THE FEASIBILITY OF USING LOW COST MODELLING TECHNIQUES TO RELATE RIVER WATER QUALITY AND DIFFUSE LOADS TO A RANGE OF LAND USES

CE Herold • PH van Eeden

WRC Report No 796/1/01

Water Research Commission

THE FEASIBILITY OF USING LOW COST MODELLING TECHNIQUES TO RELATE RIVER WATER QUALITY AND DIFFUSE LOADS TO A RANGE OF LAND USES

FINAL REPORT

to the

WATER RESEARCH COMMISSION

By

C E Herold and P H van Eeden

STEWART SCOTT (Consulting Engineers)
PO Box 784506, Sandton, 2146

WRC Report No. **796/1/01**
ISBN No **1 86845 753 2**

EXECUTIVE SUMMARY

AIM

The objectives of the research programme were as follows:

1. Develop and test a low cost methodology for estimating river flows and pollutant loads at ungauged sites.

2. Evaluate the potential for a low cost, low technology, non-site specific methodology that can be used to relate in-stream water quality to type and intensity of land use.

3. Determine if the proposed methodology warrants further research and development.

MOTIVATION

Non-point source pollution has been identified as a priority research area. There is limited information available locally to support non-point source quantification for management purposes. One of the major obstacles in increasing the database is the high cost associated with conventional methods of monitoring diffuse source loads.

The conventional rigorous approach calls for extremely expensive continuous flow gauging in conjunction with high frequency water quality sampling. A high level of technical competence is also required to maintain equipment and process the large volumes of data. This has tended to limit research to a few small site specific study catchments. Since small study catchments are typically dominated by relatively unique local land use characteristics, it is difficult to apply the research results to wider catchment areas. Construction of the large number of monitoring stations that would be required to derive reliable relationships between land use and river water quality will in all probability never be realised in time to be of use in predicting the effects of rapid urbanisation. Moreover, the constraints imposed by the difficulty in locating suitable flow gauging sites often prevent the optimal selection of sub-catchments (which should be land use driven).

This raises the need to develop a low cost, low technology methodology that can be used to estimate diffuse source loads and develop useful relationships between in-stream river quality and land-use. The research project is aimed at meeting this need.

APPROACH

The study is to be addressed in two phases, the first of which is encapsulated in this study and comprises a preliminary investigation. The second phase would involve further research to prove and develop the methodology. Execution of the second phase will depend on the outcome of the preliminary investigation.

Phase 1 was confined to developing the methodology, with only limited water quality sampling and testing. The second phase will address the development of a database, additional sampling and analysis and the development of relationships for a range of type and intensity of land use.

TASKS

The specific study tasks include:

- Identify land uses and sampling points
- Water quality sampling
- Calibrate hydrological model
- Develop flow / water quality relationships
- Evaluate methodology
- Document research findings

SELECTION OF SAMPLING POINTS

Budget constraints limited the number of samples that could be analysed. Sample points were chosen below areas with relatively uniform land use development.

Data was collected at six sites in the Rietspruit catchment. Predominant land uses include:

- low density residential / CBD
- high density residential (two sites)
- industrial
- rural (dry land cultivation)
- undeveloped.

WATER QUALITY SAMPLING

Ten samples were taken at each site over a six-month period from March to September 1997. Some of the sample dates were planned to coincide with wet weather conditions, thereby ensuring that the limited number of samples covered a reasonable range of flow conditions. The key water quality variables that were sampled included pH, electrical conductivity, sulphate, nitrate, ammonia, total phosphate, faecal coliform, dissolved oxygen and temperature. The later two variables were used to calculate the percentage oxygen saturation. Very crude estimates of flow were made at each site. However, most of these estimates are considered to be highly inaccurate and were not used in further processing. This had no detrimental effect, since the methodology is based on the assumption that definitive flow data is not available at the monitoring stations.

The small database prevented an assessment of the effect of sampling frequency and sample size on the model results. It also prevented partitioning the sample to test the effect of deriving separate relationships for wet and dry flow ranges.

HYDROLOGICAL MODELLING

The daily time step NACL model was used to simulate flows at each monitoring site. A correction was applied to the data to make the modelled flows coincide with those observed at the catchment outlet (RW station R6).

As expected, spatial and temporal variations in rainfall prevented the model from accurately replicating the flows that occurred on each day on which water quality sampling took place. The actual accuracy of flow estimation at the sampling points could not be assessed, since only very inaccurate flow observations could be made.

WATER QUALITY RELATIONSHIPS

Both natural and power regressions between modelled flow and concentration were derived for key parameters at each site. The water quality variables for which these regressions were derived included:

- electrical conductivity
- sulphate
- nitrate

- ammonia
- total phosphate
- faecal coliform
- percentage oxygen saturation.

The NACL model was used to simulate a large number of daily flows for the period October 1965 to September 1997. This period covers a wide range of hydrological conditions, including severe droughts and extreme floods.

ESTIMATION OF RANGE OF CONCENTRATIONS

The most appropriate regression equation for each water quality variable was used to estimate a regressed concentration corresponding to each simulated daily flow. Normalised random noise was then added to the regressed values to allow for the considerable variation that was attached to the observed data. Provision was made for specifying an applicable range of flow conditions (in this case close to the flow range of the observed data). Allowance was also made to filter out unreasonably high or low concentrations.

The range of concentrations thus generated for each water quality variable and station was also presented in the form of a duration curve. The results were tabulated to show key statistical properties, including mean, standard deviation, regression constants, standard error, correlation coefficient and selected percentile values (98%, 95%, 90% high values plus the median).

CONCLUSIONS

Application of methodology

The strength of the methodology was demonstrated by means of an example of the impact of an assumed new high density urban development. The potential for using the methodology in conjunction with GIS based land use data was also discussed.

Advantage of methodology

The use of a regression equation results in severe damping of the generated concentrations of non-conservative pollutants. The use of this methodology (or mean values) inevitably leads to gross understatement of the true range of concentrations that arise. The methodology put forward in this report holds the advantage of taking account of the semi-stochastic manner in which concentrations vary about the regressed values.

Use of modelled flow data (suitably corrected against observations at a downstream reference gauge) has the advantage of freeing the practitioner of the need to confine water quality sampling to sites were flows can be gauged. Application of this methodology means that monitoring sites can now be chosen close to source areas, thereby monitoring relatively homogeneously developed areas. The main requirement is that sufficient samples should be gathered to cover a range of flow conditions. It has been demonstrated that by varying sampling dates according to weather conditions it is possible to monitor a wide range of flows with relatively few samples.

Partitioning of data

The sample size was too small to permit partitioning of the data into low and high flow conditions. Nor was it possible to investigate the impact of sample size and sampling frequency on the key statistical properties (such as mean, standard deviation and standard error). There is merit in investigating both of these effects.

Limitations of methodology

The following factors are limitations on the methodology and the extent to which it could be tested in the study:

- Relatively infrequent and short sampling records can miss first flush events at the onset of rain.

- The effect of temporal and spatial variation in rainfall on the accuracy of flow estimation at upstream sites needs to be evaluated, along with its effect on the required water quality sampling period.

- In some instances the results can be sensitive to the choice of regression equation (section 5), particularly when the sample size is small.

- A linear regression often gives a superior representation to an exponential regression. However, applying random noise to a linear regression can result in negative estimated concentrations. Correction by eliminating such values has the effect of increasing the average of the generated values. This implies the need to investigate other types of regression equation.

Preliminary results

A methodology has been developed successfully and the results of its application are promising. However, the small water quality database prevented assessment of the unknowns listed above. Hence further investigations are necessary to determine the value of the method, before it can be regarded as a final or accepted technology.

RECOMMENDATIONS

- **Determination of required sample sizes**

  It is desirable to test the methodology that has been derived against fuller and longer data sets to determine the number of samples required to achieve a reasonable representation of the key statistical parameters. The effect of partitioning the data into low and high flow ranges should also be investigated.

- **Assessment of required sampling frequency**

  The effect of sampling frequency on the results should be investigated, including evaluation of the effect of the first flush at the onset of rainfall events.

- **Investigation of flow modelling time step**

  The use of daily flow models is constrained by the effort and time required to calibrate the models. However, monthly flow modelling is widespread in southern Africa. The potential exists to apply the same methodology to coarser monthly time step flow modelling data. The overall benefit of being able to use monthly flow modelling would be very far reaching.

- **Evaluation of flow modelling accuracy**

  The impact of spatial and temporal rainfall distribution on the accuracy of simulated flows at sites upstream of the reference gauge needs to be investigated.

- **Investigate use of other statistical distributions**

The current study used only linear and power regression curves. Further investigation of other types of regression is recommended. In particular, alternative methods of handling the addition of normalised random noise to the regression line should be investigated. This is because this can result in concentrations that lie outside the plausible range and therefore have to be filtered out, thereby affecting the statistical properties.

- **Development of national database**

The results thus far achieved indicate that there is much merit in developing a national data base of key statistical parameters associated with different land uses. This proposal is linked to the application of the methodology discussed in this report. However, even without the new technology, there is a pressing need to consolidate data on the effect of land use on water quality.

- **Minimisation of flow gauging requirements**

The methodology described in this report can be used effectively to reduce flow gauging requirements in water quality studies. This would enable water quality monitoring sites to be selected primarily on the basis of upstream land use and access, rather than on the suitability of the site for flow gauging. The techniques that have been developed can also be used to maximise the use of data that has already been accumulated at a number of sites that are not suitable for flow gauging.

- **Testing of scenarios**

A serious weakness of most water quality studies that have been carried out to date lies in the tendency to rely on flow/concentration regressions to explain catchment export. This inevitably leads to severe damping of the resulting scenarios of water quality concentrations. This points to the need to take better account of the variability of observed concentrations about the mean or regression line. This is especially important with regard to non-conservative constituents, which can typically show a very wide variation. The methodology that has been presented provides a means of taking this variation into account. Hence it has great potential to be applied in water quality studies to test scenarios (such as assessment of the impact of planned management options or scenarios of projected land use development).

However, use of the new methodology in this manner should not proceed until some of the key outstanding issues have been addressed in follow up studies.

- **Use of random number processing in other applications**

The methodology that has been developed for dealing with the variation in observed concentrations has great potential for application in a simplified, easy to calibrate river routing model. A simple decay model was developed and used successfully for the Environmental Impact Assessment for ERWAT's proposed new Welgedacht water care works. Proposals have been made for extending this simple but robust technology to include the methodology discussed in this report.

- **Evaluation of dependence on hydrological model**

It is desirable to test the dependence of the results on the choice of hydrological model.

ACKNOWLEDGEMENTS

The research in this report emanated from a project funded by the Water Research Commission and is entitled:

"The feasibility of using low cost modelling techniques to relate river water quality and diffuse loads to a range of land uses"

The Steering Committee responsible for this project consisted of the following persons:

Mr H M du Plessis	Water Research Commission (Chairman)
Ms U Wium	Water Research Commission (Committee Secretary)
Mr K C Pietersen	Water Research Commission
Ms M Hinsch	Department of Water Affairs
Ms T Belcher	Department of Water Affairs
Mr M Keet	Department of Water Affairs
Prof AHM Görgens	Stellenbosch University
Dr P H van Eeden	ERWAT: Laboratory Services
Dr C E Herold	Stewart Scott (Project Leader)

The financing of the project by the Water Research Commission and the contribution of the members of the Steering Committee is acknowledged gratefully.

This project was only possible with the co-operation of many individuals and institutions. The authors therefore wish to record their sincere thanks to the following:

Rand Water, in particular Mr R Heath for the provision of stream flow data.

ERWAT, who made available the services of Dr P van Eeden for field data collection.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Aim

The objectives of the research programme were as follows:

- Develop and test a low cost methodology for estimating river flows and pollutant loads at ungauged sites.

- Evaluate the potential for a low cost, low technology, non-site specific methodology that can be used to relate in-stream water quality to type and intensity of land use.

- Determine if the proposed methodology warrants further research and development.

1.2 Motivation

Most catchments in South Africa are dominated by surface washoff and therefore have a high potential for generating non-point (diffuse) source pollution. Consequently, the Water Research Commission has identified non-point source pollution as a priority research area. One of the recent projects supported by the Water Research Commission is a study conducted by Sigma Beta / IWQS (WRC Project No. K5/696, 1997). In Phase I of this study (a situation assessment of the current state of knowledge of non-point source problems in South Africa) it was concluded that much of the work has been research orientated and that there is limited data to support non-point source quantification for management purposes. Also, previous studies have often been site specific and/or limited to a small area. Some of the recommendations of the Sigma Beta/IWQS study were:

- Information (at an appropriate level of detail) must be presented for the purposes of catchment management.

- The cost-effectiveness of practices, including new techniques to address non-point source production, must be estimated.

One of the major obstacles in addressing the above needs is the high cost associated with conventional methods of monitoring diffuse source loads from different land uses. The conventional rigorous approach calls for continuous flow gauging in conjunction with high frequency water quality sampling, both of which are usually extremely expensive and require a high level of technical competence to maintain equipment and process the large volumes of data. The requisite field instrumentation is also vulnerable to vandalism. These factors have tended to limit research to a few small site specific study catchments. Since small study catchments are typically dominated by relatively unique local land use characteristics, it is difficult to apply the research results to wider catchment areas. A large number of monitoring stations would be required to derive reliable relationships between land use and river water quality. Economic realities mean that construction of these works will in all probability never be realised in time to be of use in predicting the effects of the rapid urbanisation associated with the Reconstruction Development Plan (RDP). Moreover, the constraints imposed by the difficulty in locating suitable flow gauging sites often prevent the optimal selection of sub-catchments (which should be land use driven).

This raises the need to develop a low cost, low technology methodology that can be used to estimate diffuse source loads and develop useful relationships between in-stream river quality and land-use. The research project reported on in this document is aimed at meeting this need.

1.3 Simplified approach

At the heart of the simplified approach is a methodology that combines low cost flow estimation at an appropriate level of precision with judicious water quality sampling aimed at estimating diffuse source loads from different land uses. These can then be used as input to a suitable model that can be used to simulate the routing of the pollutant loads derived from both point and diffuse sources, in-stream decay processes and ultimately the impact of expected development and intended management.

At a later stage it may be possible to find a means of simplifying the routing process. At first it was thought that this could be achieved by means of compiling scalable concentration frequency distribution curves and other relevant statistical properties for each pollutant that can be regarded as typical for the land uses under investigation. However, in the interim the research team has become convinced that a much better approach would be to simplify the modelling of the decay and routing process. Promising and practical proposals in this regard have been put forward to the WRC and the DWAF.

Irrespective of the approach (simple or complex) later adopted for simulating the routing and in-stream decay process, a simple but robust method of estimating the diffuse load export from developed areas is required. The results of the study could form the start of a database that can supply information on non-point source quantification at an appropriate level of detail for catchment management purposes. This database could then be used in conjunction with GIS information on land use to estimate the diffuse export from different portions of catchment of concern.

1.4 Basis for simplified approach

The following two factors have a critical effect on the cost-effectiveness of the methodology adopted for assessing the impact of diffuse source washoff from land use development on water quality:

- the appropriate level of precision in flow gauging; and
- the appropriate duration and frequency of water quality sampling.

1.4.1 Appropriate level of precision in flow gauging

Achievement of an appropriate level of precision in the gauging of flow rates is especially pertinent in the case of non-conservative pollutants (which are of most concern with regard to the rapid expansion of low cost housing projects associated with the RDP). Such pollutants are subject to high in-stream decay rates and do not always correlate well to flow rate. For example, faecal coliform counts can vary by orders of magnitude from one sample to another, even for similar flow conditions (due sometimes to local environmental changes that affect the decay rate, or to inaccuracies in sample analyses). Such effects obviously greatly reduce the accuracy with which the instantaneous load can be estimated *even with perfect flow gauging*. Moreover, natural decay in the delivery system between the land use in question and the point of monitoring further reduces the accuracy with which non-point source loads can be estimated. This means that in the case of non-conservative pollutants the conventional approach of providing expensive structures to make precise measurements will not necessarily result in accurate diffuse source load estimation.

The approach adopted in this study is based on the use of a daily time step rainfall-runoff model to simulate the runoff from a number of selected key sub-catchments reflecting a range of type and intensity of catchment development. The observed runoff record at the established downstream Rand Water flow gauging station at R6 was used to adjust the simulated runoff from the entire catchment and from each of the smaller upstream sub-catchments, thereby greatly improving the estimation of the daily flow. This was used to improve the accuracy of flow estimation at each of the upstream sampling points, without having to incur the cost of expensive new gauging structures and instrumentation. It also permitted more appropriate selection of sub-catchments based solely on land use, rather than on the constraints imposed by the availability of suitable flow gauging sites.

The spatial and temporal variation in rainfall still affects the accuracy with which river flows were estimated at each upstream monitoring point. However, the correction procedure using the downstream weir record should serve to preserve the general trend in the relative magnitude of the flow at the time when each sample is taken. While the accuracy with which the load can be estimated at the time of taking any individual sample will still be limited, the general relationship between flow rate and water quality should be reasonably preserved. The reliability of the relationships derived between flow and pollutant load would steadily improve as the number of samples that are collected increases. This should not significantly increase the water quality sampling requirements, since the semi-stochastic high variability of non-conservative pollutant concentrations for ostensibly similar flow conditions also requires a reasonably large number of samples to be taken before reliable load estimates can be made.

Any loss in accuracy should be more than compensated for by the ability to estimate the runoff from several sub-catchments at an appropriate level of accuracy, all of which can be selected according to land use criteria, rather than the availability of flow gauging sites. This opens the possibility of choosing water quality monitoring sites in close proximity to upstream catchment developments of interest.

1.4.2 Appropriate duration and frequency of water quality sampling

In addition to accurate flow gauging, the conventional approach to studying non-point source pollution usually calls for intensive high frequency sampling, yielding a large amount of data from which accurate estimates of pollutant loads can be derived. However, intensive high frequency sampling can prove to be expensive. As a result, budget constraints usually allow only a few events to be monitored. Since the magnitude and timing of each hydrological event is highly variable and can result in wide variations in water quality response, the results are usually difficult to extrapolate to other unseen hydrological events in the study catchment itself, let alone to other catchments. Moreover, equipment failure, human error, exceedence of the rated capacity of the flow gauging structure or overloading of laboratory facilities often lead to gaps in the record that prevent accurate load estimation despite the sophistication of the technology employed.

Due consideration needs to be given to how the data will be used in practice before determining the appropriate duration and frequency of water quality sampling. For conservative pollutants, such as salts, a time series of pollutant loads can be used as input to load based dynamic models to estimate the change in water quality likely to arise from a given option for a wide range of hydrological conditions. However, in the case of non-conservative pollutants, the complexities of decay processes have generally confined model development and application to steady state models (such as the EPA supported QUAL2E model) that are concentration, rather than load driven. Accurate load estimation is therefore of less importance than the determination of concentrations for specific critical (usually low flow, or small runoff event) conditions.

In view of the above considerations, it is difficult to justify high frequency sampling to estimate non-conservative pollutant diffuse loads (the sampling and analysis of which have a wide band of uncertainty). Instead, sustained low frequency (weekly or fortnightly) sampling could be aimed at covering a range of hydrological conditions. Sampling dates could also be adjusted in response to weather conditions to ensure that sufficient samples are obtained corresponding to high flow events. This will provide the basis for building up relationships between in-stream water quality and river flow and seasonal factors, which can in turn be used to estimate diffuse source loads from the full flow record.

1.4.3 Characterisation of in-stream water quality

Water quality management studies are aimed primarily at achieving receiving water quality objectives. In this regard the pollutant concentration duration curve (i.e. the percentage time that a given pollutant concentration can be expected to be exceeded) is of most concern. Water and pollutant mass balance simulation modelling has been used with some success to predict the effect of upstream catchment development on salinity. However, this approach is much more difficult to apply in the case of non-conservative pollutants where decay processes play an important role.

The second objective of this study was aimed at developing direct relationships between non-conservative pollutant concentration frequency distribution curves and the type and intensity of land use. The intention was that as the data base is built up over time from this and other studies, suitable factors could be developed for scaling the duration curve characteristics. Scaling can be according to factors such as the density of land use development, type and length of river reach between the land use and the point of interest, and the relative contribution to the runoff from the developed area. Suitable relationships could then be sought for other relevant statistical properties, such as the mean concentration and the standard deviation.

However, it soon became apparent that there are a large number of possible combinations of upstream land use, intensity of development and distance downstream at which the samples are taken. Against this, the Phase 1 study obtained data for only six monitoring points. It was considered imprudent to attempt to establish such complex inter-relationships from so small a sample.

Further crystallisation of the ideas promoted the development of what is considered a much more practical method of estimating downstream impacts. This calls for the generation of a range of possible diffuse source washoff scenarios for each land use. A simple river routing and decay model has then been proposed as a vehicle for simulating a range of downstream pollutant concentrations associated with both the range of hydrological events and the management option under consideration.

1.4.4 Data base development

The results of the second phase of the study will form the start of a database that can supply information on non-point source quantification at an appropriate level of detail for catchment management purposes. The data base, as well as the land-use classification and the monitored river water quality at various points along the river would be entered in a GIS, thereby enabling evaluation of the possible effect of changing land-use on river water quality.

The information stored in the data base would be confirmed as experience is gained from longer monitoring records and subsequent studies on other catchments. In time the database can also be extended to include coverage of a wider range of water quality variables and land uses.

1.5 Phasing of project

The current Phase 1 study has been confined to developing the methodology, with only limited water quality sampling and testing. The second phase will address the development of the database, additional sampling and analysis and the development of relationships for a range of type and intensity of land use. But it is important to ensure that the right type of data is stored in the database. Typically literature sources site some estimate of the total annual export from different land uses. However, these estimates also carry an underlying range of estimation uncertainty, which is usually not given.

Subject to approval of the positive findings of this preliminary investigation, the WRC will decide on proceeding with the second phase of the study.

1.6 Expected benefits

The current study has put forward a novel cost saving approach that bypasses the need to monitor flow rate at every sampling point. But at the same time the probabilistic approach provides a means of taking due account of the variability of the observed data.

Initiation of the proposed database will provide an important longer term benefit to water quality practitioners by giving them a means of unravelling the data that is accumulated from various studies of land use impacts.

Finally, by linking river water quality to land-use using GIS techniques, a what-if situation for a particular area can be assessed, i.e. the effect of a change in land-use on river quality can be evaluated. This aspect should prove to be particularly valuable for planning purposes.

1.7 Tasks

The specific study tasks of the Phase 1 study are discussed briefly in the following sections.

Task 1 : Identify land uses and sampling points

This task involved examination of land uses within the Rietspruit / Natalspruit catchment and the identification of sub-catchments characterised by relatively homogeneous land use. Land use information was obtained from Local Authorities and various other sources. Appropriate water quality sampling points were then selected, close to the affected areas.

Task 2: Water quality sampling

A low cost water quality sampling programme was initiated, based on infrequent grab samples. This was aimed at monitoring the runoff from a range of different land-uses and establishing flow / water quality relationships. Steps were taken to obtain coverage of a range of both dry and wet weather flow conditions. Sampling dates during the wet months were adjusted to ensure that a reasonable coverage of higher flow events were monitored by the limited number of samples that were collected. Despite this, the small size of the experimental sampling regime would most probably have missed the first flush at the onset of rain events. A few key indicator water quality variables were selected. These included electrical conductivity, sulphate, nitrate, ammonia, total phosphate, faecal coliform and dissolved oxygen. Temperature was also measured to permit the calculation of the oxygen deficit.

Task 3: Calibrate hydrological model

The NACL daily time step model was used to simulate the runoff from the entire Rietspruit catchment upstream of Rand Water's flow gauging weir R6. The model layout was disaggregated into sub-catchments corresponding to the sub-catchments selected in Task 1. Due account was taken of known point source wastewater discharges.

Task 4: Develop flow / water quality relationships

The hydrological model was used to simulate river flows at all points of interest in the study catchment for the period during which water quality monitoring took place. The observed flows at station R6 were compared with those simulated to derive correction factors, which were used to adjust the simulated flows throughout the system. The adjusted flows were in turn used to estimate the pollutant loads at each point at the time of sampling. Suitable regression relationships between adjusted flow and sample water quality were derived. These relationships were used in conjunction with a longer simulated flow record to calculate daily loads for a wide range of hydrological conditions. These were next used to estimate the annual loads at each monitoring point.

Task 5: Evaluate methodology

The methodology developed for estimating river flows and diffuse source pollutant loads and for relating in-stream water quality to type and intensity of land use has been developed and tested on the available sample data. The procedures have been shown to be feasible, and recommendations made on how they can be further improved.

2. IDENTIFICATION OF LAND USES AND SAMPLING POINTS

2.1 Study catchment

The Rietspruit catchment was selected for the study area. This choice was governed by the following factors:

- The catchment is highly developed, with a number of significant land uses, including high and low density urban, industrial, mining and agricultural development, as well as an undeveloped nature conservation area.

- A previous study of the Klip River System *(Stewart Scott, 1996)* showed that this area contributes significantly to the pollution load in the Klip River. However, the pollution load in the Natalspruit/Rietspruit catchment area has never been quantified.

- A flow gauging station (R6) is located at the outlet of the catchment.

- Water quality is of great interest in this catchment.

- Physical and other constraints limit the availability of flow gauging sites within the catchment.

- ERWAT, who were to do the water quality sampling and analyses, are active in the catchment.

A map of the study area is given in Figure 2.1.

2.2 Land uses

Land use information was obtained from available GIS mapping. Figure 2.1 shows the main land uses within the study area. The Rietspruit and Natalspruit catchments were examined to identify smaller sub-catchments dominated by relatively homogeneous land use. Within this constraint, an attempt was also made to locate catchments that were large enough to contribute significant runoff down a defined watercourse with reasonable road access.

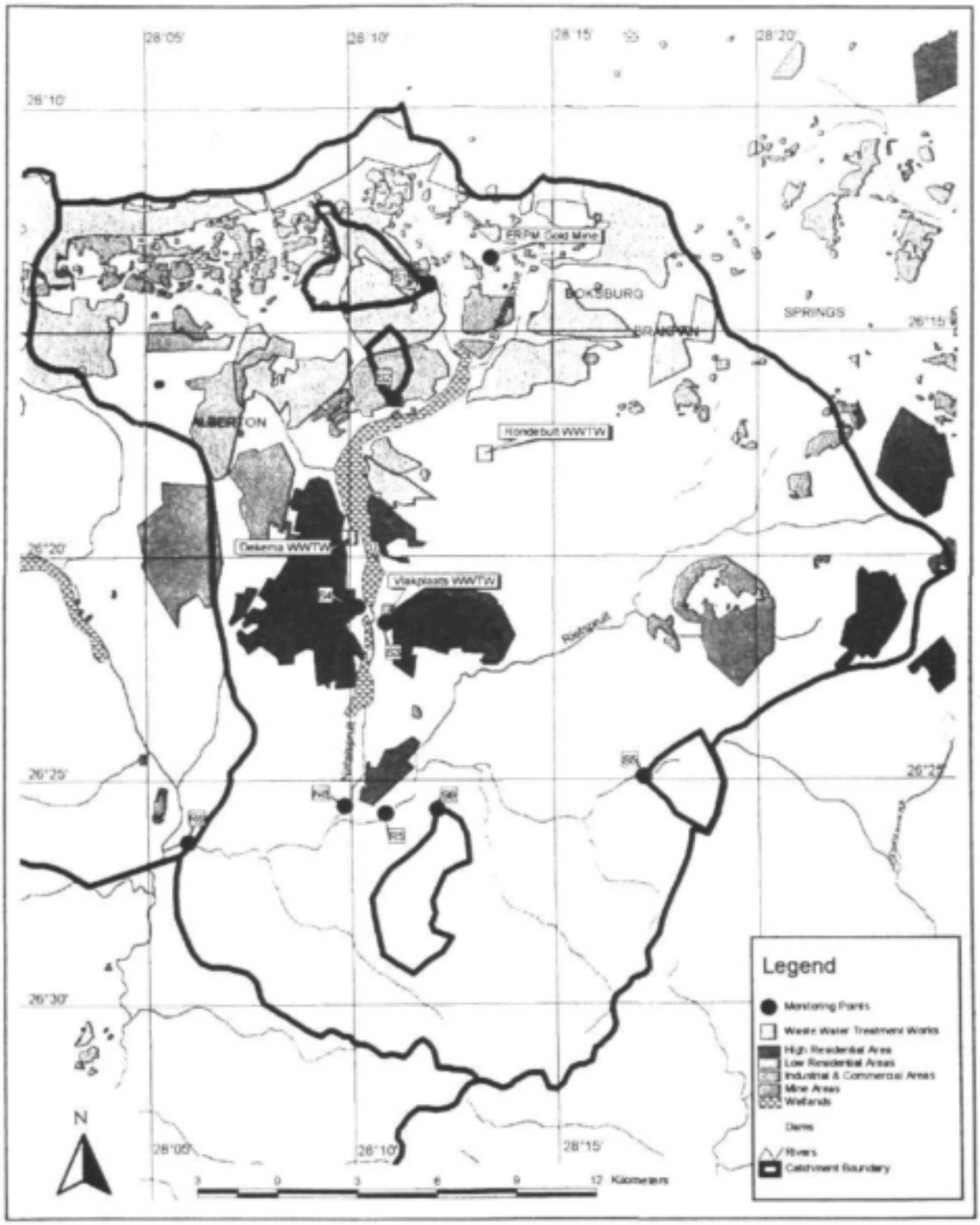


Figure 2.1 : Map of study area

A preliminary identification of potential sites was based on available mapping and the experience of the project team. Thereafter, sites were visited and samples taken at promising sites. An attempt was also made to estimate the flow at each site and arrangements made for similar crude estimates to be made when future samples were collected by ERWAT.

It was recognised that the method of flow estimation was highly inaccurate. The flow estimates are therefore considered unreliable. This is especially true of low flow conditions when flow velocities were very low and effective section widths and depths were hard to estimate. Alterations to river courses during the study period also prevented flow estimation at some sites.

The inability to accurately gauge flows was anticipated from the outset. The methodology under investigation is not dependent on accurate flow gauging at the sampling points. The methodology is aimed at overcoming this data deficit.

2.3 Selection of sampling points

Following analysis of the preliminary samples and the examination of the sites, six sampling sites were chosen. The main selection criterion was the coverage of a range of land uses. These included:

- Largely undeveloped (dominated by the Suikerbosrand Nature reserve)
- Cultivated farmland
- Mixed low density residential and central business district
- Industrial
- Formal high density residential
- Informal high density residential.

A range of climatic and geographical regions would also have been desirable, but this was not possible given the low cost nature of the study.

The chosen monitoring stations are shown in Figure 2.1. Table 2.1 shows the main characteristics of the monitoring points and their upstream catchments.

Table 2.1 : Description of monitoring sites

Description	Monitoring station					
	S1	S2	S3	S4	S5	S6
	12.74	3.30	1.89	2.37	13.98	8.30
Low density residential (km^2)	5.70	0.90	0.00	0.00	0.00	0.00
High density residential (km^2)	1.43	0.00	1.61	2.25	0.00	0.00
Central Business District (km^2)	5.00	0.00	0.00	0.00	0.00	0.00
Industrial (km^2)	0.00	2.40	0.00	0.00	0.00	0.00
Dry land farming (km^2)	0.00	0.00	0.00	0.00	13.98	0.00
Victoria lake (km^2)	0.61	0.00	0.00	0.00	0.00	0.00
Undeveloped (km^2)	0.00	0.00	0.28	0.12	0.00	8.30
Impervious area (km^2)	3.04	1.04	0.42	0.58	0.00	0.00

The areas given in Table 2.1 were obtained from GIS overlays of different land uses. The impervious (paved areas) were calculated from land use data combined with the percentages derived from DWAF report PC000/00/16396 *(Stewart Scott, 1999)*. The percentage of land estimated as being paved for the propose of the Vaal River System Analysis Update study were as follows:

Low density residential :	12.9%
High density residential :	25.8%
Industrial / Commercial :	38.7%

Brief descriptions of the six monitoring stations are given below.

2.3.1 S1: Jan Frederik Street, Germiston

Co-ordinates: (26°13'58" S, 28°11'55" E)

This monitoring station is located in a stream in parkland below the greater part of the Germiston central business district (CBD), just upstream of its confluence with the Elsburgspruit. Low-density residential areas dominate the lower portion of the catchment, with the CBD located in the upper third of the catchment. The northeastern fringe of the catchment may just include a small portion of some mining areas, although this boundary is difficult to distinguish. An industrial area called "South Germiston (Industrial East)" appears on the northern edge of the catchment. Victoria Lake is located at the southwestern extremity of the catchment.

Rain occurred on the day when the sites were visited (10 March 1997). The site comprises a natural open channel. When the site was inspected the flow was super critical, with a surface velocity of 2.5 m/s, a flow depth of about 0.20 m and a top width of 1.6 m. The water was observed to be clear and free of reeds. The channel shape was relatively regular and simple methods could yield a rough estimate of the flow.

2.3.2 S2: Wadeville, corner Lamp and Barracuda Streets

Co-ordinates: (26°16'31" S, 28°11'06" E)

This monitoring station is located in an open channel downstream of Lamp Street. The stream drains the industrial area of Wadeville. The far northern portion of the catchment includes the low-density residential areas at the southern end of Germiston. On 10 March 1997 the flow was reasonably strong, with a surface velocity of about 0.35 m/s, an average depth of 0.65 m and surface width of 3.1 m. The water was muddy. Reliable flow estimation at this point would require current metering, and even then the channel shape was not regular. Hence flow velocity and depth would vary across the section.

2.3.3 S3: Channel below Vosloorus

Co-ordinates: (26°21'26" S, 28°10'46" E)

S3 is located in a shallow open concrete drain below Vosloorus, located on the property of ERWAT's Vlakplaats Waste Water Treatment Works (WWTW). The bottom of the canal was covered by algae and sewage fungus. There was also some foam, possibly due to washing activities in Vosloorus. The shallow flow depth (80 mm at its deepest) and the growth in the channel (which caused an irregular flow path) militate against reliable flow estimation. On 10 March 1997 the surface velocity was estimated at about 0.54 m/s and the surface flow width at 1.6 m. Flow estimates at this point are considered to be very inaccurate.

2.3.4 S4: Eastern Katlehong

Co-ordinates: (26°21'03" S, 28°10'07" E)

This monitoring point is located in a drainage stream near the eastern edge of Katlehong, on the western side of the Natalspruit opposite the northern edge of ERWAT's Vlakplaats WWTW property. The original site was at a gabion weir, which was later disrupted by diversion works associated with road construction. The site is located within an area dominated by high-density informal settlement. The upstream catchment comprises high-density residential areas.

Although most of the flow was passing over the gabions (which could be treated as a pseudo broad-crested weir), there also appeared to be some leakage. The flow depth over the gabions was shallow (40 to 45 mm) and irregular. The entire weir crest was not accessible. Reeds also obstructed the approach to the gabion weir. Hence accurate assessment of the flow depth was not possible. Flow estimation ceased entirely after construction on the new road commenced.

2.3.5 S5: Kliprivier / Heidelberg road R550

Co-ordinates: (26°25'38" S, 28°12'13" E)

Station S5 is located at a road bridge. Although the bridge has 4 spans of 2.4 m width, flow was passing through only two spans. The stream flows northwards under the road. The northern portion of the stream drains part of the Suikerbosrand Nature Reserve, after which it flows through a farmed area. The catchment was assessed as being largely unimpacted.

On 10 March 1997 the water was observed to be clear, with lots of Bullrush (*Typha Capensa*) growth and some macrophytes. The surface flow velocity was very low (only 0.06 m/s) and consequently very difficult to measure. The irregular channel shape also created difficulties in measuring both the depth and width of flow. This difficulty was compounded by the low velocity, which prevented a reasonable estimate of the effective flow width and depth. Flow estimation is therefore considered to be completely unreliable.

2.3.6 S6: R103 Heidelberg road

Co-ordinates: (26°25'38" S, 28°12'13" E)

Station S6 is located on a secondary road about 1 km due east of the point where the Kliprivier/Heidelberg road crosses the N3 motorway. The small road bridge on the secondary road at which the site is located runs parallel with the N3, about 900 m northeast of the N3. The catchment is dominated by agricultural use (mainly maize), with some cattle grazing the area.

