of residents in the area. And Mancini et al. (2003) suggests that the local community plays a pivitol role in the lives of its members by promoting their physical, social, psychological, and spiritual well-being.

6. Conclusion

In this study, we investigated the involvement of the local community surrounding Kinabalu National Park in rural tourism development and the benefits of tourism for the host community. It is clear from this case study that local participation is a natural outcome of completed tourism projects and that tourism is an effective respons to many rural problems. Furthermore, the benefits of tourism encourage local communities to participate in tourism development programs. In Malaysia's rural communities, where there is a strong adherence to cultural and religious observances, tourism development training for local residents should include the inculcation of cultural tolerance and an appreciation of cultural differences.

In conclusion, even if local people regard a tourism development in their area positively, government and other stakeholders should still emphasize the benefits of community involvement. To this end, the government should aim to create opportunities and programs to help the community embrace a range of tourism products relevant and beneficial to the development of the local area. Furthrmore, stakeholders should look to provide monetary assistance, perhaps through crediting or micro-loan systems, to empower the local communities to realize the opportunities created by the rural tourism development.

References

Azman, N., Halim, S. A., Liu, O. P., & Komoo, I. (2011). The Langkawi Global Geopark: local community's perspectives on public education. *International Journal of Heritage Studies*, *17*(3), 261–279. http://dx.doi.org/10.1080/13527258.2011.557863

Backhaus, N. (2003). Non-place jungle: The construction of authenticity in National parks of Malaysia. *Indonesia and the Malay World*, *31*(89), 151–160. http://dx.doi.org/10.1080/13639810304438

Beedie, P., & Hudson, S. (2003). Emergence of mountain-based adventure tourism. *Annals of Tourism Research*, *30*(3), 625–643. http://dx.doi.org/10.1016/S0160-7383(03)00043-4

Campbell, L. M. (1999). Ecotourism in rural developing communities. *Annals of Tourism Research*, *26*(3), 534–553. http://dx.doi.org/10.1016/S0160-7383(99)00005-5

Chaudhry, P., & Gupta, R. K. (2010). Urban greenery and its sustainable extension strategies in hot arid region of India. *International Journal of Sustainable Society*, *2*(2), 146–155. http://dx.doi.org/10.1504/IJSSoc.2010.033627

Chaudhry, P., & Gupta, R. K. (2010). Urban greenery and its sustainable extension strategies in hot arid region of India. International *Journal of Sustainable Society*, *2*(2), 146–155. http://dx.doi.org/10.1504/IJSSoc.2010.033627

Claiborne, C. (2010). Community participation in tourism development and the value of social capital—*T*he case of Bastimentos, Bocas del Toro, Panamá (Master's thesis, University of Gothenburg, Gothenburg, Sweden). Retrieved from http://ejournal.narotama.ac.id/files/Community%20Participation%20in%20Tourism%20Development%20and%20the%20Value%20of%20Social%20Capital.pdf

Din, K. (1993). Dialogue with hosts: An educational strategy towards sustainable tourism. In M. Hitchcock, V. King, & M. Parnwell (Eds.), *Tourism in South-East Asia* (pp. 328–336). London, United Kingdom: Routledge.

Eruera, A. (2008). Rural tourism development in the eastern Hokianga area (Master's thesis, Auckland University of Technology, Auckland, New Zealand). Retrieved from http://aut.researchgateway.ac.nz/bitstream/handle/10292/540/ErueraA.pdf?sequence=4&isAllowed=y

Fleischer, A., & Pizam, A. (1997). Rural tourism in Israel. *Tourism Management*, *18*(6), 367–372. http://dx.doi.org/10.1016/S0261-5177(97)00034-4

Frederick, M. (1993). Rural tourism and economic development. *Economic Development Quarterly*, 7(2), 215–224. http://dx.doi.org/10.1177/089124249300700207

Gale, T. E. (2006). Finding meaning in sustainability and a livelihood based on tourism: An ethnographic case study of rural citizens in the Aysén region of Chile (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database.

Geoffrey, L. R., & Nepal, S. K. (2010). Dependency and development in northern Thailand. *Annals of Tourism Research*, *37*(4), 947–968. http://dx.doi.org/10.1016/j.annals.2010.03.005

Ghaderi, Z., & Henderson, J. C. (2012). Sustainable rural tourism in Iran: A perspective from Hawraman village. *Tourism Management Perspectives*, *2*, 47–54. http://dx.doi.org/10.1016/j.tmp.2012.03.001

Gurung, C. P., & DeCoursey, M. A. (2000). Too much too fast: lessons from Nepal's Lost Kingdom of Mustang. In P. M. Price, M. F. Price, & F. M. Zimmerman (Eds.), *Tourism and Development in Mountain Regions* (pp. 239–253). New York, NY: CABI Publishing.

Hall, C. M. (1994). *Tourism and politics: Policy, power and place*. Hoboken, NJ: John Wiley & Sons.

Hall, C. M. (2008). *Tourism planning: Policies, processes and relationships*. Upper Saddle River, NJ: Pearson Education.

Hall, D. (2004). Rural tourism development in southeastern Europe: Transition and the search for sustainability. *International Journal of Tourism Research*, *6*(3), 165–176. http://dx.doi.org/10.1002/jtr.482

Harrison, D., & Schipani, S. (2007). Lao tourism and poverty alleviation: Community-based tourism and the private sector. *Current Issues in Tourism*, *10*(2–3), 194–230. http://dx.doi.org/10.2167/cit310.0

Hasegawa, M., Ito, M., & Kitayama, K. (2006). Community structure of oribatid mites in relation to elevation and geology on the slope of Mount Kinabalu, Sabah, Malaysia. *European Journal of Soil Biology*, *42*, S191–S196. http://dx.doi.org/10.1016/j.ejsobi.2006.07.006

Horowitz, L. S. (1998). Integrating indigenous resource management with wildlife conservation: A case study of Batang Ai National Park, Sarawak, Malaysia. *Human Ecology*, *26*(3), 371–403. http://dx.doi.org/10.1023/A:1018752115074

Hummelbrunner, R., & Miglbauer, E. (1994). Tourism promotion and potential in peripheral areas: The Austrian case. *Journal of Sustainable Tourism*, *2*(1–2), 41–50. http://dx.doi.org/10.1080/09669589409510682

Inskeep, E. (1991). *Tourism planning: An integrated and sustainable development approach*. London, United Kingdom: Van Nostrand Reinhold.

Iorio, M., & Corsale, A. (2010). Rural tourism and livelihood strategies in Romania. *Journal of Rural Studies*, *26*(2), 152–162. http://dx.doi.org/10.1016/j.jrurstud.2009.10.006

Jaafar, M., Kayat, K., Tangit, T. M., & Yacob, M. F. (2013). Nature-based rural tourism and its economic benefits: A case study of Kinabalu National Park. *Worldwide Hospitality and Tourism Themes*, *5*(4), 342–352. http://dx.doi.org/10.1108/WHATT-03-2013-0016

Keane, M. (1992). Rural tourism and rural development. In H. Briassoulis and J. van der Straaten (Eds.), *Tourism and the Environment* (pp. 43–55). New York, NY: Springer.

Khalifah, Z., & Tahir, S. (1997). Malaysia: Tourism in Perspective. In F. M. Go, & C. L. Jenkins (Eds.), *Tourism and Economic Development in Asia and Australasia* (pp. 176–196). Cassell: London

Lacher, R. G., & Nepal, S. K. (2010). From leakages to linkages: Local-level strategies for capturing tourism revenue in Northern Thailand. *Tourism Geographies*, *12*(1), 77–99. http://dx.doi.org/10.1080/14616680903493654

Lama, W.B., & Sattar, N. (2004). Mountain tourism and the conservation of biological and cultural diversity. In M.F. Price, L. Jansky, & A.A. Iatsenia (Eds.), *Key issues for mountain areas* (pp.111–148). Tokyo: United Nations University Press.

Lane, B. (1994). What is rural tourism? *Journal of Sustainable Tourism*, *2*(1–2), 7–21. http://dx.doi.org/10.1080/09669589409510680

Liu, A. (2006). Tourism in rural areas: Kedah, Malaysia. *Tourism Management*, *27*(5), 878–889. http://dx.doi.org/10.1016/j.tourman.2005.05.007

Lo, M. C, Songan, P, Mohamad, A. A. & Yeo, A. W. (2013). Rural tourism and destination image: Community perception in tourism planning. *The Macro-theme review, 2*(1), 102-118.

MacDonald, R., & Jolliffe, L. (2003). Cultural rural tourism: Evidence from Canada. *Annals of Tourism Research*, *30*(2), 307–322. http://dx.doi.org/10.1016/S0160-7383(02)00061-0

Mancini, J. A., Martin, J. A., & Bowen, G. L. (2003). Community capacity. In T. P. Gulotta & M. Bloom (Eds.), *Encyclopedia of Primary Prevention and Health Promotion* (pp. 319-330). NY: Kluwer Academic/Plenum.

Maryati, M., Azizah, H., & Arbain, K. (1996). Terrestrial ants (Hymenoptera: Formicidae) of Poring, Kinabalu Park, Sabah. *Tropical Rainforest Research—Current Issues, 74*, 117–123. http://dx.doi.org/10.1007/978-94-009-1685-2_11

May-Ling, S., Ramachandran, S., Shuib, A., & Afandi, S. H. M. (2014). Barriers to community participation in rural tourism: A case study of the communities of Semporna, Sabah, Malaysia. *Life Science Journal*, *11*(11), 837–841.

Moss, L. A., & Godde, P. M. (2000). Strategy for future mountain tourism. In P. M. Price, M. F. Price, & F. M. Zimmerman (Eds.), *Tourism and Development in Mountain Regions* (pp. 323–338). New York, NY: CABI Publishing.

Mowforth, M., & Munt, I. (2008). *Tourism and sustainability: Development, globalisation and new tourism in the third world* (3rd ed.). New York, NY: Taylor & Francis.

Nepal, S. K., & Chipeniuk, R. (2005). Mountain tourism. Toward a conceptual framework. *Tourism Geographies: An International Journal of Tourism Space, Place and Environment*, *7*(3), 313–333. http://dx.doi.org/10.1080/14616680500164849

Nunkoo, R., & Gursoy, D. (2012). Residents' support for tourism: An identity perspective. *Annals of Tourism Research*, *39*(1), 243–268. http://dx.doi.org/10.1016/j.annals.2011.05.006

Okazaki, E. (2008). A community-based tourism model: Its conception and use. *Journal of Sustainable Tourism*, *16*(5), 511–529. http://dx.doi.org/10.1080/09669580802159594

Okech, R., Haghiri, M., George, B. P., George, B., & Korstanje, M. (2012). Rural tourism as a sustainable development alternative: An analysis with special reference to Luanda, Kenya. *Cultur: Revista De Cultura e Turismo*, 6(3), 36–54.

Petric´, L. (2003). Constraints and possibilities of the rural tourism development with the special stress on the case of Croatia. Ersa 2003 Congress. University of Jyva¨skyla¨, Finland. Retrieved from

http://ideas.repec.org/p/wiw/wiwrsa/ersa03p105.html, retrieved in October 2008

Scheyvens, R. (1999). Ecotourism and the empowerment of local communities. *Tourism Management*, *20*(2), 245–249. http://dx.doi.org/10.1016/S0261-5177(98)00069-7

Scheyvens, R. (2002). *Tourism for development: Empowering communities*. Upper Saddle River, NJ: Pearson Education.

Sharpley, R. (2000). Tourism and sustainable development: Exploring the theoretical divide. *Journal of Sustainable Tourism*, *8*(1), 1–19. http://dx.doi.org/10.1080/09669580008667346

Sharpley, R. (2002). Rural tourism and the challenge of tourism diversification: The case of Cyprus. *Tourism Management*, *23*(3), 233–244. http://dx.doi.org/10.1016/S0261-5177(01)00078-4

Sharpley, R., & Sharpley, J. (1997). *Rural tourism. An introduction.* Mumbai, India: International Thomson Business Press.

Sheldon, P. J., & Abenoja, T. (2001). Resident attitudes in a mature destination: The case of Waikiki. *Tourism Management*, *22*(5), 435–443. http://dx.doi.org/10.1016/S0261-5177(01)00009-7

Shukor, N. (2001). Elevational diversity patterns of small mammals on Mount Kinabalu, Sabah, Malaysia. *Global Ecology and Biogeography*, *10*(1), 41–62. http://dx.doi.org/10.1046/j.1466-822x.2001.00231.x

Simpson, F., Chapman, M., & Mahne, L. (1998). *Partnership approaches to tourism and rural development in post-socialist Europe: The experience of Notranjski Kras, Slovenia*. Paper Presented at the Rural Tourism Management: Sustainable Options International Conference, Auchincruive, Scotland, Sep. 1998.

Stepp, J. R. (2000). Mountain ethnobiology and development in highland Chiapas, Mexico: Lessons in biodiversity and health. *Mountain Research and Development*, *20*(3), 218–219. http://dx.doi.org/10.1659/0276-4741(2000)020[0218:MEADIH]2.0.CO;2

Su, B. (2011). Rural tourism in China. *Tourism Management*, *32*(6), 1438–1441. http://dx.doi.org/10.1016/j.tourman.2010.12.005

Suh, Y. K., & Gartner, W. C. (2004). Preferences and trip expenditures—A conjoint analysis of visitors to Seoul, Korea. *Tourism Management*, *25*(1), 127–137. http://dx.doi.org/10.1016/S0261-5177(03)00056-6

Telfer, D. J., & Sharpley, R. (2007). *Tourism and development in the developing world*. New York, NY: Routledge.

Thongma, W., Leelapattana, W., & Hung, J.-T. (2011). Tourists'satisfaction towards tourism activities management of Maesa community, Pongyang sub-district, Maerim district, Chiang Mai province, Thailand. *Asian Tourism Management*, *2*(1), 86–94.