On 10 March 1997 a very heavy growth of Bullrushes was observed in the stream and in the bridge openings. The flow was extremely sluggish (surface flow only 0.06 m/s), with a flow depth of approximately 0.22 m. Only two of the four 2.4 m bridge spans showed any sign of surface flow. Estimation of surface flow was found to be adversely affected by wind effects at such low velocity Flow estimation at this point is considered to be highly inaccurate.

3. WATER QUALITY SAMPLING

3.1 General considerations

A water quality sampling programme was initiated at the six designated sampling points. This programme was aimed at monitoring the runoff from a range of different land-uses. Sampling was carried out over a six-month period. The length and intensity of the sampling programme was constrained by the limited objectives and budget for the Phase 1 study. In effect ten samples were taken at each of the six sampling sites. This limitation arose from the relatively small budget available for sample collection and analysis. As a result the sampling programme is likely to have missed the "first flush" of pollution. Hence the results would tend to under-estimate such concentrations at the onset of runoff events. This could be important with regard to the toxicological impact. This could affect the determination of the ecological reserve. However, this impact may not be a great in high-density urban areas, where peak concentrations are frequently associated with sewer spillage, even during dry weather. *Even the small experimental water quality sample is expected to partially cover* these conditions. Thus the random noise added to the regressions (see Section 6) should to some extent account for this semi-random fluctuation in pollutant export. The constraint on the number of samples was not considered an over-riding obstacle to the testing of the methodology. This is because the objective of the first phase is to develop the methodology, rather than to provide a definitive analysis of the data.

The best use was made of the available samples by ensuring that a range of wet and dry weather events was sampled. A reasonable representation of wet weather events was achieved by timing such sampling trips to occasions when significant rainfall events occurred.

3.2 Variables

A range of representative water quality variables was sought.

The variables selected are given in Table 3.1.

Table 3.1 : Water quality variables sampled

Parameter	Symbol	Units
	EC	mS/m
Sulphate	SO_4	mg/l
Nitrate (as N)	NO_3	mg/l
Ammonia (as N)	NH_3	mg/l
Total phosphate (as P)	PO_4	mg/l
Faecal coliform	FC	N/100 ml
Dissolved oxygen	DO	mg/l
Temperature	T	°C

Electrical conductivity was chosen as an indicator of conservative salts. Salinity is a primary concern in most developed catchments and also affects some natural catchments, particularly in more arid regions. Salinity impacts are also evident downstream of significant irrigation areas. Sulphate also falls within this category, but was added since it is a major ion associated with mining activities. It is also subject to soil absorption and release processes and precipitation. It can also be released into the atmosphere by veld burning and biological processes. Hence the net effective retention time for sulphate can be significantly different from that for TDS in general. Soil adsorption / desorption process also tend to reduce the variation in sulphate concentrations between wet and dry periods. (As soil moisture concentrations rise, more sulphate is adsorbed by soil particles; with desorption occurring when the soil moisture is diluted by the ingress of rainwater). Aside from these considerations, sulphate is a major anion in areas affected by mining pollution.

Nitrate, ammonia and phosphate were chosen since they are major nutrients that can promote eutrophication, which is a major water quality problem. These nutrients behave in a different manner from salts, in that they are subject to in-stream decay processes. Phosphate can also behave differently from the nitrogen compounds, in that it can be contained in particulate form on catchment surfaces and in stream sediments. Hence, although phosphate appears to be subject to decay processes, it is not a true non-conservative pollutant because it generally tends to remain in the environment, even though sedimentation and various other processes can render it either temporally or (for all practical purposes) permanently unavailable to the water environment. Ammonia is also an indicator of incomplete oxidation of nitrogen compounds. As such, its presence can point to the impact of human and animal waste.

Faecal coliform was chosen as an indicator of human and animal contamination. In most developed catchments most of the faecal coliform count is made up of *E.coli*. Faecal coliform was chosen since it is easier and less expensive to analyse. There are many other measures of human and animal waste. However, faecal coliform is the most commonly measured parameter and is a good overall indicator, making it a natural choice for testing.

Dissolved oxygen was chosen as the oxygen deficit is indicative of the presence of unoxidised metals (e.g. from mining operations), biologically active suspended and particulate matter, or faecal contamination. Its behaviour can be affected by the propensity of water bodies to re-dissolve oxygen from the air.

The time of site visits could have an influence on both dissolved oxygen and temperature. The fist site (station S1) was visited early in the morning, when low temperature and high dissolved oxygen could be expected. The last site (station 6) was visited at noon, when higher temperature and lower dissolved oxygen levels would occur. These diurnal effects were not explicitly accounted for by the methodology.

Although temperature was measured, a reliable regression with flow is not expected. However, temperature, together with altitude, was used to calculate the saturated oxygen content for the water at each sampling point. The oxygen deficit was then calculated by subtracting the observed dissolved oxygen values. The oxygen deficit (expressed as a percentage) is considered a more meaningful parameter to regress against flow, since it is a direct measure of the water quality constituents that exercise an oxygen demand on the water body.

pH was measured, but used purely as an indicator of general water quality. No attempt was made to regress pH against flow, as the relationship was expected to be poor. On all but one occasion pH levels also did not present a problem and were therefore of little interest.

3.3 Period

From the outset of the project a sample period of 6 months was planned for. The initial intention was to take roughly two samples per month (i.e. 12 samples per site). However, the actual number of samples taken was dependent on the number of sampling sites. In the end 10 samples were taken at 6 different sites between March and September 1997.

This period did not span the wettest part of the summer. However, by selecting sampling periods to coincide with the largest available storm events it was possible to obtain data for some sizeable storm events. Based on the model simulation data for 11 688 days from October 1965 to September 1997, the runoff at station S1 on 10 March 1997 was exceeded for only 3% of the time, while three other events were exceeded for between 14% and 26% of the time. Hence four (40%) of the samples corresponded with wet weather events, the remainder being more akin to dry weather conditions. Thus a reasonable spread of dry and wet weather events was achieved.

3.4 Water quality data

The water quality data collected at the six sampling points is summarised in Appendix A.

4. HYDROLOGICAL MODELLING

4.1 Choice of model

The NACL suite of daily time step rainfall / runoff models was chosen for the following reasons:

- A reasonable spread of daily rain gauges is readily available for the catchment. A significant spread of finer time step rainfall records is not available for this catchment. Nor are such records long enough to represent a full range of hydrological conditions. Such data will also be available for few other catchments throughout the country. Hence there is little point in choosing an hourly, or finer, time step model, since this would defeat the purpose of developing a methodology that can be widely applied.

- Apart from the general scarcity of the requisite data, the data requirements for an hourly or shorter time step model are extremely onerous.

- The NACL model was originally developed and tested for the Klip River catchment, including the Rietspruit *(Herold, 1981)*. These models were also used in a number of applications in this area. Moreover these models have recently been recalibrated for the Rietspuit as part of the DWAF's Vaal River System Analysis Update study *(Stewart Scott, 1999)*.

- The project team is familiar with the working of the NACL model. As such they are well placed to apply it.

4.2 Model layout

The NACL model *(Herold, 1981)* comprises two main components. The first (NACL01) is a daily time step catchment rainfall-runoff model, based on that developed by *Pitman (1976)*. NACL01 includes modules for simulating the surface and sub-surface daily salt balance. The salinity modelling functions are not described in detail, since only the flow modelling component results were used in this study. NACL01 is run for each sub-catchment of interest and the results stored.

NACL02 is the second model component. This model accepts as input the daily output from NACL01 for each sub-catchment. NACL02 routes the runoff and associated salt load through a system of river reaches to the catchment outlet. Allowance is made for point input to be entered at the head of each defined channel reach, along with independent mine discharge inputs. Riparian irrigation, channel bed losses and point abstractions are also provided for. This routing module takes account of channel slope, channel and wetland cross sectional properties and allows for separate friction factors for the main channel and the flood plain. The model includes a number of salinity related processes. However, these have little bearing on the current discussion, since only the flow output data has been used.

Figure 4.1 shows the model layout used to represent the Rietspruit catchment upstream of Rand Water (RW) gauge R6. This model layout is similar to that used in calibrating the model. The main difference is that the Natalspruit and Rietspruit catchments have been sub-divided and extra river reaches included representing the six small new monitoring points.

The catchment was divided into 8 sub-catchments. This was done partially to represent the main features of the river system and to differentiate the inputs from the six new sampling points. These had to be represented by different sub-catchments, because each has its own impervious area (as defined in Table 2.1). 11 channel reaches were defined to link up the main river channels.

Flow gauging station N8 comprises a Parshall flume set in an old road causeway. Unfortunately this weir has long been inoperative since the old causeway has not been maintained. Moreover, flow recording equipment is prone to vandalism. Although various proposals for re-instating flow gauging at this site have been put forward, these have been rejected on the grounds that site security is jeopardised by the proximity of large informal settlements. Two hijackings of DWAF personnel have already occurred at this site even when only water quality sampling is taking place at the road bridge. The risk to staff would be compounded by the need to leave the road to service flow gauging equipment.

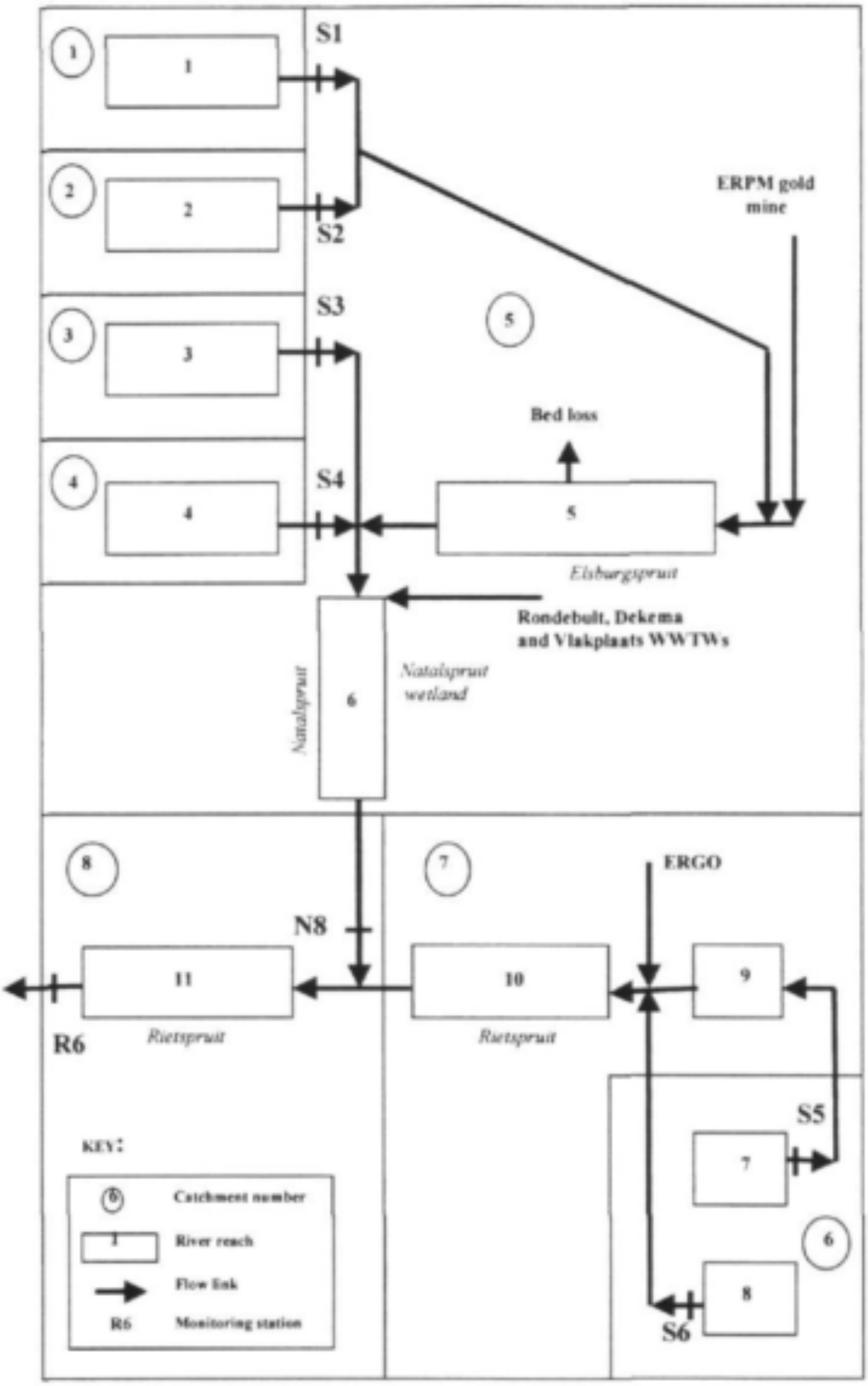


Figure 4.1 : NACL model layout for the Rietspruit

A notable feature of the system is the bed loss of water from reach 5, downstream of the ERPM gold mine. This loss is thought to be to dolomitic groundwater, with ingress to underground mine workings. Reach 6 represents the Natalspruit wetland, with an area of 7.4 km^2

Rand Water station R6 is the only serviceable flow gauging station in the Rietspruit catchment. R6 comprises a Crump weir. Flow gauging at this station appears to be reliable, although the record is relatively short. This problem was overcome when the model was calibrated for the Vaal River System Analysis Update study by taking account of the longer record available at other flow gauging stations in the Klip River catchment. Consistent model calibration parameters were sought for all of the Klip River catchments, including the Rietspruit.

4.3 Point sources

Major point sources comprise:

- ERPM gold mine
- ERWAT's Rondebult WWTW
- ERWAT's Dekema WWTW
- ERWAT's Vlakplaats WWWTW
- leakage from the ERGO plant and tailings dam
- leakage from sewers.

While the last two are not true point sources, they were treated as such for the purpose of hydrological modelling.

4.4 Irrigation

Estimated irrigation areas for each river reach are given in Table 4.1.

Table 4.1 : Present day irrigation (km^2)

{PRIVATE }					River channel reach					
1	2	3	4	5	6	7	8	9	10	11
0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	4.4	17.6	4.4

4.5 Model calibration

The model was calibrated for the period October 1977 to September 1995 *(Stewart Scott, 1999)*. While this was the length of the record for the longest maintained stations in the Klip River catchment, that for station R6 was only available for the 4½-year period from April 1991. The purpose of the present report is not to give a detailed discussion of the calibration of the flow simulation model. This process is highly complex, because the calibration of all the flow gauges in the Klip River (each with different periods of available record) has to be assessed. Model calibration also had to take account of the remainder of the Barrage catchment and the flow records in the Vaal River itself. Suffice it to say that model calibration for the Rietspruit was not based only on the short record available at station R6.

A summary of the modelled and observed daily flow records at station R6 are given in Table 4.2. Figure 4.2 is a plot of the modelled and observed daily flows at this station.

Table 4.2 : Comparison between modelled and observed daily flows at station R6

Parameter	Observed (m^3/s)	Modelled (m^3/s)	Error (%)
Mean	2.58	2.64	+2.36
Standard deviation	1.91	1.76	-7.49
Linear correlation coefficient	.785		
RMS error	1.21		
Sample size	1587		

4.6 Extension of flow records

Weather Bureau data was used to extend the catchment rainfall records for the Natalspruit and Rietspruit catchments. Sub-catchments 1, 2, 3, 4 and 5 all shared the same rainfall data file (for the Natalspruit). The remaining sub-catchments (6 and 7) made use of the Rietspruit catchment rainfall data file.

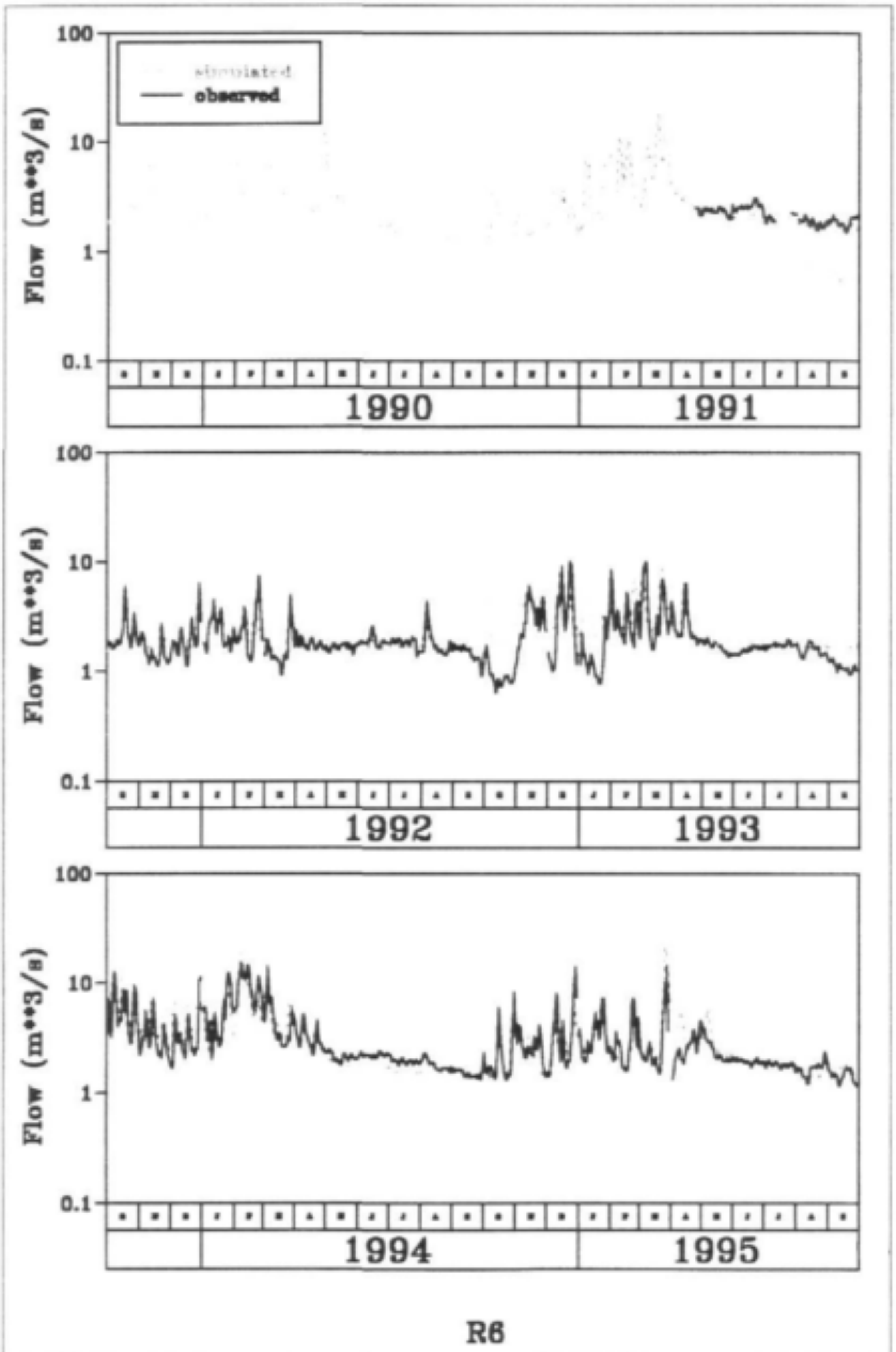


Figure 4.2 : Modelled and observed daily flows at station R6

The effluent flow records for the major point sources were extended for the next two years, to the end of September 1997.

No attempt was made to re-calibrate the model for the last two years up to the end of September 1997. This is because the small change in the model calibration resulting from two years of extra data does not warrant the extra effort required to recalibrate the model for the entire Klip River catchment.

The model as calibrated using the 1997 water quality data was used to simulate the daily flow record for the period October 1965 to September 1997.

4.7 Flow correction

Rand Water provided 1997 flow data for station R6 on the lower Rietspruit for the period during which the water quality data was gathered. The observed flow at R6 on each sampling day is compared with that modelled in Table 4.3.

Table 4.3 : Comparison between modelled and observed daily flows at station R6 on days when water quality sampling took place

Date	Observed (m^3/s)	Modelled (m^3/s)	Difference (%)
10/03/1997	11.83	15.59	+31.8
04/04/1997	15.47	14.18	-8.3
07/05/1997	3.08	4.83	+56.8
22/05/1997	3.02	3.68	+21.9
10/06/1997	6.02	4.17	-30.7
25/06/1997	4.48	3.75	-16.3
10/07/1997	4.21	3.24	-23.0
20/08/1997	2.84	2.38	-16.2
08/09/1997	3.50	3.08	-12.2
18/09/1997	3.73	3.07	-17.7
Average	**5.82**	**5.78**	**-0.4**

Table 4.3 shows that for the period in question there was a reasonable comparison between the modelled and observed flows at station R6. Although the average error is very small, the error on any one-day can be much larger. This is an inevitable and expected result of the

temporal and spatial distribution of the rainfall over the catchment. The actual variation in rainfall intensity cannot be replicated by the spatially averaged rainfall data used for each sub-catchment. A daily time step model also cannot replicate shorter duration fluctuations in rainfall input. Hence on a sample-by-sample basis the modelled flows can vary quite widely from the actual values. However, as the sample size increases, the modelled mean should approach that observed. Moreover, as the sample size corresponding to each incremental range of flow increases, the modelled average for each flow range should also tend towards that observed (provided the model calibration is valid).

In the general case an adjustment can be made to the simulated flows to account for the difference between the modelled and observed means at the reference gauge (in this case R6). However, this is considered unnecessary, since the model has already been calibrated against the observed data. One of the key objectives of the calibration was to obtain good correspondence between the modelled and observed means. Hence, provided the calibration has been carried out properly, the error between modelled and observed means should be insignificant. This is bore out by Table 4.3, which shows very close fit between the modelled and observed means, even for the small flow sample corresponding to the days when water quality samples were taken.

Since the sample size is small (and generally always will be smaller than the desirable) an attempt was made to reduce the error on each occasion when sampling occurred. Ideally this could be achieved by using the estimated observed flows made at each site on occasions when samples were taken. However, as discussed in section 2, the observed flows were generally highly unreliable. This finding was not unexpected, given that small upland catchments near to the pollution source areas are generally characterised by poorly defined and shallow streams devoid of flow gauging structures. This results in very large errors. In view of this problem, the hydrological model was used to estimate the flows. The flow estimates were refined by applying a correction to the modelled flows, based on the modelled and observed results at the gauged point (i.e. at station R6). This was achieved by applying a correction factor based on the last column of Table 4.3 to the upstream modelled flows. For example, on 10 March the modelled flow at R6 was 31.8% larger than that observed. A factor of 0.759 was therefore applied to reduce the modelled flow. Applying the same correction to the remaining flows meant resulted in a reduction of the modelled flow at station S1 from 0.287 m^3/s to 0.219 m^3/s.

The methodology could be further refined by taking account of both point and catchment runoff inputs. For example, it could be assumed that the entire error at R6 was attributable to inaccuracy in the catchment runoff. However, this would imply that the point discharge gauging is perfect and does not vary between recording dates, which is not likely to be true. It also implies that the estimation of bed, irrigation and evapotranspiration losses is perfect. Clearly this approach would require estimation of the error associated with every flow component of the river flow (both inputs and losses). A method of dealing with inconsistencies would also have to be developed. (For example, as soon as each component is dealt with separately it is possible to arrive at "correction" factors that lead to negative flows.)

In the absence of definitive information regarding the relative contribution of each flow component to the overall error, the simplifying assumption has been made that the percentage error applies equally to all simulated flow components. This produces a robust model that cannot lead to negative flows. The flow adjustment has been applied merely to speed up the process by somewhat reducing the error applicable on each day that samples were taken. The degree to which this reduces the number of samples required to arrive at a reliable concentration/flow regression has not been assessed, since the observed sample size was already small and could not be split further for this purpose. Nor has any attempt been made to refine the method used to distribute the error to upstream monitoring points.

The adjusted flows were non-dimensionalised prior to comparison with the water quality. This was done by expressing the flows as mm per day (by dividing the flows in m^3/s by the catchment area in km^2 and multiplying by 86.4). This approach was adopted to facilitate using the regressions to estimate the impact of future catchment development or project the results to estimate the contribution of other catchment areas.

5. WATER QUALITY RELATIONSHIPS

5.1 Approach

The hydrological model was used to simulate river flows at all points of interest in the study catchment. The modelled flows at R6 were compared with those observed in order to derive daily adjustment factors. These factors were used to adjust the modelled flows at stations S1 to S6, thereby providing revised estimates of the flows at these monitoring points. The adjusted flows were then compared with the sampled water quality variables. Various regressions between adjusted flow and sample quality were then tested for each station and variable.

5.2 Regression types

Both linear and power regressions were tested.

The linear regression is of the form:

$$C = A \cdot Q + B \quad \text{(1)}$$

The power regression was of the form:

$$C = A \cdot Q^B \quad \text{(2)}$$

where:

C = concentration of water quality variable (appropriate units)

A,B = regression constants

Q = flow (m^3/s)

Means, standard deviations, correlation coefficients, standard errors and 95% confidence and prediction intervals were also calculated for each regression type, variable and sample point.

Ideally the same type of regression equation should be selected for all stations for the same water quality parameter. This is based on the expectation that the same underlying processes will control the accumulation, mobilisation and decay of a given water quality variable. However, in some instances a significantly better correlation could be obtained for some stations using a different regression equation. Sometimes this may have resulted from the small sample size, which could cause one extreme high or low value to favour another type of regression at a given station. Typically this was found to apply in cases of low correlation. However, the nature of the catchment development could also lead to a legitimate shift from one type of regression line to another. Cognisance was also taken of the reasonableness of the regression results. Hence, as a general rule an attempt was made to select one type of regression line for a variable, but judgement was applied where necessary to select another type of regression where deemed necessary for some monitoring stations or parameters.

There is merit in partitioning the data between high and low flows, to arrive at two sets of regression equation. The rationale behind this approach is that during low flow conditions the concentration of pollutants (particularly those in particulate form derived from urban environments) can remain high and even increase as low flows increase. Base flow concentrations can also be highly variable, depending on antecedent flow conditions and fluctuating local conditions (such as intermittent sewer overflows). As the discharge increases beyond a certain point the concentrations will tend to decline due to the exhaustion of accumulated pollutants. Hence a dual set of regressions for high and low flow conditions is well justified. Consequently the programs used to carry out the water quality analyses have been designed to allow for partitioning of the data according to flow range. However, this has not been done with the current test data set, since the sample size is already small (10 samples). Partitioning of the data would further reduce the sample size (especially that for high flow conditions) thereby rendering the analysis worthless.

5.3 Regression results

Table 5.1 gives a summary of the linear correlation coefficients achieved using both linear and power regression equations. Plots showing the chosen regression equations and the confidence and prediction intervals are given in Appendix B.

Table 5.1 : Linear correlation coefficients (r) obtained from different regression types

Variable	Parameter	Sampling station					
		S1	S2	S3	S4	S5	S6
EC	N	10	10	10	10	10	10
	Linear	-0.5205	-0.0430	-0.3267	-0.0980	-0.4902	-0.4775
	Power	-0.6110	-0.0118	-0.4487	-0.0981	-0.6981	-0.6672
SO_4	N	10	10	7	10	10	10
	Linear	-0.5732	-0.1680	0.0641	-0.2151	0.3250	-0.0187
	Power	-0.6644	-0.0931	0.0811	-0.1604	0.0573	0.1609
NO_3	N	10	10	7	10	10	10
	Linear	0.1781	0.1134	-0.0999	0.3715	-0.0656	-0.2126
	Power	0.1348	0.1258	-0.3184	0.5998	0.2368	0.0056
NH_3	N	10	10	7	10	9	9
	Linear	-0.1792	0.2506	-0.1077	0.0863	-0.0360	0.0123
	Power	-0.4089	0.2264	-0.2783	0.3851	0.1433	0.1507
Total P	N	10	10	7	10	10	10
	Linear	-0.3333	-0.4635	0.0787	0.2882	-0.4358	-0.4413
	Power	-0.7174	-0.7396	0.4214	0.1523	-0.7002	-0.5560
Faecal coliform	N	9	9	7	9	9	9
	Linear	-0.1772	0.7172	0.7078	0.2902	-0.2091	-0.2075
	Power	0.0884	-0.0605	0.7939	0.2644	-0.2652	0.0437
Oxygen saturation	N	10	10	7	10	10	10
	Linear	0.5454	0.6991	-0.3821	0.1569	-0.0595	0.1924
	Power	0.7721	0.6572	-0.4170	0.0903	0.1180	0.3060

The results for each water quality variable are discussed in the following sections.

5.3.1 Electrical conductivity

A power regression was found to give the best overall results for electrical conductivity. At station S2 a linear regression produced slightly better results. However, the correlations at this station were very poor and the results were not considered significant enough to alter the choice of equation type. A power regression with a negative exponent is also thought to be more likely to mimic the variation of EC concentration with flow.

Further details are given in Table 5.2.

Table 5.2 : Statistical details for electrical conductivity (mS/m)

Parameter	S1	S2	S3	S4	S5	S6
Average	76.9	49.8	51.6	34.8	43.9	33.1
Standard deviation	18.1	13.0	12.7	7.3	10.1	14.6
Maximum	103.0	81.0	63.0	51.0	57.0	67.0
Minimum	58.0	33.0	27.0	25.0	26.0	19.0
Regression type	power	power	power	power	power	power
Regression constant A	61.327	48.269	42.291	33.233	30.997	19.031
Regression constant B	-.12454	-.00257	-.12302	-.01740	-.17958	-.26805
Correlation coefficient, R	-0.6110	-.0118	-.4487	-.0981	-.6981	-.6672
Standard Error, SE	14.67	13.05	11.11	7.367	6.911	11.73
Sample size, N	10	10	7	10	10	10

The relatively small variation in the mean electrical conductivity recorded at the six sites was unexpected. In particular, the average for station S5 (which is sited below the Suikerbosrand Nature Reserve) was higher than expected.

5.3.2 Sulphate

A power regression was found to give the highest average correlation coefficient. Sulphate behaves in a manner similar to electrical conductivity, as it is a salt that behaves in a predominantly conservative fashion. However, it has been observed that during wet conditions the sulphate proportion of the electrical conductivity in the diffuse source washoff can rise (*Herold et al, 1997*). This is attributable to the tendency of catchment soils to adsorb sulphate as the soil moisture concentration rises (thereby reducing the potential increase in sulphate concentration during dry periods). Conversely, desorption of sulphate occurs as the soil moisture concentration is diluted during wet conditions. This means that the variation in sulphate concentrations is damped by adsorption / desorption processes and consequently tends to be much lower than that for electrical conductivity. Hence sulphate does not behave as a normal conservative constituent. For this reason a linear regression was sometimes found to produce a better correlation than a power regression. Moreover, in some instances use of a power regression resulted in an abnormally high upper boundary for the 95% confidence interval. For example, Figure B.2.12 indicates 95 percentile values up to 350 mg/l for a dry land farming catchment. This is considered unrealistic, and the linear regression was favoured (see Figure B.2.11), even though it yielded a lower correlation.

The improbable result also sometimes occurred that as the flow increases the regression curve results in increasing sulphate concentrations. It is unlikely that this trend can hold true indefinitely, as dilution should occur at higher flow rates. This aberration is probably attributable to the small sample size and a cluster of sample points for low flows that are poorly correlated. (This could also arise from the use of data with a limited range comprising only low flows. However, although the sample size is small, the data includes a reasonable range of high flows.)

Further details for the chosen regression types are given in Table 5.3

A high mean sulphate concentration at station S1 could be expected since the upper portion of the catchment is affected by mining development. However, once again the values recorded at station S5 were higher than expected. This could be due to a natural sulphate yield from the weathering of rocks in the Suikerbosrand Nature Reserve catchment. However, it could also be related to the relatively low soil moisture storage in the thin soils of the upland portion of this catchment. This could mean that this catchment could show a more rapid response to enhanced atmospheric sulphate deposition, than would be the case for the deeper soils of the more low lying areas.

Table 5.3: Statistical details for sulphate (mg/l)

Parameter	S1	S2	S3	S4	S5	S6
Average	276.7	78.4	19.0	36.2	38.3	24.3
Standard deviation	102.8	62.3	15.3	20.9	30.6	15.5
Maximum	449	242	49	83	94	53
Minimum	153	37	5	13	9	2
Regression type	power	linear	linear	linear	linear	linear
Regression constant A	182.22	-22.018	13331	-9.894	30.075	-1.8799
Regression constant B	- 21927	86.19	18.94	39.98	30.07	24.81
Correlation coefficient, R	-0.6644	-0.1680	0.0041	-0.2151	0.3250	-0.0387
Standard error	82.30	61.84	15.33	20.45	28.93	15.45
Sample size, N	10	10	7	10	10	10

5.3.3 Nitrate

A power regression was found to give slightly better overall results than was the case for a linear regression. However, the correlation coefficients were low, as can be expected from a non-conservative pollutant. Moreover, ammonia tends to decay to nitrate, further reducing the effectiveness with which it can be regressed. At such low correlations the choice of regression type should not be dictated solely by a marginal increase in the overall correlation coefficient. A power regression (since it uses the logarithms of the concentrations) was found to lead to unreasonably high natural concentrations for the upper boundaries of the 95% prediction interval (see Figure B.3.8). A linear regression was adopted since this resulted in more reasonable upper prediction intervals.

Further details are given in Table 5.4

Table 5.4: Statistical details for nitrate (mg N/l)

Parameter	S1	S2	S3	S4	S5	S6
Average	.81	2.19	1.33	.96	.206	.232
Standard deviation	.36	1.18	1.62	1.42	.126	.113
Maximum	1.27	3.80	4.80	4.83	.40	.51
Minimum	.10	.28	.15	.05	.02	.05
Regression type	linear	linear	linear	linear	linear	linear
Regression constant A	.14007	.27900	-.34641	1.1618	-.02508	-.10899
Regression constant B	.759	2.090	1.478	.518	.213	.261
Correlation coefficient, R	.1781	.1134	-.0999	.3715	-.0656	-.2426
Standard error	.354	1.170	1.608	1.322	.126	.139
Sample size, N	10	10	7	10	10	10

5.3.4 Ammonia

As with nitrate, a linear regression was found to give more realistic results for the first three stations. A power regression was found to give better results for the last three stations. In some instances a single isolated high value at relatively low flow may be an outlier (e.g. see Figure B.4.10), but the sample size is too small to confirm this. However, the high values could also have been influenced by the excreta from cattle or the natural decay of plant material (in the case of station S5) or sewer spillage (as indicated by the presence of sewage fungus at station S3). As such the data (and hence the variation in concentration) would be quite legitimate and should be taken into account. Hence, for the purpose of this exercise all of the data was used. The low correlation coefficients were expected, as they are consistent with this non-conservative pollutant.

Further details are given in Table 5.5.

The results indicate high ammonia values in the drainage from Vosloorus. However, ammonia levels in the stream draining Katlehong (station S4) appear to be comparable to those at the two least developed catchments (stations S5 and S6). This result is encouraging, given that sampling point S4 was located in an area dominated by informal housing.

Table 5.5: Statistical details for ammonia (mg N/l)

Parameter	S1	S2	S3	S4	S5	S6
Average	1.18	4.85	13.29	1.72	1.71	1.26
Standard deviation	1.79	7.90	17.21	1.66	2.71	1.42
Maximum	6.2	24.1	47.51	5.60	8.90	5.00
Minimum	.2	.7	.21	.63	.70	.70
Regression type	linear	linear	linear	power	power	power
Regression constant A	-.69635	4.1367	-3.9787	1.9015	1.3511	1.8825
Regression constant B	1.43	3.38	14.98	.27475	.14379	.11838
Correlation coefficient, R	-.1792	.2506	-.1077	.3851	.1433	.1507
Standard error	1.748	7.651	17.114	1.674	2.817	1.465
Sample size, N	10	10	7	10	9	9

5.3.5 Total phosphorus

The high variation of this non-conservative pollutant and the small sample size resulted in variation in the choice of regression type at each site. Further details are given in Table 5.6.