Tosun, C. (2000). Limits to community participation in the tourism development process in developing countries. *Tourism Management*, *21*(6), 613–633. http://dx.doi.org/10.1016/S0261-5177(00)00009-1

Wild, C. (1994) Issues in Ecotourism. In C.P. Cooper and A. Lockwood (Eds.) *Progress in Tourism, Recreation and Hospitality Management* (pp. 12–21). Chichester: John Wiley.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Desorption Isotherms of the Koumiss and Shubat Clots Enriched by Various Additives

Azret Shingisov[1], Ravshanbek Alibekov[1], Saparkul Erkebaeva[1], Zeinep Nurseitova[1], Gulbagi Orymbetova[1], Gulzhan Kantureeva[1] & Elvira Mailybaeva[1]

[1] M. Auezov South Kazakhstan State University, Kazakhstan

Correspondence: Ravshanbek Alibekov, M. Auezov South Kazakhstan State University, 160012, Tauke Khan avenue 5, Shymkent city, Kazakhstan. E-mail: ralibekov@hotmail.com

Received: December 8, 2014 Accepted: January 14, 2015 Online Published: July 31, 2015

doi:10.5539/mas.v9n8p409 URL: http://dx.doi.org/10.5539/mas.v9n8p409

Abstract

The fermented mare's milk – Koumiss and camel's milk – Shubat are traditional national dairy products in Kazakh cuisine. Concentrated dairy products in the form of clots have a quantity content of useful chemical compounds. In the presented study it was investigated the desorption isotherms of the koumiss and shubat clots enriched by vegetative additives. It was established, that the interface between weakly and strongly bound moisture can be: for the koumiss clots enriched by juices: of a carrot a_w =0,56, u_x = 0.69 kg/kg; of a pumpkin a_w =0,58, u_x = 0,80 kg/kg; of a beet a_w =0,59, u_x = 0,81 kg/kg; as well as for shubat clot, enriched by juices: of a carrot a_w =0,62, u_x =0,73 kg/kg; of a pumpkin a_w =0,65, u_x = 0,67 kg/kg and of a beet a_w =0,63, u_x =0.66 kg/kg.

Keywords: dried dairy products, fermented, camel's milk, mare's milk, koumiss, shubat

1. Introduction

1.1 Koumiss and Shubat

One of the advanced directions in the development of healthy foodstuff with therapeutic actions is fermented dairy products on the bases of mare's milk – koumiss and camel's milk – shubat, enriched by vegetable origin bioactive substances with vitamins and microelements (Faye and Konuspayeva, 2012). According to the variability of milk composition, nutritional and medicinal properties (true or postulated) could be potentially an important added value for producers and dairy sector (Konuspayeva et al, 2009).

The chemical compositions of mare's and camel's milks are specified for geographical origin. Personal data from Kazakhstan showed significantly higher fat matter and total protein contents, but lower lactose content compared to other references from Central Asia (Malacarne et al, 2002). As well as protein content in mare's milk is higher than in human milk and lower than in cow's milk; casein concentration in mare's milk is intermediate between the other two milks. Fat content is lower in mare's milk compared to human and cow's milk. Distribution of di- and tri-glycerides in mare's and women' milk is similar (Saitmuratova et al, 2001).

In addition microelements play important role in the formation of milk structural proteins and affect the quality and food value of dairy products. Microelements (especially Cu, Fe, Zn and Ni) can act as catalysts of certain chemical reactions. Deficiencies of them are causes of the various pathological changes in the human organism (Alibekov et al, 2014).

1.2 Vacuum-Sublimation Drying Method

Milk is extremely perishable, yet for a number reasons it is desirable to preserve it for later consumption. Today, drying is most important method for preservation of foodstuff (Varnam & Sutherland, 1994). For the maximum product saving of original quality, the most promising preservation method is the vacuum-sublimation drying. However, this method requires large energy expenditures. Therefore, a research of new ways for reducing the energy costs of this drying method is one of important tasks in the Food industry (Shingisov, 2011).

Heat and mass transfer processes occurring during drying of food products in the frozen state in a vacuum are complex nature because the process of evaporation of moisture from the product surface is necessary to connect with the structure and phenomena occurring inside it. From the viewpoint of the thermodynamics of the drying process some features of products in a frozen state in a vacuum can be divided into the following groups:

- The first group of features of heat and moisture transfer can be attributed the phenomenon of interaction of the mobile phase with a dry skeleton of the product itself. Status of the mobile phase in the product is determined by the energy and moisture forms of communication that must be considered for the selecting the drying method and its calculation;

- The second group of features of heat and moisture transfer phenomena includes energy and mass of the movable phase within the product. These phenomena depend on the structure of the product. As it is known, a large part of the open pores are interconnected, and therefore the amount of moisture in the pores determines the ratio between the amount of evaporated and diffused from the deep layers of the product water;

- The third group of features of heat and moisture transfer includes physical phenomena arising from the transfer of moisture from the surface of the product to its environment (air). The intensity of the mass and energy transfer in this case is determined by the state of the boundary layer.

Study of the state of the boundary layer during the drying a lot of scientific papers are devoted (Kamovnikov et al, 1985). However, still there is no sufficiently clear view not only of the physical transport mechanism, but also on the qualitative impact of moisture transfer on heat transfer (Bykov, 1987). The mechanism of transfer of moisture, kinetics of transport and energy costs in the process of freeze-drying depend not only on moisture content, but also the structure of the product and forms moisture binding dry carcass material (Flaubenbaum et al, 1986).

Analysis of a several papers (Izbasarov, 1999) devoted for the studying of patterns of changes in the kinetic curves shows that the various forms of drying rate curves indicate the different nature of the relationship with the product moisture. However, it is not possible to determine the nature of the product a significant impact on the intensity of the internal moisture transfer and the rate of evaporation of moisture, as well as to establish the total effect of the moisture content; the physicochemical properties of drying and drying curves describe by theoretical predictions (Kauhcheshvili, 1985). In this regard, presently semi-empirical laws and model for different product classes are useful. For intensify of the freeze-drying process requires strict control of all relevant factors (Key, 1983). Currently, industrial freeze-dryers are often operated solely on the basis of experience of staff (Riedel, 1961). This means that except the temperature and pressure of the residual air in the sublimator, operators do not use any other parameters. In such a situation freeze-dryers are served with deliberately exaggerated amounts of energy consumption (Ginzburg and Savina, 1982).

Control of the drying process is a challenging task, and in some cases impossible. Thus, control of the drying process by controlling the residual pressure of the air in the sublimator can not always be achieved. The reason is the variations in the residual pressure in sublimator not always adequate response for the changing in the water state of the product due to a number of reasons. These reasons can include variations in the product structure in the process of drying that leads for a changing in the thermodynamic parameters, such as the chemical potential (μ), water activity (AW), the binding energy of moisture (L) (Kutsakova & Bogatyrev, 1987).

Thus, on the basis of the above mentioned analysis it has concluded the following:

- Basic directions of improving of technique and technology of sublimation drying are the study of the interaction of the product with its environment (air) based on the analysis of static and drying kinetics;

- Accumulated scientific and practical potential in thermodynamics of moist air can be used to directly control of the process of sublimation drying.

- For intensify the drying process, in the laboratory of the food production technology of M.Auezov' SKSU has been developed a new device for the freezing, based on the intensification of the drying process by increasing the heat transfer coefficient on the air side.

The created device provides a grid configuration of freeze drying product with a plurality of through holes. As a result, in the product increases the contact area between the product and air, thereby repeatedly increases the coefficient of heat transfer from the surface of the product to its environment (Kamerbayev, 2002).

In order to reduce energy consumption during a drying of liquid foods has developed a new technology of water condensation contained in the compositions of koumiss and shubat enriched by herbal supplements from other components before drying with preserving the original quality. In this technology, for the separating of water as a serum from compositions of koumiss and shubat, there were used protein coagulated ingredients (rennet, pepsin, calcium chloride) and apple vinegar, as components reducing the activity of the aqueous medium. In the separation of clot and serum the serum yield for koumiss enriched by plant origin juices was 69-72% and 60-63% for shubat of the total product weight (Shingisov, 2008).

According to the currently existing theory drying regularities of heat - and mass transfer processes during the vacuum freeze-dried foods are described by curves of drying rate that is defined as the slope of the tangent drawn through the point of the experimental curve drying rate, corresponding to a certain moisture content of the material (Varnam and Sutherland, 1994). However, this method does not fully describes the regularities of heat - and mass transfer processes in their vacuum-sublimation drying, especially at the end of the process, when the drying curve asymptotically approaches the line of equilibrium moisture (Shervud et al, 1982).

It is known that the character of the heat - mass transfer processes during the vacuum-sublimation drying depends on the thermodynamic state of the water contained in the drying product and characterized by a water activity (Ginzburg, 1973). In this regard, for the disclosure of the character of the heat - and mass transfer processes were studied the water activity in koumiss and shubat enriched by vegetative additives in the vacuum vacuum-sublimation drying (Shingisov, 2011).

2. Methods.

2.1 Used Raw and Equipment

As research raw materials were used a fresh diurnal koumiss of the farm "SAP" and shubat of the farm "Gulmairam" those is situated in the “South Kazakhstan Oblast” (SKO). In the capacity of additives for the enriching koumiss and shubat were taken a carrot, a pumpkin and a beet from local producers of the SKO. Juices from plant material were obtained by the household (Braun) brand juicer. For enrichment of koumiss and shubat by vegetable juices and their thickening and separation of the clot from the serum it was used the technology developed by the laboratories of the Kazakhstan Research Institute of Processing and Food industries. As per this technology at the separation of the clot and serum the output for a koumiss enriched by plant origin vegetable juices was 69-72% and for a shubat - 60-63% of the total product weight.

For the determining of water content in the samples of the koumiss and shubat clots enriched by vegetative juices it was used the method of convective drying in the sublimation chamber at the 102 ÷ 105 ° C. Water content of the koumiss clots enriched by juices: of a carrot $W_i = 88,3 \pm 0,02\%$, of a pumpkin $W_i = 87,8 \pm 0,02\%$ and a beet $W_i = 89,2 \pm 0,02\%$; of the shubat clots, enriched by juice: of a carrot $W_i = 86,8 \pm 0,02\%$, of a pumpkin $W_i = 87,8 \pm 0,02\%$ and a beet $W_i = 88,2 \pm 0,02\%$.

The experimental laboratory vacuum sublimation apparatus consists of: sublimation camera of circular section with diameter 0.5 m and a length 0.7 m for the product drying; the vacuum pump AVPR-60D of the "Hydromeh" company (Russia) to create a residual pressure at the sublimation chamber within $5 \cdot 10^{-4}$ kPa; and low-temperature refrigeration unit (Russia) for the ensuring a drying temperature in the sublimation chamber and traping of water steam existed from the sublimation chamber. In the sublimation chamber air temperature varied from minus 5 °C until minus 15 °C by a thermo-regulating valve of the SPORLAN company (Russia).

At the beginning of the experiment, the researched samples were enriched by vegetative additives koumiss and shubat clots, as well as a distilled water were frozen at the -18 ° C ÷ -20 ° C in the refrigerator of the INDESIT company during 8-10 hours in a specially designed cylindrical form with a strict observance of the same thickness (thickness δ = 0.006 m, diameter d = 0,135 m and mass m = 0,01 kg). Then by assistance of the “Testa-625” device (Germany) with an umbrella for determining water activity, the water activity of the samples was determined. Furthermore the product samples and a distilled water ice were placed in freeze chamber in a chessboard pattern. During the drying process the installation was periodically breaking (an interval per 1 hour.), finally the product was removed, and after that a mass and a water activity were measured.

The temperature measuring at the freeze chamber was provided by a temperature sensor ZET 7020 Termo TC - 485 with an accuracy of 1,5 ± 0,01 / t / (Russia), and the mass measurement - with weights VLKT - 500 with an accuracy of ± 1% weighting (Russia). Drying the clot samples were done until a final humidity at the 268K, 263K and 258K.

2.2 Method of the Mathematical Treatment of Experimental Results

The product moisture changing was calculated at the current time by the formula:

$$W_p = \left(G_j - (G_j - G_p)\right) / G_j \quad (1)$$

where, W_p - a product moisture at the measuring time,%;

G_i – an initial weight of the sample, g;

G_p- a sample weight at the measuring time, g.

The experiments were performed three times for the each drying temperature. In order to avoid gross experiment

errors, the experimental data were processed by mathematical statistics. For this purpose the uniformity of the dispersion estimates was determined by using the Kohren criterion based on the following formula:

$$G = \frac{\left[S^2(y_{uk})\right]_{\max}}{\sum_{u}^{n} S(y_{uk})} \tag{2}$$

where $[S^2(y_{uk})_{max}$ - maximal dispersion of each exit;

$\sum_{u}^{n} S(y_{uк})$ - sum of dispersion of all exits.

The hypothesis of dispersion uniformity is accepted, if the estimated value of the Kohren criterion G_p is less than the table value G_t, i.e.

$$G_p < G_т \tag{3}$$

Checking of the hypothesis about the significance of the coefficient "b" was conducted according to the estimated dispersion:

$$\varepsilon(b_i) = t(p,f) \cdot S(b_i) \tag{4}$$

$$S^2(b_i) = \frac{S^2\left(\bar{y}_i\right)}{N} = \frac{\sum_{u}^{N}\sum_{K}^{M}(y_{uк} - y_u)}{N \cdot (m-1)n \cdot N} \tag{5}$$

where, t(p,*f*) – table value of the Student criterion for a given probability (ζ=0,95) and quantity of freedom degrees f_j ; $S^2\left(\bar{y}_i\right)$ - average value for the entire experiment reproducibility dispersion of average exit value in each row; N- dispersion number.

Coefficient "b_i" is considered significant if

$$|b_i| > \varepsilon(b_i) \tag{6}$$

f $t_{кр}$ is determined for freedom quantity f

$$f = N(m-1) \text{ ,} \tag{7}$$

where, N- experiments quantity

m - number of cycles (in our case m = 3)

equation approximations are provided by using the "Excel-2007" software

3. Results and Discussions

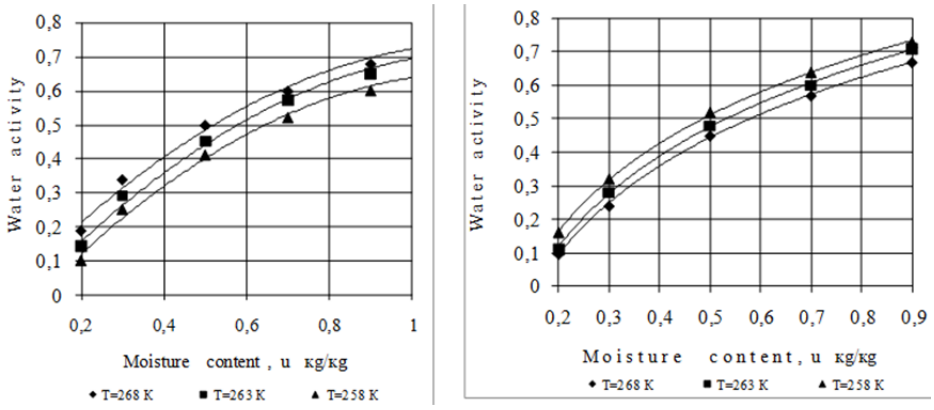
3.1 Analysis of the Drying Processes

Analysis of drying kinetics curves (Figures 1, 2, and 3) shows that the samples of koumiss and shubat clots as other foodstuff consist in three periods: the establishment of the operational parameters, constant drying and falling drying rates.

The period of the setting operational parameters has been begun with the volume vacuumization and the residual pressure reduction till the 50 ÷ 60 Pa. This period was during 8-12 minutes.

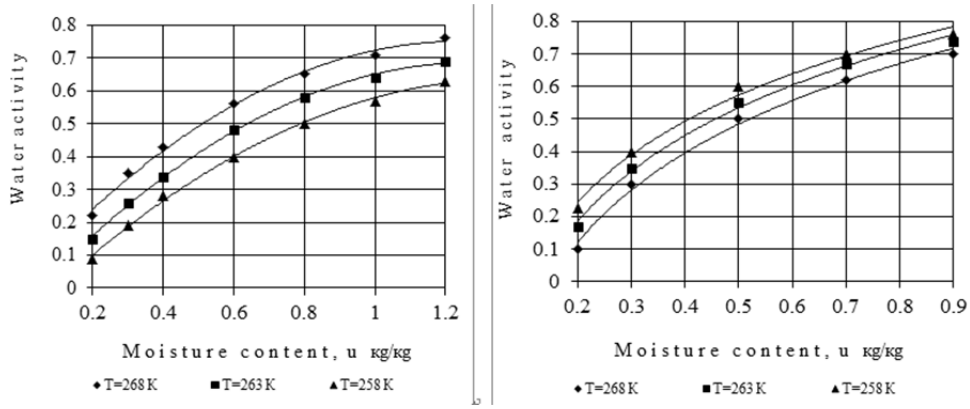
The period duration of the continuous drying of the fermented milk product's clot was 5,20 ÷ 6,0 hours.

The spread of the drying failing speed time for koumiss clots enriched by vegetative additives, were 3,0 ÷ 3,40 hours, and for shubat clots enriched by the same additives, were 2,0 ÷ 2,30 hours.



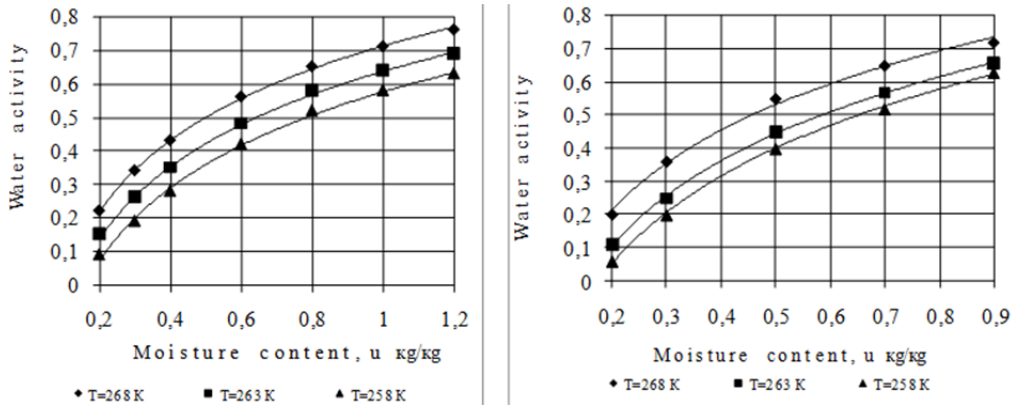
(a) (b)

Figure 1. Desorption isotherms of the koumiss (a) and shubat (b) clots enriched by carrot



(a) (b)

Figure 2. Desorption isotherms of the koumiss (a) and shubat (b) clots enriched by pumpkin



(a) (b)

Figure 3. Desorption isotherms of the koumiss (a) and shubat (b) clots enriched by beet

From the shown plots (Figures 1, 2 and 3) can be seen there are four zones with boundaries at three points on the curves of the desorption isotherms.

I zone is an area till the first point of inflection characterized by formation of stable hydrate complexes with high bond energy and moisture strict orientation of the molecule relative to the water and the dry carcass material and practically does not depend on temperature. The zone is located in the range for shubat clot enriched by carrot juice $0 < u_x < 0{,}072$ kg/kg; pumpkin juice $0 < u_x < 0{,}074$ kg/kg; beet juice $0 < u_x < 0{,}081$ kg / kg.

II zone is region of multilayer adsorption of moisture that in contrast to the first zone is characterized by the mobility of water molecules and flexible molecular chains of the protein complex and take the most favorable conformations. This zone corresponds to the interval for clot shubat enriched by juice: carrot 0,072 < u_x <0,36 mg / kg; pumpkin 0,074 < u_x <0,32 kg/kg; beet 0,081 < u_x <0,34kg/kg.

III zone is characterized by the diffusion of water molecules in the intermolecular space corresponds to a protein complex and moisture of mezo and macro capillars (for shubat clots enriched by juice: carrot 0,36 < u_x <0,70 kg/kg; pumpkin 0,32 < u_x <0,82 kg/kg; beets 0,34 < u_x <0,91 kg/kg). The binding energy of the moisture with a dry carcass material is significantly reduced, the internal energy increases, hence, the mobility of water molecules, significantly increases the probability of microbial growth.

IV zone is after the third singular point. The moisture in this zone has a small binding energy of water (close to pure water) due to the attenuation, the flexibility of the molecular chains of the protein complex and further significantly increases moisture sorption leads to separation of the protein products of these chains and essentially to their diffusion in the solvent.

In terms of selecting the optimum storage of the dried product is great interest in the second zone. Based on the results of the experiment can be determined that the maximum value of the moisture content for storage conditions can take a value equal to u_x = 0,32-0,36 kg/kg of dry matter, which corresponds to a water activity equal to a_w = 0,321 ÷ 0,325.

Based on these studies we can conclude that the process of removing moisture in freeze dried product in the following sequence.

In the first stage removes loosely bound water from the ice crystals on the surface of the product. In the second phase due to the temperature difference between the surface of the front drying and all the frozen sample some of the weakly bound water is removed from the low value of the energy due to the frame. The rest (in the lower layers) continues to sublimate from the inner layers to the front surface drying. After this period, when almost all the frozen moisture and a certain amount of tightly bound water is removed and the sample temperature rises to about 0 ° C, the value of residual moisture reaches 15-17% of dry weight. In the third period, when the temperature reached 36ºC drying, the water fraction is removed, the associated connections and strong frame product, whereby its residual moisture is reduced to 4,0÷4,5 %.

Analysis of the Figures 1, 2 and 3 shows that the moisture desorption isotherms of the enriched by vegetative additives koumiss and shubat clots, have a complicated form and by the character of desorption isotherms' curves the koumiss clots differ from the shubat clots. This distinction is explained by the different structure of the product and the nature of the moisture diffusion process in the investigated products[14].

The type of desorption isotherms' curve has an evidence, that the clots of koumiss and shubat enriched by vegetative additives relate to the colloid-porous materials.

Whereas the conventional dividing a moisture of the product into weakly and strongly bonded, we may assume that the interface can be: for a koumiss clots enriched by juice: of a carrot a_w =0,56 unit portion, u_x = 0.69 kg/kg; of a pumpkin a_w =0,58 unit portion, u_x = 0,80 kg/kg; of a beet a_w =0,59 unit portion, u_x = 0,81 kg/kg; for a shubat clot, enriched by juice: of a carrot a_w =0,62 unit portion, u_x =0,73 kg/kg; of a pumpkin a_w =0,65 unit portion, u_x = 0,67 kg/kg and of a beet aw =0,63 unit portion, u_x =0.66 kg/kg.

3.2 Mathematical description of drying desorption isotherms.

The experimental data treatment of the water activity value dependence upon the koumiss and shubat moistures enriched by vegetative additives can be satisfactorily described by the logarithmic form:

$$Y = A \cdot \ln x + C$$

where: A and C- constant coefficients from the product type dependent.

A and C coefficients are indicated at the table 1.

Table 1. A and C coefficients values for koumiss and shubat clots enriched by vegetative additives.

№	Product	Temperature, K	Coefficients values A	C	Approximation authenticity
		koumiss clots enriched by vegetative additives			
1	carrot	268	0,3227	0,7183	0,9892
		263	0,3372	0,6877	0,9967

		258	0,3266	0,6342	0,9981
2	pumpkin	268	0,302	0,710	0,9915
		263	0,309	0,637	0,9971
		258	0,305	0,566	0,9965
3	beet	268	0,307	0,710	0,9947
		263	0,308	0,638	0,9884
		258	0,311	0,577	0,9986
		shubat clots enriched by vegetative additives			
1	carrot	268	0,366	0,774	0,9893
		263	0,383	0,748	0,9967
		258	0,367	0,708	0,9851
2	pumpkin	268	0,350	0,810	0,9845
		263	0,372	0,787	0,9976
		258	0,383	0,742	0,9978
3	beet	268	0,347	0,771	0,9962
		263	0,368	0,700	0,9897
		258	0,377	0,662	0,9934

4. Conclusion

Thus in the result of the investigations was determined the koumiss and shubat clots enriched by carrot, pumpkin and beet juices can be relevant to the colloid-porous materials.

Following the conventional division of moisture in the product into weakly and strongly bonded, we may assume that the interface can be: for koumiss clots enriched by juice: of carrots a_w =0,56 unit portion, u_x = 0.69 kg/kg; of pumpkin a_w =0,58 unit portion, u_x = 0,80 kg/kg; of beet a_w =0,59 unit portion, u_x = 0,81 kg/kg; for shubat clot, enriched by juice: of carrots a_w =0,62 unit portion, u_x =0,73 kg/kg; of pumpkin a_w =0,65 unit portion, u_x = 0,67 kg/kg and of beet aw =0,63 unit portion, u_x =0.66 kg/kg.

For the practical application is important the optimum storage condition of the water content in the dried product. On the base of the experimental results, we can predicate that the maximum moisture content for storage conditions can take a value equal to u_p = 0,32-0,36 kg / kg, which corresponds with a value of a water activity equal to a_w = 0,321 ÷ 0,325 unit portion.

Symbols and abbreviations

W_i - water content in the product before drying,%;

δ - thickness of the product, m;

d - diameter of the product, m;

m – product mass, kg;

a_w - water activity, unit portion;

u - product specific moisture,%;

W_d - moisture content by dried substance,%;

Indexes of designation

i – initial;

f – finished;

p- product;

References

Alibekov, R. S., A. A. Utebaeva, K. A. Urazbayeva, Usenova, S. O., & Ermolaeva, E. A. (2014). Sensory evaluation in the standardization of tomato juice with various functional additives. *Herald of Kazan Technological University, 17*(7), 208-212, Kazan, Russia.

Bykov, B. A. (1987). Kinetics of the aggregation of soil particles in the pseudo clotting and driving layer, Diss. Cand. of Tech. Sciences, p. 235, Sverdlovsk, Russia.

Faye, B., & Konuspayeva, G. (2012). The sustainability challenge to the dairy sector – The growing importance of non-cattle milk production worldwide. *International Dairy Journal, 24*(2), 50-56.

http://dx.doi.org/10.1016/j.idairyj.2011.12.011

Flaubenbaum, B. L., Tanchev, S. S., & Grinshin, M. A. (1986). Bases of the canning of the food products. *Agroporomizdat,* p. 494, Moscow, Russia.

Ginzburg, A. S. (1973). Theory and techniques bases of food products drying. Eds. *Food Industry,* p. 528, Moscow, Russia.

Ginzburg, A. S., & Savina, I. M. (1982). Mass transfer characteristics of food products, p. 280. *Light and Food industry,* Moscow, Russia.

Izbasarov, D. S. (1999). Drying of food vegetable materials, p. 312. Almaty, Kazakhstan.

Kamerbayev, A. U. (2002). Thermodynamic bases of hydratation and drying of polycomponented food systems. Doctor of Tech. Sc. abstract of a thesis, specialty: 05.18.12. (p. 25). Almaty, Kazakhstan.

Kamovnikov, B. P., Malkov, L. S., & Voskoboinikov, V. A. (1985). Vacuum and sublimate drying of food products. Eds. *Agropromizdat,* p. 285. Moscow, Russia.

Kauhcheshvili, E. I. (1985). Physico-technical bases of the food products processing. *Agropromizdat,* p. 256, Moscow, Russia

Key, R. B. (1983). Introduction to the industrial drying technology. English transl. *Science and technics,* p, 262, Minsk, Belarus.

Konuspayeva, G., Faye, B., & Loiseau, G. (2009). The composition of camel milk: A meta-analysis of the literature data. *Journal of Food Composition and Analysis, 22*(2), 95-101. http://dx.doi.org/10.1016/j.jfca.2008.09.008

Krisher, O. (1961). Scientific bases of the drying technics, Foreign literature, p. 540, Moscow, Russia.

Kutsakova, B. E., & Bogatyrev, A. I. (1987). Intensification of heat and mass transfer during food products drying, Agropromizdat, p. 412, Moscow, Russia.

Malacarne, M., Martuzzi, F., Summer, A., & Mariani, P. (2002). Protein and fat composition of mare's milk: some nutritional remarks with reference to human and cow's milk. *International Dairy Journal, 12*(11), 869-877.

Riedel, L. (1961). Zum Problem des gebundenen Wassers in Flelsch. *Kältetechnik,* p. 300.

Saitmuratova, O. Kh., Sulaimanova, G. I., & Sadykov, A. A. (2001). Camel's milk and Shubat from the Aral Region Chemistry of Natural Compounds; (Volume 37, Issue 6, pp. 566-568).

Shervud, T. (1936). Drying of the solids, Gosletizdat, p.250, Moscow, Russia.

Shervud, T., Pigford, R., & Uilki, Ch. (1982). Mass transfer. English transl., Chemistry, p. 695, Moscow, Russia.

Shingisov, A. U. (2011). The optimization of vacuum-sublimation drying parameters of cultured milk foods by using microwave energy. Journal of Food Industry, pp. 22-24. Moscow, Russia.

Shingisov, A. U. (2008). Definition of thermodynamic exponents of condensed domestic fermented milk products, enriched by vegetative additives, at drying. *Journal of Dairy Industry, 5*, 85-86. Moscow, Russia.

Varnam, A. H., & Sutherland, J. P. (1994). Milk and Milk Products: Technology, Chemistry and Microbiology. Eds. Routledge, Chapman & Hall, Incorporated, p. 451.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Features of Vertically Integrated Agribusiness Corporations in Western Europe Countries

Zakharova Elena Nikolaevna[1], Kerashev Anzaur Aslanbekovich[1] & Mokrushin Aleksandr Aleksandrovich[1]

[1] Adyghe State University

Correspondence: settar S Keream, University of Anbar- Univrsity Malaysia Pahang, E-mail: zahar-e@yandex.ru/ kerashev@mail.ru/mokrushin_alex@inbox.ru

Received: December 29, 2014 Accepted: January 12, 2015 Online Published: July 30, 2015

doi:10.5539/mas.v9n8p417 URL: http://dx.doi.org/10.5539/mas.v9n8p417

Abstract

In order to ensure the region's population with food promising direction of development of agro-industrial complex is the vertical integration. The article summarizes foreign experience of vertically integrated structures formation in the agro-industrial complex. Special emphasis is made on the methods and ways of small agriculture producers' integration into agro-business. Authors studied the indexes of the current condition and agro-industrial complexes (AIC) development tendencies in the EU countries, agro-holdings dynamics, as well as the structure of used farmland in the EU countries. Authors also present the results of the largest multinational corporations' activity, formed due to active integrating processes in agro-business sphere. Authors underline the peculiarities of integration bonds formation in the EU countries' agro-industrial complex. The author believes that the further development of the vertical integration depends on the process of institutional mechanisms perfection. They provide the enlargement of small farms' participation in the present-day AIC market.