Table 5.6: Statistical details for total phosphorus (mg P/l)

Parameter	S1	S2	S3	S4	S5	S6
Average	.46	.27	1.28	1.67	.16	.22
Standard deviation	.87	.26	1.22	4.45	.19	.20
Maximum	2.80	.90	3.81	14.30	.60	.60
Minimum	.04	.06	.20	.01	.01	.01
Regression type	power	power	Linear	linear	linear	linear
Regression constant A	.03278	.07137	.20596	2.8164	-.25635	-.27844
Regression constant B	-.92660	-.99082	1.194	.594	.235	.292
Correlation coefficient, R	-.7174	-.7396	.0787	.2882	-.4358	-.4413
Standard error	.837	.219	1.216	4.260	.175	.180
Sample size, N	10	10	7	10	10	10

The high disparity between phosphate levels in the two undeveloped catchments (S5 and S6) and those measured at the remaining sites indicates the impact of development on phosphate production. The particularly high values at station S3 and S4 could be indicative of the use of detergents for clothes washing.

5.3.6 Faecal coliform

A power regression was found to lead to unreasonably high faecal coliform counts for the upper boundary of the 95% prediction interval (see Figure B.6.6). In any event, a linear regression was found to give better correlations for most stations. For these reasons a linear regression was adopted for all stations for faecal coliform. Further details are given in Table 5.7.

Table 5.7: Statistical details for faecal coliform (N/100 ml)

Parameter	S1	S2	S3	S4	S5	S6
Average	20 432	12 451	112 676	239 415	353	968
Standard deviation	59 843	21 644	133 363	716 470	993	1941
Maximum	180 000	63 000	298 000	2 150 000	3000	5900
Minimum	0	60	13	20	0	2
Regression type	linear	linear	linear	linear	linear	linear
Regression constant A	-22 370	31 061	202 555	439 475	-597.62	-1 202.7
Regression constant B	29 199	623	26 429	59 085	525	1309
Correlation coefficient, R	-.1772	.7172	.7078	.2902	-.2091	-.2075
Standard error	58 895	15 083	94 209	685 629	971	1 899
Sample size, N	9	9	7	9	9	9

The results show much higher faecal coliform counts for the developed catchments (station 1 to 4) than was the case for catchments 5 and 6. An average count of 353 per 100 ml seems quite high for station S5. However, closer examination of the data (see Figure B.6.9) shows that this result was swayed by a single sample value of 3000 counts per 100 ml. The somewhat higher average at the rural station S6 may have resulted from the cattle that were observed to frequent the site. This level is to be expected for an agricultural drainage site.

5.3.7 Dissolved oxygen

Dissolved oxygen was measured in the field. However, it was thought that a regression with oxygen deficit or percentage saturation should yield better results. Hence temperature was also measured in the field. The altitude of the stations was then used to estimate the oxygen saturation. The observed dissolved oxygen content was then subtracted from this figure and the result expressed as a percentage of full saturation. Both linear and power regressions were chosen for the various sites. Further details are given in Table 5.5.

Table 5.8: Statistical details for oxygen saturation (%)

Parameter	S1	S2	S3	S4	S5	S6
Average	82.7	59.1	41.6	69.9	70.3	45.6
Standard deviation	11.2	12.9	17.1	26.1	10.1	13.7
Maximum	101.4	79.7	64.8	92.0	85.1	63.8
Minimum	63.6	39.2	21.8	5.9	59.1	17.8
Regression type	power	linear	linear	linear	power	power
Regression constant A	95.362	18.807	-14.037	8.9842	71.832	53.654
Regression constant B	.09260	52.48	47.56	66.52	.01726	.12190
Correlation coefficient, R	.7721	.6991	-.3821	.1569	.1180	.3060
Standard error	7.30	9.21	15.82	25.75	10.13	13.37
Sample size, N	10	10	7	10	10	10

High dissolved oxygen levels at S1 are consistent with this site being below a largely low-density residential area and the relatively steep channel slope. Reductions in dissolved oxygen at the industrial site (S2) and the high-density residential sites (S3 and S4) are also to be expected. Although not alarmingly low, the 70% oxygen saturation at station S5 below the undeveloped catchment is unexpectedly lower than that observed at S1. This may be attributable to the flat terrain and consequent low flows. The presence of a few small farm dams upstream of the site may also have contributed to a reduction in aeration of the water. A ready explanation for the low values (46% saturation) at station S6 is hard to find, although generally low flow velocity and the presence dense reed growth around this bridge would have affected oxygen levels.

6. ESTIMATION OF RANGE OF CONCENTRATIONS

6.1 Generation of sequences

6.1.1 Regression with flow

The NACL daily time step model was used to simulate daily flows at stations S1 to S6 and R6 for the period October 1965 to September 1997. The procedure described in Section 4 was then used to obtain adjusted daily flows at monitoring stations S1 to S6. The regression equations given in Section 5 were next used to calculate the regressed concentrations on each day that flows were modelled (i.e. 11668 days). Provision was made for the user to specify maximum and minimum flow rates (to avoid applying the regressions to flows that were outside the range of the observed data).

6.1.2 Addition of random noise

A random number generator was used to provide a random number for each sample point on each day. These random numbers were then normalised (to conform to a normal distribution) to yield normalised random numbers. The normalised random number for each day simulated was then multiplied by the standard error (see Section 5) for the observed data set. This provided the normalised random noise, which was then added to the regressed values for each day. By this means 11668 estimates of the concentration at each point were generated by this means. Allowance was made for the user to specify upper and lower bounds for the calculated concentrations. This was necessary to prevent the generation of negative concentrations or levels that are unnaturally high.

Figure C.1 (see Appendix C) is an example of the resulting range of phosphate concentrations at station S2, calculated from the adjusted simulated daily flows for the period October 1965 to September 1995. This period covers a wide range of possible hydrological conditions. Extreme droughts and floods are included in this period, along with many more normal periods.

Alternatively, the methodology can be (and has been) used to estimate a range of concentrations for a fixed flow rate. In this case the 95% low flow for station S2 is 0.0015 m^3/s (i.e. 0.0393 mm/day). Any other percentile low flow can be chosen, in place of the 95% low flow. (For example, the 7-ay average low flow with a recurrence interval of 10 years is often used.) However, it is very difficult to specify a single flow rate to be used to assess the effect of an option on water quality. However, Appendix B *shows that high concentrations (which are often related to* failure of the system, such as overflowing sewers) cannot be linked invariably to a particular low flow discharge. The inability to link the resulting concentration statistics to a probability of occurrence is considered a serious weakness of the methodology of adopting a fixed flow rate for analysis.

6.1.3 Generation of duration curves

Figure C.2 shows the result of ranking the phosphate concentrations contained in Figure C.1, to form a duration curve. This is a powerful tool that can be used to pick out a range of key percentile values that can be used to assess the impact of expected new developments or the implementation of planned management options.

Table 6.1 is an example of the statistical properties of the sequence of concentrations generated by the above methodology.

Table 6.1 : Statistical properties of simulated phosphate concentrations at station S2

Parameter	Concentration (mg/l)	Load (t/d)
Mean	0.517	0.0383
Standard Deviation	0.546	0.0412
Minimum	0.01	0.002
Maximum	3.05	0.960
2% high	0.87	0.173
5% high	0.65	0.118
10% high	0.50	0.084
50% (median)	0.17	0.025

6.1.4 Consequence of not including random noise

Figure C.3 shows the impact on the results of failing to add the random noise to the regression line. Use of the regression line alone with no allowance for the observed variance about the regression damps the results very severely, with the result that the peak values are much too low. This clearly demonstrates the importance of taking proper account of the variance observed in nature.

6.2 Example of application of methodology

The main purpose of this study was to develop a usable methodology, rather than to generate a large volume of data. Faecal coliform was chosen to illustrate the use of the methodology. The results for stations S4 and S5 have been used for this purpose. Faecal coliform at these two stations have been chosen for the following reasons:

- Faecal coliform washoff from informal / low cost housing developments is one of the most common problems associated with rapid urbanisation;

- Faecal coliform correlations are generally poor (linear correlation coefficients of only .29 and .21 at stations S4 and S5 respectively). As such they are typical of non-conservative pollutants and are most in need of a methodology that can simulate the wide variation in observed concentrations.

- The results can be used to illustrate the possible effect of urban development of an erstwhile rural catchment.

In short, if the methodology works for faecal coliform, it can be applied to just about anything.

In order to illustrate the methodology the hypothetical situation was chosen based on the assumption that 50% of the catchment upstream of station S5 would be developed into a low cost housing estate similar to that commanded by station S4 at Katlehong.

The hydrological model was used to simulate this hypothetical option, taking due account of the increased runoff due to the new paved area. The generated faecal coliform values before and after the hypothetical urban development have been ranked to produce the duration curves given in Figure C.3.

Table 6.2 gives the corresponding statistical properties at S5 before and after the new development.

Table 6.2 : Statistical properties of simulated faecal coliform counts at station S5, before and after the hypothetical development of a new high-density urban settlement (N/100 ml)

Parameter	Before development	After development
Mean	461	143 487
Standard Deviation	1 068	735 894
Minimum	0	0
Maximum	2 709	2 483 095
98% high	2 461	1 640 193
95% high	2 110	1 280 063
90% high	1 802	1 113 618
50% (median)	438	11 932

The approach used in this example can be applied for a range of management options.

It is important to observe that successful use of the methodology does not depend on the goodness of fit of the regression equation used. What is important is that:

(a) the regression used should result in a reasonable normally distributed scatter of values around the regression line; and

(b) the sample size is large enough to yield a reasonable representation of the regression line and the random variation around it.

With non-conservative pollutants, the latter becomes more important, since much of the variation is better explained by the random noise.

6.3 Extension to ungauged catchments or new conditions

The same methodology could be used to extend the results to any ungauged catchment within the Rietspruit catchment. This could easily be achieved by using the GIS overlays to determine the areas under each type of development above any selected portion of the catchment. This was not done, since the illustration of the effect of the hypothetical development of the S5 sub-catchment is sufficient to demonstrate the potential use of the methodology. Moreover, extension to the entire Rietspruit catchment would require the use of a water quality routing model to account for in-stream decay process between the source areas and the catchment outlet. The use of such a routing model falls outside the scope of the present study.

7. DATA BASE REQUIREMENTS

7.1 Overview

The methodology put forward in this report requires some additional information to enable the characteristics of the observed data to be replicated by a second user. This additional information is relatively easy for the original researcher to compile from their raw data, but virtually impossible for a second party to achieve, except at great cost and repetition of work. Proposals are put forward for standardising these data requirements and the development of a national database. The idea is that the key statistical requirements will be accumulated from more detailed studies that are regularly conducted by researchers and water quality practitioners around the country. In some instances it might be possible for researchers to process their data sources at relatively low cost to provide the required information for earlier studies. As the results of projects are accumulated, the database will become a valuable source that future researchers can draw on to arrive at estimates of the likely response of catchment areas to different land use developments.

The results of the second phase of the study will form the start of a database that can supply information on non-point source quantification at an appropriate level of detail for catchment management purposes. The data base, as well as the land-use classification and the monitored river water quality at various points along the river will be entered in a GIS, thereby enabling evaluation of the possible effect of changing land-use on river water quality.

The information stored in the data base can be confirmed as experience is gained from longer monitoring records and subsequent studies on other catchments. In time the database can also be extended to include coverage of a wider range of water quality variables and land uses.

7.2 Data base requirements

A detailed evaluation of the data base requirements is to be covered during Phase 2 of the study. This section therefore contains only a rough evaluation of the requirements for a national database.

The purpose of the database is to accumulate appropriate data from previous and future studies. The data that is accumulated should be in a form that can easily be interpreted by users to estimate the impact of projected land use changes and the effect of existing developments that have not yet been monitored and studied. In order to fulfil these requirements the data must be relevant to the methodology that will be used. In this regard the data quoted in the literature tends to be in a largely unusable form, with each practitioner quoting various statistical properties (usually an average annual load).

In order to arrive at an annual load, each practitioner should have at his or her disposal a comprehensive observed data set. The objective is to carry out relatively simple additional processing of the databases to provide useful information with which to populate the national database.

Fulfilment of this ideal will require some measure of agreement on the parameters to be stored. It might also be desirable to make available an easy to use software package available to water quality practitioners to facilitate the processing of databases to produce the required statistical properties.

Some of the key values that would be useful for application of the methodology described in this report include (for each water quality variable):

- A and B linear regression constants
- A and B power regression constants
- mean concentration (natural)
- mean concentration (log)
- standard error of observed concentrations (natural)
- standard error of observed concentrations (log)
- applicable range of flows used
- applicable range of concentrations
- indication of type of regression that yielded the best results
- correlation coefficient achieved
- type of land use (percentages of area in each specified category)
- land use management
- sanitation type and age of sanitation system

- age of sanitation system
- Origin of waste loads(e.g. domestic or industrial)
- locality and climate
- geology and catchment soil types
- catchment area
- slope

Most of the above information could be obtained using standard software that could be made available to practitioners.

The foregoing list of requirements is by no means fixed, but could require modification in the light of discussions held with a range of practitioners. The success of the project would depend on the co-operation of those carrying out investigations of land use effects. It would also be essential for an authority (such as the Department of Water Affairs and Forestry or the WRC) to take ownership of the database and to ensure that it is populated with data as further investigations take place. In this regard the WRC and other employers could require future researchers in this field to provide the necessary information as one of their research products.

8. CONCLUSIONS

8.1 Water quality data

A limited amount of water quality data has been collected at six sites in the Rietspruit catchment. Predominant land uses include:

- low-density formal residential / CBD
- high-density and informal residential
- industrial
- rural (dry land cultivation)
- undeveloped (nature reserve).

Ten samples were taken at each site over a six-month period from March to September 1997. Some of the sample dates were planned to coincide with wet weather conditions, thereby ensuring that the limited number of samples covered a reasonable range of flow conditions. The key water quality variables that were sampled included pH, electrical conductivity, sulphate, nitrate, ammonia, total phosphate, faecal coliform, dissolved oxygen and temperature. The later two variables were used to calculate the percentage oxygen saturation. Very crude estimates of flow were made at each site. However, most of these estimates are considered to be highly inaccurate and were not used in further processing. This had no detrimental effect, since the methodology is based on the assumption that definitive flow data is not available at the monitoring stations.

8.2 Flow modelling

The daily time step NACL model was used to simulate flows at each monitoring site. A correction was applied to the data to make the modelled flows at the catchment outlet (RW station R6) coincide with those observed.

8.3 Water quality relationships

Both natural and power regressions between modelled flow and concentration were derived for key parameters at each site. The water quality variables for which these regressions were derived included:

- electrical conductivity

- sulphate
- nitrate
- ammonia
- total phosphate
- faecal coliform
- percentage oxygen saturation.

The NACL model was used to simulate a large number of daily flows for the period October 1965 to September 1997. This period covers a wide range of hydrological conditions, including severe droughts and extreme floods. The most appropriate regression equation for each water quality variable was then used to estimate a regressed concentration corresponding to each simulated daily flow. Normalised random noise was then added to the regressed values to allow for the considerable variation that was attached to the observed data. Provision was made for specifying an applicable range of flow conditions (in this case close to the flow range of the observed data). Allowance was also made to filter out unreasonably high or low concentrations.

The range of concentrations thus generated for each water quality variable and station was also presented in the form of a duration curve. The results were tabulated to show key statistical properties, including mean, standard deviation, regression constants, standard error, correlation coefficient and selected percentile values (98%, 95%, 90% high values plus the median).

8.4 Application of methodology

The strength of the methodology was demonstrated by means of an example of the impact of an assumed new high-density urban development. The potential for using the methodology in conjunction with GIS based land use data was also discussed. The methodology can be applied as a screening tool and be of some assistance in the determination of the water quality reserve.

8.5 Advantage of methodology

The use of a regression equation results in severe damping of the generated concentrations of non-conservative pollutants. The use of this methodology (or mean values) inevitably leads to gross understatement of the true range of concentrations that arise. The methodology put forward in this report holds the advantage of taking account of the semi-stochastic manner in which concentrations vary about the regressed values.

Use of modelled flow data (suitably corrected against observations at a downstream reference gauge) has the advantage of freeing the practitioner of the need to confine water quality sampling to sites were flows can be gauged. Application of this methodology means that monitoring sites can now be chosen close to source areas, thereby monitoring relatively homogeneously developed areas. The main requirement is that sufficient samples should be gathered to cover a range of flow conditions. It has been demonstrated that by varying sampling dates according to weather conditions it is possible to monitor a wide range of flows with relatively few samples.

The sample size was too small to permit partitioning of the data into low and high flow conditions. Nor was it possible to investigate the impact of sample size on the key statistical properties (such as mean, standard deviation and standard error). There is merit in investigating both of these effects.

8.6 Limitations of methodology

The following factors are limitations on the methodology:

- Relatively infrequent sampling can miss first flush events at the onset of rain. In some instances a long period of sampling may be required before such events are sufficiently represented in the data base on which the statistics of the regression line and random noise is based.

- Temporal variation in rainfall of shorter duration than the model time step (i.e. less than a day) and spatial variation in rainfall reduce the accuracy with which flows at distant sites can be estimated. This could be particularly significant for small sub-catchments.

- Flow estimation could be model dependent. This might affect the application of results obtained with one model to areas where another hydrological model is in use.

- In some instances the results can be sensitive to the choice of regression equation (section 5), particularly when the sample size is small.

- A linear regression often gives a superior representation to an exponential regression (see figures B.6.11 and B.6.12). However, applying random noise to a linear regression can result in negative estimated concentrations. Correction by eliminating such values has the effect of increasing the average of the generated values. This implies the need to investigate other types of regression equation.

8.7 Preliminary results

Although a methodology has been developed successfully and the results of its application are promising, the small water quality database prevented assessment of the unknowns listed above. Hence further investigations are necessary to determine the value of the method. Until such time as the major outstanding issues are addressed, the methodology cannot be regarded as a final or accepted technology.

9. RECOMMENDATIONS

9.1 Determination of required sample sizes

It is desirable to test the methodology that has been derived against fuller and longer data sets to determine the number of samples required to achieve a reasonable representation of the key statistical parameters. The effect of partitioning the data into low and high flow ranges should also be investigated. The use of longer data sets collected for other studies is required for this purpose.

9.2 Assessment of required sampling frequency

The effect of sampling frequency on the results should be investigated. This should include evaluation of the effect of the first flush at the onset of rainfall events. Attainment of these ends requires the choice of one or more monitoring stations at which long records of high frequency sampling are available.

9.3 Investigation of flow modelling time step

The use of daily flow models is constrained by the effort and time required to calibrate the models. However, monthly flow modelling is widespread in southern Africa. The potential exists to apply the same methodology to coarser monthly time step flow modelling data. This could well require a larger number of samples before a reasonable representation of the key statistical properties is achieved. However, the overall benefit of being able to use monthly flow modelling would be very far reaching.

9.4 Evaluation of flow modelling accuracy

The impact of spatial and temporal rainfall distribution on the simulated flows at sites upstream of the reference gauge needs to be investigated. This will require the use of data for a catchment with good flow records at a number of sites controlling relatively small sub-catchments. A mobile hand-held current metering could be used in some instances. However, the time taken at each site would be lengthened, thereby curtailing the number of sites that can be assessed. Since current metering could only be applied to future data, this method could not make use of longer historical records.

9.5 Investigation of other statistical distributions

The current study used only linear and power regression curves. Further investigation of other types of regression is recommended. In particular, alternative methods of handling the addition of normalised random noise to the regression line should be investigated. This is because this can result in concentrations that lie outside the plausible range and therefore have to be filtered out. This filtering can in turn affect the mean and standard deviation of the generated sequence of concentrations. The best means of handling this phenomenon should be investigated.

9.6 Development of national database

The results thus far achieved indicate that there is much merit in developing a national data base of key physical and statistical characteristics associated with different land uses. This proposal is linked to, but not dependent on, the application of the methodology discussed in this report. Even without the new technology, there is a pressing need to consolidate and standardise data on the effect of land use on water quality.

It is recommended that at least steps 9.1 and 9.2 be carried out before a national database is developed. (Completion of recommendation 9.3 is also desirable, but not essential before proceeding to develop a national database. Recommendations 9.1 and 9.2 could be carried out simultaneously as part of the same or separate projects.)

9.7 Minimisation of flow gauging requirements

The methodology described in this report can be used effectively to reduce flow gauging requirements in water quality studies. This would enable water quality monitoring sites to be selected primarily on the basis of upstream land use and access, rather than on the suitability of the site for flow gauging. The techniques that have been developed can also be used to maximise the use of data that has already been accumulated at a number of sites that are not suitable for flow gauging. It is recommended that once step 9.3 has been implemented, an early start should be made on utilising the technology. The initial focus could be on the Vaal Barrage catchment, since the NACL daily time step model has recently been re-calibrated for this catchment as part of the DWAF's Vaal River System Analysis

Update Study. Pollution of the Vaal Barrage sub-catchments is also a cause of great concern to water quality managers.

9.8 Testing of scenarios

A serious weakness of most water quality studies that have been carried out to date lies in the tendency to rely on flow/concentration regressions to explain catchment export. This inevitably leads to severe damping of the resulting scenarios of water quality concentrations. This points to the need to take better account of the variability of observed concentrations about the mean or regression line. This is especially important with regard to non-conservative constituents, which can typically show a very wide variation. The methodology that has been presented provides a means of taking this variation into account. Hence it has great potential to be applied in water quality studies to test scenarios (such as assessment of the impact of planned management options or scenarios of projected land use development). However, use of the new methodology in this manner should not proceed until some of the key outstanding issues have been addressed in follow up studies.

9.9 Use of random number processing in other applications

The methodology described in Section 6 for dealing with the variation in observed concentrations has great potential for application in a simplified, easy to calibrate river routing model. A simple decay model was developed and used successfully for the Environmental Impact Assessment for ERWAT's proposed new Welgedacht water care works *(Stewart Scott, 1998)*. Proposals have been made for extending this simple but robust technology to include the methodology discussed in Section 5.

9.10 Evaluation of dependence on hydrological model

It is desirable to test the dependence of the results on the choice of hydrological model. Such testing may prove difficult, since a great deal of time and effort goes into calibrating daily time step models. The cost of calibrating more than one model in each catchment could prove costly. The results may not be sensitive to this, since the objective of all hydrological model calibrations is to replicate the observed data. Because such models employ similar algorithms to simulate hydrological processes, there should not be much difference between the flow results obtained, especially if the same person calibrates each model. The differences in flow estimation by each model should in any event be much less significant to the methodology than the observed variation in water quality. Consequently this recommendation has been given a low priority.

10. REFERENCES

Herold, C.E. (1981): A model to simulate daily river flows and associated diffuse-source conservative pollutants. Report No. 3/81, Hydrological Research Unit, University of the Witwatersrand, Johannesburg. March 1981.

Herold, C.E., Taviv, I. and Pitman W.V. (1997): *"Modelling the long term effect of atmospheric deposition on the salinity of catchment runoff with special reference to the Vaal Dam catchment"*. Water Research Commission (project K5/697), draft report by Stewart Scott, Sandton.

Pegram, G.C. and Görgens, A.H.M. (1999): *"A guide to nonpoint source assessment to support water quality management of surface water resources in South Africa"*. Water Research Commission (project K5/696), draft report by Sigma Beta Consulting Engineers, Stellenbosch.

Pergram, G.C., Quibell, G and Görgens, A.H.M. (1997): "A situation assessment of nonpoint sources in South Africa". Water Research Commission Report, Pretoria.

Pitman, W.V. (1976): A mathematical model for generating daily river flows from meteorological data in South Africa. Report No. 2/76, Hydrological Research Unit, University of the Witwatersrand, Johannesburg. March 1976.

Stewart Scott (1996): *"Water quality impact assessment of Johannesburg's southern wastewater treatment works on the Klip River"*. Report to Greater Johannesburg transitional Metropolitan Council, Johannesburg.

Stewart Scott (1998): *"Comprehensive environmental impact report for a new regional water care works at Welgedacht"*, Report to East Rand Water Care Company, July 1998.

Stewart Scott (1999): *"Hydro-salinity model calibration: Vaal Barrage catchment Part (a): Daily analysis"*. Report No. P C000/00/18196, Vaal River System Analysis Update study, Directorate of Project Planning, Department of Water Affairs and Forestry, Pretoria.

APPENDIX A

WATER QUALITY SAMPLE DATA

Table A.1 : Observed water quality at station S1

Station : S1 - Jan Frederick Street Germiston (page 112, DZ 119) - Residential + CBD
ERWAT observations k:\common\wrcdiff\wqdata\s1.raw 11.84 km2 26°123'58" S 28°11'55" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	0800	7.50	62	1.04	0.04	0.28	179	6.12	0	20.6	0.530	.2886	.2189
1997	4	4	0740	7.22	58	0.67	0.04	0.92	190	7.49	95	18.4	0.620	.0949	.1035
1997	5	7	1015	7.60	103	0.86	0.10	6.20	412	6.17	180000	16.0	0.042	.0388	.0247
1997	5	22	0752	7.90	101	0.54	0.98	0.20	449	5.20	155	13.8	0.040	.0099	.0081
1997	6	10	0814	7.50	85	1.02	0.15	+0.70	360	6.80	79	12.0	0.000*	.0199	.0287
1997	6	25	0924	8.20	63	+0.10	+0.10	+0.70	252	7.01	400	11.2	0.069	.0199	.0237
1997	7	10	0926	7.70	66	1.27	+0.05	+0.70	247	6.90	-1	10.8	0.046	.0099	.0129
1997	8	20	1047	7.30	99	1.20	0.30	+0.70	317	6.18	159	14.4	0.025	.0089	.0106
1997	9	8	0923	7.60	63	0.90	+0.05	+0.70	153	8.02	2400	15.4	0.320	.0798	.0907
1997	9	18	0844	8.00	69	0.50	2.80	+0.70	208	6.41	600	15.0	0.067	.0092	.0112

NOTE:
+ Denotes value below detection limit, when half of detection limit was used.
-1 Denotes missing data.
* Current velocity too low to estimate flow, but noticable flow.

Table A.2: Observed water quality at station S2

Station : S2 - Opposite c/o Lamp and Baracuda Sreets, Wadeville (page 156, EF118) - Industrial
ERWAT observations k:\common\wrcdiff\wqdata\s2.raw 3.40 km2 26°16'31" S 28°11'06" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	0900	8.00	49	2.18	0.06	13.63	58	5.71	63000	20.7	0.470	.0798	.0605
1997	4	4	0820	7.82	54	3.11	0.08	1.52	48	4.65	60	18.7	0.260	.0212	.0231
1997	5	7	1043	7.20	49	0.28	0.24	5.00	38	3.37	1700	17.3	0.150	.0097	.0062
1997	5	22	0824	7.70	[illegible]	1.86	0.13	24.10	37	4.70	4400	14.7	0.190	.0031	.0025
1997	6	10	0841	7.60	45	1.78	0.25	+0.70	47	5.30	3000	14.0	0.000*	.0043	.0062
1997	6	25	0959	7.80	40	0.40	+0.10	+0.70	77	3.94	2200	14.1	0.130	.0039	.0047
1997	7	10	0950	3.70	[illegible]	3.38	0.37	+0.70	242	4.60	-1	12.8	0.140	.0033	.0043
1997	8	20	1112	8.30	[illegible]	3.80	0.50	+0.70	116	3.14	2600	15.1	0.120	.0019	.0023
1997	9	8	0953	7.60	[illegible]	2.90	0.10	+0.70	82	5.79	1500	15.8	0.240	.0198	.0225
1997	9	18	0908	7.90	[illegible]	2.20	0.90	+0.70	39	5.30	33600	15.7	0.150	.0024	.0029

Table A.3: Observed water quality at station S3

Station : S3 - Vlakplaats Wastewater Treatment Works internal stormwater - low cost residential
ERWAT observations k:\common\wrcdiff\wqdata\s3.raw 1.77 km2 26°21'26" S 28°10'46" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	0930	7.80	47	0.35	0.71	0.21	11	1.94	298000	21.6	0.046	.0400	.0303
1997	4	4	0910	7.67	63	0.15	3.81	47.51	20	1.71	242000	19.7	0.085	.0111	.0121
1997	5	7	1150	7.80	52	0.70	0.88	20.70	5	4.31	218000	12.8	0.000*	.0062	.0039
1997	5	22	0916	7.90	60	0.78	0.72	+0.70	17	4.30	520	10.7	0.000*	.0021	.0017
1997	6	10	0928	8.00	49	0.73	1.85	17.00	5	2.00	4200	9.0	0.000*	.0029	.0042
1997	6	25	1059	-1	-1	-1	-1	-1	-1	-1	-1	-1	0.000	.0026	.0031
1997	7	10	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0.000	.0022	.0029
1997	8	20	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0.000	.0015	.0019
1997	9	8	1104	7.40	27	4.80	0.80	+0.70	49	5.19	26000	15.3	0.013	.0198	.0113
1997	9	18	1021	8.40	63	1.80	0.20	6.20	26	4.32	13	15.3	0.000*	.0014	.0017

Table A.4 : Observed water quality at station S4

Station : S4 - Katlehong (Turn left after crossing spruit on main road) - low cost residential
ERWAT observations   k:\common\wrcdiff\wqdata\s4.raw 2.44 km2 26°21'03" S 28°10'07" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	1030	7.90	25	0.41	0.14	0.63	30	6.00	1000	20.2	0.048	.0500	.0375
1997	4	4	0950	7.84	34	4.83	0.04	2.40	26	5.79	60	17.4	0.000*	.0210	.0229
1997	5	7	1249	7.90	34	0.39	0.01	5.60	44	3.36	20	17.8	-1	.0379	.0050
1997	5	22	0958	8.10	32	0.07	0.05	+0.70	21	7.20	630	10.7	-1	.0024	.0020
1997	6	10	0955	7.80	31	1.07	1.05	1.70	13	7.30	1000	11.0	-1	.0034	.0049
1997	6	25	1143	7.90	27	+0.10	0.21	+0.70	26	0.51	233	12.3	-1	.0031	.0037
1997	7	10	1136	8.10	38	1.19	0.19	+0.70	83	6.40	-1	11.5	0.013	.0026	.0034
1997	8	20	1311	8.10	40	0.50	0.30	+0.70	60	6.95	1200	18.2	0.000*	.0013	.0015
1997	9	8	1200	7.80	51	1.00	14.30	3.40	33	5.75	2150000	15.5	-1	.0187	.0213
1997	9	18	1117	8.50	36	+0.05	0.40	+0.70	26	6.42	590	20.4	-1	.0017	.0021

Table A.5 : Observed water quality at station S5

Station : S5 - Kliprivier/Heidelberg Road R550 bridge at small farm dams - farmland below nature reserve
ERWAT observations   k:\common\wrcdiff\wqdata\s5.raw 13.70 km2 26°25'38" S 28°12'13" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	1110	8.00	26	0.22	0.03	0.00	9	5.00	0	19.2	0.042	.0960	.1237
1997	4	4	1015	7.83	44	0.15	0.01	1.62	94	5.17	15	17.4	0.220	.1489	.1625
1997	5	7	1309	7.80	48	0.27	0.07	8.90	14	4.95	3000	13.2	0.084	.0261	.0166
1997	5	22	1020	8.00	45	0.02	0.04	+0.70	9	5.70	6	9.4	0.044	.0183	.0150
1997	6	10	1017	8.20	33	0.32	0.10	+0.70	22	7.70	7	9.0	0.000*	.0246	.0355
1997	6	25	1207	8.20	37	+0.10	+0.10	+0.70	28	7.69	1	9.7	0.150	.0246	.0293
1997	7	10	1207	8.10	40	0.33	0.25	+0.70	67	5.70	-1	8.8	0.130	.0185	.0240
1997	8	20	1329	8.00	53	0.20	0.60	+0.70	79	6.47	13	16.0	0.000*	.0121	.0144
1997	9	8	1222	8.10	57	0.40	+0.05	+0.70	26	6.02	100	14.3	0.250	.0118	.0134
1997	9	18	1138	8.20	56	+0.05	0.40	+0.70	35	4.87	34	14.7	0.300	.0070	.0085

Table A.6 : Observed water quality at station S6

Station : S6 - R103 Heidelberg (Turn right out of R550, 600m) - Rural cultivated
ERWAT observations   k:\common\wrcdiff\wqdata\s6.raw 8.77 km2 26°25'00" S 28°17'09" E

Year	Mon	Day	Time	pH	EC	NO3	TPO4	NH4	SO4	DO	F.coli	T	O.Flow	M.Flow	P.Flow
				-	mS/m	mg/l	mg/l	mg/l	mg/l	mg/l	N/100ml	°C	m3/s	m3/s	m3/s
1997	3	10	1230	7.80	19	0.23	0.14	0.00	27	3.67	700	20.2	0.008	.0960	.0728
1997	4	4	1125	7.72	24	0.11	0.01	1.42	20	3.48	105	17.9	0.005	.0864	.0923
1997	5	7	1329	7.30	27	0.51	0.04	5.00	53	1.53	5900	11.6	0.000*	.0150	.0096
1997	5	22	1045	7.20	24	0.19	0.23	+0.70	2	2.60	2	7.8	0.000*	.0111	.0091
1997	6	10	1046	7.90	22	0.33	0.17	+0.70	21	6.10	4	7.0	0.000*	.0148	.0214
1997	6	25	1226	8.00	28	+0.10	+0.10	+0.70	36	5.73	3	6.9	0.000*	.0148	.0177
1997	7	10	1227	7.90	33	0.20	0.33	+0.70	11	5.10	-1	5.8	0.000*	.0111	.0144
1997	8	20	1349	7.70	42	0.20	0.60	+0.70	13	3.97	4	9.6	0.000*	.0074	.0088
1997	9	8	1242	7.60	67	0.40	+0.05	+0.70	42	4.01	1800	12.0	0.014	.0073	.0083
1997	9	18	1156	8.00	45	+0.05	0.50	+0.70	18	4.11	200	11.0	0.008	.0039	.0047

APPENDIX B

REGRESSION PLOTS

	Description	Page

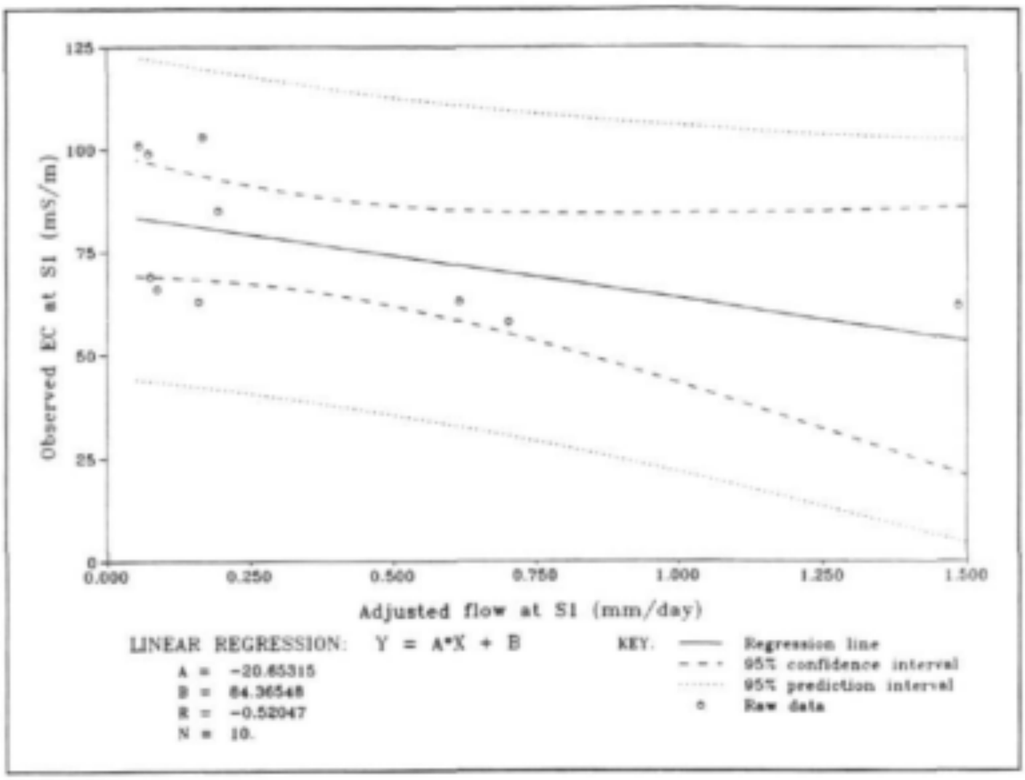


Figure B.1.1 : Linear regression between flow and electrical conductivity at station S1