Keywords: agro-industrial complex (AIC), agribusiness, vertically integrated structure, agro-holding, agro-industrial integration, farm households

1. Introduction

1.1 The Concept and Essence

Agro-industrial complex (AIC) is one of the main generators of employment and profit and plays the major role in food provisions manufacture and the problems of undernourishment solvation. Agro-industrial complex is a combination of industrial branches dealing with production, processing and sale of agricultural products as well as of the branches serving functioning of agricultural enterprises including agricultural machinery. Agro-chemical manufacture, melioration branch and so on.

The term 'agro-business' stands for collective entrepreneur activity in the sphere of agro-industrial complex (AIC), including supply of agricultural raw materials, manufacture and processing of agricultural products and its distribution among the final consumers.

The main condition of agro-business activation can become the development of integration processes. Joining up of the enterprises, participating in different kinds of activity in AIC field, allows to use various kinds of resources (manufacturing, financial, investing, human resource, raw materials and others) more effectively, to use the scientific technical potential, to provide competitiveness of agricultural production and food products.

Integration processes development in agro-industrial complex is a global regularity. Developed countries experience shows that at present only large agro-manufacturing joints, including different forms of manufacturing and trade cooperation (agro-firms, agro-corporations, agro-holdings, agro-concerns and so on), can operate properly.

1.2 Reasons of the Necessity for the Integration Process in the Field of AIC

Various directions and forms of integration processes have been developed in Western Europe countries, still the vertical agro-industrial integration has become the most widely spread in the AIC sphere. One of the advantages of this integration type is the enhancement of information exchange, gaining control over critical resources, e.g. over products net cost, noticeable increase of output product quality.

In the conditions of severe competition, the system of promotion of agricultural production to the consumers obtains a huge meaning. Therefore, some structures are formed in which the manufacturer's vertical integration is actualized with logistics and realization organizations.

At present, the vertical integration processes are being actively developed in the Russian Federation as well. In many cases, large vertically integrated corporations make a decisive contribution to the formation of territorial budgets, investment attraction, filling local markets, development of socio-economic infrastructure. That's why the studying of foreign experience of vertically-integrated transnational corporations' construction and functioning in agro-industrial complex appears especially relevant.

The aim of the presented article is to summarize the world practice and to find out the peculiarities of vertically integrated agro-industrial structures development in the Western European countries.

1.3 Methods

Approximately up to 50-60-ies of XX century the agricultural manufacture of Western European countries existed mainly in the form of scattered into small fragments (parcels) of natural and later—of petty economies -style peasant farms.

Industrial revolution and formation of large-scale trade, financial and industrial capital resulted in concentration tendency appearing in all the economics spheres including agricultural one.

Historically, the first integration direction in Western European agriculture was vertical integration, which start was connected with buyers of agricultural products appearing in the villages. They had a certain capital and a possibility to buy a ready set of agricultural products, sort the goods out and sell them. Simultaneously, the buyer informed the producers about the market demands and consumers' preferences. Consequently, he managed the specialization of agricultural manufacture and influenced the technological processes.

The integration interaction developed with time. New processes were introduced into integration interaction, including agricultural raw materials processing and ready production selling to the consumers. The manufacturer of the agricultural production as an element of integrated agro-industrial system was economically dependent in everything connected with specialization, selection, supply and selling, technological politics etc. With this, he formally remained autarkic and legally independent.

Thereby, large integrated structures bring small-scale producers under their influence sphere, subdue them, and organize them in necessary scales.

At the present stage of integration processes in AIC sphere, the most intense development is with the vertical bonds of agricultural manufacturers with industrial enterprises. A vertically integrated company is a holding, which structure of assets includes the shares of participation in enterprises, responsible for two or more stages of manufacturing-commercial circuit, in the scope enough for implementation of factual control over their activity.

Integrated systems are formed on the base of certain principles:

- Technological connection of the subjects providing the volume increase and products competitiveness improvement.
- Necessary financial and material support in joint manufacture organization and activity.
- Existence of management company (integrator), which has the experience of business set-up, investment resources management and possibilities of loans attraction.
- Active participation in production-financial processes of all the structural subdivisions of integrated units under centralized management.
- Creation of economic conditions, providing all the integrated organizations with equal possibilities to achieve good results in their production and household activity.
- The process of involving agricultural producers into vertically integrated structure can be actualized in two ways.

The first way is when agro-industrial integration has the form of a concern, i.e. a trust with the single property and centralized management. In this case, two or more different producing processes can be united (for example, keeping the milk cattle flock and dairy products production). Equally, it can be a joint of enterprises, which have a common value chain. For example, breeding livestock is performed at different farming households, while the products processing and meat products are made at a meat-processing plant, which is an enterprise-integrator

The second way is to sell the finished products through specialized and universal trade commissary units. In this

case the enterprise-integrator can manage the production volumes and control the production-selling activity of the primary agricultural producer (a farmer).

Depending on the integration growth direction, there are two kinds of vertical integration distinguished in the global practice, both of which are applied in agro-industrial complexes:

- Backwards vertical integration – it supposes control acquirement or control increase over suppliers, i.e. agricultural enterprises-producers.
- Direct vertical integration is expressed in the company expansion by means of control acquirement or control increase over the structures, distributing and selling the products.

A consequence of the interaction of contradictory trends of globalization and localization is coming to a regional economic space other regional corporations, formation in agro regions transregional territorial industrial clusters and inter-horizontal and vertically integrated structures for integrating local structures into broader productive systems - inter-regional, national, cross-country, global - and transforming regions into economic entities of the world economy.

Hence, the vertical agro-industrial integration strengthens the competitive advantages in the same sphere either concerning the raw material sources or concerning the direction of ready products promotion and selling to the final consumers.

The main methods of vertical agro-industrial integration are:

- Consolidation – amalgamation of the integrated enterprise with the mother company as a branch or a subsidiary.
- Grouping – formation of a special structure consisting of the integrated enterprise and the basic company which has the property commitments (shares exchange), or management commitments (including another enterprise directors into the board of directors), or various services commitments (informing/advertising, technologies provision, etc.).
- Target extension – integration is reached by means of supply contract extension aiming at setting the steady long-term relationships.
- Franchising – passing the rights to use the brand, the business-model of the main company to the integrated enterprise, as well as passing other types of property necessary for setting-up and running the business.

The most widely spread form of agro-industrial integration is contracting, i.e. signing up the contracts between the immediate agricultural producers (farmers) and non-agricultural enterprises. Prices, volumes, terms of products supply, sides' obligations concerning materials supply and production means, production quality indexes as well as technological conditions of production are fixed in the contract [11]. As a rule, the contracts are negotiated for the term of 5-10 years, which provides the partners with stable long-term business relationships.

Another direction of agricultural production concentration is agricultural cooperation, which supposes the uniting of separate peasant households in separate kinds of manufacture processes. Organizational-economical result of the agricultural cooperation and that of the vertical integration is the same. The difference lies in diametrically different direction of these processes: vertical integration is actualized top down, (often ignoring the interest of primary producers), while agricultural cooperation is a bottom-up process and takes into consideration the maximum of peasant households interest.

2. Results

2.1 Development Indicators of the Agricultural Sector in the EU Countries

EU is the leading producer of milk, swine, cattle and corn in the world. Table 1 shows the main products, produced in the EU, and their share in the global production.

Table 1. Main kinds of agricultural products, made in the EU, and their share in the global production

Product	EU countries		The whole world		EU share in the global production, %
	Volume of goods production billion $	Production volume, billion, tons	The volume of commodity output billion $	Production volume, billion tons	

Milk	45,07	150996	187	625754	24,1%
Pork	34,10	22184	167	108507	20,4%
Beef	21,53	7971	169	62737	12,7%
Paltry	15,05	10567	132	92730	11,4%
Gapes	12,31	21543	38	67067	32,1%
Wheat	12,25	136081	79	671497	20,3%
Potato	7,53	54564	49	365365	14,9%
Tomatoes	5,55	15159	59	161794	9,4%
Chicken egg	5,51	6648	55	66373	10,0%
Apples	4,61	10982	32	76379	14,4%
Corn	1,56	59358	54	872792	6,8%

In 2012 EU agricultural households produced about 159,4 billions euro of added value, which makes 1,4% of the total EU economy added value. Nevertheless, in the period of 2002-2012 the input of agriculture decreased from 1,8% to 1,2%.

11,3 million of people older than 15 years are employed in the EU agriculture (1,1 million people of them are older than 65 years), which makes 5,2 % of all employed people in the EU countries. The agricultural census fixed that 23,5 million people out of 25,5 million people working in the agricultural sphere permanently, were either the owners or the owners' family members.

There's a striking contrast between the regions in western and eastern EU parts from the point of view of productivity. The highest working efficiency (over 45 000 EUR) is fixed in the Netherlands and in different regions of France and Great Britain which are specialized in growing plants. The low agricultural labour productivity (less than 5000 EUR) was demonstrated by Poland, Bulgaria, Romania, Portugal and Slovakia.

According to Eurostat the general amount of agro-holdings in EU countries in the period of 2003-2010 years fell by 18,5% and by the end of year 2010 made 12,2 thousand of units (Figure1).

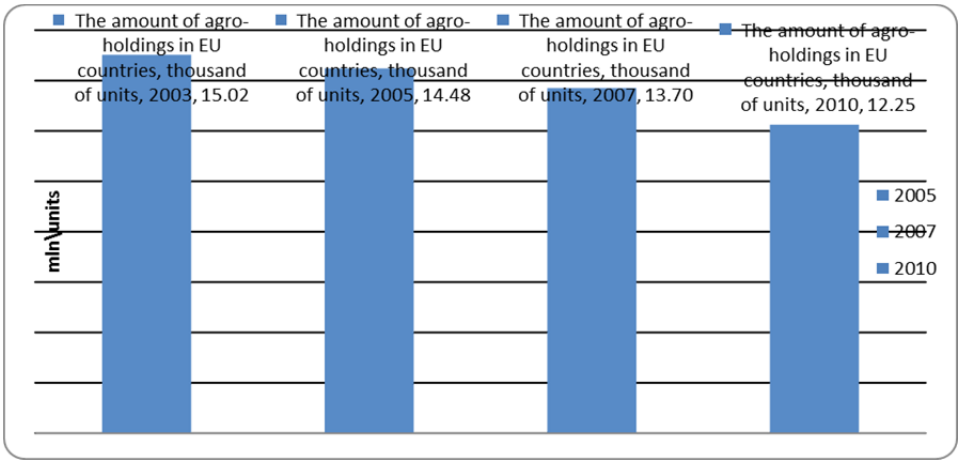


Figure 1. The dynamics of the agro-holdings number in EU countries in 2003-2010

The most rapid decrease (-35%) was demonstrated by livestock-breeding agro-holdings, taking 56% of all agricultural land: their number by the end of the year 2010 was 6,7 mln of units.

Every fourth agro-holding (24,9%), in EU households specializes in growing field crops (for example, grain, oilseeds, vegetables). Also even one from five agricultural holding (20.1%) specializes in permanent crop (vineyards, olive groves and orchards). The holdings, specializing in cattle pasture (milk cows, sheep and other ruminant animals), swine, poultry breeding holdings and mixed livestock-breeding agro-holdings take almost a half of all the EU agrarian (46,7 %).

For the same period there was an even greater reduction (35%) in the number of livestock farms, occupying 56% of EU agricultural lands (at the end of 2010 the number of farms was 6.7 million units). At the same time the territory of all used agricultural grounds is 174,1 mln. hectares, which makes 40% of the general EU territory.

The average size of one household is 14,2 hectares. Half of the grounds used in EU agriculture is cultivated in 4 countries: France (16,0 %), Spain (13,6 %), Great Britain (9,7 %) and Germany (9,6 %). One fourth of other territories of agricultural grounds (23,3%), are in Poland, Rumania and Italy, (figure 2).

Still one can observe harsh contrasts in the structure of agriculture in EU countries. On the one hand, there's a huge amount (6.0 mln. or a half of all the agro-holdings) of very small households (less than 2 hectares) which take a small share (2.5 %) of the total grounds used for agriculture. And on the other hand, a small amount (2,7 % of all the agro-holdings) – are very large farms (more than 100 hectares) which cultivate almost a half (50,2 %) of all the EU agricultural grounds.

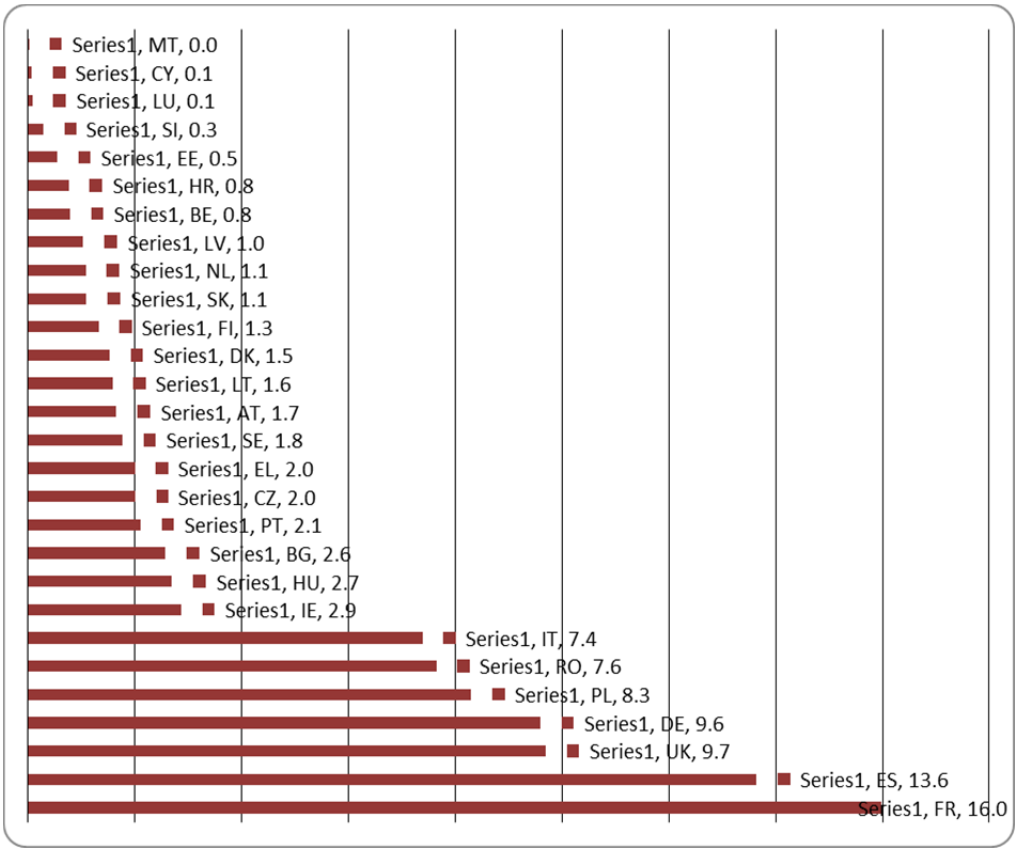


Figure 2. The structure of used agricultural grounds in EU countries-28, %

However, there are significant contrasts in the structure of agriculture in the EU. On the one hand, there are a large number of agricultural holdings (6.0 million.), consisting of very small farms (of less than 2 ha each) that occupy a small fraction of the total land area (2.5%). On the other hand, a small number of agricultural holdings (2.7%) are very large farms (with an area of over 100 hectares of farmland), treated with almost the half (50.2%) of farmland in the EU.

Almost a third (31.5 % or 3.9 mln.) of all the agro-holdings in EU was in Rumania (figure 3).

These agro-holdings can be characterized as small ones: three fourth of Rumania private grounds were 2.0 hectares big. Italian agricultural enterprises can be called small too (their share in the total number of EU households is 13,2%) while Polish households make 12,3 %– their size is on average 10,0 hectares.

At the same time there are quite a number of states –EU members in which larger agro-holdings are more typical. The majority of German households (53,9%), French (54,4%), Dutch (55,1 %), Finnish (57,0%), Irish (57,8%), UK (61,4 %) and Luxemburgish (65,5 %) were larger than 20 hectares. The average size of an agricultural holding in Great Britain is indeed 6 times larger (90,4 hectares) than in EU on average in 2010 while the average size of holdings in the Check Republic was even larger (152,4 hectares) with the small number of really huge farms.

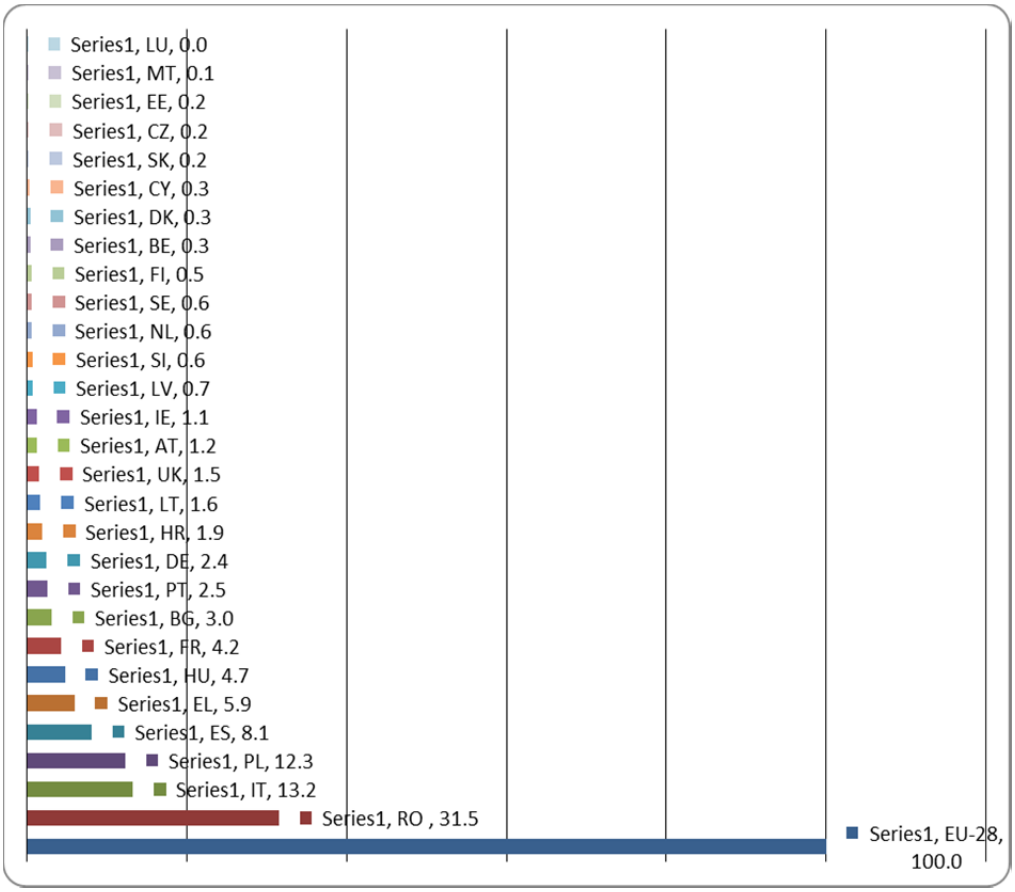


Figure 2. The amount of agro-holdings in EU-28 countries in 2010

This contrast is reflected on holdings' economical results. 5,5 mln. of enterprises (44.6 %) out of 12.2 mln. of agro-holdings in EU had a standard output of less than 2 000 euro and provided only 1,4 % of total agricultural economical manufacture in 2010. The standard output of agricultural products (crops or livestock) in average monetary terms makes the volume of agricultural production in farmers' prices, in euro per hectare or per one beast. On the contrary, 1.9 % of households which had a standard output of more than 250 000 EUR, made almost a half (47.8 %) of all agricultural economical results.

2.2 The results of the development of the integration processes in the field of agribusiness in EU

As a result of active integration processes in the agro-business sphere multi-profile transnational joint ventures are being organized. They consist of groups of enterprises, situated on the territory of different countries but under the united management of mother companies such as Unilever (the Netherlands /Great Britain), Anheuser-Busch InBev (Belgium), Nestle (Switzerland), Danone (France), Imperial Tobacco Group, SABMiller, British Foods (the UK), Pernod Ricard (France), Heineken Holding (the Netherlands), Carlsberg (Denmark). Ten largest agro-industrial multinational corporations are presented in Table 2.

Table 2. Top-10 largest agro-industrial multinational corporations in 2013

Brand	Country of origin	sales, billion. $	Profit billion. $	Assets, billion. $	Market value, billion. $	Staff, thousand people
"Nestle"	Switzerland	100,6	11,6	134,6	233,8	339
Unilever	The Netherlands/Great Britain	67,7	5,9	59,9	117,8	172

Danone	France	27.5	2,2	38.1	44,2	102,4
Associated British Foods	The United Kingdom	19,8	0,9	16,2	22,9	106,2
British American Tobacco	The United Kingdom	24,7	6,2	43,9	103,2	87,5
Heineken Holding	The Netherlands	24,3	3,9	46,7	43,5	76,2
SABMiller	The United Kingdom	16,9	4,3	55,2	84,3	77,1
Imperial Tobacco Group	The United Kingdom	23,7	1,07	44,4	34,1	37,2
Diageo	The United Kingdom	15.34	2,67	28.69	40.93	25,7
Pernod Ricard	France	10,4	1,5	33,4	33,1	18,3

As Table 2 shows, five (not considering the English-Dutch company Unilever) out of 10 largest agro-industrial multinational corporations are from the United Kingdom of Great Britain and Northern Ireland.

The initiator of corporate tradition to control the whole production circuit from growing the raw materials to selling the products to the final consumer is the Unilever Corporation. The company was founded in 1929 by uniting the margarine Union of Holland 'Margarine Uni' and the British soap-producing 'Lever Brothers'.

The first vertically-integrated company Unilever owned the enterprises of processing seeds into oil, packing business, trucking industry, retail net.

Agro-industrial multinationals' structure includes national vertically-integrated companies, distributed according to geographical zones. As a rule, the agricultural sector of multinational corporations' agribusiness relies on the raw material base of the country-recipient.

For example, the Swiss global corporation "Nestle" uses the regional model of interaction with the local farmers in the placement countries. The basis of this model was developed as long ago as 1870. It includes: creation of centers of milk collection providing the quality control and milk safety, its cooling and weighing on the electronic balance; payment for the milk; free technical and veterinary support of milk cattle; providing transport and infrastructure nets; microloans for the farmers (25 mln US dollars per year); in-time payment for every milk delivery.

Danone company, specializing in fermented milk products, realized the strategy of absorption of foreign enterprises, aspiring to increase their economic activity and financial productivity indexes by means of vertical concentration.

The organizational structure is largely decentralized. National subsidiaries follow the same corporate strategy and vision but they possess enough independency in making decisions concerning products adaptation to the local consumers' demands in taste, texture and packing.

One of the key elements of the corporation business-strategy is the development of partner relationship with the suppliers. Thus, for example, Danone company uniting with the French Institute of live-stock breeding 'Institut de l'Elevage' and dairy farmers became a company under the name "Le lait de nos éleveurs" (milk of our dairy farms) aiming at public awareness increase concerning the production quality. One more example is the program "Vertical growth" launched in 2012 by the Group of Companies Danone in Russia. Within this program Danone experts provide their suppliers with consulting services, regularly organize studying trainings in Milk Business Academy (MBA). More than 150 managers of dairy farms from different Russian regions have passed this course. Different programs of motivation, including celebration of "Supplier's Day", are implemented.

It is evident from the above-described information that in the conditions of the agricultural market globalization there's such a widely spread form of vertical agro-industrial integration as a transnational corporation, which influences greatly the development of AIC branches and global economy on the whole.

Discussion

The structure of agriculture in EU countries differs greatly under the influence of such factors as geology,

landscape, climate and natural resources. It also depends on the variety of regional events, infrastructure and social customs. The formation of integration bonds in EU countries has a number of regional and industry-sector peculiarities.

For example, in Great Britain the most widely spread form of agro-industrial integration is industrial-trade corporations.

In Western Europe countries, including France, the most widely-spread form of integration relationship is a holding. Integration processes in France primarily touch upon the agricultural enterprises, those of food industry and the sphere of food trade.

Cooperatives in dairy sector have the highest degree of vertical integration, which is connected with the necessity of immediate milk processing. Large cooperatives specializing in dairy products are mostly located in the west of France. Milk cooperatives of other regions supply milk to the plants of other enterprises—cooperative unions (mostly of the union "Sodiaal") or private companies.

Vertical integration in the grain sector happens both with sale and processing of farm products and with supply of manufacture means. Feed compounds manufacture is the most widespread sector of all the grain-processing sectors. A deeper grain processing (flour-grinding, malt production, manufacture of bakery semi-finished products) is spread only in large agro-holdings which structure contains processing subsidiaries.

A harsh contrast between large vertically integrated holdings and cooperatives with a short manufacture chain is present in meat industry sector. Vertically integrated holdings realize the whole chain including fodder production, cattle slaughter, and meat processing and selling. This kind of integration is typical of poultry and swine farming. For meat cattle sector, cooperatives with a short production cycle (primarily cooperatives of cattle sale) are most typical.

Government plays an important role in formation of integration bonds. The Supreme Council for orientation and coordination of agriculture and food industry development is the organ of inter-sector coordination in food complex of France. It includes Ministry of Finance representatives, manufacturers in the sphere of processing and realization of agriculture products and farmers. Its functions are the development of measures concerning enhancement of food sub-complexes organization and management, pursuing the coherent agricultural and food policy.

In Germany as well as in the countries of central Europe special cooperative state unions are created to smooth the relationships between agricultural enterprises, companies selling agricultural products. Their activity sphere touches upon consultancy work on tax law, cooperation matters, cooperative system subjects representation before legislation organs and society, experience exchange. In order to extend the long-term warranty on keeping rather high purchase prices for agricultural goods in Germany they created the Farmers Union.

Processing enterprises are the integrators in formation of partner relationship between the agriculture and other German trades. It's around them, where the agro-industrial integration is put into practice.

In the countries of Central Europe (Denmark, Norway, Sweden, Finland) the dominating form of agro-industrial manufacture is an agricultural cooperation.

In the Finnish and Swedish agriculture there are Central unions of agricultural manufacturers which work in the close mutual collaboration and provide farmers, authorities, economy, professional unions and population representatives with information. Their activity sphere includes negotiation and development of agreements with the government concerning the problems of agricultural goods supply volume planning, food prices and other questions concerning farms profits.

Agricultural cooperatives have the decisive role in working out and performing the state agrarian policy. Subsides to support agriculture are transferred through them. It so happened that agricultural cooperatives have become the most effective form of Scandinavian countries' agro-business integration. They provide the farms with necessary manufacture conditions; guarantee the sales and efficient introduction of agro-industrial innovations.

This way, the results of research of integration processes development peculiarities in the sphere of EU countries AIC let us state that the most appropriate coordination form of large scale manufactures and small scale farming based on the private land property, is the form of agricultural cooperatives.

It should be noted that in many vertically integrated agro-industrial structures problems of providing organizational and administrative and economic unity, equal benefit of economic relations, responsibility and interest of the participants in the final result of combining the activities of the integrated form, as well as the

deepening of integration relations are remain unsolved. Overcoming the existing antagonism of economic interests, as well as the disparity of barter relations between agricultural, processing and trading enterprises belonging to the agro holding structure requires scientific study of the system of the inter-economic of the above mentioned integration partners.

Institutional mechanism of AIC diversification is definitely in creation of large vertically integrated agrarian corporations as the most effective form of household organization in AIC. At present, the strengthening of integration processes in agro-industrial complex of European countries is essential for agro-business development. The successfully arranged bond guarantees getting additional cost in agricultural sectors, creating working places and increasing the level of manufacturers' profit.

Petty farmers can't remain the food products only, they'll have to take up an additional role of entrepreneurs and work behind the subsistence farming limits. Besides, agro-industrial companies need reliable local raw material suppliers to increase their international competitiveness.

Modern agro-food system implies intense competitive pressure on all the participants of the cost creation chain. The necessity to respond to the consumers' demands, processers', retailors', transporters' demands imposes stricter and harder demands on the agricultural suppliers from the point of view of quality, terms, processing, transporting. Small producers, representing the main part of global agriculture, nowadays face the growing problem of following these demands.

The question of small farmers' participation in agro-industrial system is not only the question of the system effectiveness but is also one of the key questions of rural inhabitants' welfare level. Exactly for these reasons, the perfection of institutional mechanisms providing the extension of small farms' participation in the modern AIC market must get all the attention of Food and Agriculture Organization of the United Nations [17].

One of such mechanisms is the contracting of agriculture, which includes several aspects, which deserve a separate study and further enhancement: economic, institutional, social, legislative and others.

The advantages from the contract include extension of the approach to the market for the small farms, credits and technologies availability, risk management improvement, agricultural population employment perfection and successful commercial farming.