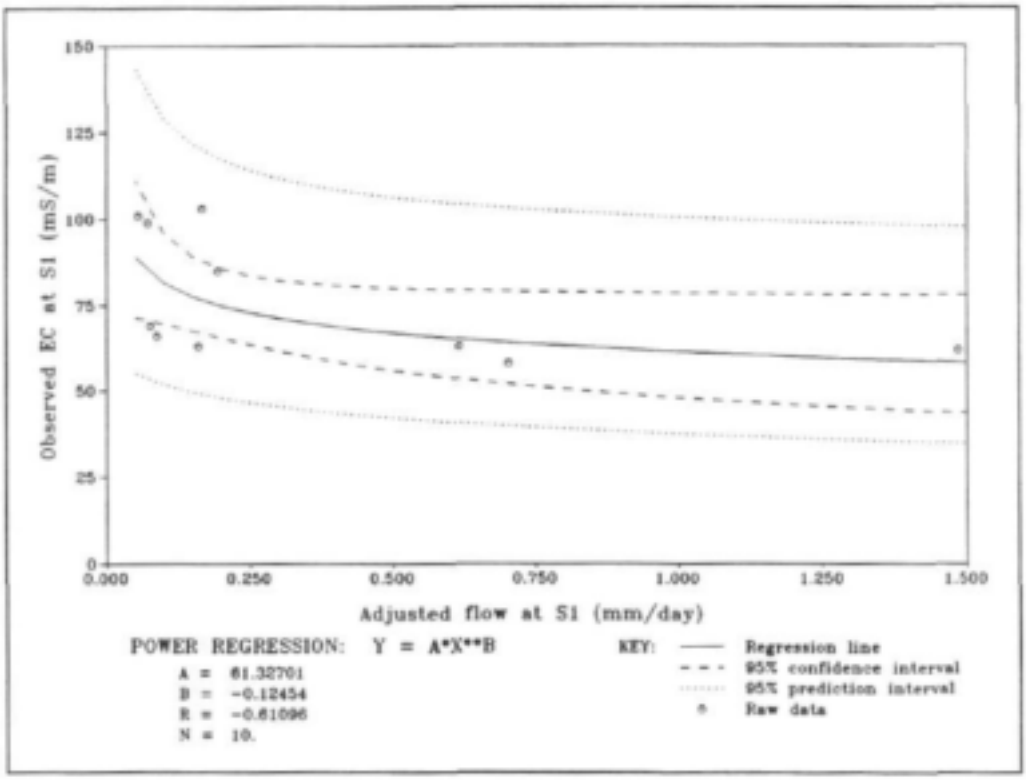


Figure B.1.2 : Power regression between flow and electrical conductivity at station S1

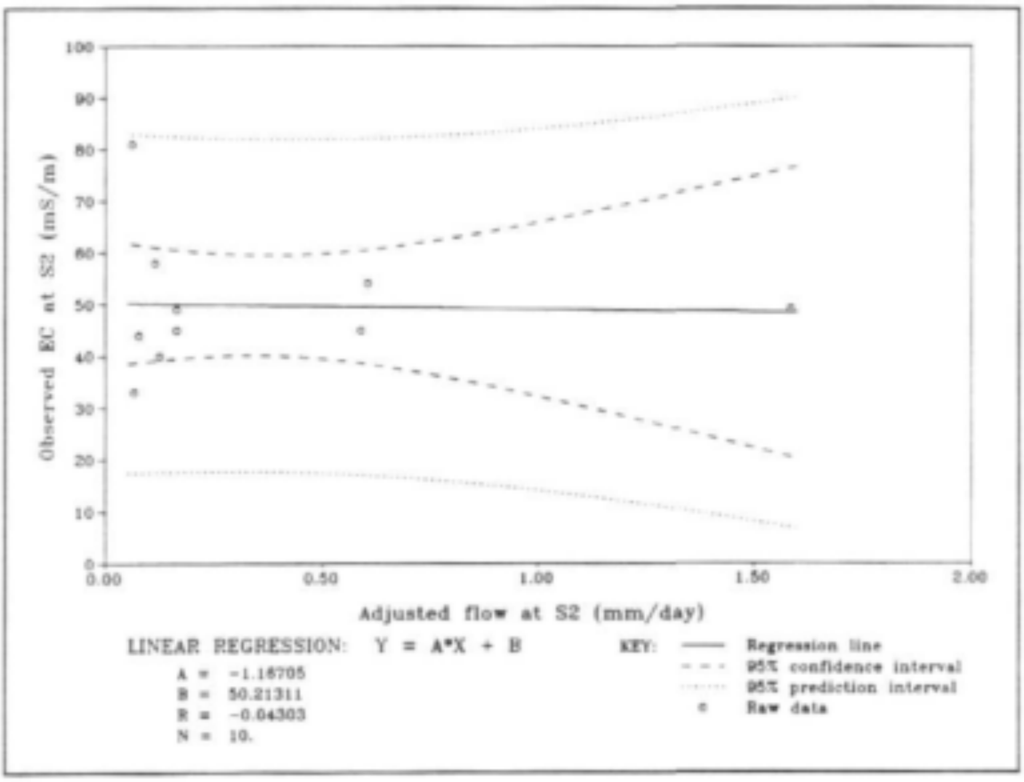


Figure B.1.3 : Linear regression between flow and electrical conductivity at station S2

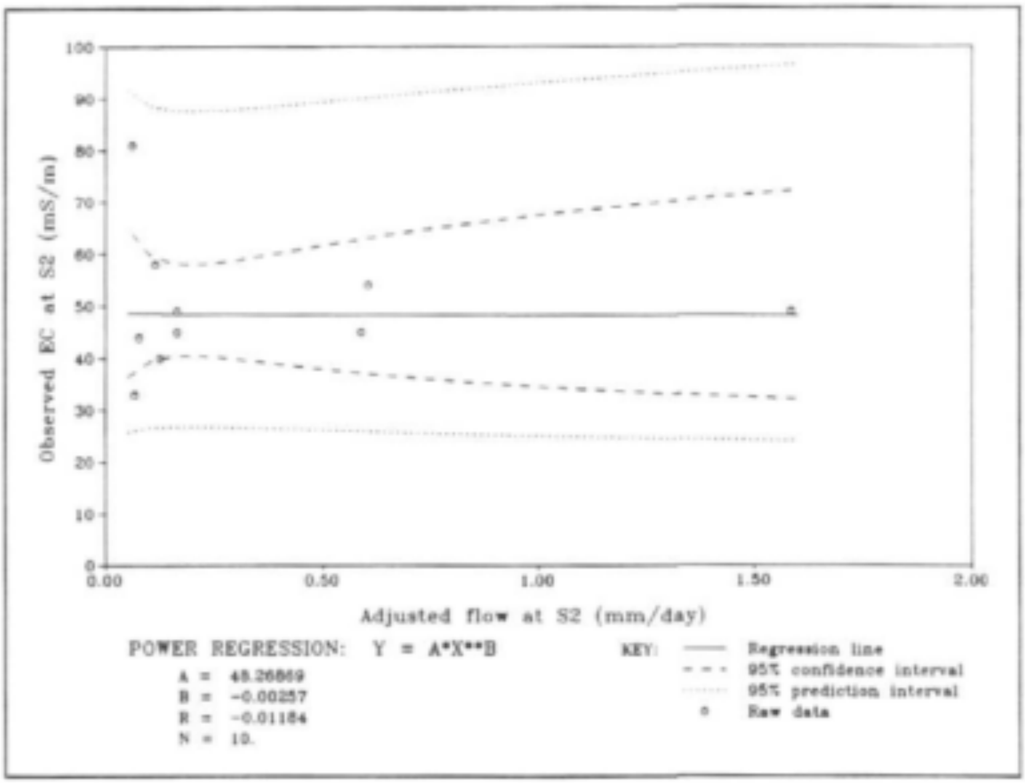


Figure B.1.4 : Power regression between flow and electrical conductivity at station S2

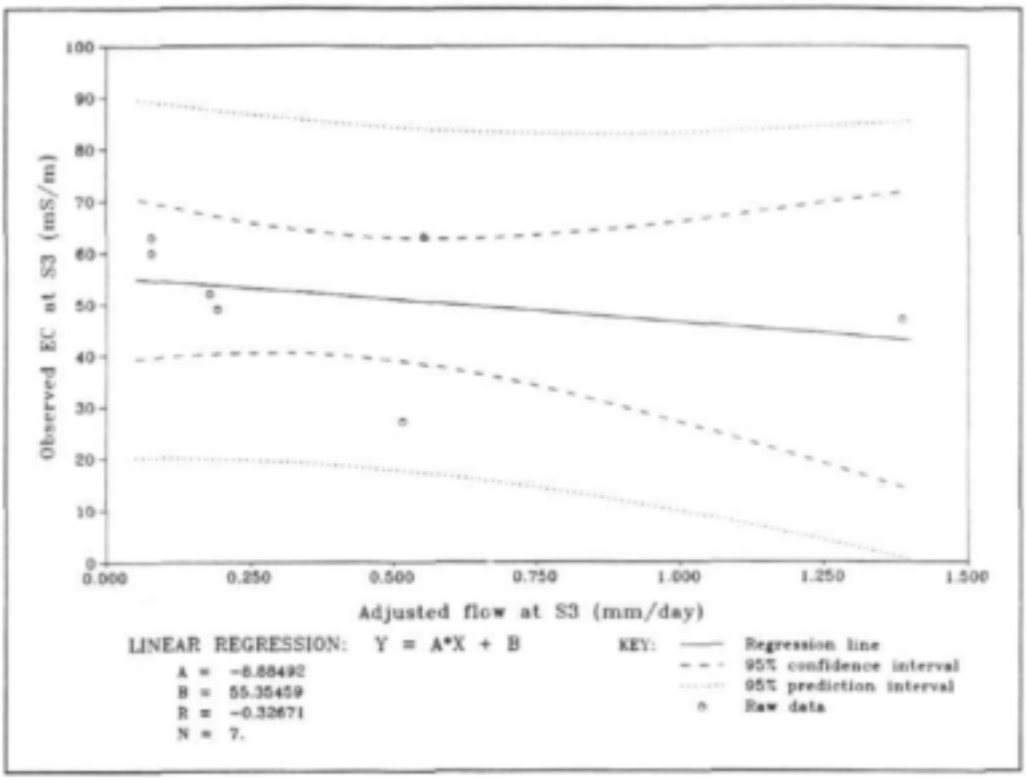


Figure B.1.5 : Linear regression between flow and electrical conductivity at station S3

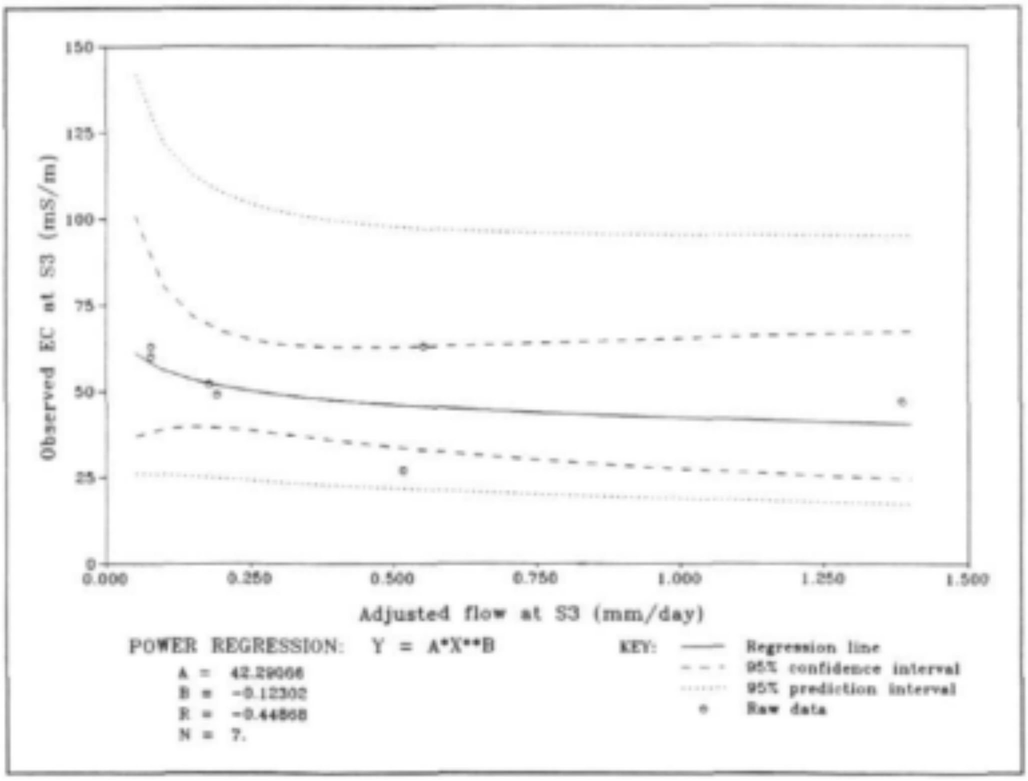


Figure B.1.6 : Power regression between flow and electrical conductivity at station S3

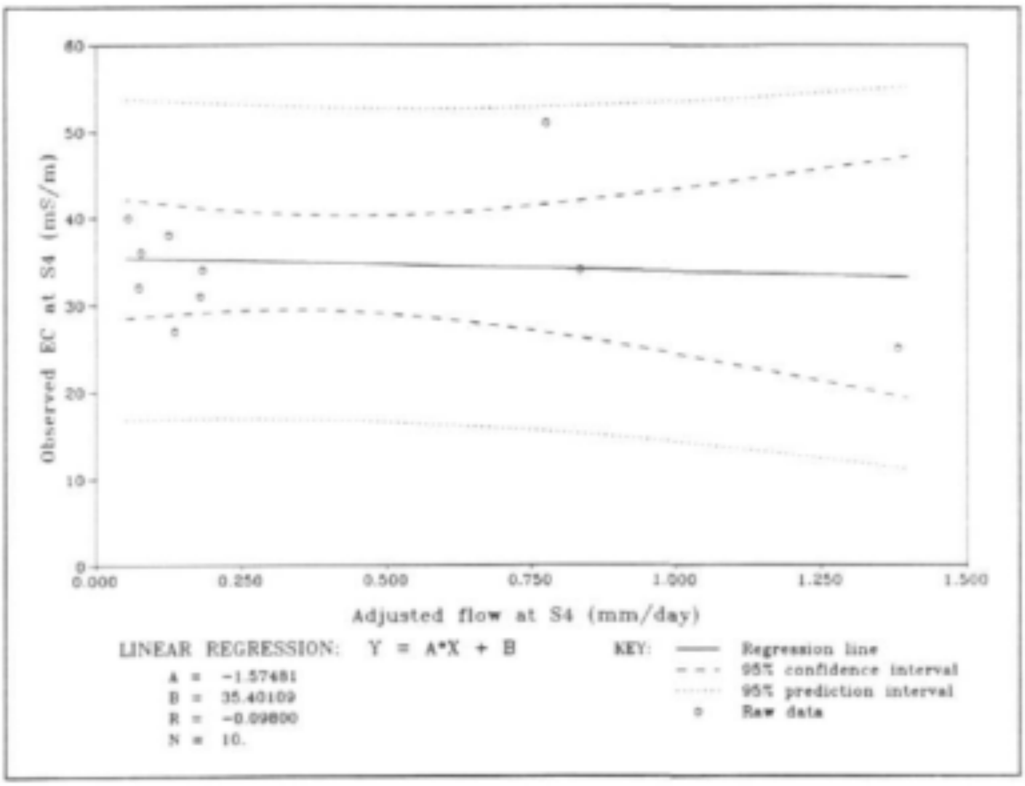


Figure B.1.7 : Linear regression between flow and electrical conductivity at station S4

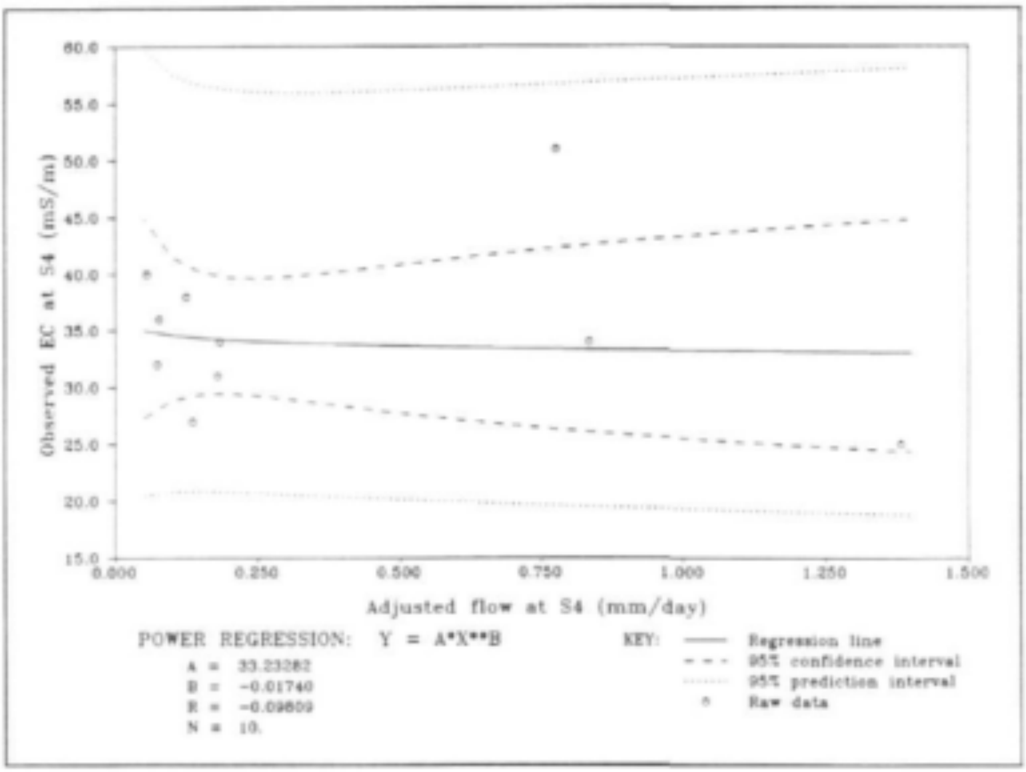


Figure B.1.8 : Power regression between flow and electrical conductivity at station S4

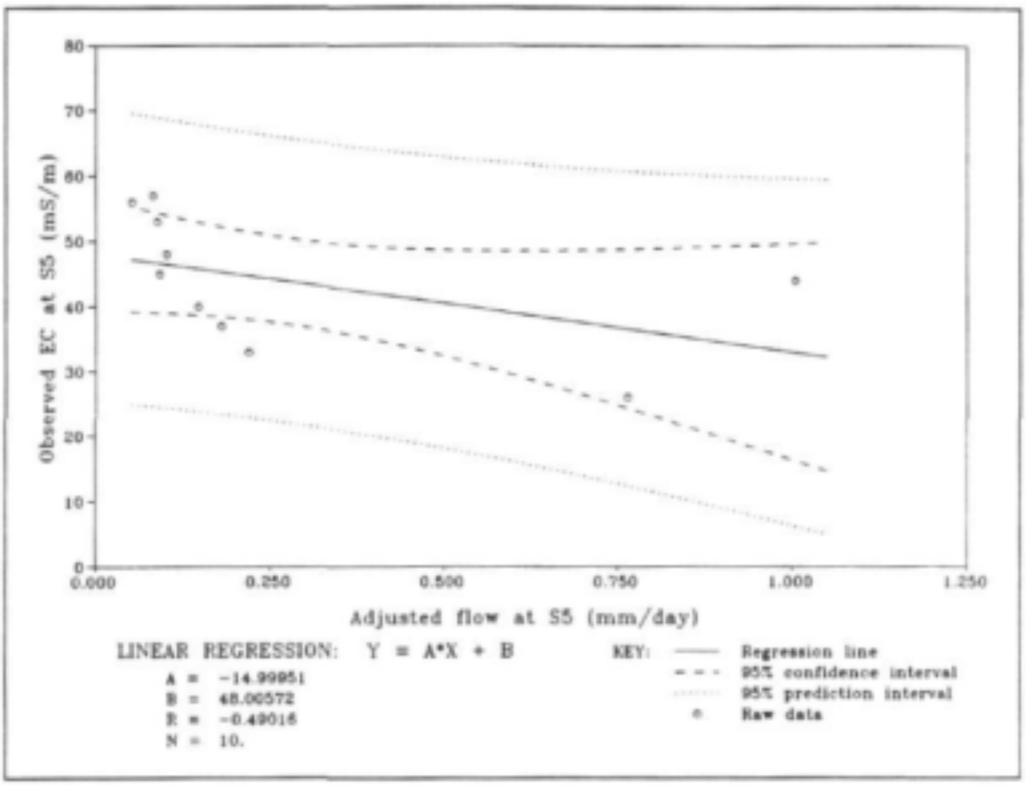


Figure B.1.9 : Linear regression between flow and electrical conductivity at station S5

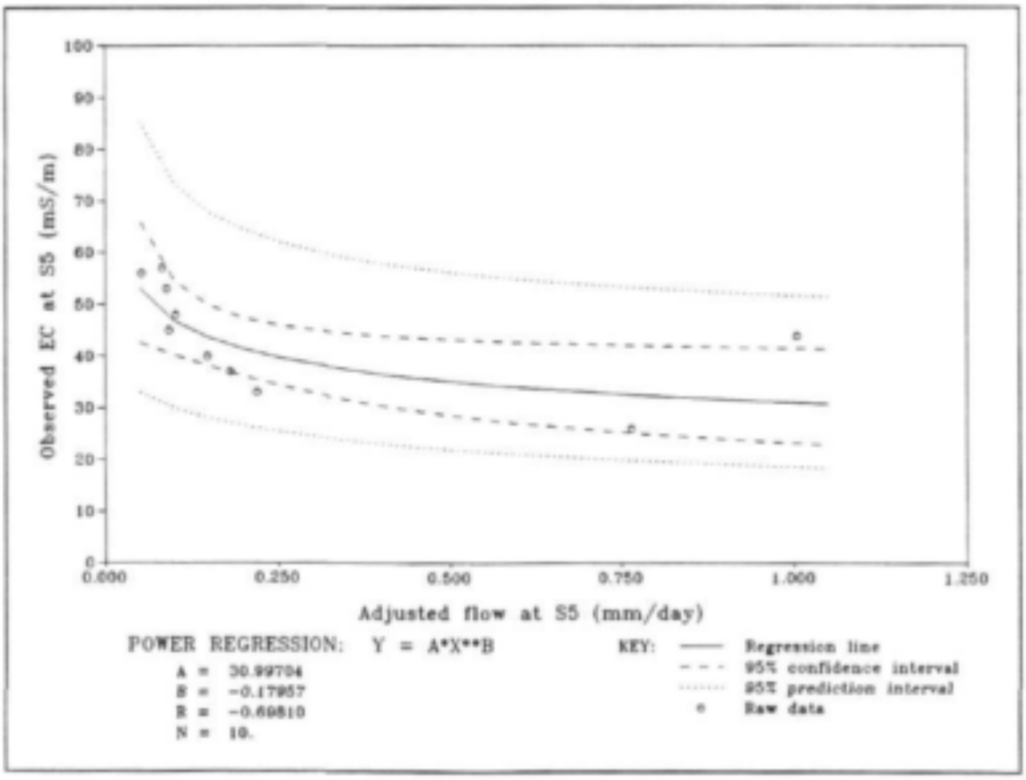


Figure B.1.10 : Power regression between flow and electrical conductivity at station S5

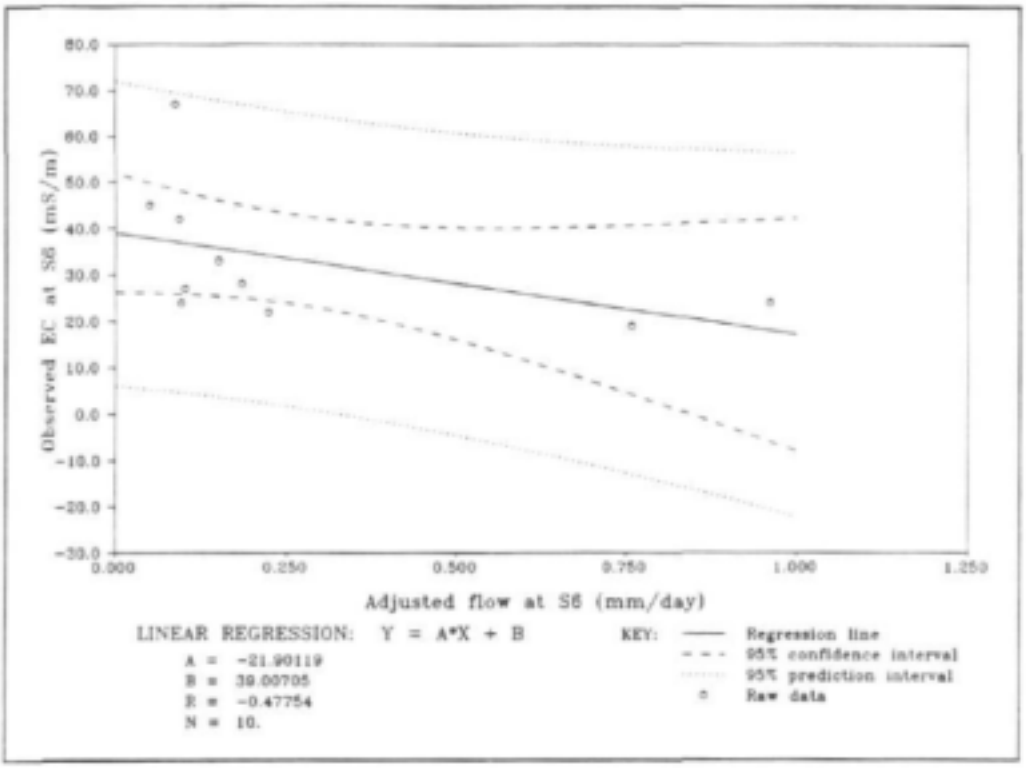


Figure B.1.11 : Linear regression between flow and electrical conductivity at station S6

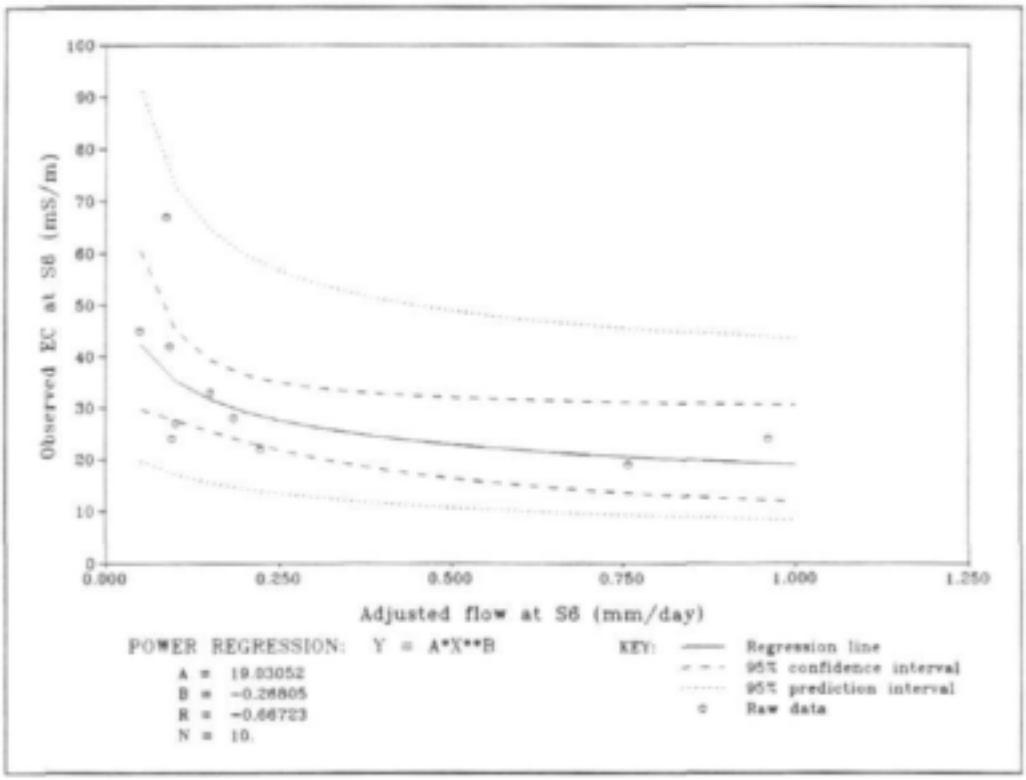


Figure B.1.12 : Power regression between flow and electrical conductivity at station S6