Now there are five basic models of the interaction of farms and integrator-companies that can define the selection of appropriate mechanisms for the development of contract farming. Enlarged characteristics of basic models of interaction are presented in Table 3.

Table 3. Models of interaction and their characteristics.

Characteristics	informal model	Intermediary model	Multilateral model	Centralized model	nucleus estate model
Investment\credit	never	seldom	seldom	sometimes	sometimes
Agroconsulting service	never	sometimes	always	sometimes	sometimes
Use of conracts	never	sometimes	always	always	sometimes
Farm groups	never	sometimes	sometimes	sometimes	sometimes
Growing management	never	seldom	always	sometimes	often
Centralized production\processing	sometimes	sometimes	sometimes	sometimes	always
Postharvest logistics (packaging, transport)	sometimes	seldom	seldom	often	always

Informal model - is the most short-term and speculative of all models of interaction with the risk of default for the integrator company, and for the farm. However, it depends on the situation: the interdependence of parties of the contract or long-term trust relationship can reduce the risk of opportunistic behavior. Features of the model:

- Small firms enter into a simple, informal contract for seasonal production with smallholders.
- Often, success depends on the availability and quality of external contacts
- Typical products: requiring minimal processing / packaging, vertical coordination; for example, fresh fruit / vegetables on the local markets, sometimes the main crops.

Intermediary model involves the integrator company to conclude subcontracts with intermediaries (buyers of

agricultural products, the company aggregators) who work with farmers on the basis of formal or informal agreements. This model can work, if it has good incentives structure and adequate control mechanisms. However, this model has drawbacks for vertical coordination and provision incentives for farmers (integrator company may lose control on production processes, quality control and regularity of supplies, farmers can not benefit from technology transfer; there is also the risk of price distortions and decrease of farmers' level income).

Multilateral model This model can evolve from a centralized or nucleus estate model, for example, after the privatization of parastatals. It includes a variety of organizations, such as government agencies, along with private companies and sometimes financial institutions. Vertical coordination is carried out at the discretion of the integrator company. Proper attention should be paid to the possible risks of political interference. Except the agreement with agrofirm used in this model agreement with the third partyis used, which are service providers (for example, education, loans, deposits, logistics).

Centralized model In this model, the integrator company interacts with a large number of small, medium and large farmers. Participation of integrator can vary from a minimum contribution of a resource (such as certain types of crops) to the control of the most aspects of production (for example, from soil preparation to harvest). Communication and coordination is carried out strictly vertical. Quantities (quota), quality and delivery times are determined at the beginning of the season. Production processes and quality are strictly controlled, sometimes it is directly carried out by employees of the integrator company.

Nucleus estate model In this model, the integrator company provides centralized production and processing on the own real property, complementing their capacity through direct contracts with farmers. Keeping real estate requires a significant investment in land, machinery, staff from the integrator. Model of the nucleus estate usually guarantee deliveries to ensure cost-effective use of installed refining capacity. In some cases, this model is used for research, breeding demonstration purposes and / or as a collection point.

Each of these models of interaction between the integrator company and farmers has a number of advantages and disadvantages (Table 4).

Table 4. Advantages and disadvantages of models of interaction between the integrator company and farmers

Model	Advantages	Disadvantages
informal model	Virtually no investments from integrator company in technical / financial support for farmers; Low operating costs; high level of supply flexibility	Limited control over production (products, variety, quality, etc.); High risk of supply disruptions; Strong buyer competition.
Intermediary model	Risk reduction, provided effective control; Minimum investment in technical / financial support; Slight improvement of supply chain management; Low cost of switching to new partners	Buyer opacity for farmers Marginal production control (volumes, quality)
Multilateral model	Limited investment and reduce costs by sharing the costs between the partners; Risk reduction (compared to commercial production) due to geo distribution of agricultural contractors	High risk of side sales; No primary production, polnaya dependence small producers; High transport costs
Centralized model	Provides a high level of quality control and production volumes; Close interaction with farmers holding back third-party sales	The high level of investment for technical assistance, as well as pre- and post-harvest logistics

Estate nucleus	high level of control over the supply chain; Simplified technical assistance / extension Low risk of supply disruptions	Requires a significant investment in production; Higher risks associated with the harvest; Limited flexibility in the choice of agricultural contractors

The model chosen for the initial phase with the passage of time can and should be adapted and possibly changed to consolidation and extension of scale, respectively, to the higher stage of development.

3. Conclusion

Thus, the summary of integration processes practice in AIC of EU countries let us make the following conclusion:

- In the integration process of agricultural, processing and sale companies, the system of contracting has become the most widespread in EU countries.
- In the majority of cases the direct initiators and coordinators are non-agricultural companies, at this the object of integration is agriculture.
- There is active development of various forms and structures of agro-industrial unions (from cooperation to agro-industrial multinationals), which affords to consolidate and use the available resources more effectively.
- Vertically integrated structures with a complete circuit are formed—from the agricultural products producer to its selling to the customers.
- In many European countries, the mechanisms of governmental regulation and support are actively applied in the attempts to develop the agro-industrial market.

Vertical integration allowed to extend the companies activity in AIC. Large agro-industrial corporations of EU countries were formed due to the development of agro-holdings and cooperatives.

References

Aleshina, T. N. (2010). *Agro-industrial integration in the AIC system*. Moscow.

Anikina, B. A., & Rodkina, T. A. (2014). *Logistics. Theory and practice. Managing the supply chains*. Moscow: Prospect.

Ayusheeva, A. O. (2013). *Formation of integrated structures of the regional agro-industrial complex: problems and perspectives.* monography. Publish House CENS.

da Silva, C. A., & Rankin, M. (2013). *Contract farming for inclusive market access. Food and agriculture organization of the united nations*. Retrieved from http://www.fao.org/uploads/media/cf.pdf

European Commission. (2013). *Eurostat regional yearbook.* European Union.

Eurostat. (2010). *The results of agricultural census of the year.* Retrieved from http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Agricultural_census_2010_-_main_results

FAOSTAT. (2014). *Food and Agricultural commodities production. Official site of Statistics Department of Food and Agriculture Organization of UN FAOSTAT*. Retrieved from http://faostat.fao.org/site/339/default.aspx

Financial Times Global 500 Rating. Retrieved from http://www.ft.com/indepth/ft500

Food agricultural UN organization (2014). The situation in food and agriculture sphere. Official site FAO from http://www.fao.org/3/a-i4036r.pdf, free

Food agricultural UN organization (2014). The situation in food and agriculture sphere. Official site FAO. Retrieved from http://www.fao.org/3/a-i4036r.pdf

Implementation of the CAP reform. Agriculture and rural development. Retrieved from http://ec.europa.eu/agriculture/cap-post-2013/implementation/index_en.htm

Kerashev, A. A., & Mokrushin, A. A. (2011). The strategic management of interaction between vertically integrated corporations and regional economic systems of the South of Russia, Bulletin of Adyghe State

University. Episode 5: The Economy, 3.

Maltseva, I. S. (2013). Cooperation and agro-industrial integration in AIC. *Syktyvkar: SLE.*

Mokrushin, A. A. (2008). Challenges for intra-economic interaction of subjects of agroholding structures. *Scientific Journal KubSAU, 35*(1).

Mokrushin, A. A. (2011). The strategic aspects of the interaction of vertically integrated corporations with regional economic system, Bulletin of Adyghe State University. *Episode 5: The Economy, 2.*

Semykin, V. A., & Safronov, V. V. (2009). Development effectiveness increase in the agro-industrial complex on the base of realization of its activity diversification strategy. *Kursk state agricultural academy Vestnik, 3.*

Turyansky, A. V., & Anichin, V. L. (2010). *Agricultural cooperation and agricultural integration*: Textbook. 2-nd edition corrected and extended. Belgorod: BelSAA Publish house.

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Development of Models and Methods of Data Analysis for Enhancing Efficiency of the Processes of Quality Management Systems

Alla Vladimirovna Kuzminova[1] & Valeriy Valentinovich Gurov[1]

[1] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia

Correspondence: Alla Vladimirovna Kuzminova, National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 115409, Russia, Moscow, Kashirskoe highway, 31, Russia E-mail: AVKuzminova@mephi.ru/VVGurov@mephi.ru

Received: January 9, 2015 Accepted: March 20, 2015 Online Published: June 25, 2015

doi:10.5539/mas.v9n8p429 URL: http://dx.doi.org/10.5539/mas.v9n3p429

Abstract

This article deals with the education system, along with the other systems, which should include monitoring, measurement and analysis of the processes required to improve the quality management system. In this system, a student is considered as a product being created together with its input and output parameters and characteristics. Due to investigation of the nature of relations between the training process parameters, there are being developed methods and mathematical models that describe regularities of the system in order to enhance efficiency of quality management processes in education.

Keywords: academic success, cluster analysis, regression analysis, maximum likelihood estimation, principal components method, levels of quality

1. Introduction

Avalanche-like increasing flow of information makes it difficult for a contemporary person to qualitatively accept the necessary data, to process, comprehend, preserve, and create new knowledge.

In a current situation of information crisis, in order for a person to find his/her place in the sphere of material and socio-cultural production, special load must be transferred to the sphere of continuous training and education.

Teaching a person to navigate in this rapidly changing world requires time, but a significant part of the knowledge obtained, as well as the skills in its processing a person receives in the course of study in an institution of higher education. In this regard, the task of selecting applicants able to master more and more considerable volumes of scientific and technical knowledge and successfully complete training programs is faced.

In accordance with the ISO 9000 global standards (ISO Standards; State Standard of the Russian Federation, 2006), an organization shall carry out monitoring, measurement and analysis of the processes necessary to improve the system of quality management. Increase in recording and analysis of the quality of human resource development is possible if based on the established requirements for the evaluation of educational processes described in the National Standard GOST R 52614.2-2006, State Standard of the Russian Federation, 2006 (State Standard of the Russian Federation, GOST R 52614.2-2006). This standard includes guidelines for institutions providing educational services; it proclaims quality management systems in education. An information level of quality analysis system (Recommendations on Standardization, 2002, http://www.complexdoc.ru/ntdtext/541946/; The ISO/IEC/IEEE 42010 Website, http://www.iso-architecture.org/ieee-1471/; Cheremnikh, Semenov, & Ruchkin, 2003) is being developed. Parameters of product's conformity with the levels of quality required may serve as the input data for such analysis. While considering a university student within the educational system as a product being created, the education quality data in the form of examination marks for each term, and the psycholinguistic parameters of their written language, are analysed (Kuzminova, 2013; Sticht, 1973).

Problems of education quality are solved while measuring, analysing and improving processes of education system. These educational institution processes may be demonstrated through students training processes, determination of the process quality, and final evaluation of the achievement rate with awarding academic degree to a graduating student in accordance with his/her diploma.

To enhance the efficiency of educational process, on the agenda there will be certainly raised a question concerning an individual tutor-navigation for a trainee, which depends not only on the features of his/her specialization and scores obtained during training, but also on the psychological personality characteristics, such as responsibility, purposefulness, ability to work and create a new theoretical or practical product.

To address one of the priority problems concerning formation of mechanisms for education quality evaluation, systems of evaluating vocational training quality are to be developed. Urgency of the problem under consideration is confirmed by a significant amount of studies performed in this area (Marukhina, & Berestneva, 2003; Educational Testing Service ETS, http://www.ets.org/gre). It is also determined by the necessity of excellence in a reduction in the number of applicants in connection with the demographic problems in the country. These processes are taking place against the backdrop of the increasing amount of information that you want to learn. The solution to this problem is particularly important in the transition to a multi-stage system of higher education in Russia.

Thus, the development of models, methods and means of data analysis is quite urgent in order to enhance the efficiency of quality management system processes, as well as the study of predicting methods for specialist training quality.

The analysis carried out has shown availability of a large variety of research efforts that are similar in theme. However, predicting of academic success (AcSuc) itself did not receive sufficient consideration. The reason for this is that the research issue considered in this work belongs to the interdisciplinary problems, and has not been so far clearly specified.

In this article, we will focus on the following issues:

- identification of the most significant factors affecting the determination of academic success levels;
- construction of mathematical models that makes it possible to identify availability of different-class data domains by means of investigated parameters of the educational process quality indicators;
- study of models of predicting academic success levels as an indicator of the educational process quality;
- development of predicting methods for the levels of mastering training programs by trainees;
- experimental test of the developed models and methods.

To address these issues, an analysis of the expert information data (Giarratino, & Riley, 2007; Muromtsev, 2005; Jackson, 2001) of the study sample was carried out, and on their basis academic success indicators were invented (Bogomolov, et al., 2009). Simultaneously, texts of the trainees' written works were considered. The analysis of syntactic parameters of the written language (Luria, 2002; Popov, Y., et al., 2001; Baranov, 2003; Flesch, 1948) was performed, which resulted in development of information text parameters. Analysis of the developed mathematical predicting models of academic success levels as an indicator of specialists training quality was carried out. Using computer-based methods of information processing, imaging of the results obtained during the study of constructed mathematical models of the boundary between the predictive levels was carried out. When creating an information system for determination of specialists' quality levels, these data are considered to be the system's input data. In order to determine the prognostic level of quality, mathematical models of success and information text model are created.

2. Methods

The purpose of this article is to describe the development of means that would contribute to the enhancement of quality management system efficiency in education through the predicting of the level of specialists training quality. This work demonstrates mathematical model for determining threshold between the predictive levels of specialists training quality and methods developed for their application. Models were created using methods of regression analysis and maximum likelihood estimation.

2.1 Construction of the Model for Classification of Academic Success Levels

The study of a complex system, as a rule, requires its preliminary partitioning into subsystems and determination of boundaries between them. There are many methods for solving this problem, ranging from classical statistical methods to methods of nonparametric statistics and neural networks. Each method has its advantages and disadvantages. To solve this problem, we propose a new approach that allows for the set output parameters of the process to obtain optimal in terms of established criteria boundary separating objects within the system.

The general scheme of solving the problem of the classification of quality levels with the use of this approach is described as follows. Let us assume that we have a system, each element of which is characterized by a set of

parameters $\{W, X\}$. Within the first subset of parameters $\{W\}$, using the Spearman rank correlation, a selection of the weightiest parameter W_p is carried out. Further, division of the system under investigation into groups is performed according to the dedicated parameter, using cluster analysis and applying the nearest-neighbour method. This creates clusters $\{C_m\}$, relating to different quality categories $\{QC_m\}$, which are considered in terms of the quality management system.

For the second subset of input parameters $\{X\}$ of a certain process of the system in question, carried out is the typologization of its elements on the basis of their belonging to a particular cluster $\{C_m\}$ found, defined as a class of the system quality.

As a result, predictive models of separating parameters X_k with a glance to W_p are being constructed – models of boundaries between the selections of points. Experimental test of the predictive models developed is conducted. We analyse the results, and choose the best according to certain parameters mathematical model of the boundary between predictive levels of academic success – functionality Q.