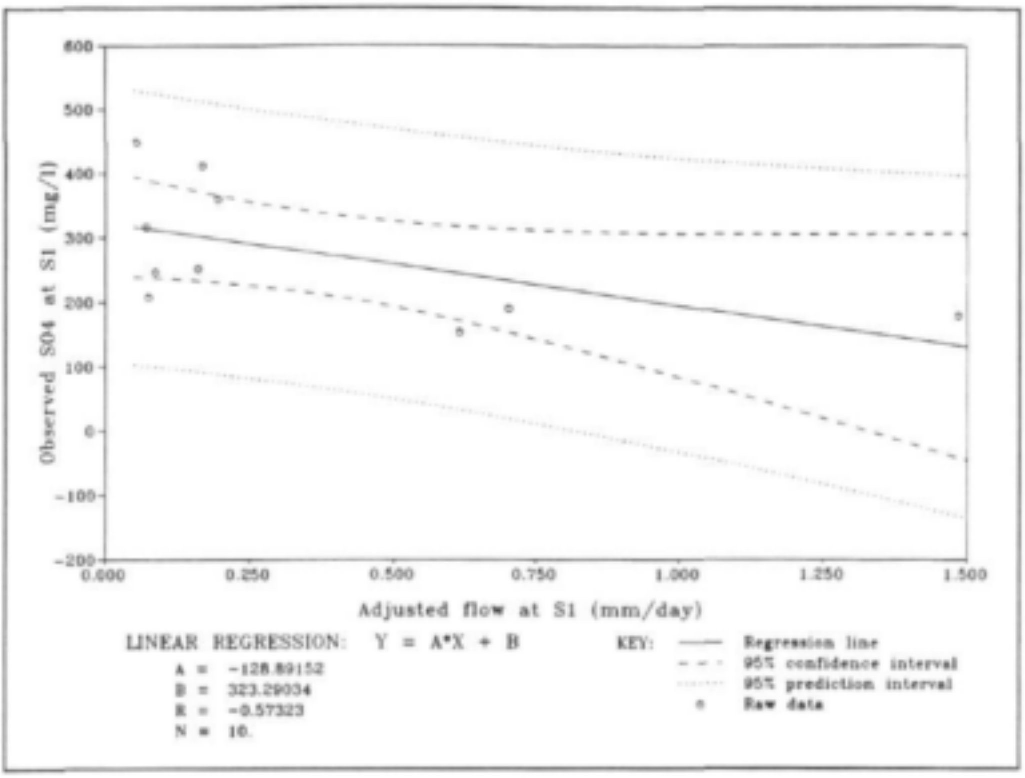


Figure B.2.1 : Linear regression between flow and sulphate at station S1

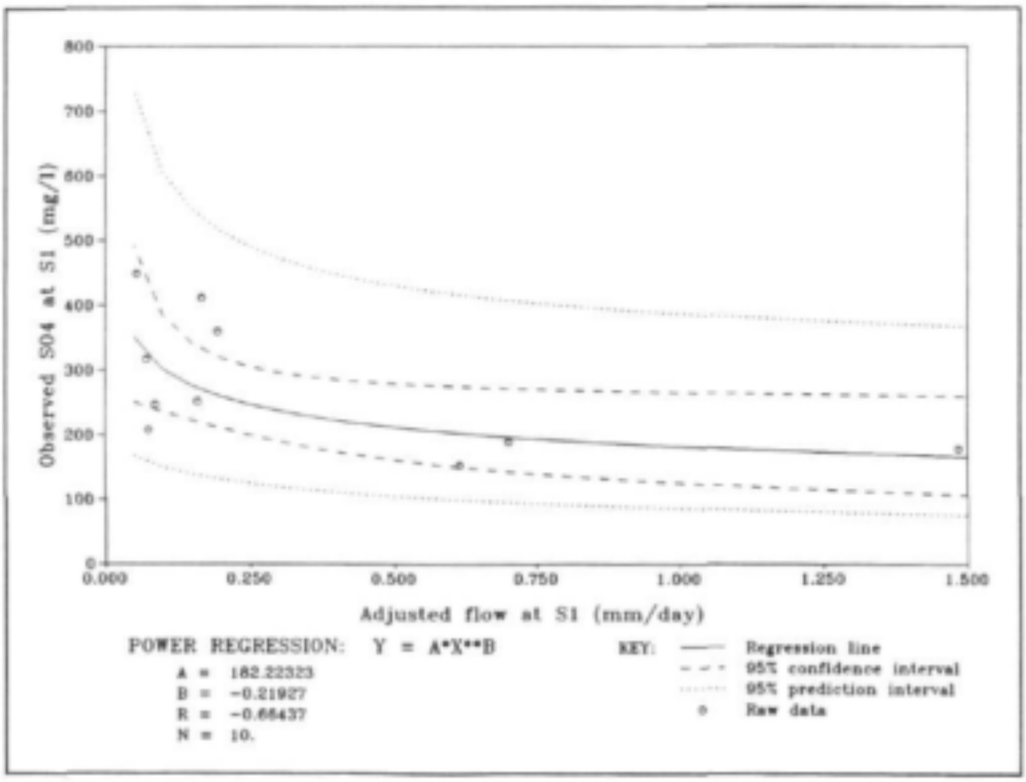


Figure B.2.2 : Power regression between flow and sulphate at station S1

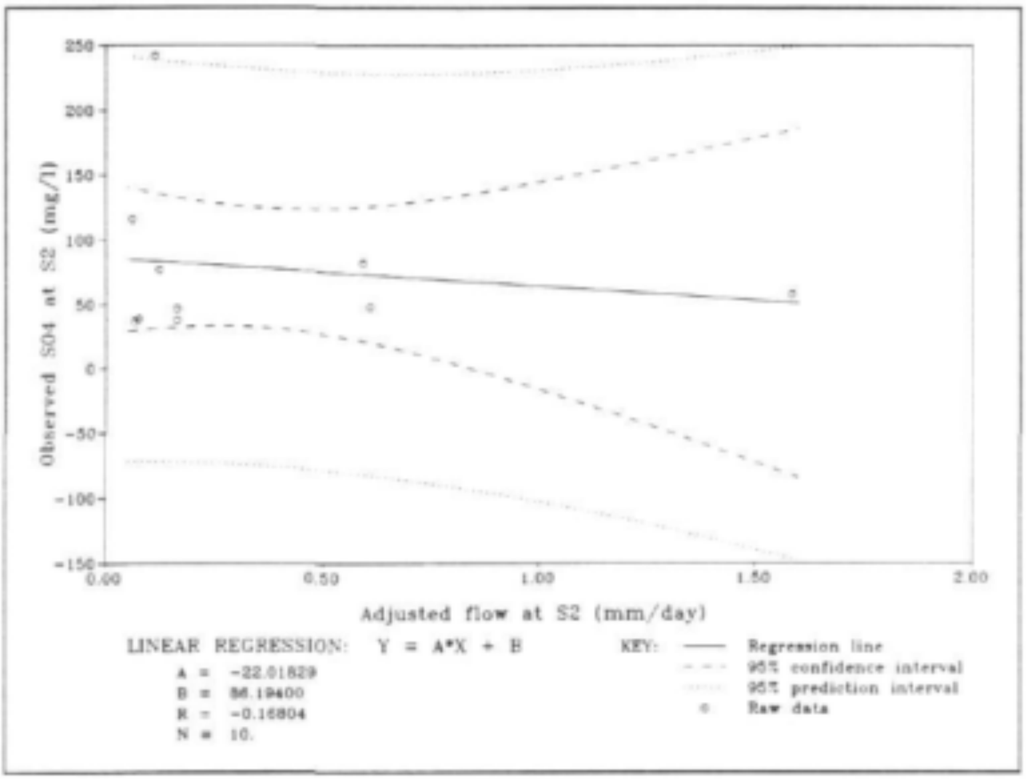


Figure B.2.3 : Linear regression between flow and sulphate at station S2

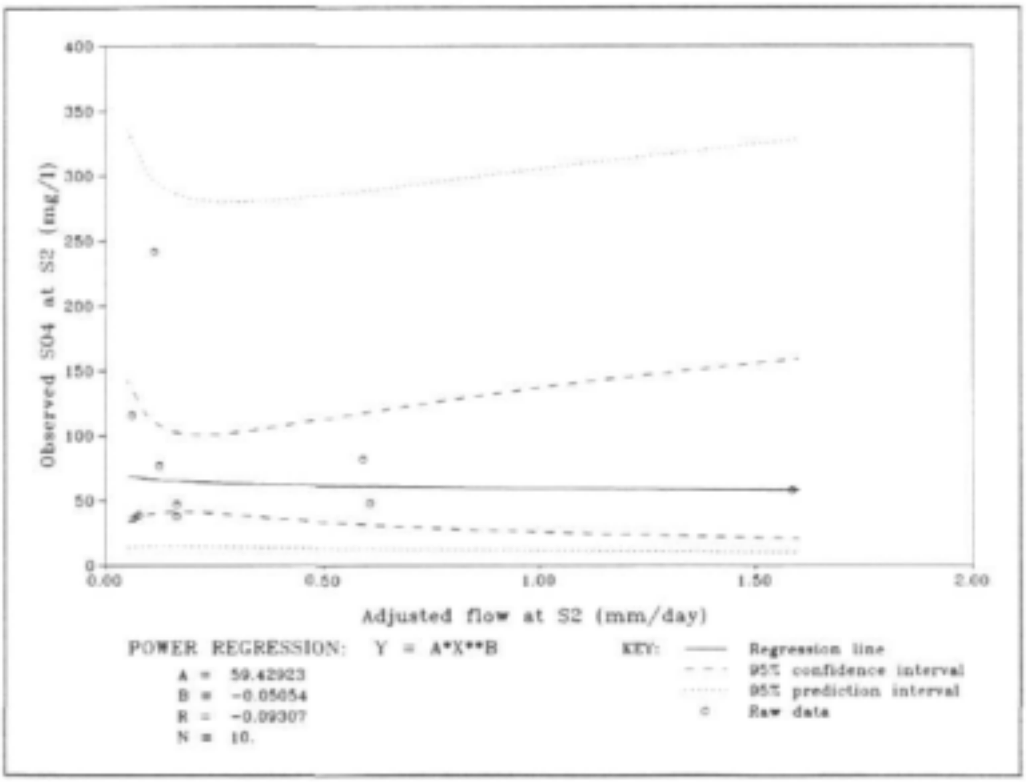


Figure B.2.4 : Power regression between flow and sulphate at station S2

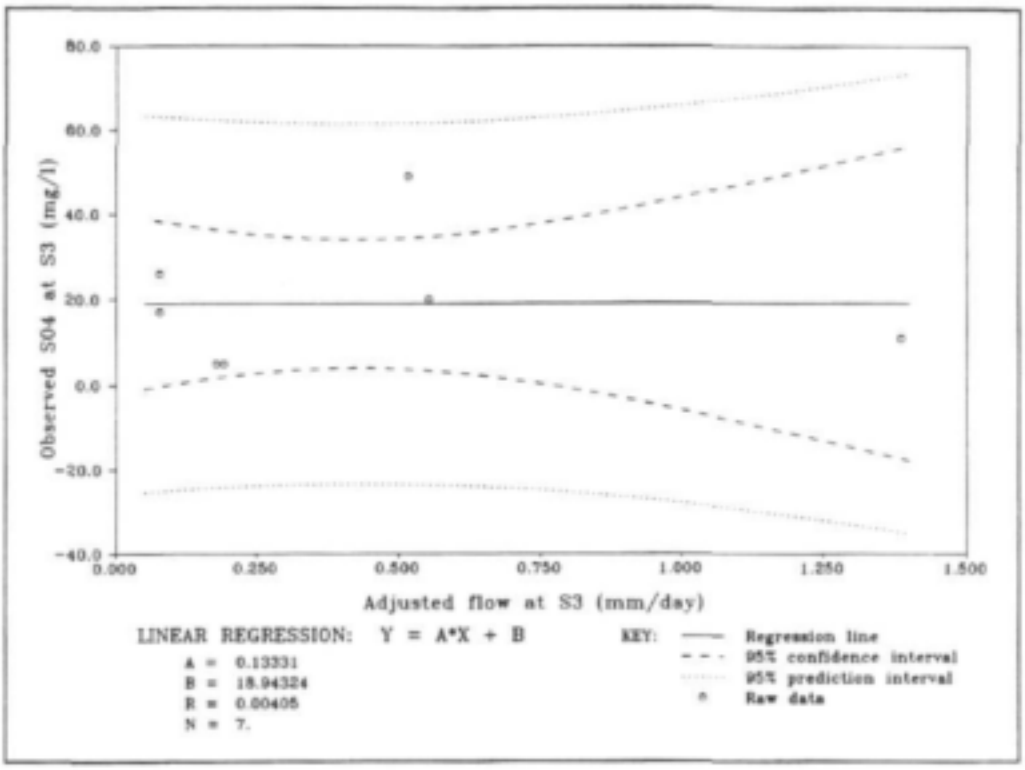


Figure B.2.5 : Linear regression between flow and sulphate at station S3

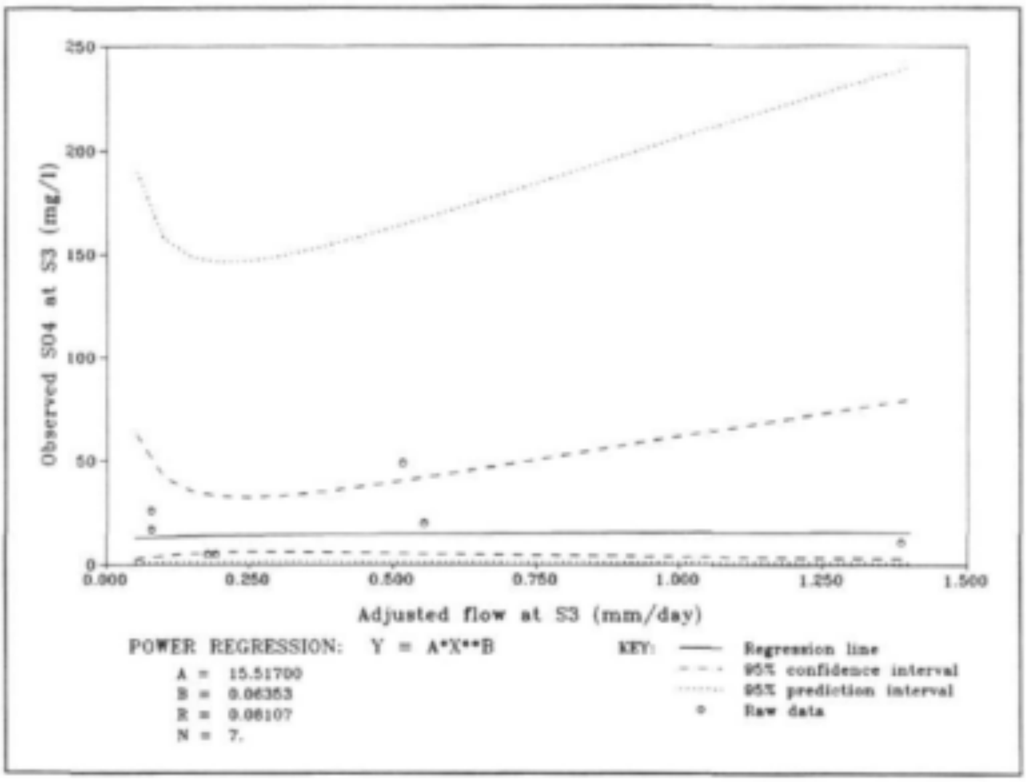


Figure B.2.6 : Power regression between flow and sulphate at station S3

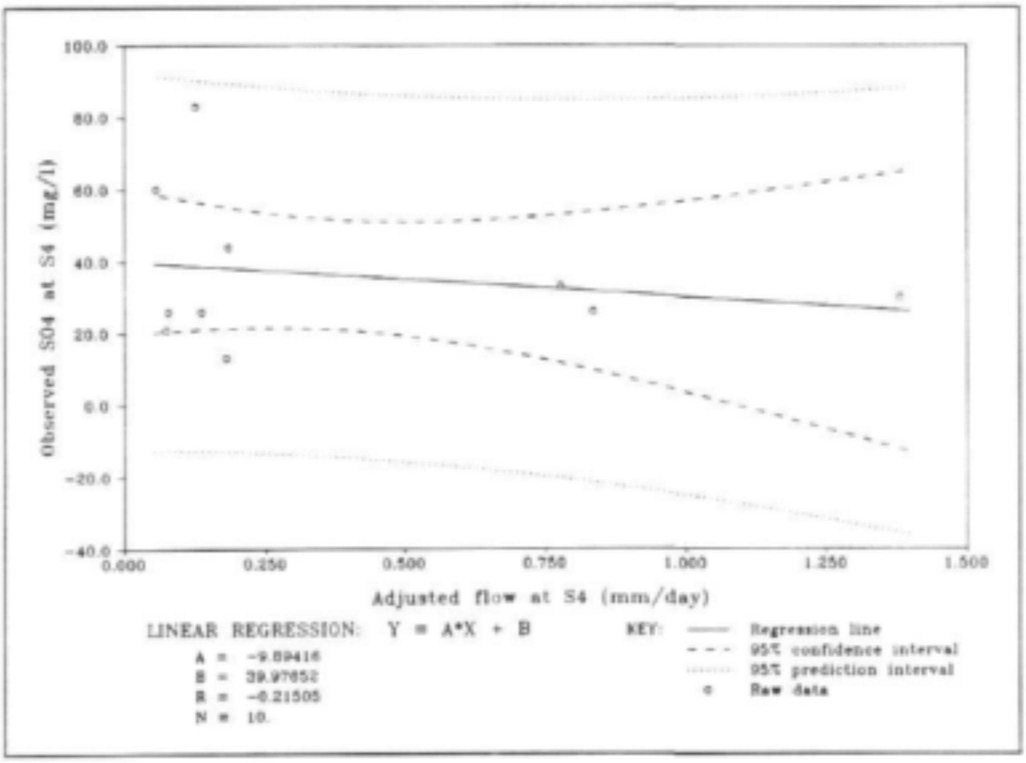


Figure B.2.7 : Linear regression between flow and sulphate at station S4

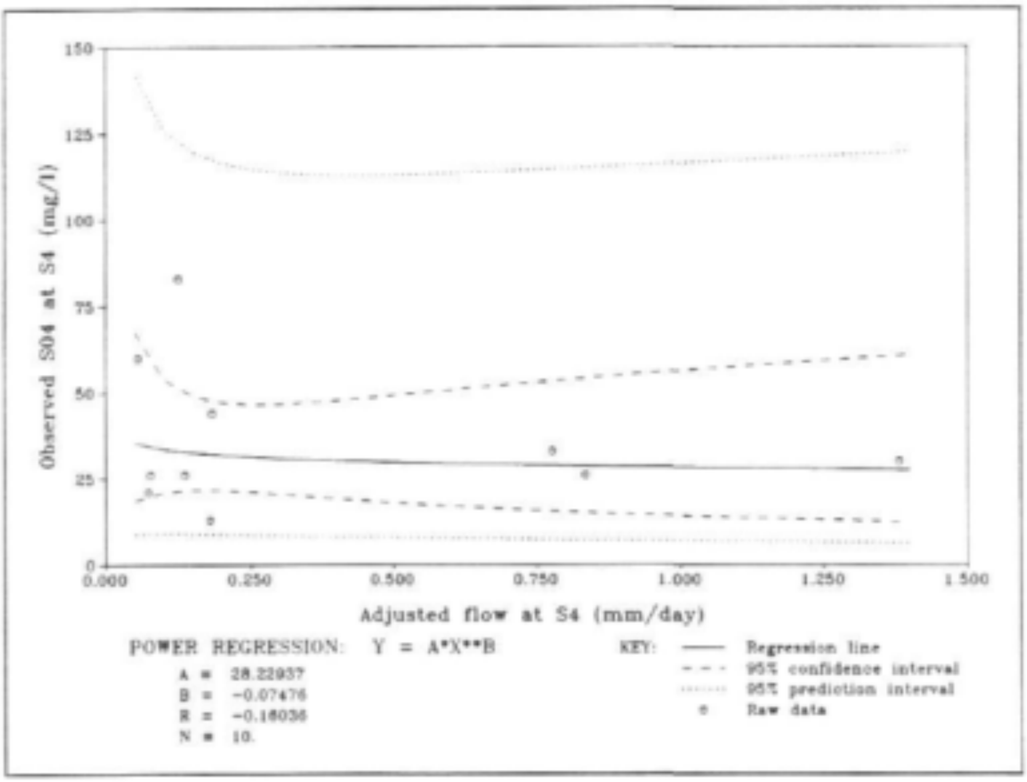


Figure B.2.8 : Power regression between flow and sulphate at station S4

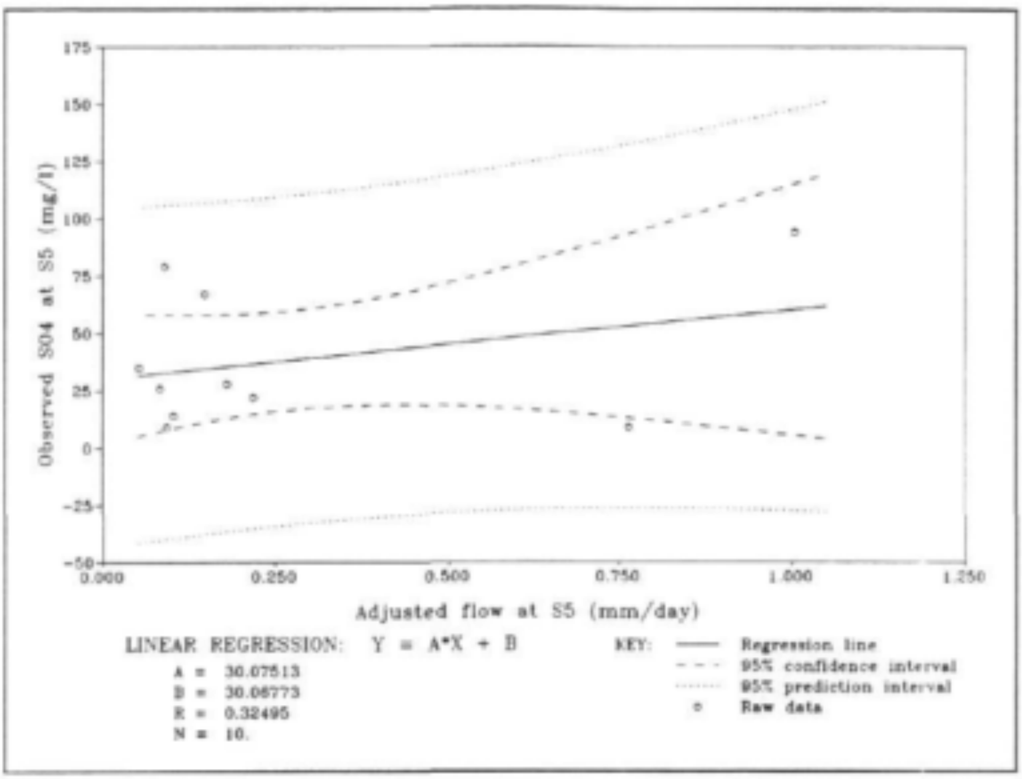


Figure B.2.9 : Linear regression between flow and sulphate at station S5

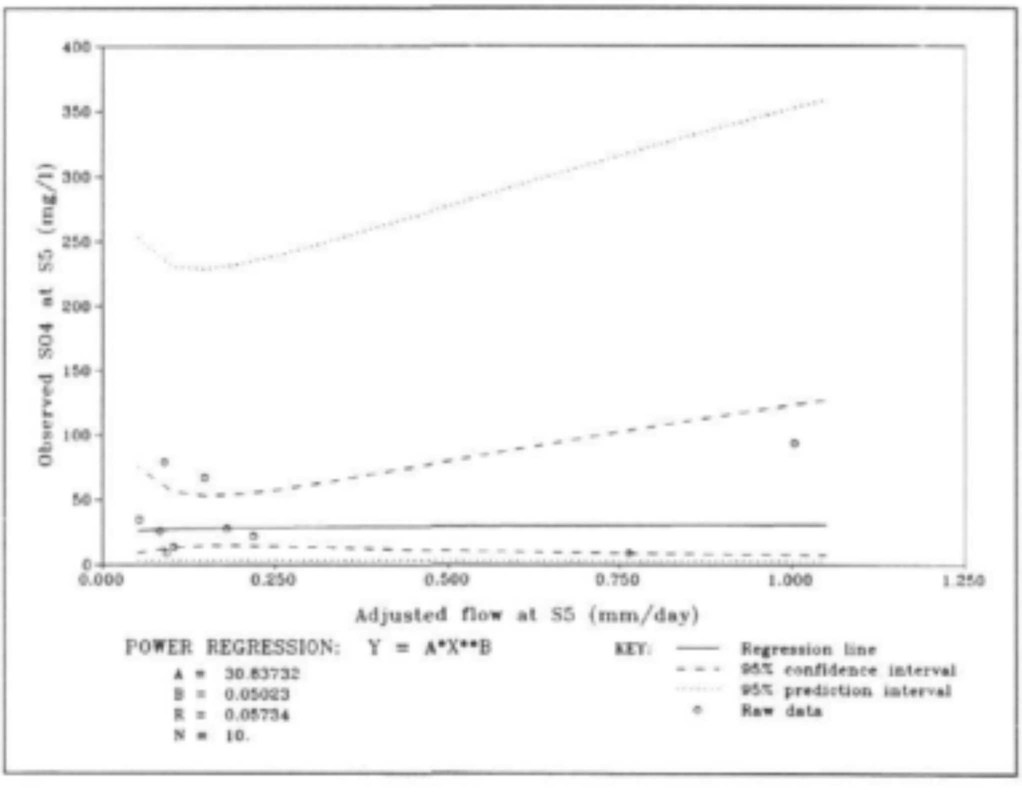


Figure B.2.10 : Power regression between flow and sulphate at station S5

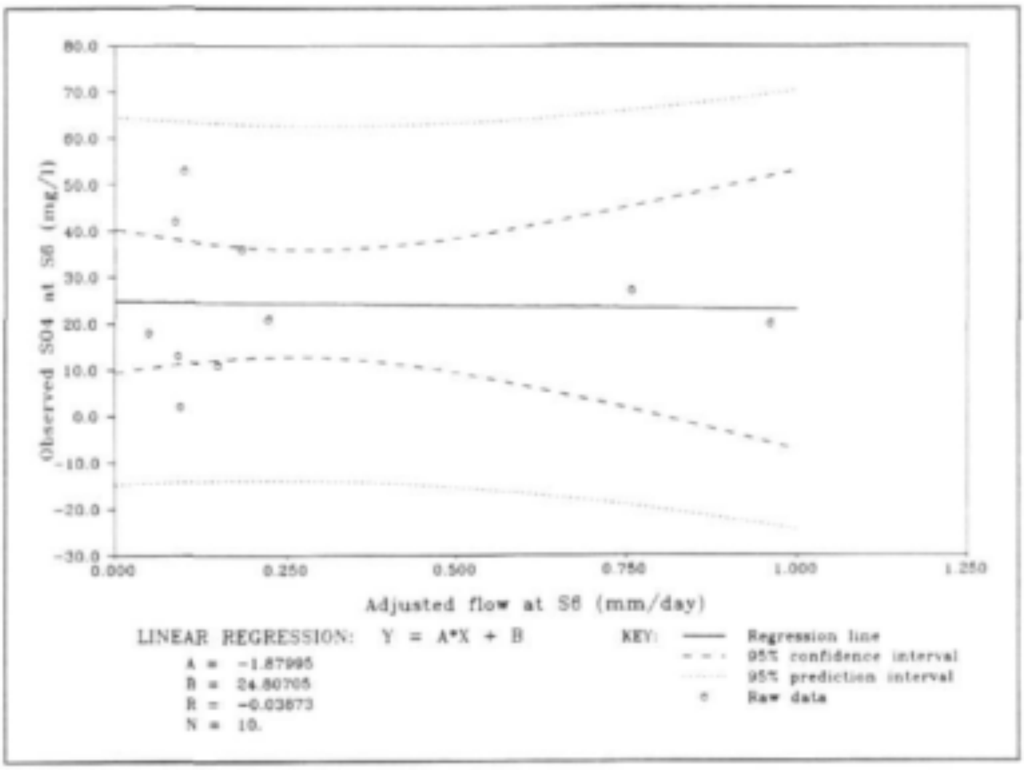


Figure B.2.11 : Linear regression between flow and sulphate at station S6

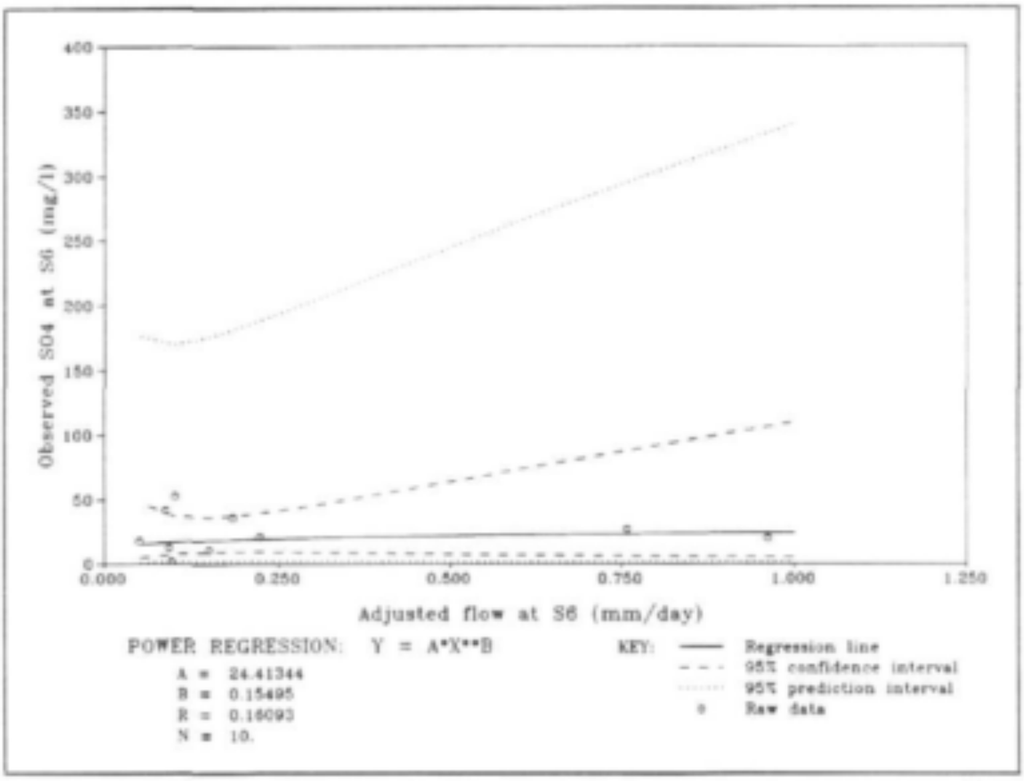


Figure B.2.12 : Power regression between flow and sulphate at station S6

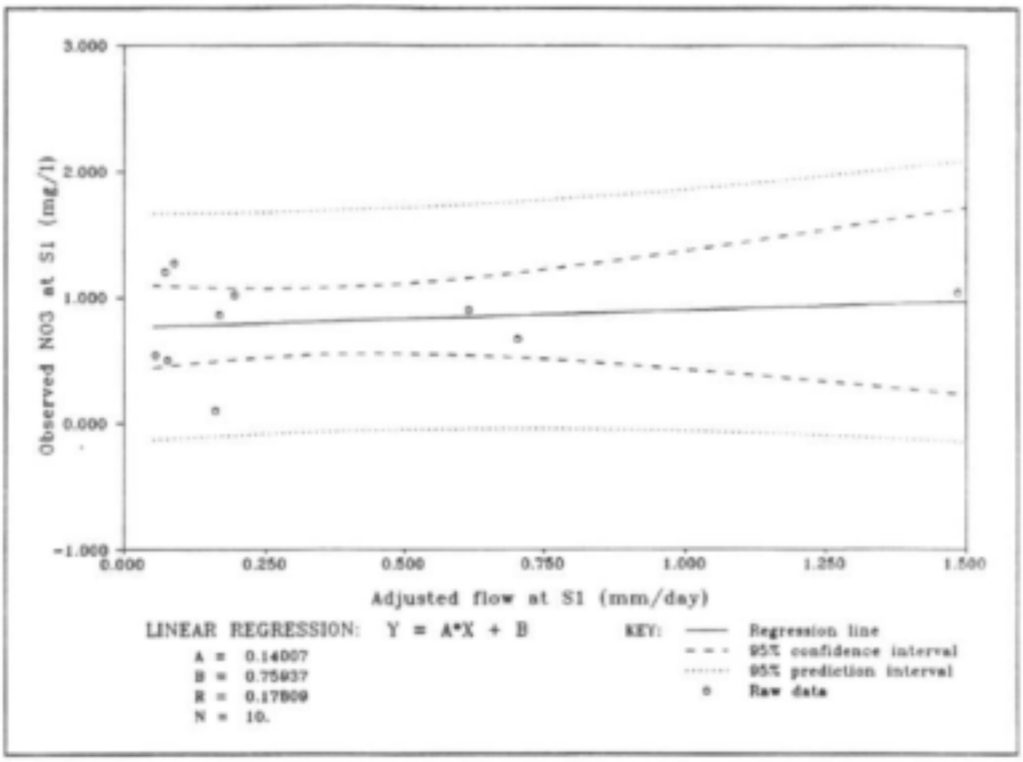


Figure B.3.1 : Linear regression between flow and nitrate at station S1

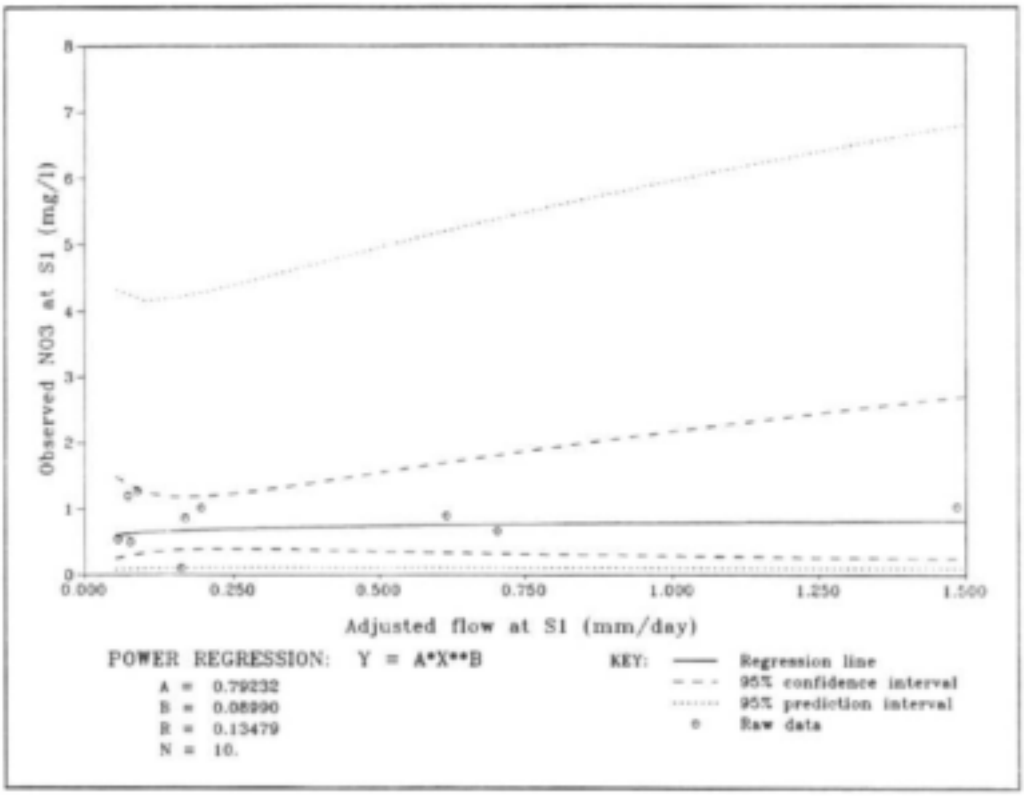


Figure B.3.2 : Power regression between flow and nitrate at station S1

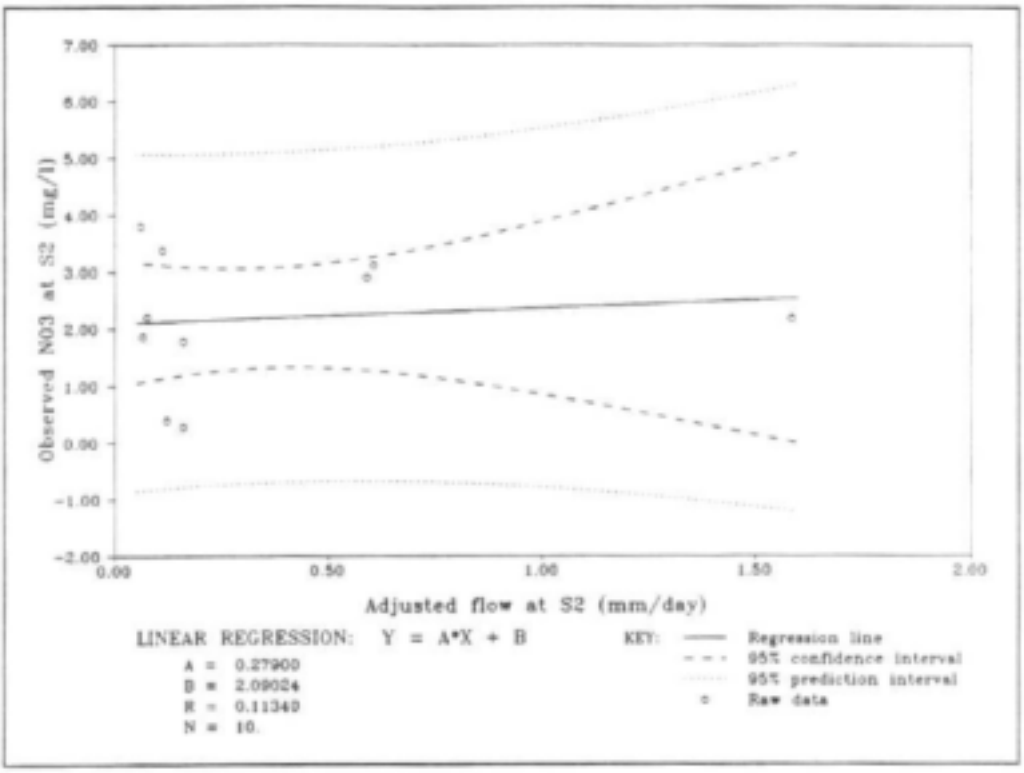


Figure B.3.3 : Linear regression between flow and nitrate at station S2