For this problem, as the input parameters of the educational process, exam results for all examinations passed during the educational period at university and parameters of penscripts obtained after the written entrance exams to university in the Russian language are selected.

The next step is the development of principles for analysing indicators of university educational process. For this purpose, a model of academic success classification is developed, as well as the models of predicting levels of academic success. When evaluating quantitative indicators of determination of the trainee's academic success levels, one should not be oriented on the trainee's grade point average obtained during the entire education period. It cannot serve as an objective indicator for evaluating the success of mastering the material during the training period, since it can be affected by the personal attitude of a teacher-expert to a particular student or by the student's individual attitude manifested to the subject ("clear – not clear") or emotional attitude to the teacher ("like – do not like"). Therefore, assignment of a trainee to the number of highly successful or slightly successful students on the basis of this indicator is impractical.

As a replacement for this seemingly obvious indicator, we propose to use 5 indicators of academic success and 3 formulas of academic success for the entire training cycle during *n* terms. For example, the indicator of academic success, describing GPA of the *i*-th trainee for the entire training cycle, is represented by the following formula:

$A_i = \sum_{j=1}^{n} S_j \Big/ \sum_{j=1}^{n} K_j$, in which the amount of scores S_j, obtained during the j-th term $(j = 1, ..., n)$ for K_j exams is calculated. After analysing indicators of academic success with the help of Spearman rank correlation, the best of these indicators is selected.

For the further analysis of the selected indicators of academic success, the cluster analysis of time series of scores obtained by graduates as a vector of *n*-dimensional space points: $\vec{W}_j = (W_1, W_2, \ldots, W_n)$ is used. For this, sorting of expert evaluations is conducted through the method of cluster analysis, for which the parameters of cluster analysis (Kim, J., et al., 1989): Euclidean distance and nearest-neighbour method (single binding) are selected. As a result of such classification, clusters characterising different classes of academic success are formed. To determine the parameters of the statistical and cluster analysis, the Statgraphics Centurion Program

(Statgraphics: The Statistical Program) was used. As a result of partitioning the entire set of points ***W***, clusters with different, in general case, number of elements are formed. These elements belong to different areas of quality categories – levels of academic success $\{QC_m\}$.

Thus, the construction of the model for classification of AcSuc levels with the implementation of a trajectory model (cluster-analysis) of academic success is carried out. By its application, it is possible to identify a dynamic component in the determination of academic success.

2.2 Mathematical models of Different-Class Data Separation

As a result of the analysis of sample distribution of random variables – developed transformational parameters of texts (x, y, z) – it becomes possible to use them in solving the task of searching for the classifying boundary of AcSuc.

Search for the boundary of different-class data separation can be performed using a variety of well-known classification models. In our work, we consider construction of a classifier using the method of multiple linear regression MR and the maximum likelihood method (MLM) – probabilistic model PM.

In the first classification model of MR, as a regression function applied is the function of three variables $f\left(x, y, z, \vec{A}\right)$ as a polynomial function of degree s $\left(s \in N\right)$, where $\vec{A} = \{A_{s00}, \ldots, A_{001}, A_{000}\}$ – are the required model parameters, which are set by the least-squares method. In the description of the model using three predictors, multiple polynomial regression model Qrg of the degree s is applied, which looks as follows (1):

$$Qrgs = f\left(x, y, z, \vec{A}\right) = A_{000} + A_{100}x + A_{010}y + A_{001}z + A_{200}x^2 + A_{110}xy + A_{101}xz + A_{020}y^2 + A_{011}yz + A_{002}z^2 + \\ + A_{300}x^3 + \ldots + A_{421}x^4y^2z + \ldots + A_{00s}z^s \quad (1)$$

If you take recognition error as a criterion, i.e. the ratio of correct answers to their total number, which determines the quality of recognition, preference should be given to a suitable embodiment of the j-th regression model of MR_j. The study sample was obtained randomly from the entire assembly of data on the university trainees.

In the event of applying the second classification model, a probabilistic model PM, the educational process data (x, y, z) obtained are considered as random variables. The performed analysis showed that the resulting sample was normally allocated by the developed psycholinguistic text parameters. For the problem of estimating entire assembly parameters by sample data, there was used one of the methods of its solving demonstrated in the work (Kendall & Stewart, 1973). Mathematical model for determining the prognostic quality level is created by means of the maximum likelihood method application. Density of the normal probability distribution of a continuous random variable for one-dimensional case is described as follows: $P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$, where a [alfa] – is the mathematical expectation of a continuous random variable, σ [sigma] – is the root-mean-square deviation of normal distribution.

And given that the function of input data is a function of three arguments, density of its distribution is constructed in the following form:

$$P(x,y,z)=\frac{1}{\left(\sqrt{2\pi}\right)^3\cdot\sigma_1\cdot\sigma_2\cdot\sigma_3}\cdot\exp\left\{-\frac{(f_1-\theta)^2}{2\sigma_1^2}\right\}\cdot\exp\left\{-\frac{(f_2-\theta)^2}{2\sigma_2^2}\right\}\cdot\exp\left\{-\frac{(f_3-\theta)^2}{2\sigma_3^2}\right\},$$

where $f_j=[x-\bar{x}]\cdot\cos(\alpha_j)+[y-\bar{y}]\cdot\cos(\beta_j)+[z-\bar{z}]\cdot\cos(\gamma_j);\ (j=1,2,3)$, $\cos^2(\alpha_j)+\cos^2(\beta_j)+\cos^2(\gamma_j)=1;$

$\cos(\alpha_k)\cdot\cos(\alpha_j)+\cos(\beta_k)\cdot\cos(\beta_j)+\cos(\gamma_k)\cdot\cos(\gamma_j)=0;\ (k\neq j)$.

Unknown parameters $\cos(\alpha_j),\ \cos(\beta_j),\ \cos(\gamma_j);\ (j=1,2,3)$ shall be determined by the principal components method.

Evaluation of the parameter θ [teta] was carried out.

Having indicated: $\sum_n\{[x_n-\bar{x}]\cdot\cos(\alpha_j)+[y_n-\bar{y}]\cdot\cos(\beta_j)+[z_n-\bar{z}]\cdot\cos(\gamma_j)\}^2=S_j^2\cdot N$ and using method of maximum likelihood, there has been obtained an optimization problem: $\sum_j\left(\frac{s_j^2\cdot N}{\sigma_j^2}+2\cdot\ln(\sigma_j)\cdot N\right)\to\min$, which solution is found as follows $\sigma_j=S_j$

Given that the observation vector consists of independent random variables, obtained is a partition of the likelihood function into contributions of individual observations. For normal samples of two classes of transformational text parameters (x,y,z), in view of belonging to different levels of quality: $P_1(x,y,z)$ and $P_2(x,y,z)$, densities of probability distributions are determined.

For the first level it is obtained as follows:

$$P_1(x,y,z)=\frac{1}{\left(\sqrt{2\pi}\right)^3\cdot\sigma_1\cdot\sigma_2\cdot\sigma_3}\cdot\exp\left\{-\frac{\{[x-\bar{x}]\cdot\cos\alpha_1+[y-\bar{y}]\cdot\cos\beta_1+[z-\bar{z}]\cdot\cos\gamma_1\}^2}{2\cdot\sigma_1^2}\right\}\cdot$$
$$\cdot\exp\left\{-\frac{\{[x-\bar{x}]\cdot\cos\alpha_2+[y-\bar{y}]\cdot\cos\beta_2+[z-\bar{z}]\cdot\cos\gamma_2\}^2}{2\cdot\sigma_2^2}\right\}\cdot\exp\left\{-\frac{\{[x-\bar{x}]\cdot\cos\alpha_3+[y-\bar{y}]\cdot\cos\beta_3+[z-\bar{z}]\cdot\cos\gamma_3\}^2}{2\cdot\sigma_3^2}\right\};\quad(2)$$

Similarly, we obtain expression for the second level of quality $P_2(x,y,z)$.

Evidently, in the area of class separating hyperplane propagation, we can observe equality of probabilities of belonging to them. As a result, the surface – boundary separating the ellipsoids of data diffusion, which belong to different levels of quality is determined.

Based on the statistical analysis data of three-dimensional ellipsoids of diffusion and the principal components method, we carry out identification of the boundary separating two ellipsoids of data diffusion from the condition of probabilities equality at the boundary. A polynomial of the following form is obtained (3):

$$QQ=a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{23}yz+2a_{13}xz+2a_{14}x+2a_{24}y+2a_{34}z+a_{44}.\quad(3)$$

At $QQ=0$ polynomial (3) is represented as an equation of a second order surface in space (x,y,z). Unreduced form of the second order surface equation is reduced to a canonical form by classical manner. Based on the

analysis of the canonical equation type, it is possible to determine the type of surface that we will call a functional of the academic success.

2.3 Method for Studying Processes of Education Quality Management and Its Approbation

Based on the developed models of predicting levels of training quality, a methodology for conducting studies of process parameters that require division of system elements into classes is developed.

Approbation of the developed models, methods and data analysis was carried out on the example of data on the NRNU MEPhI students. For this purpose, a representative subsample of students belonging to the categories of "graduates" and "students expelled for academic failures" was taken.

Analysis of indicators of academic success confirmed that the best approximation to the median estimates among the indicators of academic success was a highlighted parameter, a truncated middle-terminal score B_j, which is determined for the entire cycle of student learning, after deduction of one the best and one the worst mark in each term j.

In accordance with the developed algorithms for constructing mathematical models, a cluster analysis of the selected parameter B_j was performed. The "nearest neighbour" method was implemented. An experimental study of the model of classifying levels of academic success for the highlighted parameter B_j was carried out. As a result of the cluster analysis of university graduates sample, out of 2446 points of baseline data on examination scores identified are 2 large clusters belonging to different levels of AcSuc: categories of students are defined as "highly successful" (32%) and "slightly successful" (55%). In addition, six small clusters containing 1-2 elements are specified. These 13% of students make up the category of "moderately successful" students.

Analysis of the developed mathematical models – regression models of MR (1) and with implementation of MLM – PM models (3) is conducted.

2.4 Regression Models of Determining the Level of Quality

As the analysis of regression models to the 4th degree inclusive has shown, conducted stepwise method of improving regression models does not give any significant improvement. It is natural to assume that for the higher degrees of polynomial it is possible to obtain a model, which will more accurately explain the share of the dependent variable variation. However, this increases the bulkiness of the model itself, increases the complexity of the regression model, which is determined by the number of predictors included in it, and increases the complexity of the calculations.

For the experimental sample obtained are MR – classification regression models *Qrg*s for different degrees s of the polynomial. Thus, for the degree s=1 the following model is defined: $Qrg1 = 1.1 - 1.6x + 1.3y - 0.03z$.

And for the fourth-degree polynomial (s=4) a regression model of the following form is developed:

$$Qrg\ 4 = -7390.1 + 20291.1x + 951.7y - 1899.0z - 20268.6xx - 2071.6xy + 2949.9xz + 602.3yy -$$
$$- 389.7yz + 369.1zz + 9314.8xxx - 1163.6yyy - 43.7zzz + 1687.3xyz + 2010.4xyy - 292.7xzz -$$
$$- 70.1yzz - 633.5xxy - 1951.5xxz - 930.4yyz - 1694.3xxxx + 231.9yyyy + 0.5zzzz + 520.1xxxz +$$
$$+ 704.5xxxy + 76.1yyyx + 174.6yyyz + 32.0zzzx - 12.2zzzy - 764.4xxyz +$$
$$+ 272.7yyxz + 61.6zzxy - 838.3xxyy + 24.0xxzz + 21.3yyzz$$

For regression models, their stepwise changing is carried out by the method of successive elimination of variables from the model, which criterion of p-value is the largest, and the p-value ≥ 0.05.

Finally, completed at a certain *k* step (stp*k*) regression models of the classification boundaries are obtained. Thus,

for the linear polynomial regression of the 4th degree, a calculated by 31 steps model looks as follows:

$Qrg\ \ 4\,(stp\ \ 31\,) = 1\,.2 - 1\,.5\,xy\ - 0\,.1\,xzz\ + 1\,.1\,yyyz$.

As a result, this mathematical model allows describing the relationship between the function Qrg and 34 independent predictors.

For the selected regression model $Qrg3$, based on the analysis of about 100 models, the pairs of error functions of the I O_1 and II O_2 types are experimentally determined.

Based on their analysis, it is found out that the developed model Q_{19} – is the model with the highest sensitivity (100%), which gives true result at presence of positive outcome (it reveals positive examples better than other models):

$$Q_{19} = -6.8x^2 + 3.6y^2 - 0.2z^2 + 5.6xy + 1.2xz + 2.5yz + 14.4x - 18.3y - 3.1z - 0.8 .$$

And the Q_{22} model – is the model with the highest specificity, which gives true result at the presence of negative outcome. This model identifies negative examples rather well:

$$Q_{22} = -12.4x^2 + 6.6y^2 - 0.5z^2 + 10.3xy + 2.3xz + 4.6yz + 26.3x - 33.5y - 5.8z - 2.3 .$$

2.5 Probabilistic Model for Determining the Level of Quality

Densities of text indicators distribution (x, y, z) for two samples are defined in view of belonging to different levels of quality. For this purpose, the statistical analysis and the method of principal component analysis are carried out. Identification of boundaries out of the condition of probability equality at the boundary separating two ellipsoid of data diffusion ("slightly successful" students and "unsuccessful" trainees) is conducted. Obtained is a polynomial of the form (3), and the unreduced form of the second order equation of surface is reduced to the canonical form.

As a result, equation (4) is developed, which determines a one-sheet hyperboloid of revolution shown in Figure 1.

$$QQ_{2\backslash 3} = 1.43 + H_1^2 - H_2^2 - H_3^2 , \quad (4)$$

where H_1, H_2, H_3 – are the linear combinations of text parameters (x, y, z) $H_i = k_i x + l_i y + m_i z + n_i$, a $k_i,\ l_i,\ m_i$ ($i = 1,\ 2, 3$) – coefficients obtained as a result of the second order surface reduction to a canonical form.

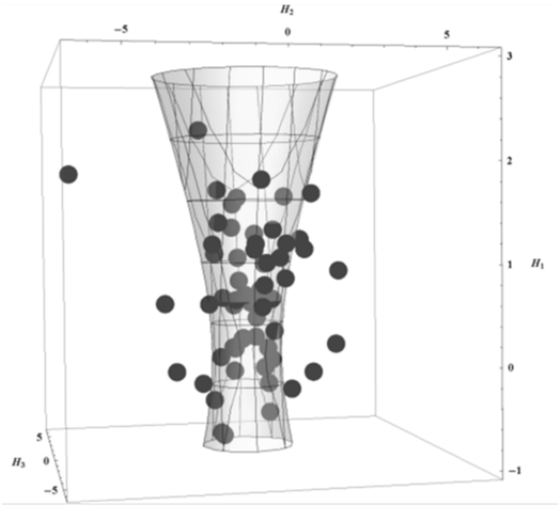


Figure 1. Classification boundary "Unsuccessful – Slightly Successful"

In such a way, a mathematical PM model of academic success at university training for "unsuccessful" and "slightly successful" graduates was constructed. At the boundary of class division, equality of probabilities of belonging to them is observed. For definiteness, it is assumed that the points at $QQ > 0$ are located inside the domain, and at $QQ < 0$ – they are located outside the domain.