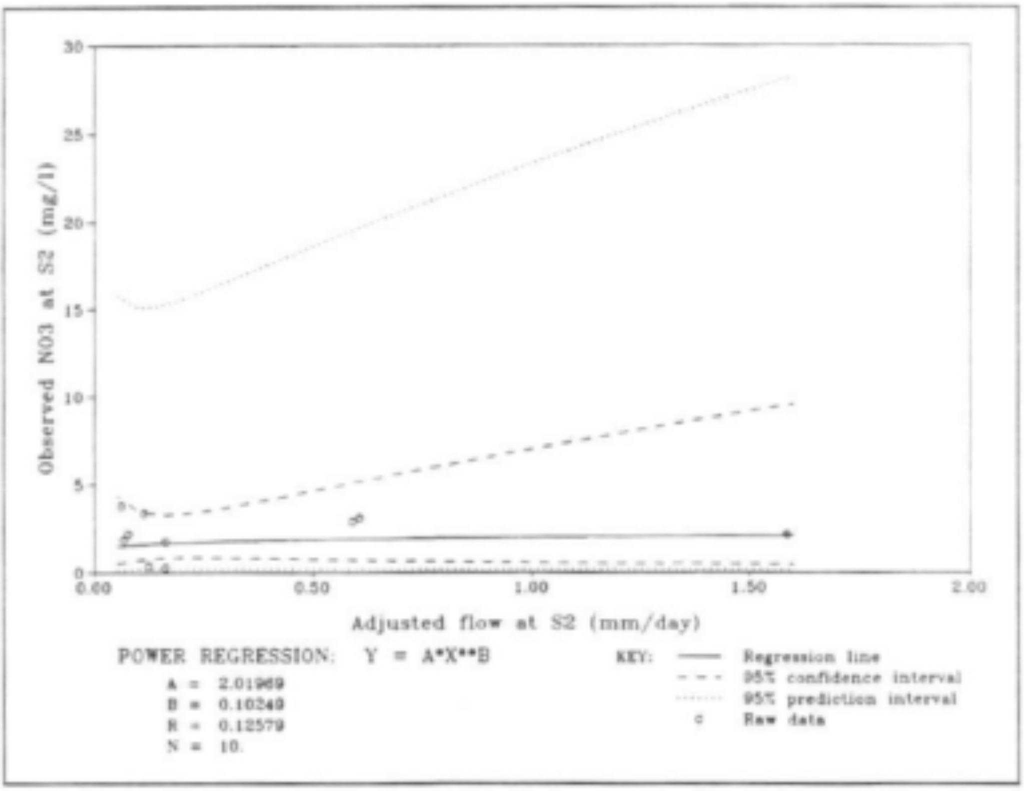


Figure B.3.4 : Power regression between flow and nitrate at station S2

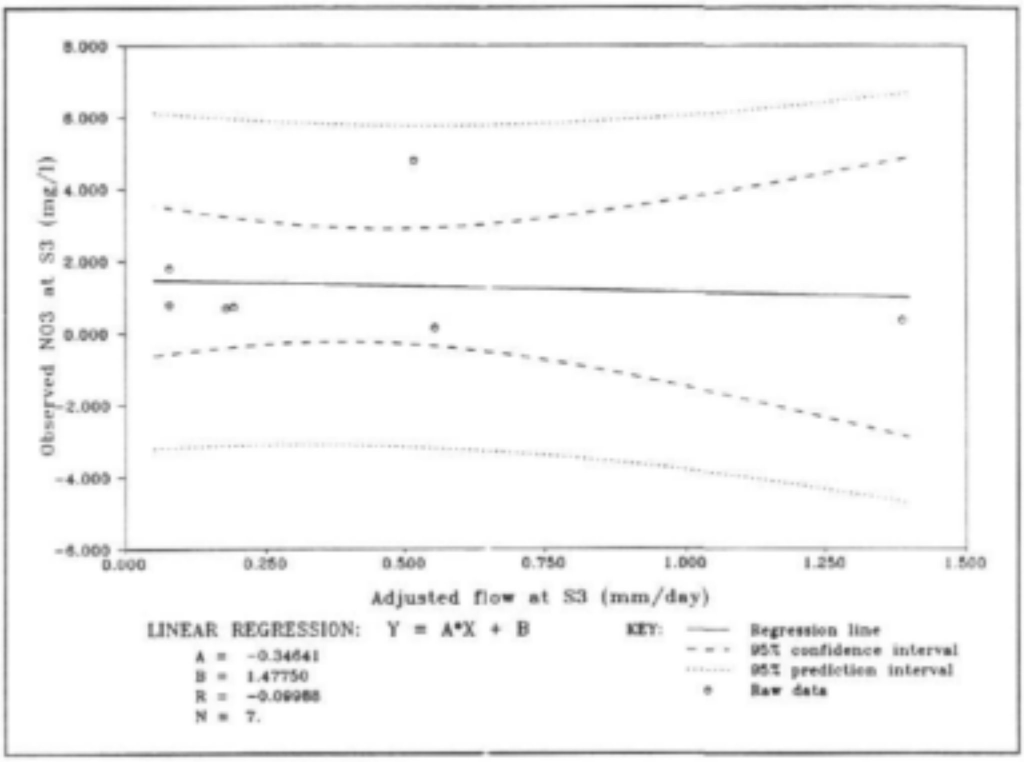


Figure B.3.5 : Linear regression between flow and nitrate at station S3

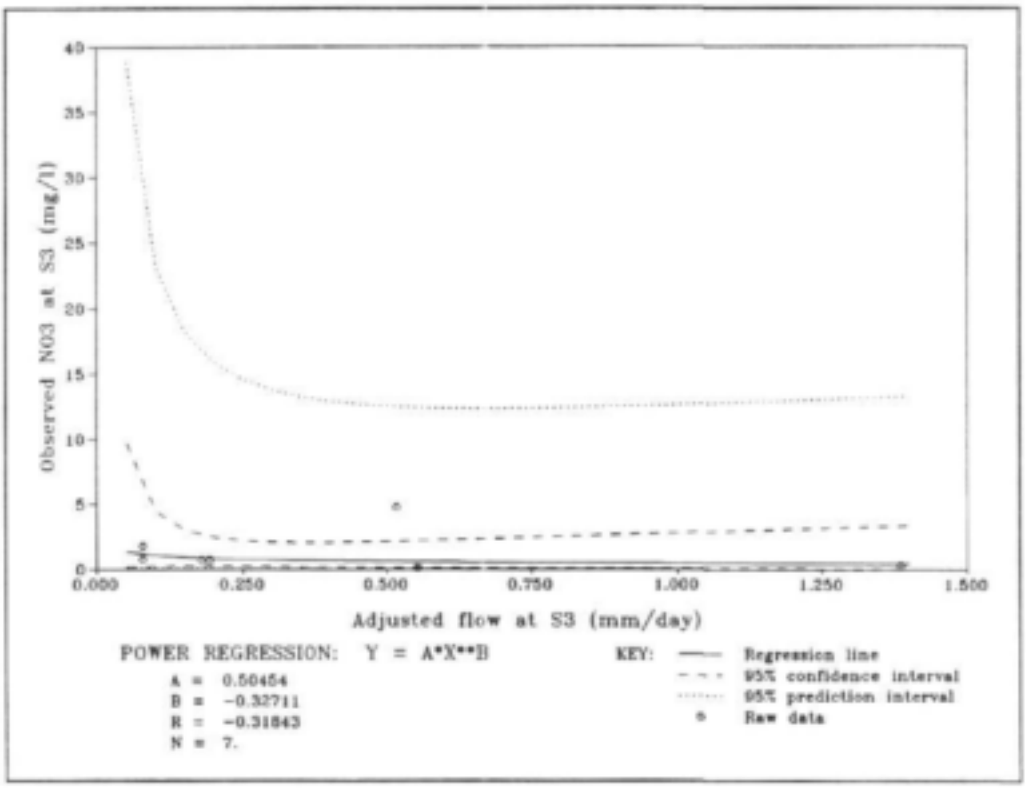


Figure B.3.6 : Power regression between flow and nitrate at station S3

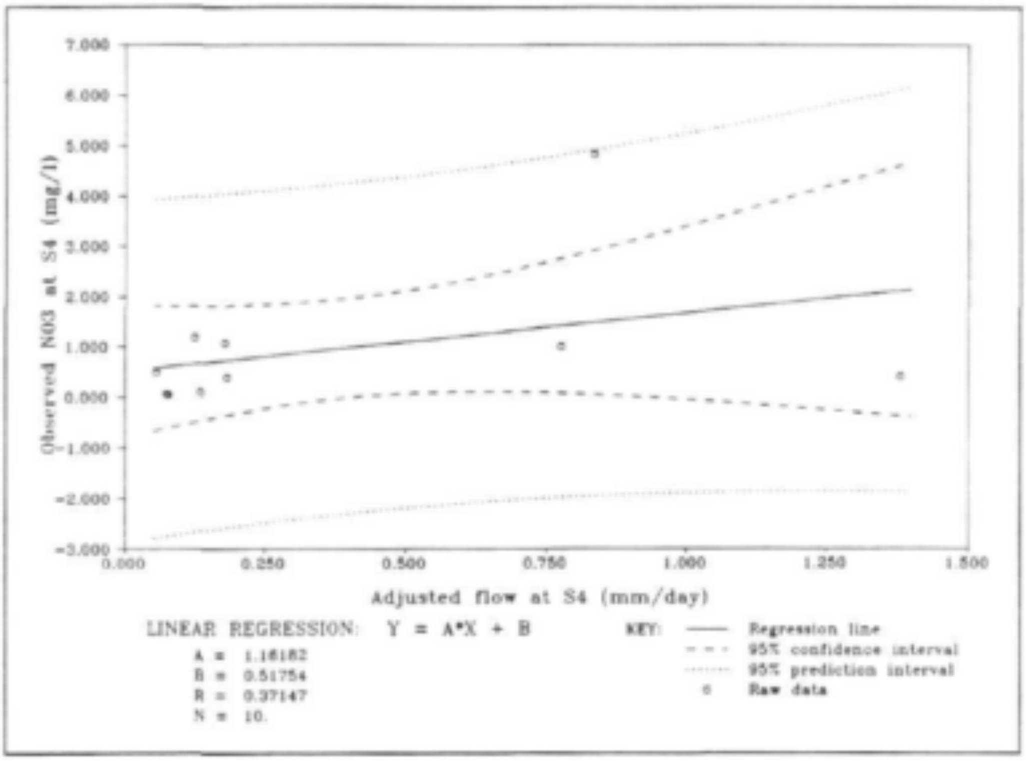


Figure B.3.7 : Linear regression between flow and nitrate at station S4

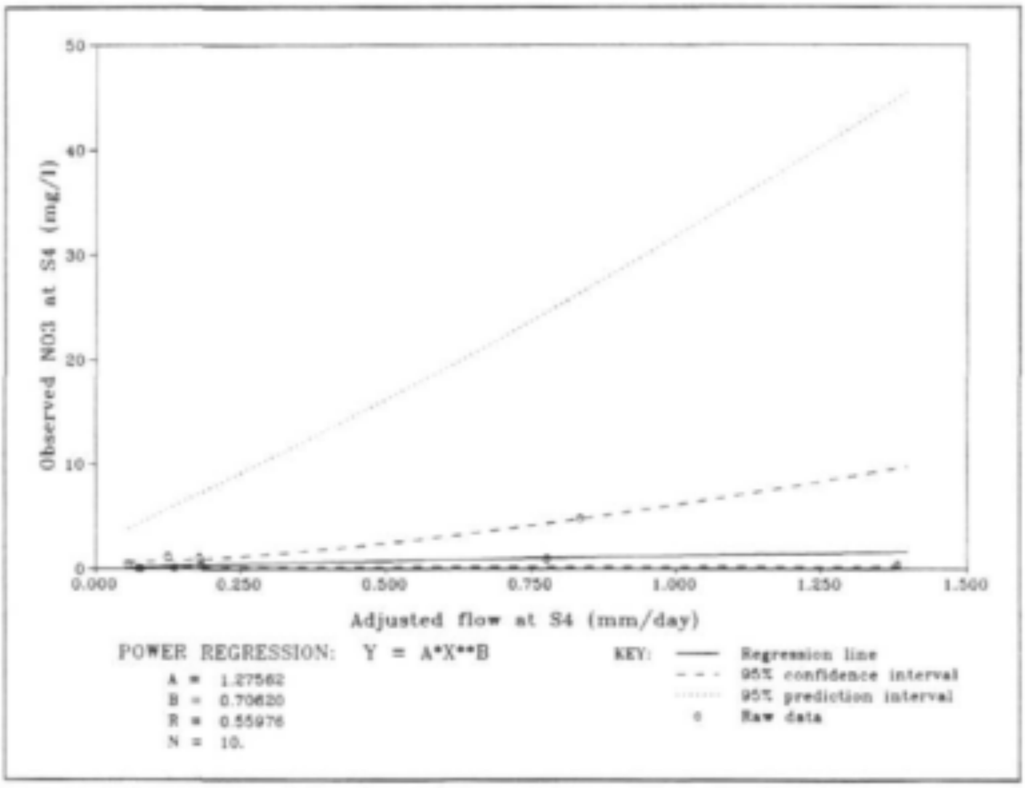


Figure B.3.8 : Power regression between flow and nitrate at station S4

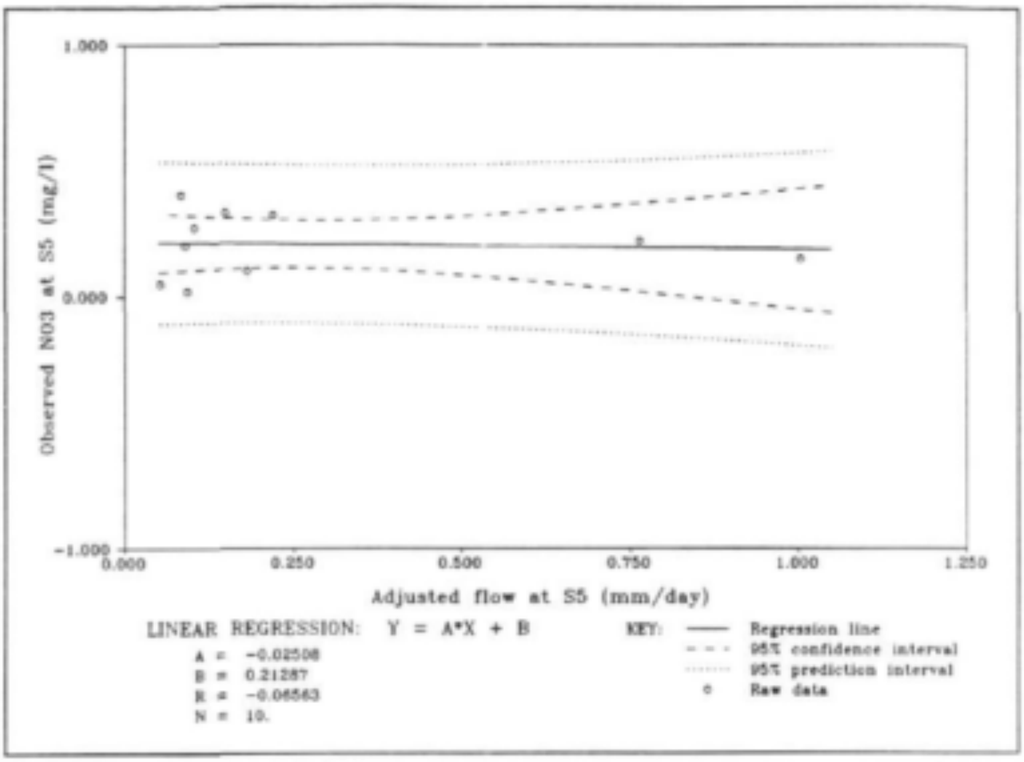


Figure B.3.9 : Linear regression between flow and nitrate at station S5

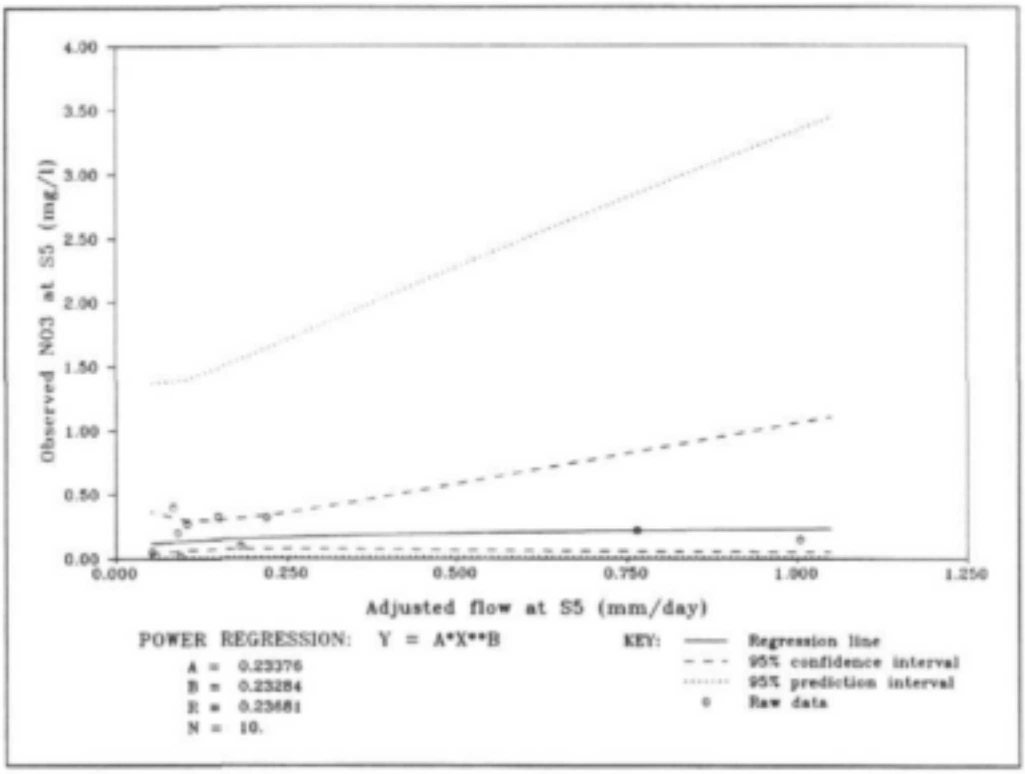


Figure B.3.10 : Power regression between flow and nitrate at station S5

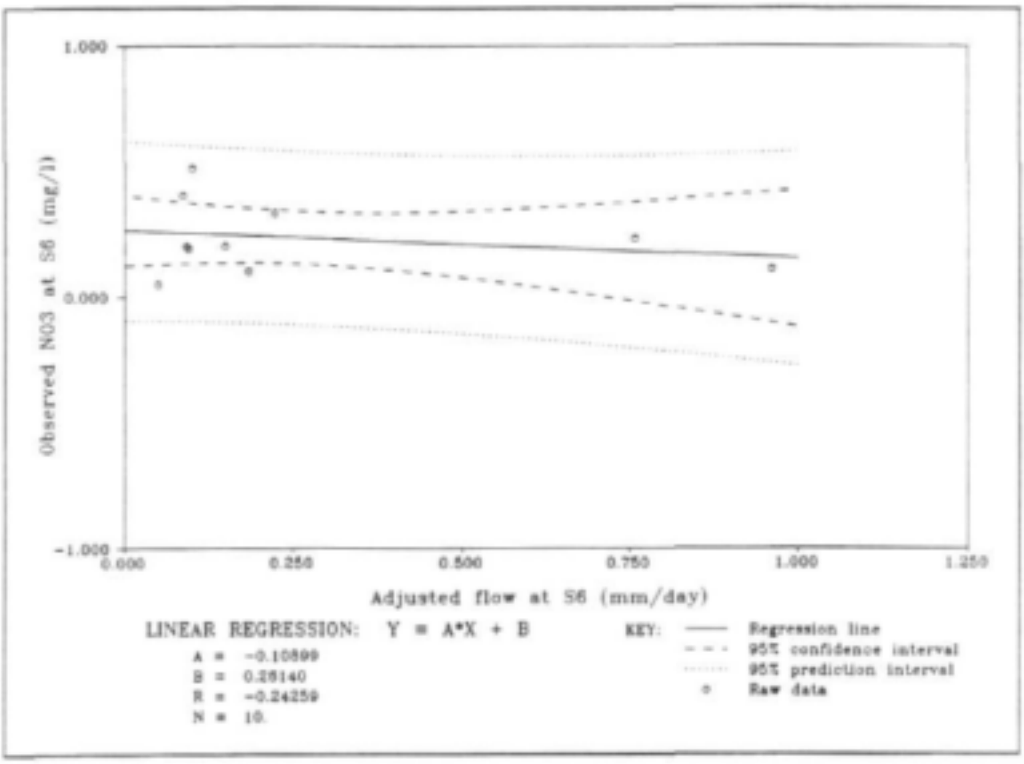


Figure B.3.11 : Linear regression between flow and nitrate at station S6

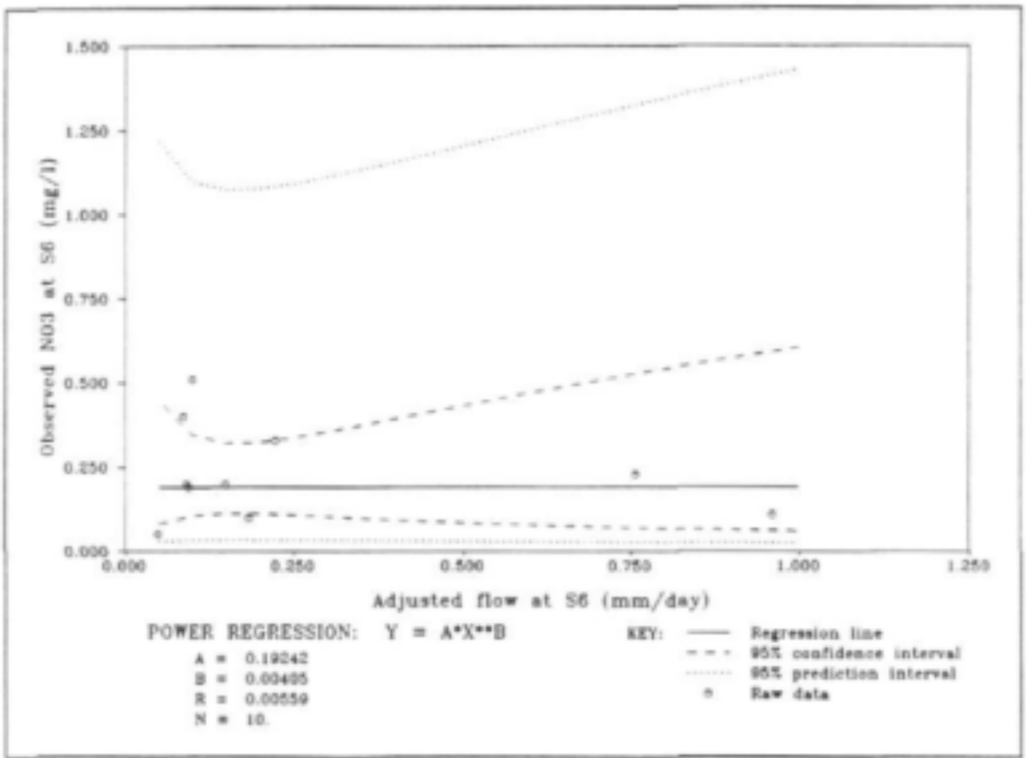


Figure B.3.12 : Power regression between flow and nitrate at station S6

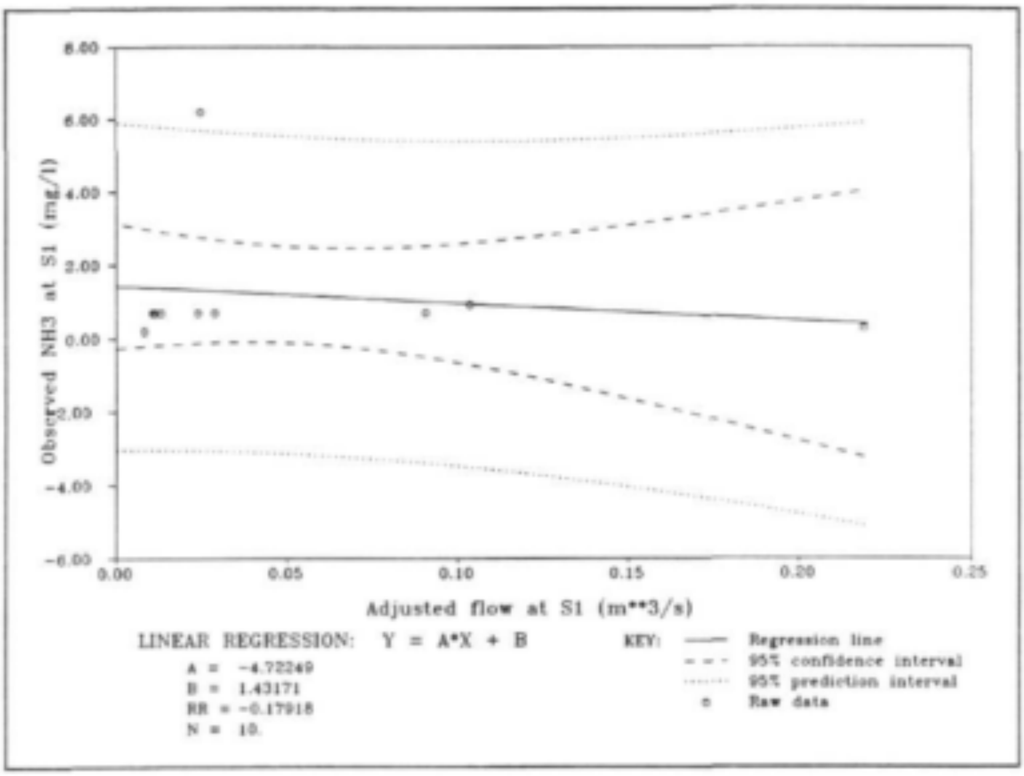


Figure B.4.1 : Linear regression between flow and ammonia at station S1

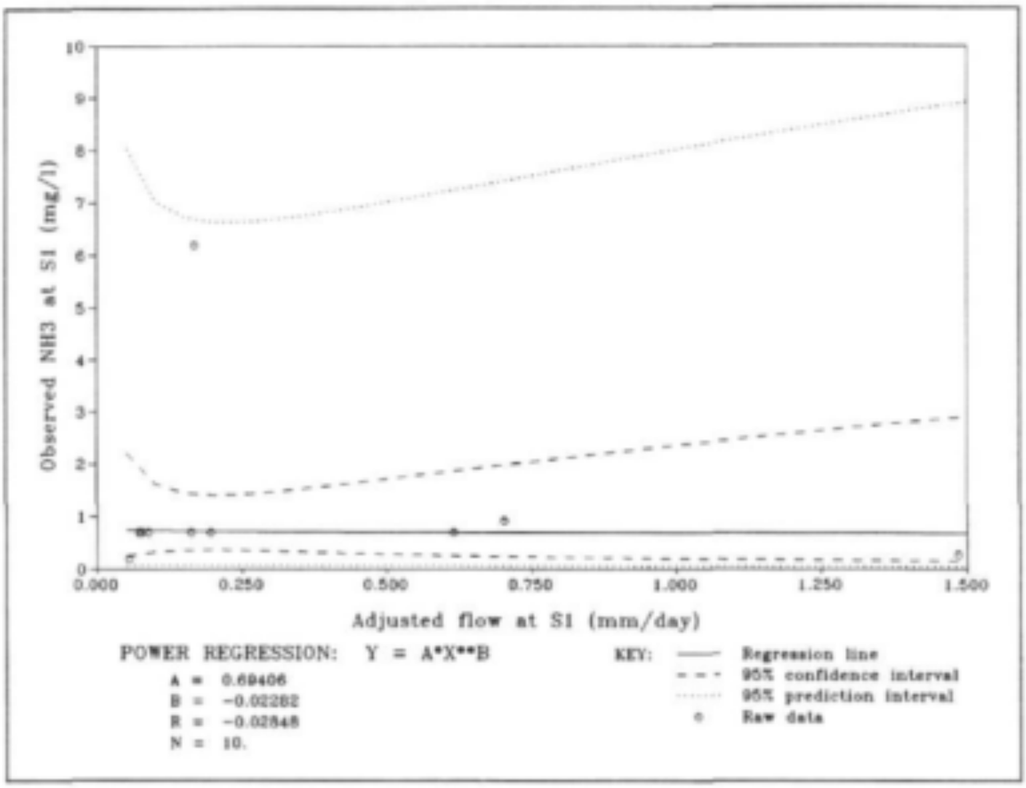


Figure B.4.2 : Power regression between flow and ammonia at station S1

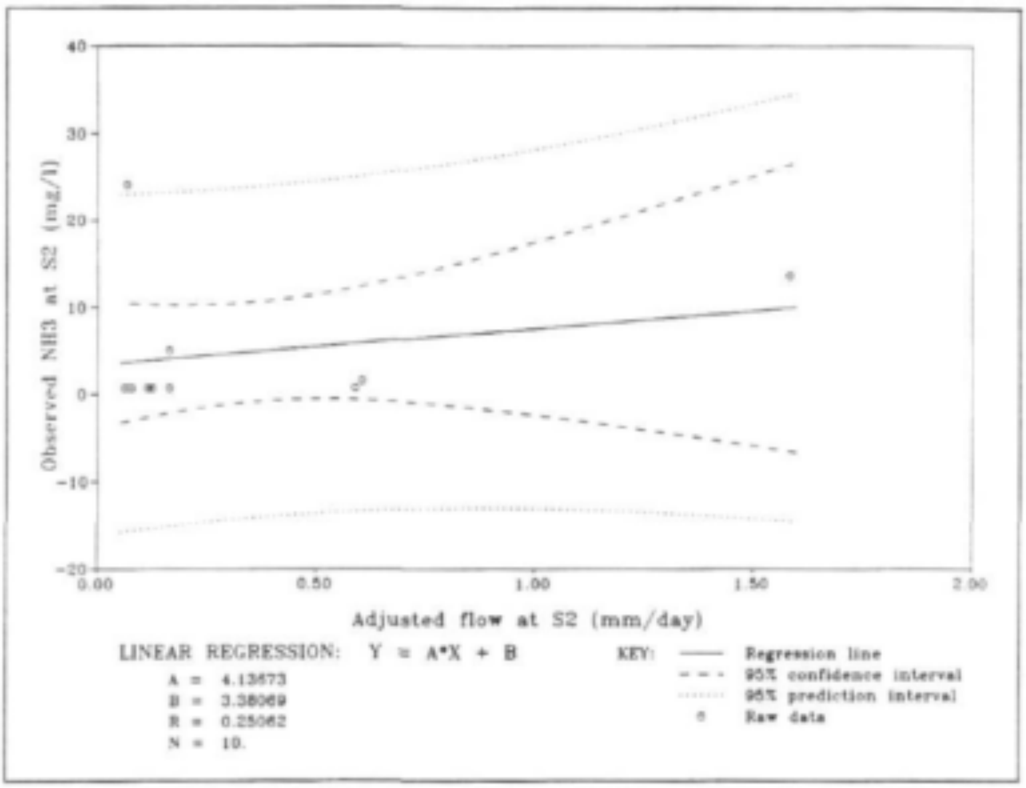


Figure B.4.3 : Linear regression between flow and ammonia at station S2

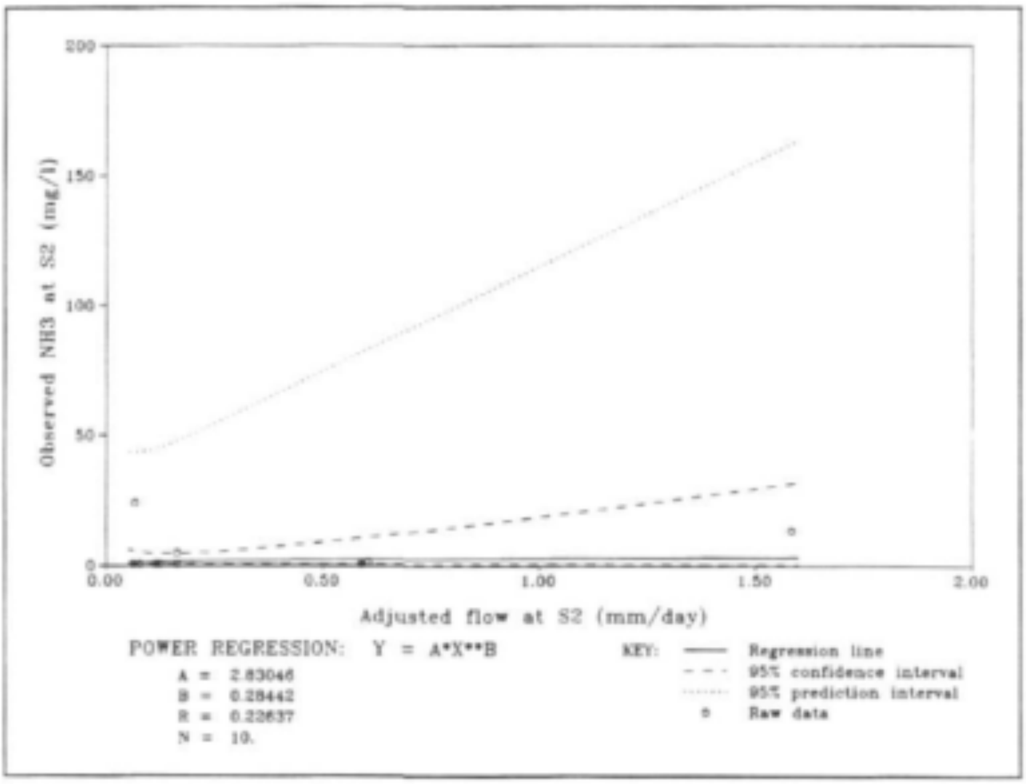


Figure B.4.4 : Power regression between flow and ammonia at station S2

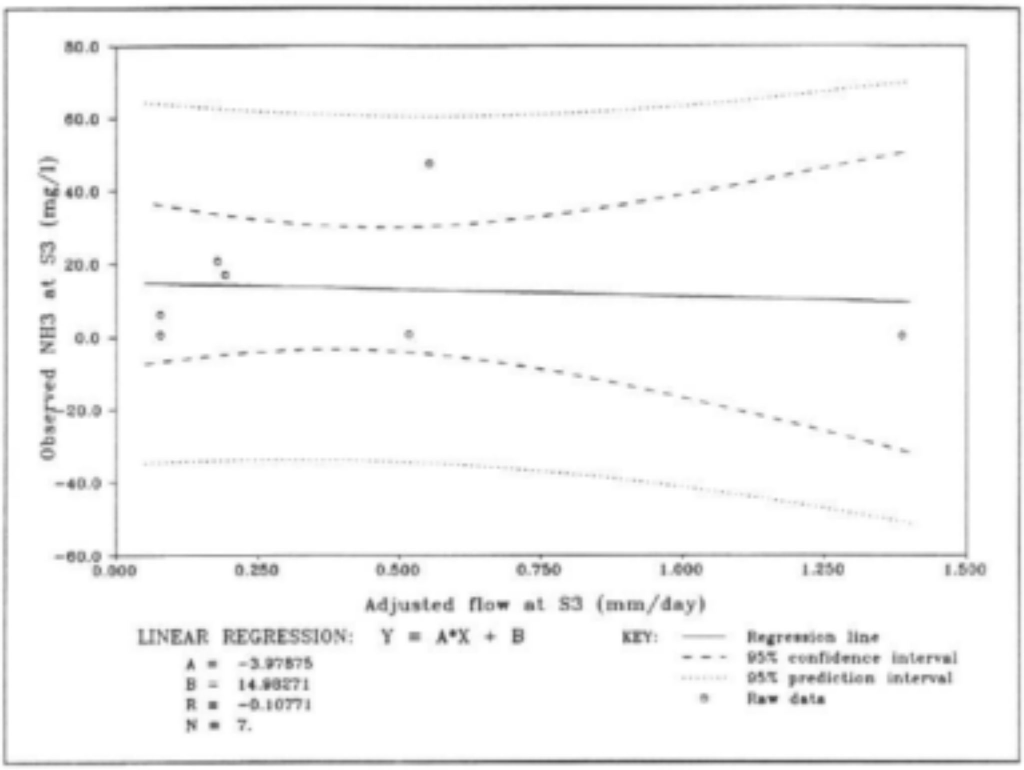


Figure B.4.5 : Linear regression between flow and ammonia at station S3

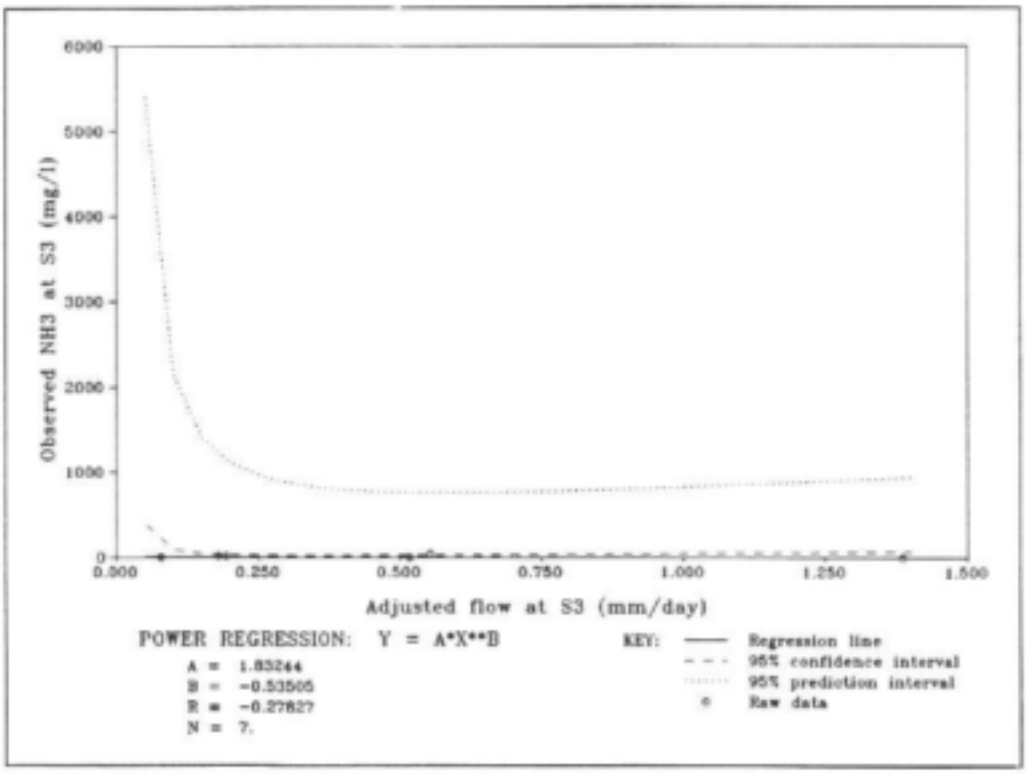


Figure B.4.6 : Power regression between flow and ammonia at station S3

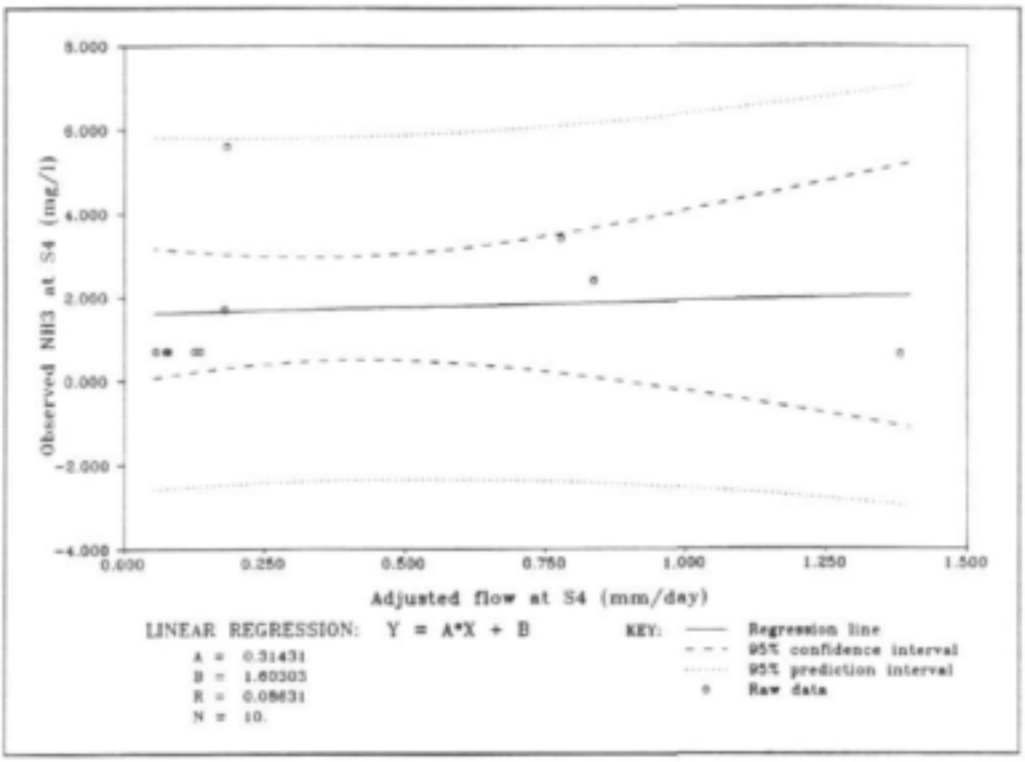