Calculations and visualization of the study results were carried out with the help of the symbolic computation software system WolframMathematica (Website of the Wolfram Company. URL: http://wolfram.com/resources/).

3. Analysis of Mathematical Models

Due to the analysis of the developed mathematical models of predicting academic success levels, a PM model, which was more effective in predicting compared with a MR model, was selected.

To check the adequacy of the developed models, methods and techniques, a set of the best essays of potential applicants written by competitions winners in Russian literature and school medal winners was subjected to a screening analysis.

Check of the hypothesis about the type of distributions for transformational indicators of texts (x, y, z) of the "Russia Medal Winners" sample (fitting criterion χ^2 and Kolmogorov-Smirnov criterion) has shown that we cannot reject the hypothesis of normal distribution at a significance level of 10.

Further, assuming that the works could be written by medal winners-applicants at entrance exams to the NRNU MEPhI, with respect to a random sample of medal winners, an approbation of the developed models and techniques was carried out.

The probability of applicants belonging to one of the classes of academic success was determined: to the level of successful students or to the level of unsuccessful students (slightly successful and unsuccessful). Thereto, a PM model developed for these levels and function of academic success $QQ_{2\backslash 3}$ were used.

It is determined that some of these AcSuc functions of the Russia Medal Winners belong to the predictive class of "unsuccessful" students at training in the NRNU MEPhI. And only 75% of the Russia Medal Winners

obtained values of the function $QQ > 0$, which allowed predicting for them a successful completion of training in the NRNU MEPhI. The performed analysis of practical data showed that those were 83% of the applicants-medal winners. That is, we see that the predicting accuracy by this model is about 90%.

4. Results

Based on the implementation of principles of the national standard for quality management system in education, methods and data analysis tools for enhancing efficiency of the educational process are developed:

1. Based on consideration of the basic concepts of national education development and works of foreign researchers, relevance of construction and analysis of efficient ways expected to improve the quality of higher education and their relation to the problem of excellence is shown.
2. The study of the "trainee" subsystem's output parameters is carried out, and an original method of beam projection classification is developed, confirming existence of systemic connection between the selected output parameters and providing for determination of classification boundaries availability between the subsystem objects.
3. The quantitative indicators of academic success are developed. It is determined that the dynamic analysis of academic success shall be carried out based on the analysis of the truncated middle-terminal scores.
4. Original models of academic success classification are created. In particular, the models of predicting levels of academic success through regression and probabilistic (statistical) approaches are developed.
5. Technique for separating surface acquisition is developed. For the first time, the functional for probabilistic predicting of academic success levels in the form of a second-order equation of the surface – the classification boundary between different-level data is acquired.
6. The verification of developed models, methods and techniques for probabilistic predicting of academic success levels in a sample of the best essays written by competition winners in Russian literature and medal winners is conducted. The adequacy of the chosen model has been proven by example of "Russia Medal Winners" and a set of applicants-medal winners, the NRNU MEPhI students. The result predicted by the model coincided with the actual result by 87%.
7. Classification boundaries between the levels of training quality are constructed. The constructed models of predicting levels of academic success ("highly successful – slightly successful – unsuccessful") has shown that the results predicted by the developed mathematical models of PM for different levels of academic success coincided with the actual results by 70 to 90% of cases.

5. Discussion

At carrying out of the system study of the "trainee" subsystem's output data, there have been developed methods and data analysis tools for enhancing the efficiency of the quality management system processes in education.

The most significant factors affecting the determination of levels of academic success are singled out; a model that allows to determine the presence of different-class data domains is constructed; mathematical models of predicting levels of academic success as an indicator of the educational process quality are developed and investigated; a method of predicting the level of training program mastering by students is developed; the experimental test of the models and techniques developed is executed. Designed functional of determining the level of training quality, boundaries separating different-type categories of students will provide for predicting quality of the educational process of the main stages of multi-level education (Bachelor, Specialist, Master), which will be expressed in cost, time, and health savings.

For evaluation of examination marks provided by expert teachers, studied is their changing in the process of acquiring knowledge for the entire training period. Based on consideration of correlation between the individual results for all examinations, both oral and written, its existence is explained by availability of a psychological component.

Regression models that to the best advantage identify "high levels of quality" and "low levels of quality", but have a low interpretation degree of variables in the models are obtained. Aggregate analysis of the obtained regression models (Gurov, & Kuzminova, 2012) is carried out. A statistically significant relationship between the variables at a confidence level of 95.0% is identified. Statistic R-Squared value demonstrated that the models obtained cannot explain more than 25% of the variability in the MR models.

In contrast to the regression models, the mathematical model developed by application of maximum likelihood method of PM allows predicting these levels to 85% of cases.

In the course of further work there shall be given the answers to the following questions:

- What other separating surfaces can be formed?
- What is the specificity of data belonging to different levels besides their affiliation to a particular cluster of academic success?

6. Conclusion

Thus, we developed mathematical models of data analysis using different methods; evidenced the great practical value of created functional; constructed mathematical models of classification boundaries between the predictive levels of academic success through the application of maximum likelihood method. Implementation of the developed model of PM allows classifying levels of quality, predicting levels of training quality at a higher level compared to the other models considered, enhancing efficiency of quality management system processes. To increase the quality of the developed models, it is permissible to explore the additional parameters of texts or bring into consideration other parameters that affect the levels of training quality.

Acknowledgements

The authors highly appreciate creative abilities of the Professor, Doctor of Technical Sciences Popov, Y.A. from the "Computer Systems and Technologies" Department of the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), owing to whom this research work was carried out, and pay tribute to his blessed memory. Authors value highly professional support by the Candidate of Technical Sciences Assistant Professor Zhigirev, N.N., the chief leading researcher of the Science Research Institute of Economics and Management in Gas Industry in study of multidimensional data, a specialist in the field of applied mathematics.

References

Baranov, A. (2003). *Introduction into Applied Linguistics.* Moscow: Editorial URSS.

Bogomolov, A., et al. (2009). Forecasting of Trainees' Academic Achievements in Special Disciplines on the Basis of Regression Equations. *Herald of the Volga region, Humanities Series, Proceedings of the Higher Educational Institutions, 1*, 124-132.

Cheremnykh, S., Semenov, I., & Ruchkin, V. (2003). *Structural analysis of systems: IDEF-technology.* Moscow: Finance and Statistics.

Educational Testing Service ETS. (n.d.). Retrieved January 8, 2014, from http://www.ets.org/gre

Flesch, R. (1948). A New Readability Yardstick. *Applied Psychology, 3*, 221-233.

Giarratino, J., & Riley, G. (2007). *Expert Systems: Principles of Design and Programming.* Moscow: "Williams Publishing Company", LLC.

Gurov, V., & Kuzminova, A. (2012). Analysis of regression models of system processes indicators, which study requires division of its elements into classes. *Natural and Technical Sciences, 2*, 314-318.

International Organization for Standardization. (n.d.). *ISO Standards.* Retrieved January 8, 2014, from http://www.iso.org/iso/ru/iso_9000

Jackson, P. (2001). *Introduction to Expert Systems.* Moscow: "Williams" Publishing Company.

Kendall, M., & Stuart, A. (1973). *The advanced theory of statistics*. Moscow: Science.

Kim, J., et al. (1989). *Factor, Discriminant and Cluster Analysis.* Moscow: Finance and Statistics.

Kuzminova, A. (2013). Application of text analysis methods at levels of quality forecasting. *Quality. Innovation. Education, 12*, 27-29.

Luria, A. (2002). *Writing and Speaking. Neurolinguistic Researches.* Moscow: The Academy.

Marukhina, O., & Berestneva, O. (2003). Forecasting of the success of students training based on non-uniform sequential pattern recognition procedure. *Computer modelling: Proceedings of the IV International Scientific and Engineering Conference in St. Petersburg*. Retrieved January 8, 2014, from http//berestneva.am.tpu.ru/Papers/KONF2009/PITER-2003.doc

Muromtsev, D. (2005). *Introduction to the Technology of Expert Systems.* St. Petersburg: SPb ITMO University.

Popov, Y., et al. (2001). *Psycholinguistic technologies of personality assessment*. Moscow: PAIMS.

Recommendations on Standardization of the Russian Federation. R 50.1.028-2001 Continuous Acquisition and Life Cycle Support Technology. Methodology of functional modelling. (2002). Moscow. Retrieved January 8, 2014, from http://www.complexdoc.ru/ntdtext/541946

State Standard of the Russian Federation. GOST R 52614.2-2006 Quality Management Systems. Guidelines on Application of GOST R ISO 9001-2001 in Education. (2006). *Federal Agency on Technical Regulating and Metrology Website.* Retrieved January 8, 2014, from http://www.gost.ru/wps/portal/pages.CatalogOfStandarts

Statgraphics: The Statistical Program. (n.d.). Retrieved January 8, 2014, from http://statgraphics.com/downloads.htm

Sticht, T. (1973). Research towards the design, development and evaluation of a job-functional literacy training program for the US Army. *Literacy Discussion, 4*, 339-369.

The ISO/IEC/IEEE 42010. (n.d.). Retrieved January 8, 2014, from http://www.iso-architecture.org/ieee-1471

Website of the Wolfram Company. (n.d.). Retrieved January 8, 2014, from http://wolfram.com/resources

Modern Applied Science; Vol. 9, No. 8; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Reviewer Acknowledgements

Modern Applied Science wishes to acknowledge the following individuals for their assistance with peer review of manuscripts for this issue. Their help and contributions in maintaining the quality of the journal are greatly appreciated.

Modern Applied Science is recruiting reviewers for the journal. If you are interested in becoming a reviewer, we welcome you to join us. Please find the application form and details at http://www.ccsenet.org/reviewer and e-mail the completed application form to mas@ccsenet.org.

Reviewers for Volume 9, Number 8

Alexey Naidenov, Institute of Economics of the Ural Branch of the Russian Academy of Sciences, Russia

Anatoliy S. Semenov, Perm State Agricultural Academy, Russia

Anna Grana, University of Palermo, Italy

Awangku Hassanal Bahar Bin Pengiran Bagul, Universiti Malaysia Sabah, Malaysia

B.Bharathi, Sathyabama University,Chennai, India

H.Yegorov, Academic Research Institute, Russia

J. Eric Jensen, Harvard Medical School / McLean Hospital, Canada

Jardi Andaki, University of Sam Ratulangi, Indonesia

Juan José Villaverde, National Institute for Agricultural and Food Research and Technology, Spain

Levent Kurt, The City University of New York, USA

Malygina Natalia, Ural Federal University, Russia

Marek Brabec, Academy of Sciences of the Czech Republic, Czech Republic

Marek Brabec, Academy of Sciences of the Czech republic, Czech Republic

Mikhail Dudin, Russian academy of Entrepreneurship, Russia

Naumov I. V., Institute of Economics of the Ural Branch of the Russian Academy of Sciences, Russia

Olga Artemenko, Kaluga branch of Bauman Moscow State Technical University, Russia

Saiful Darman, University of Tadulako, Indonesia

Sergey Subachev, Emergencies and Elimination of Consequences of Natural Disasters, Russia

Tharek Abd. Rahman, Universiti Teknologi Malaysia, Malaysia

Yating Hu, Middle Tennessee State University, USA

Yonghong Meng, Shaanxi Normal University, China

Yury Subachev, Institute of Engineering Science, Russia

Call for Manuscripts

Modern Applied Science (MAS) is an international, double-blind peer-reviewed, open-access journal, published by the Canadian Center of Science and Education. It publishes original research, applied, and educational articles in all areas of applied science. It provides an academic platform for professionals and researchers to contribute innovative work in the field. The scopes of the journal include, but are not limited to, the following fields: agricultural and biological engineering, applied mathematics and statistics, applied physics and engineering, chemistry and materials sciences, civil engineering and architecture, computer and information sciences, energy, environmental science and engineering, mechanics. The journal is published in both print and online versions. The online version is free access and download.

We are seeking submissions for forthcoming issues. All manuscripts should be written in English. Manuscripts from 3000–8000 words in length are preferred. All manuscripts should be prepared in MS-Word format, and submitted online, or sent to: mas@ccsenet.org

Paper Selection and Publishing Process

a) Upon receipt of a submission, the editor sends an e-mail of confirmation to the submission's author within one to three working days. If you fail to receive this confirmation, your submission e-mail may have been missed.

b) Peer review. We use a double-blind system for peer review; both reviewers' and authors' identities remain anonymous. The paper will be reviewed by at least two experts: one editorial staff member and at least one external reviewer. The review process may take two to four weeks.

c) Notification of the result of review by e-mail.

d) If the submission is accepted, the authors revise paper and pay the publication fee.

e) After publication, the corresponding author will receive two hard copies of the journal, free of charge. If you want to keep more copies, please contact the editor before making an order.

f) A PDF version of the journal is available for download on the journal's website, free of charge.

Requirements and Copyrights

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the authorities responsible where the work was carried out, and that, if accepted, the article will not be published elsewhere in the same form, in English or in any other language, without the written consent of the publisher. The editors reserve the right to edit or otherwise alter all contributions, but authors will receive proofs for approval before publication.

Copyrights for articles are retained by the authors, with first publication rights granted to the journal. The journal/publisher is not responsible for subsequent uses of the work. It is the author's responsibility to bring an infringement action if so desired by the author.

More Information

E-mail: mas@ccsenet.org

Website: www.ccsenet.org/mas

Paper Submission Guide: www.ccsenet.org/submission

Recruitment for Reviewers: www.ccsenet.org/reviewer

The journal is peer-reviewed
The journal is open-access to the full text
The journal is included in:

CABI
Chemical Abstracts database
DOAJ
EBSCOhost
ERA
Google Scholar
LOCKSS
Open J-Gate
Polish Scholarly Bibliography (PBN)
ProQuest
Scopus
SHERPA/RoMEO
Standard Periodical Directory
Ulrich's
Universe Digital Library

Modern Applied Science
Monthly

Publisher Canadian Center of Science and Education
Address 1120 Finch Avenue West, Suite 701-309, Toronto, ON., M3J 3H7, Canada
Telephone 1-416-642-2606
Fax 1-416-642-2608
E-mail mas@ccsenet.org
Website www.ccsenet.org/mas

ISSN 1913-1844