Figure B.4.7 : Linear regression between flow and ammonia at station S4

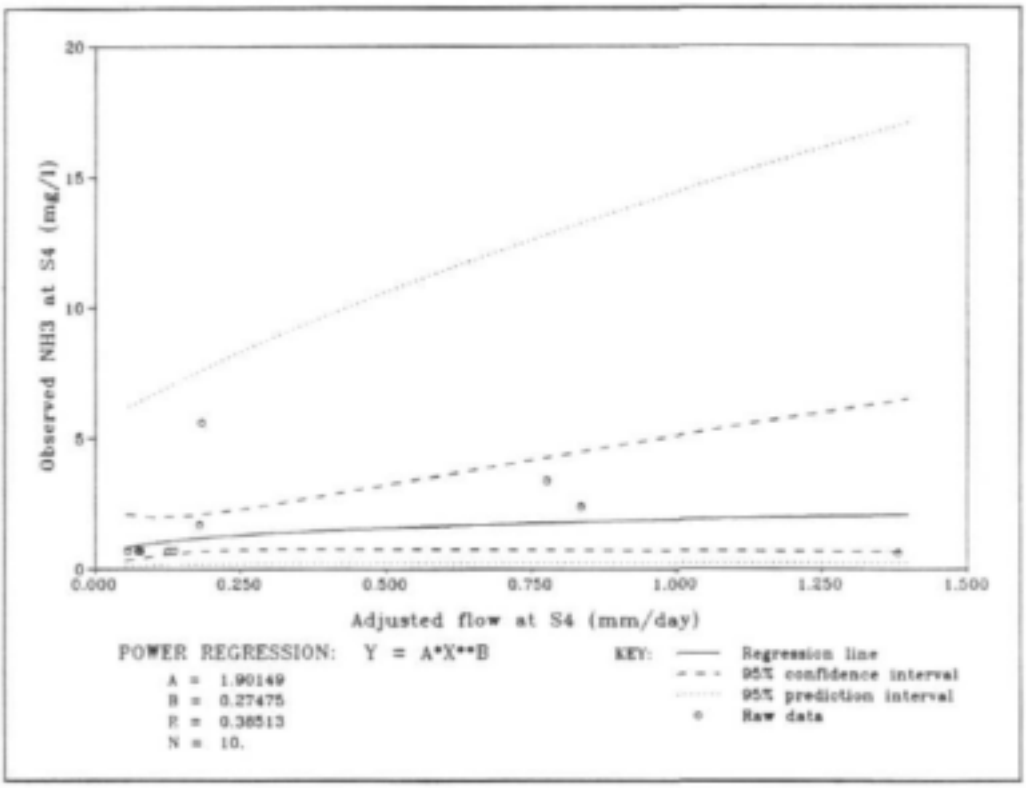


Figure B.4.8 : Power regression between flow and ammonia at station S4

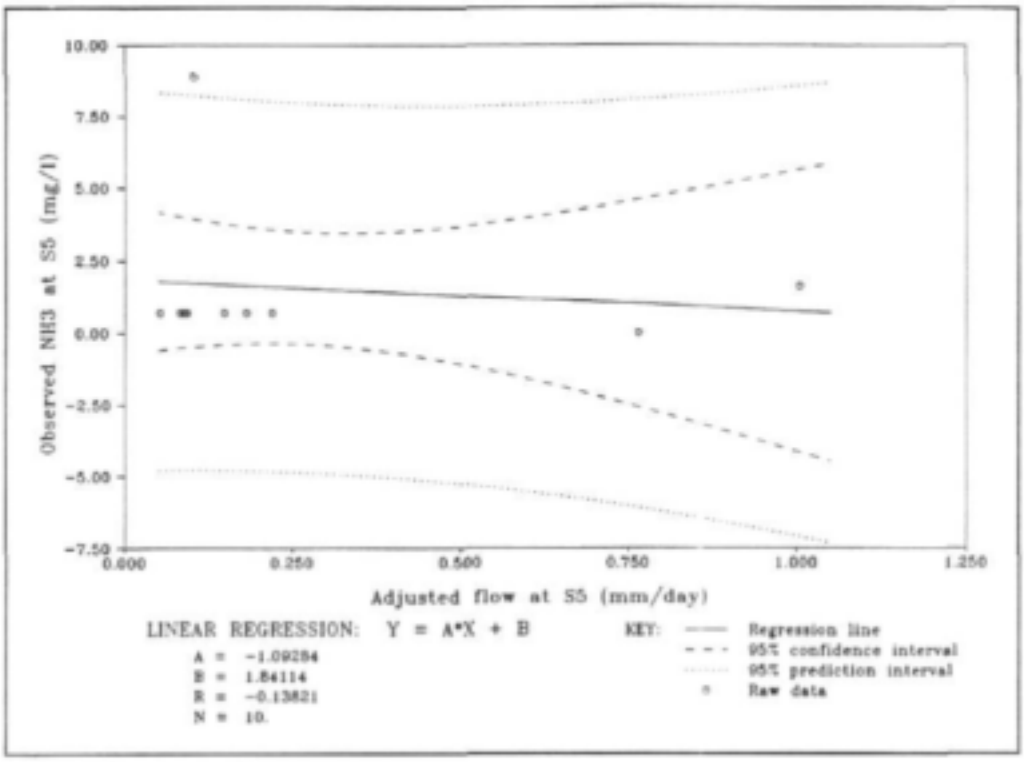


Figure B.4.9 : Linear regression between flow and ammonia at station S5

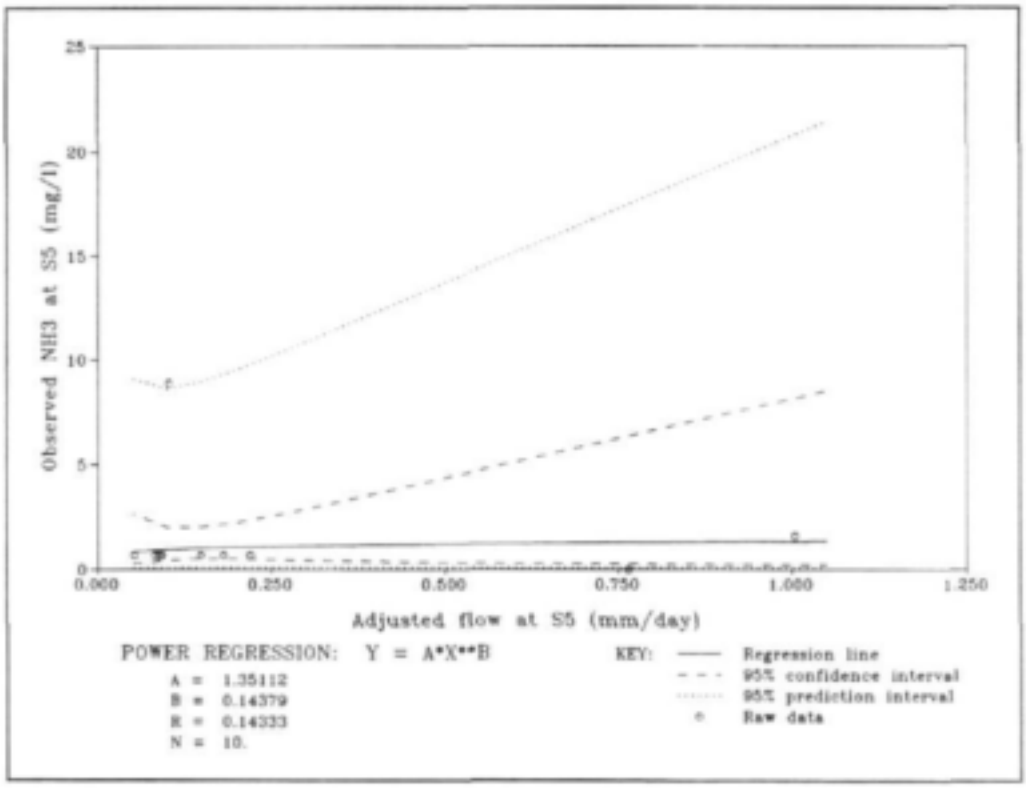


Figure B.4.10 : Power regression between flow and ammonia at station S5

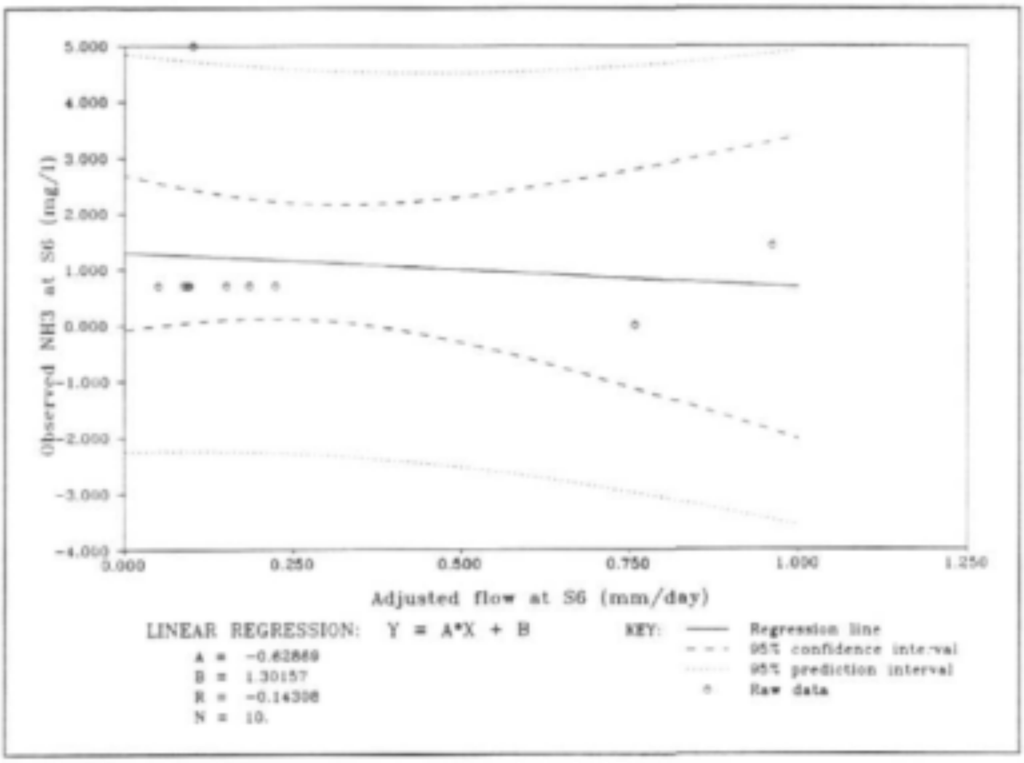


Figure B.4.11 : Linear regression between flow and ammonia at station S6

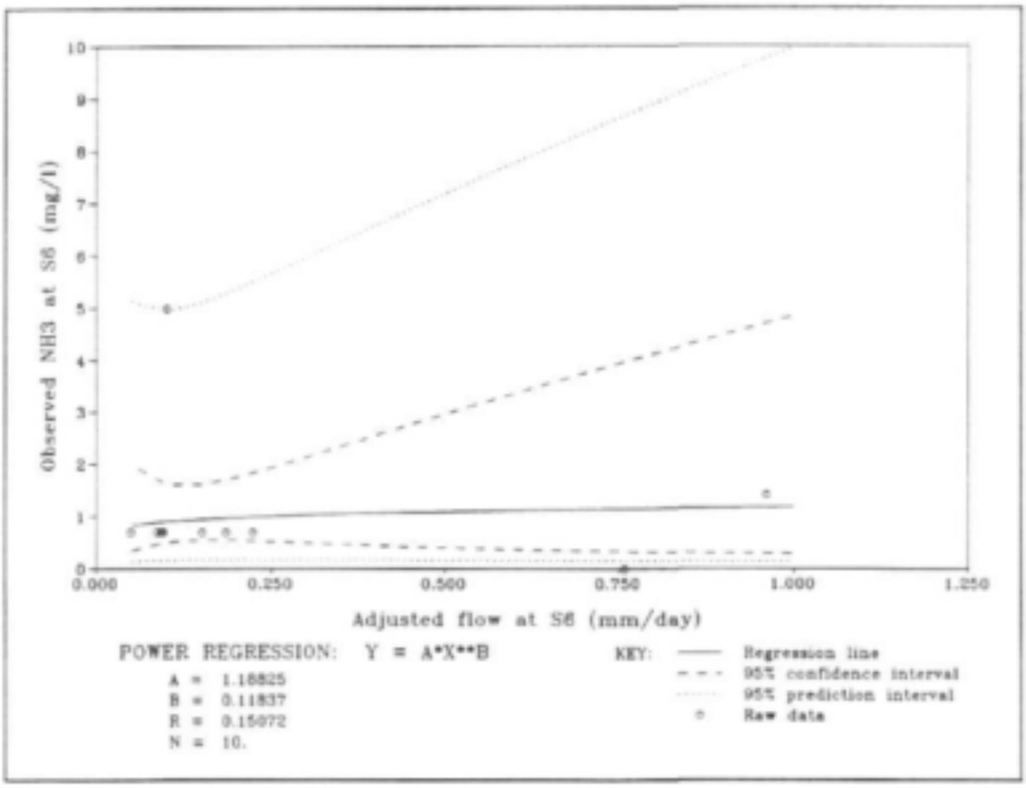


Figure B.4.12 : Power regression between flow and ammonia at station S6

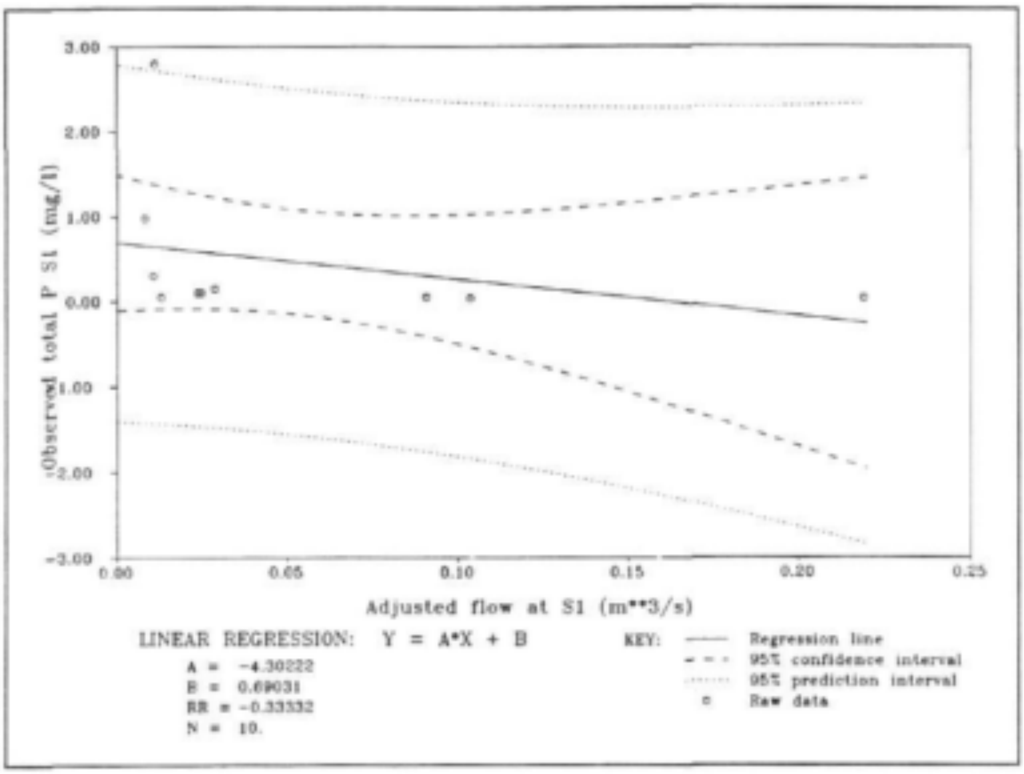


Figure B.5.1 : Linear regression between flow and total phosphate at station S1

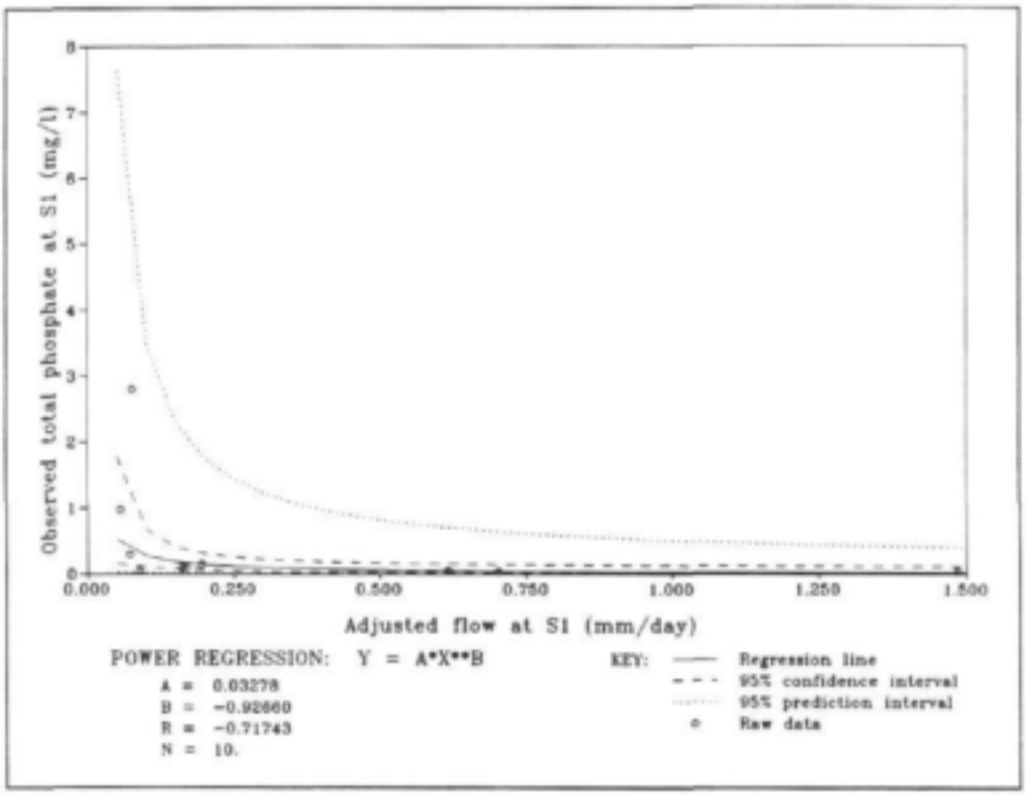


Figure B.5.2 : Power regression between flow and total phosphate at station S1

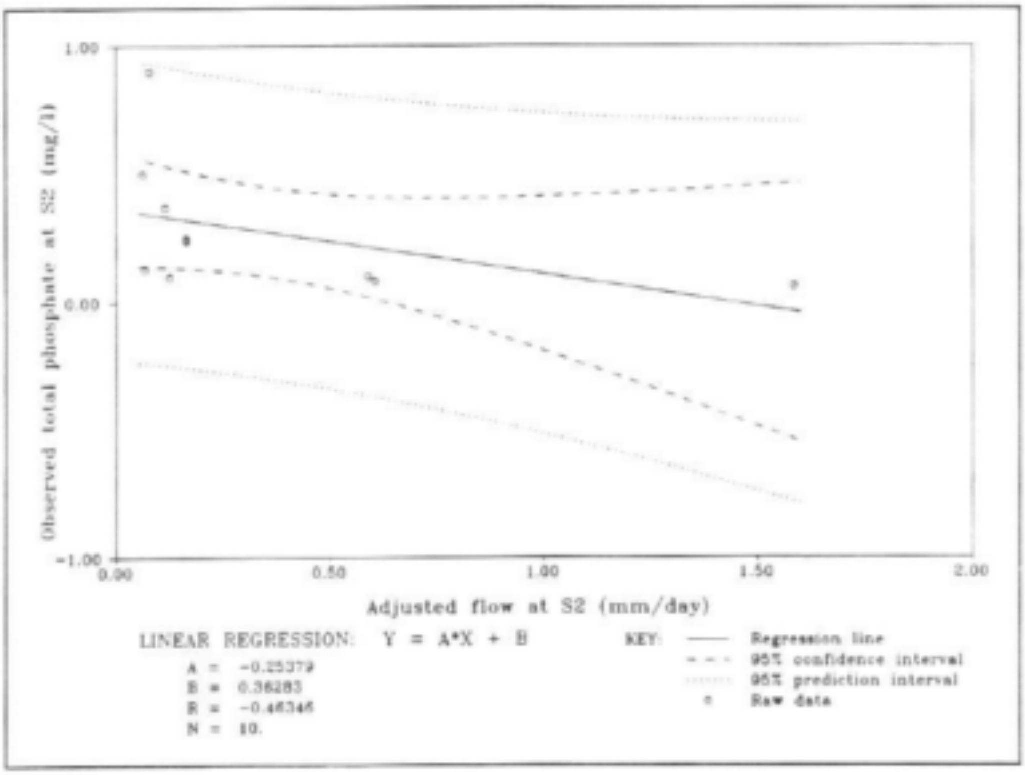


Figure B.5.3 : Linear regression between flow and total phosphate at station S2

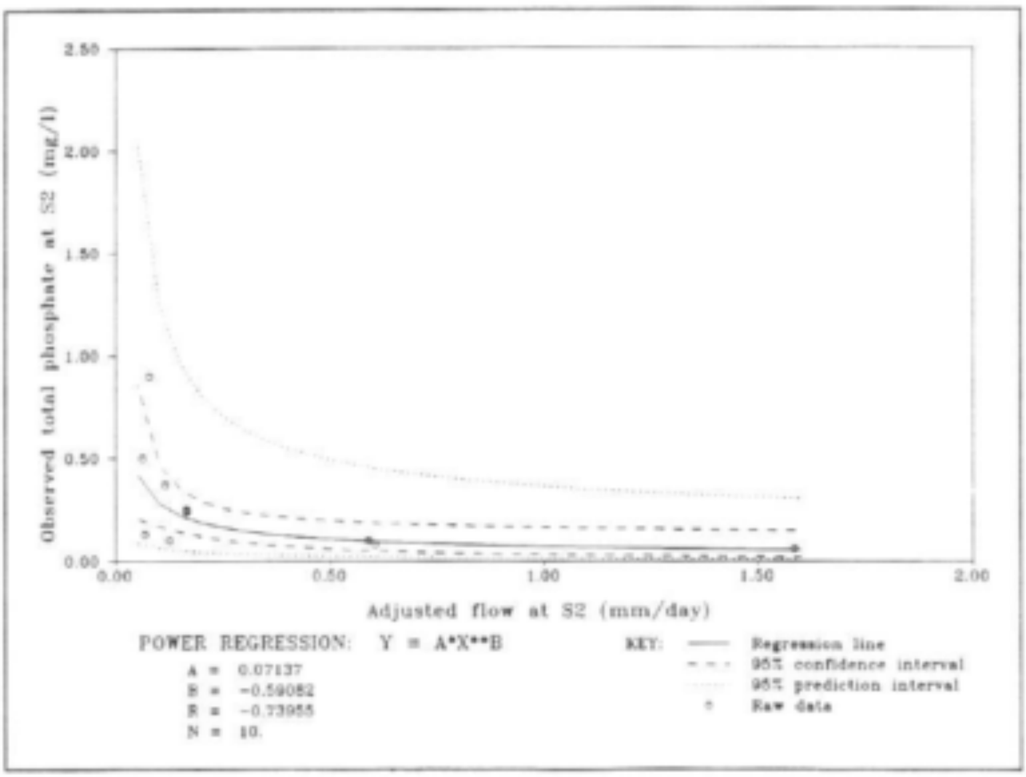


Figure B.5.4 : Power regression between flow and total phosphate at station S2

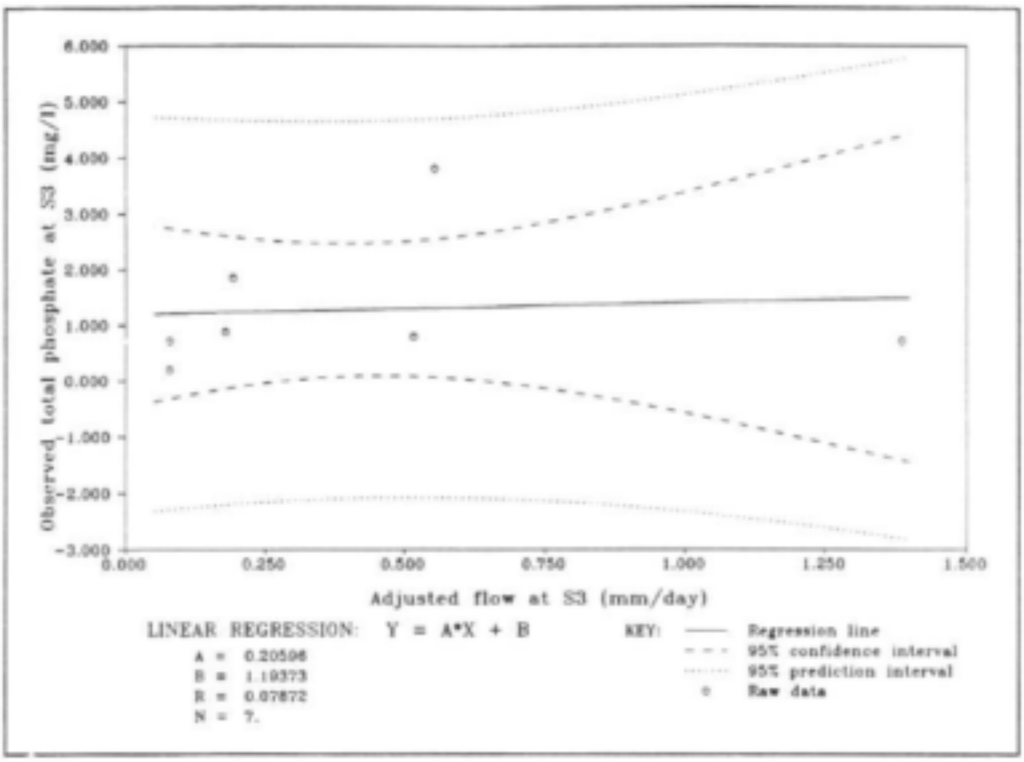


Figure B.5.5 : Linear regression between flow and total phosphate at station S3

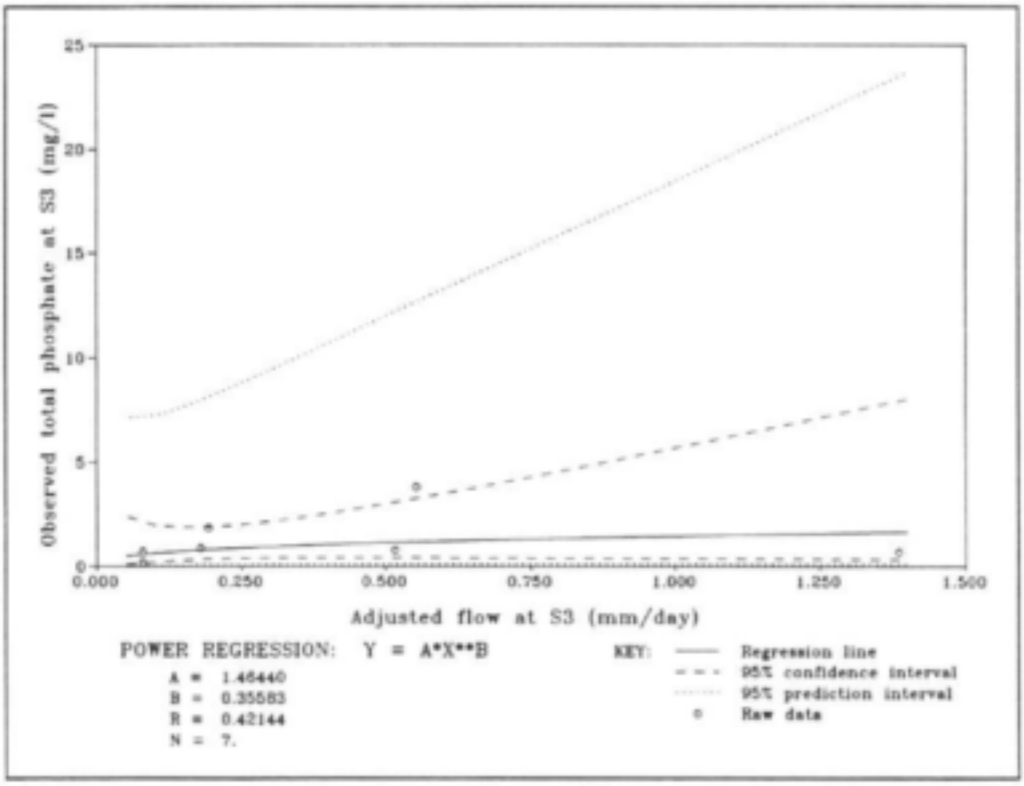


Figure B.5.6 : Power regression between flow and total phosphate at station S3

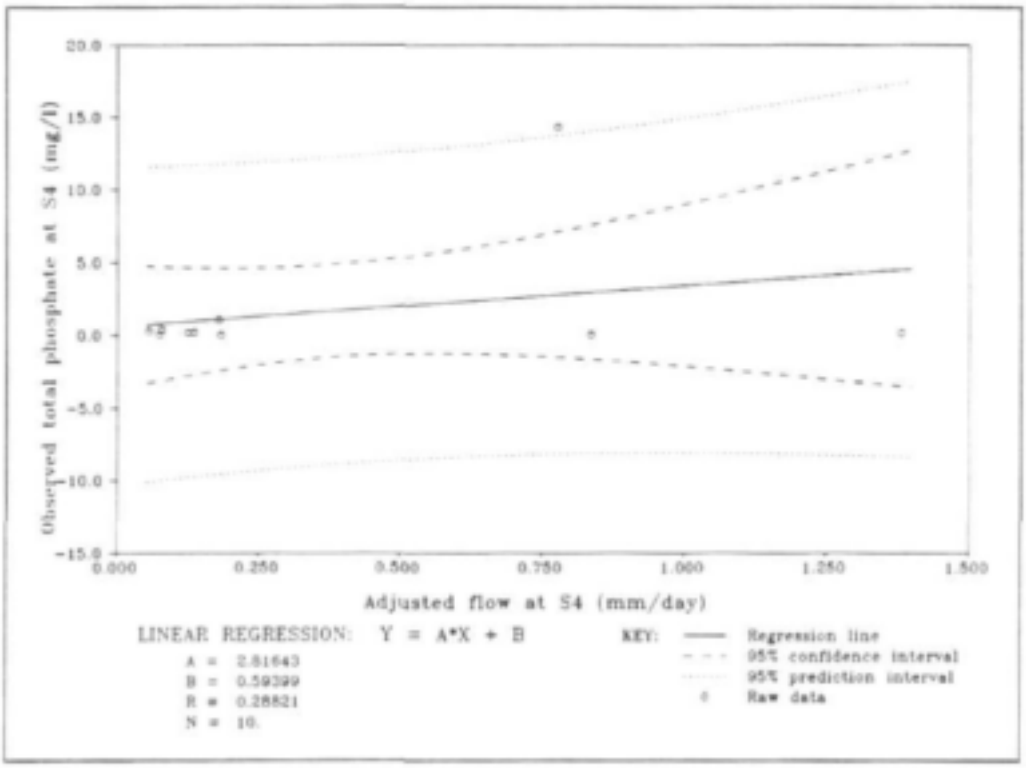


Figure B.5.7 : Linear regression between flow and total phosphate at station S4

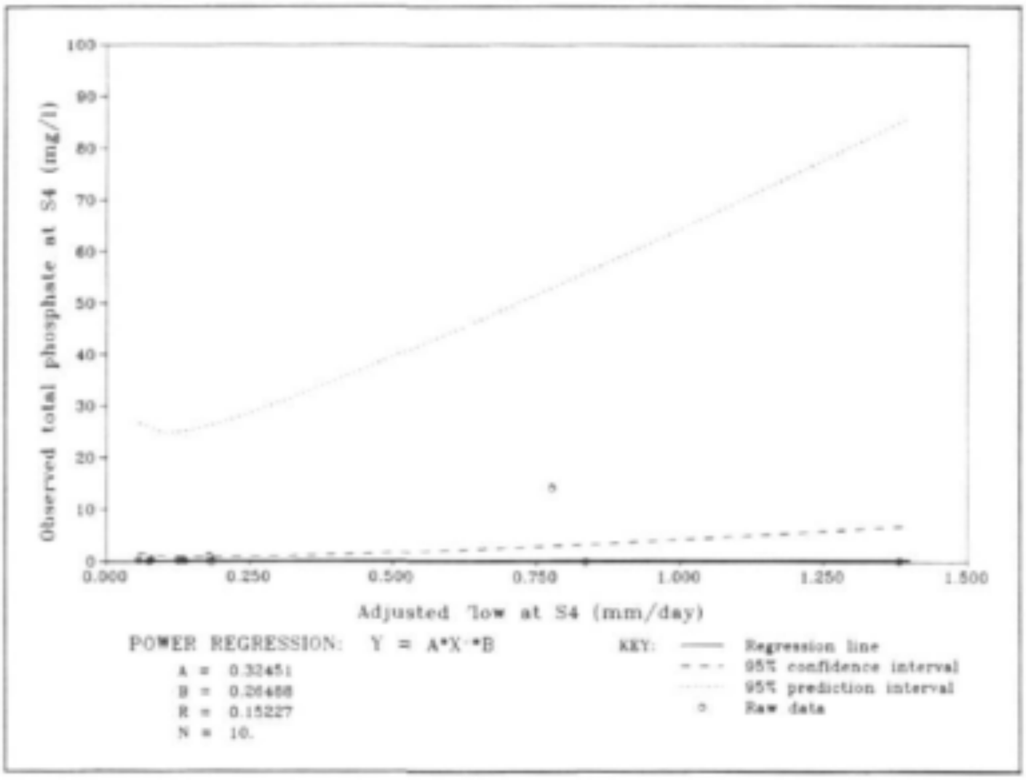


Figure B.5.8 : Power regression between flow and total phosphate at station S4

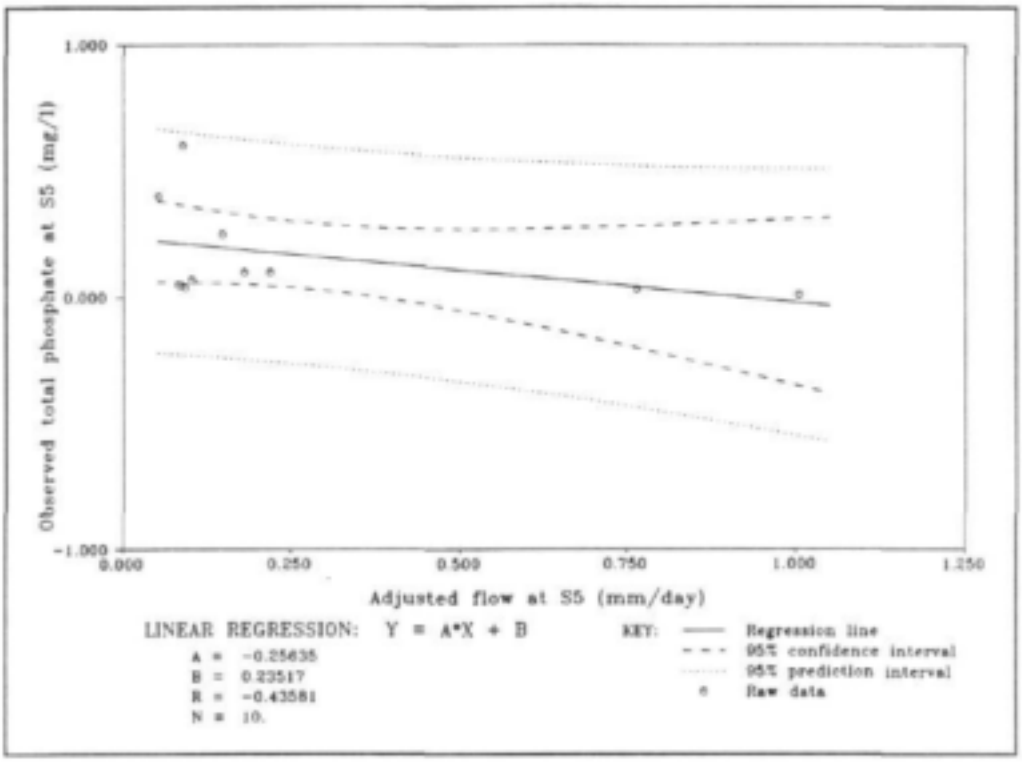


Figure B.5.9 : Linear regression between flow and total phosphate at station S5

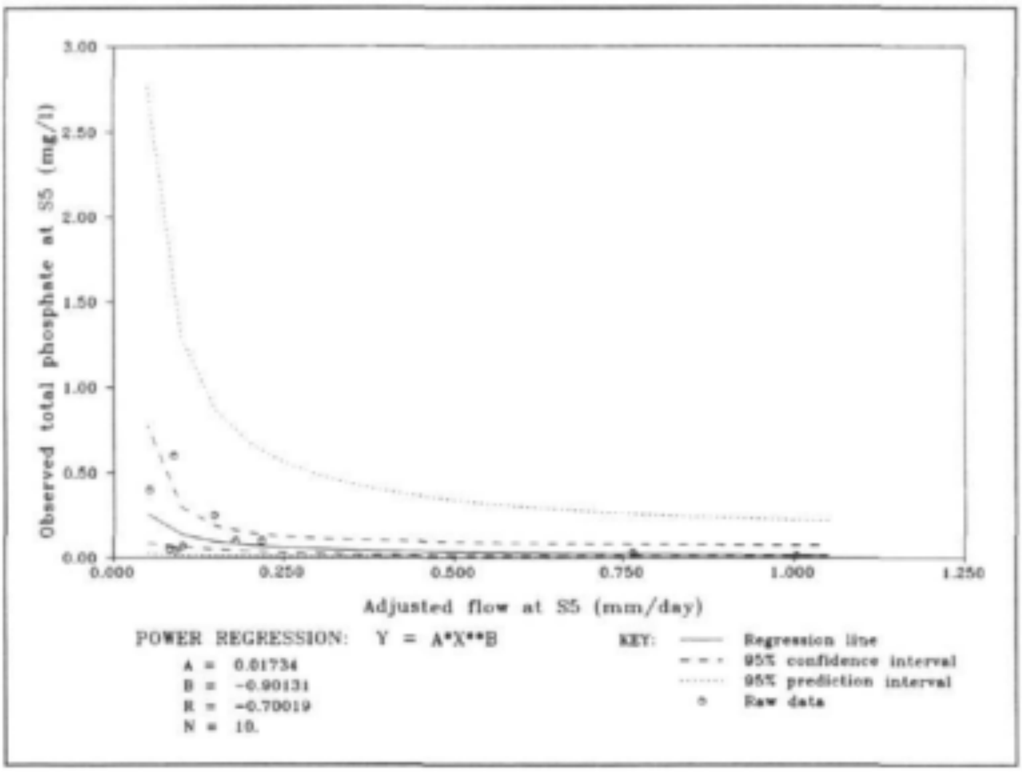


Figure B.5.10 : Power regression between flow and total phosphate at station S5

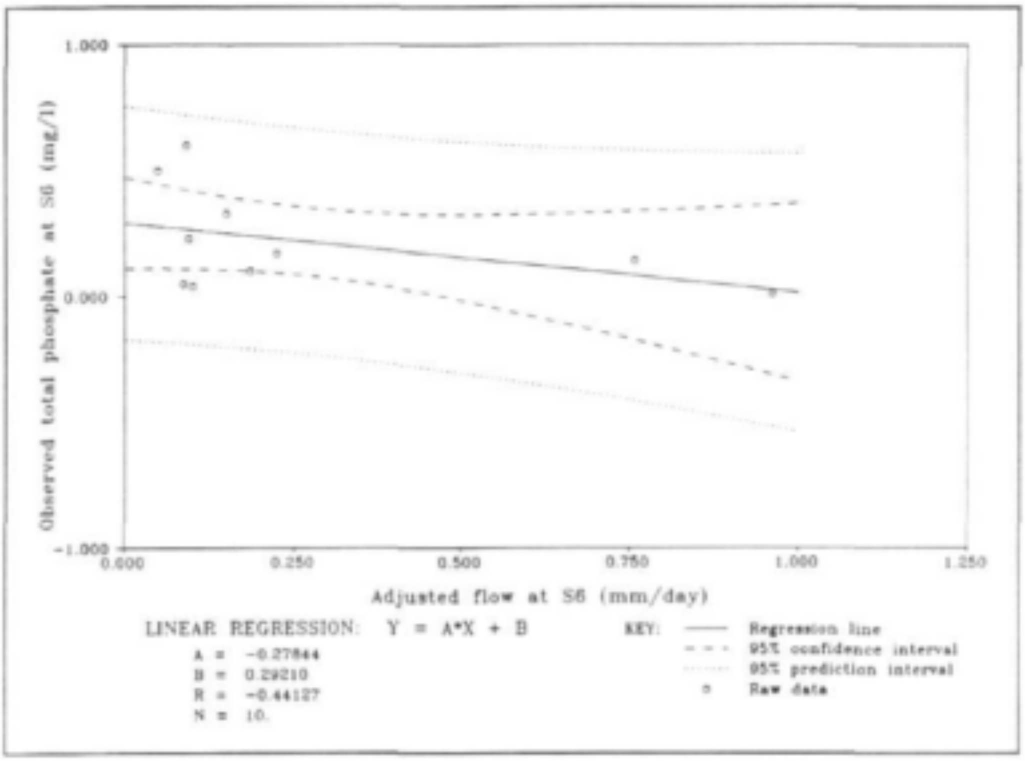


Figure B.5.11 : Linear regression between flow and total phosphate at station S6

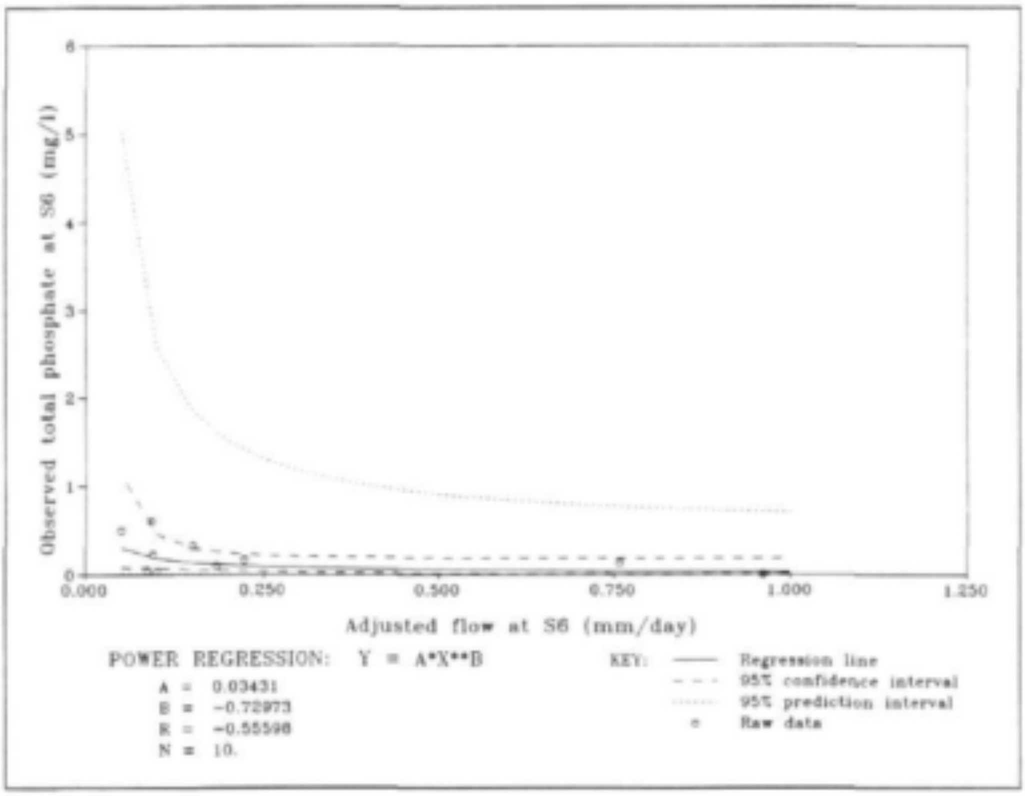


Figure B.5.12 : Power regression between flow and total phosphate at station S6

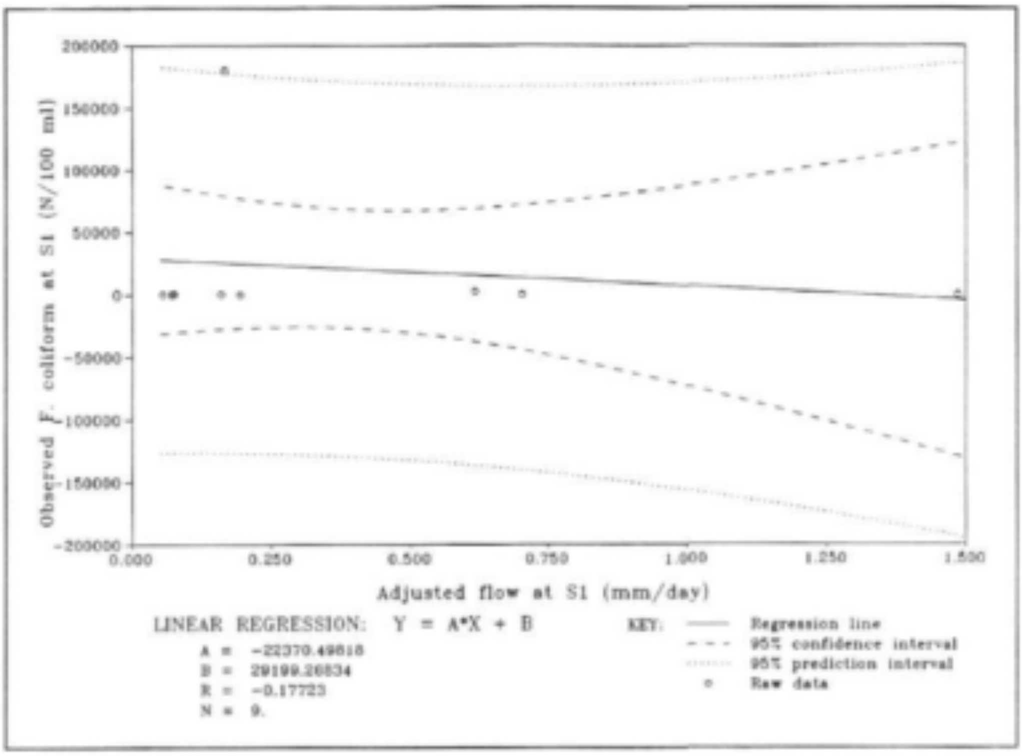


Figure B.6.1 : Linear regression between flow and faecal coliform at station S1

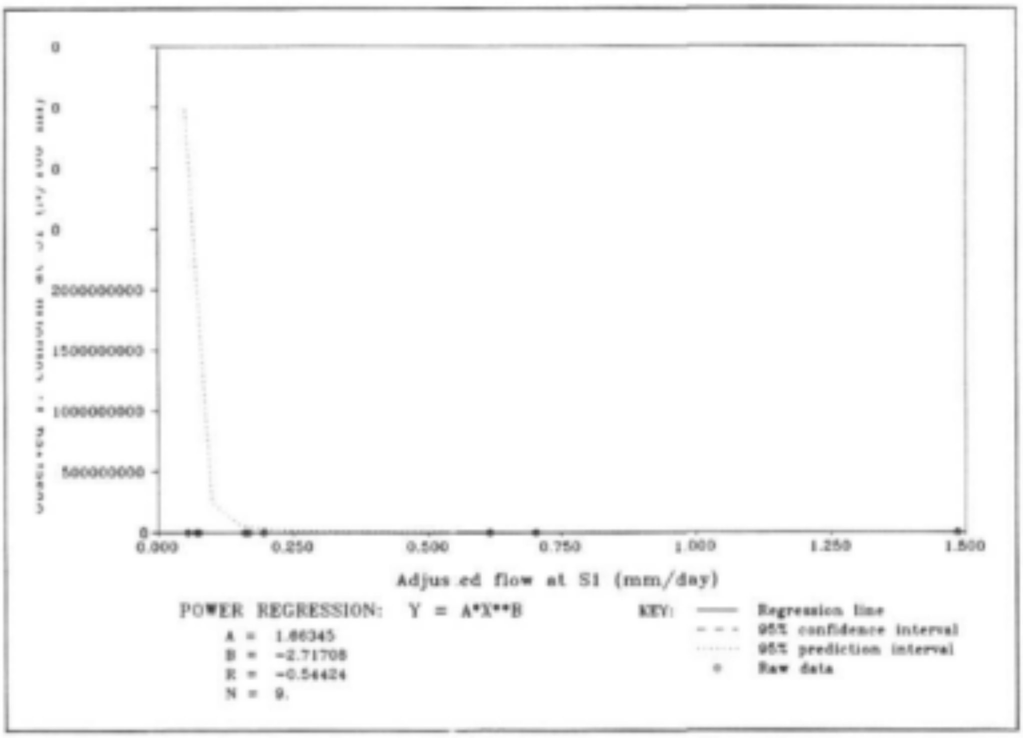


Figure B.6.2 : Power regression be:ween flow and faecal coliform at station S1

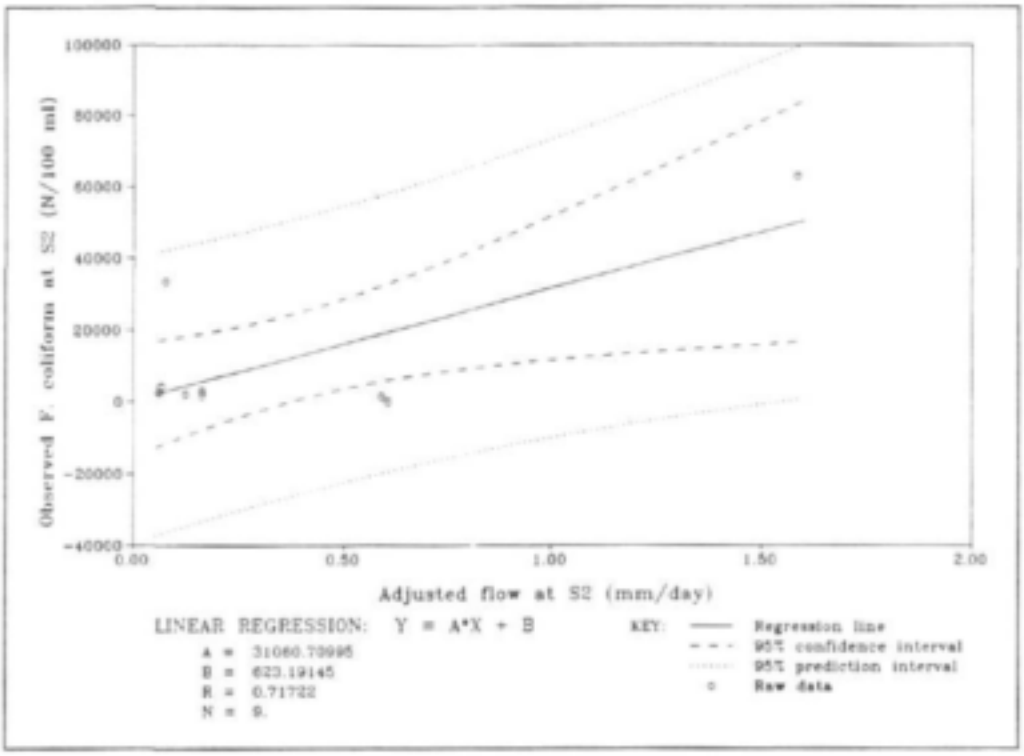


Figure B.6.3 : Linear regression between flow and faecal coliform at station S2

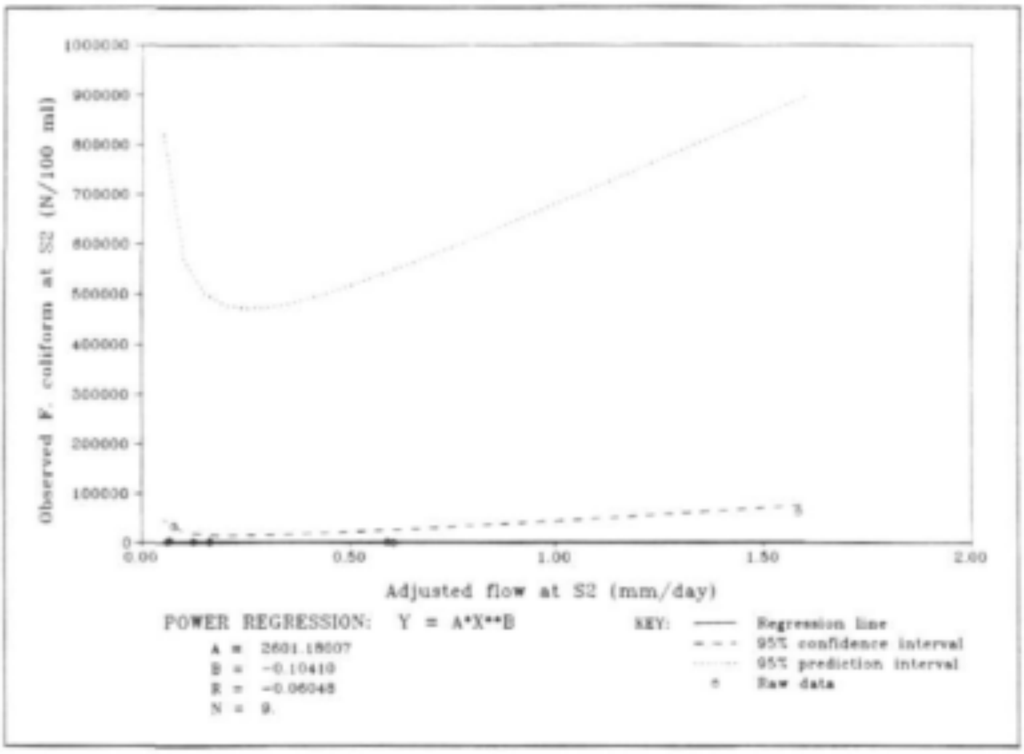


Figure B.6.4 : Power regression between flow and faecal coliform at station S2

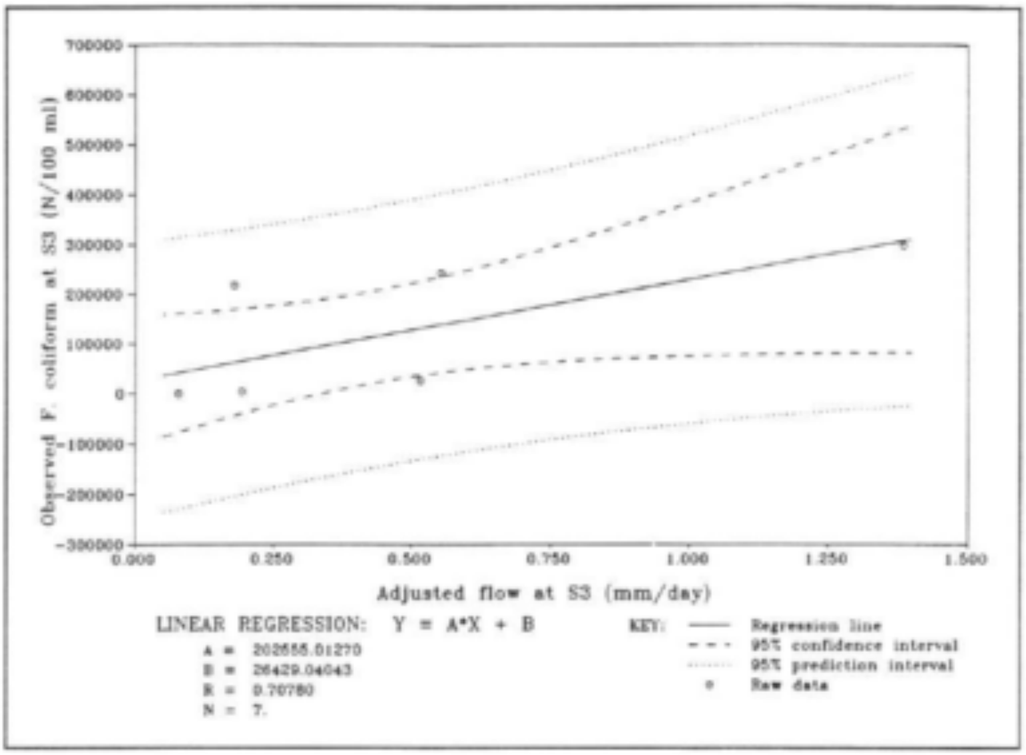


Figure B.6.5 : Linear regression between flow and faecal coliform at station S3

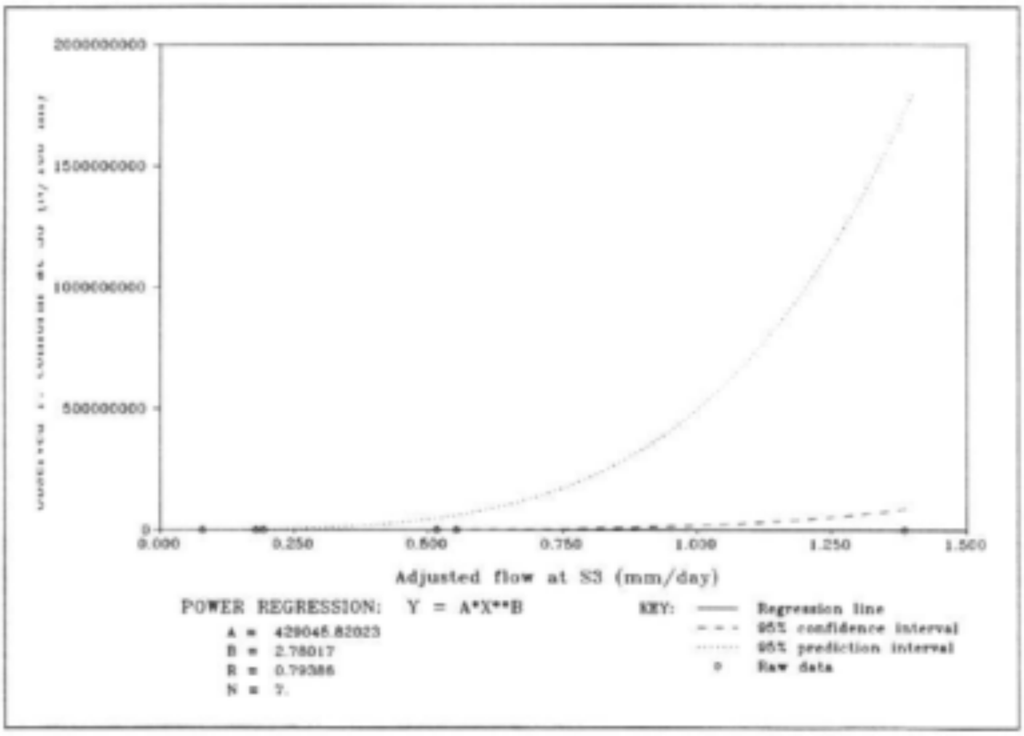


Figure B.6.6 : Power regression between flow and faecal coliform at station S3

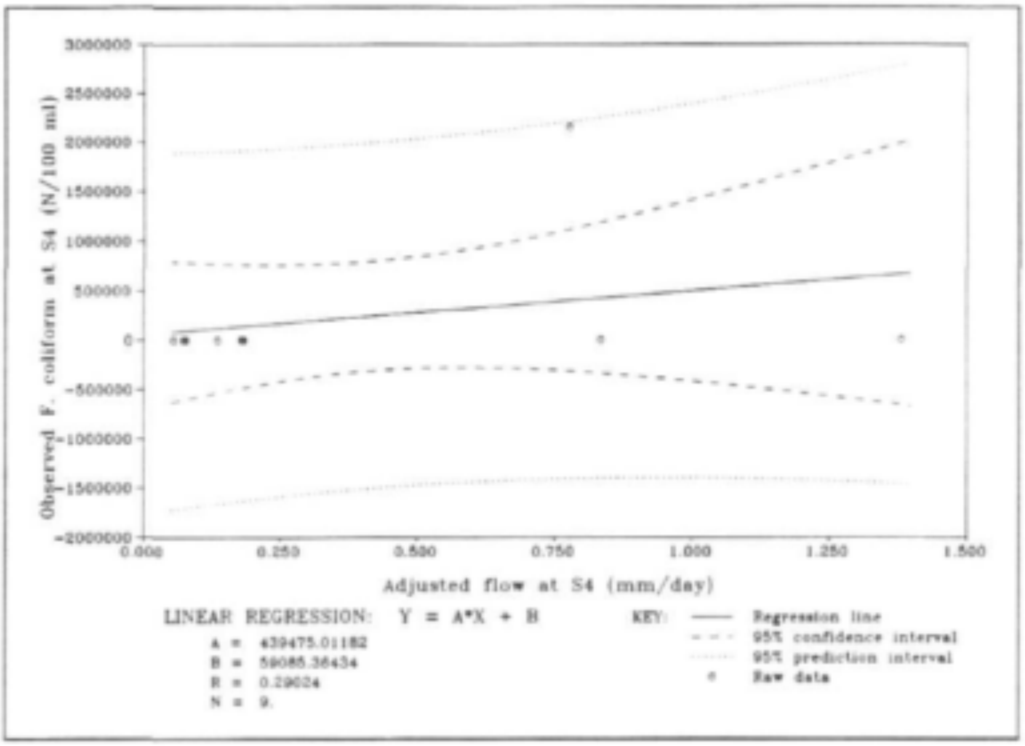


Figure B.6.7 : Linear regression between flow and faecal coliform at station S4

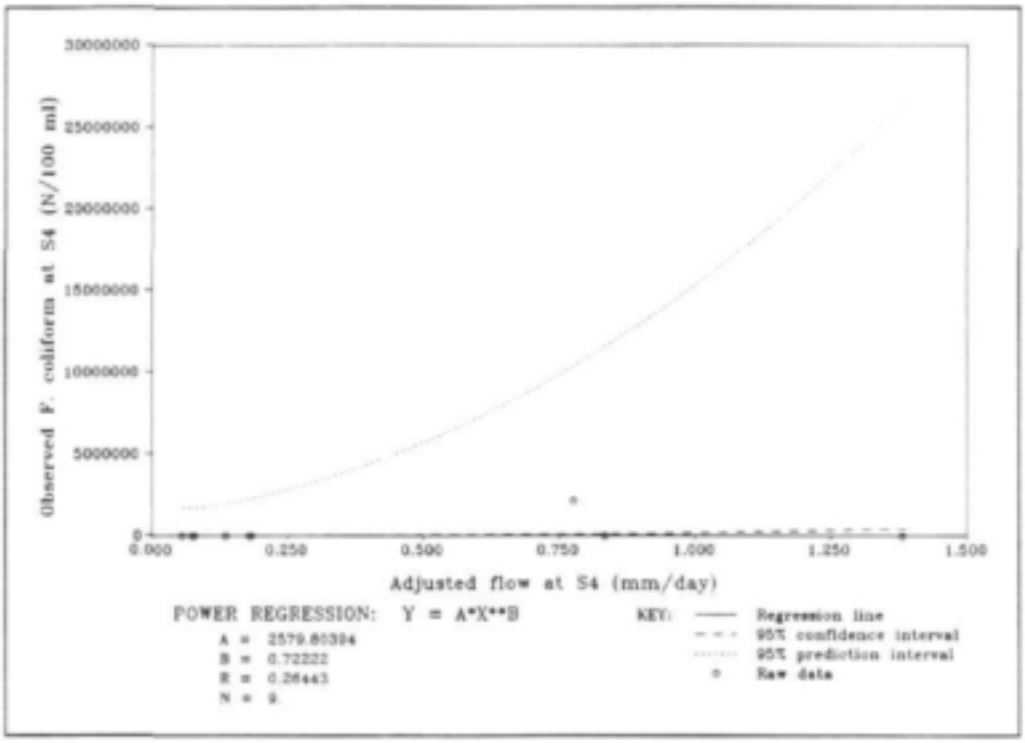


Figure B.6.8 : Power regression between flow and faecal coliform at station S4

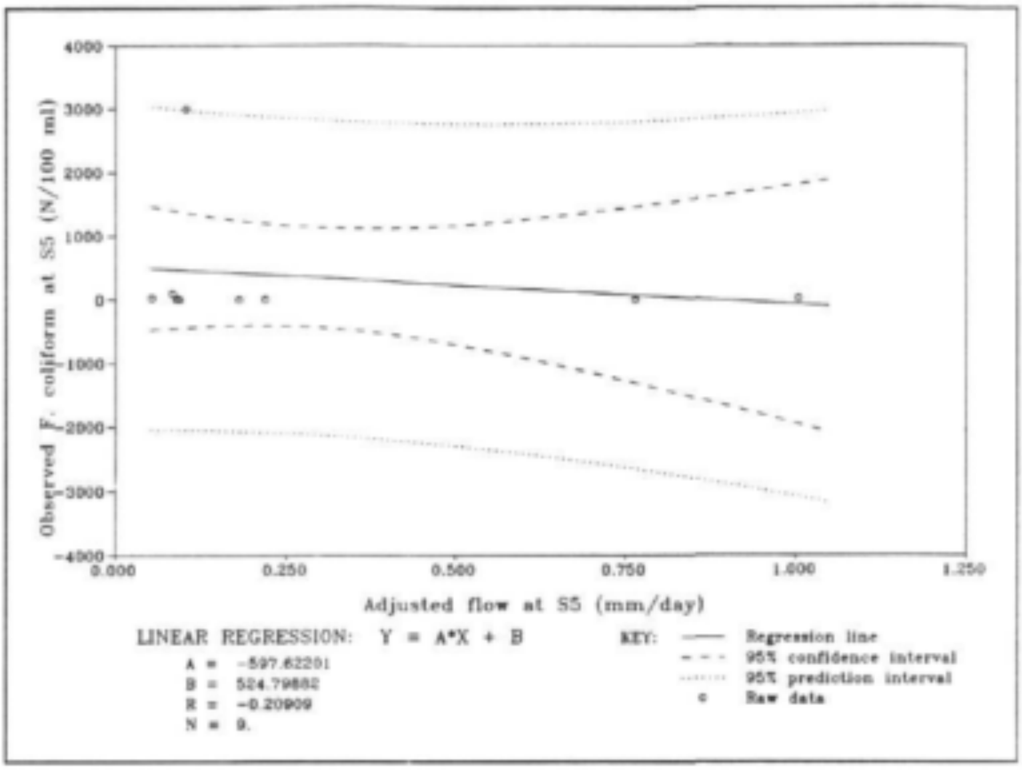


Figure B.6.9 : Linear regression between flow and faecal coliform at station S5

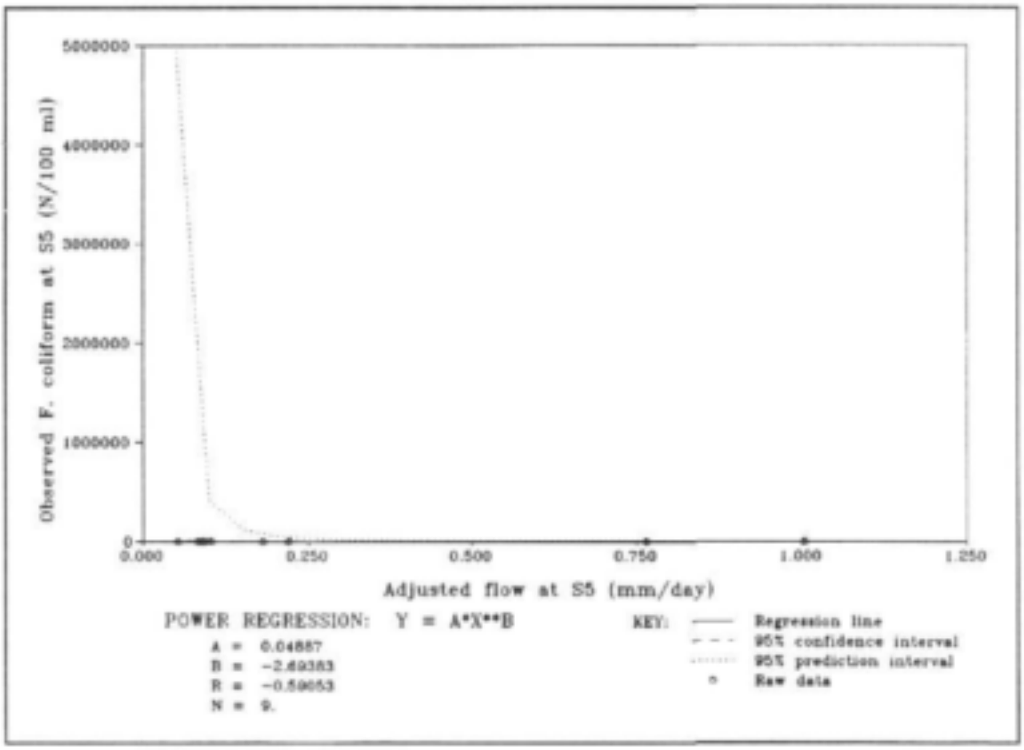


Figure B.6.10 : Power regression between flow and faecal coliform at station S5

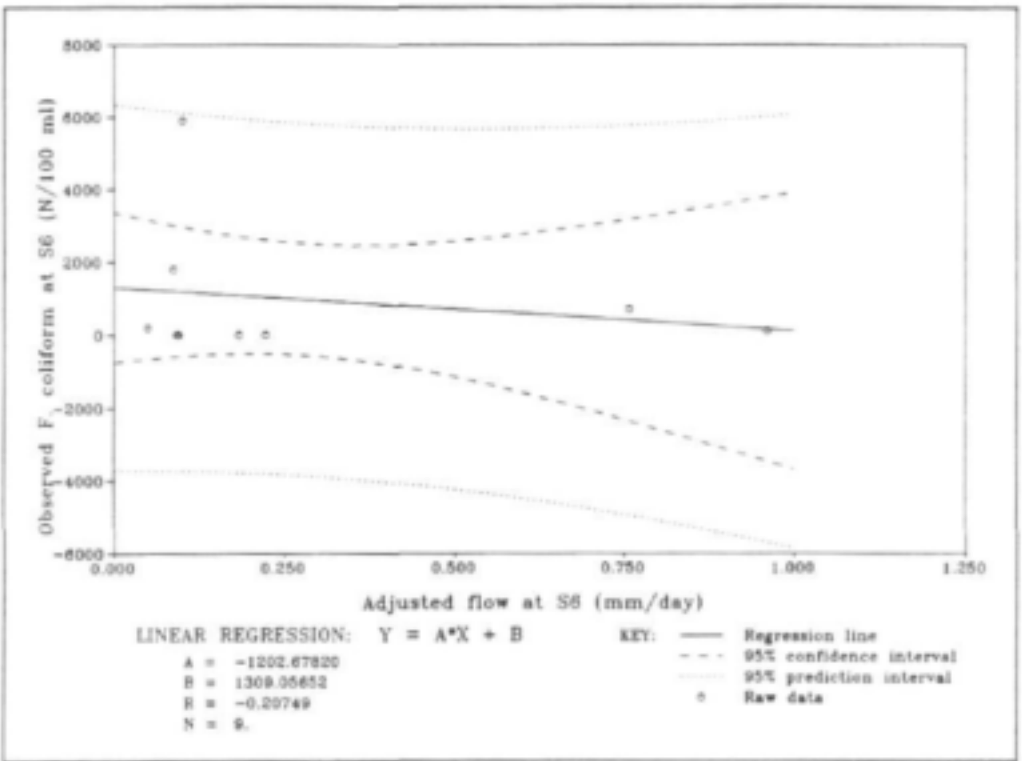


Figure B.6.11 : Linear regression between flow and faecal coliform at station S6

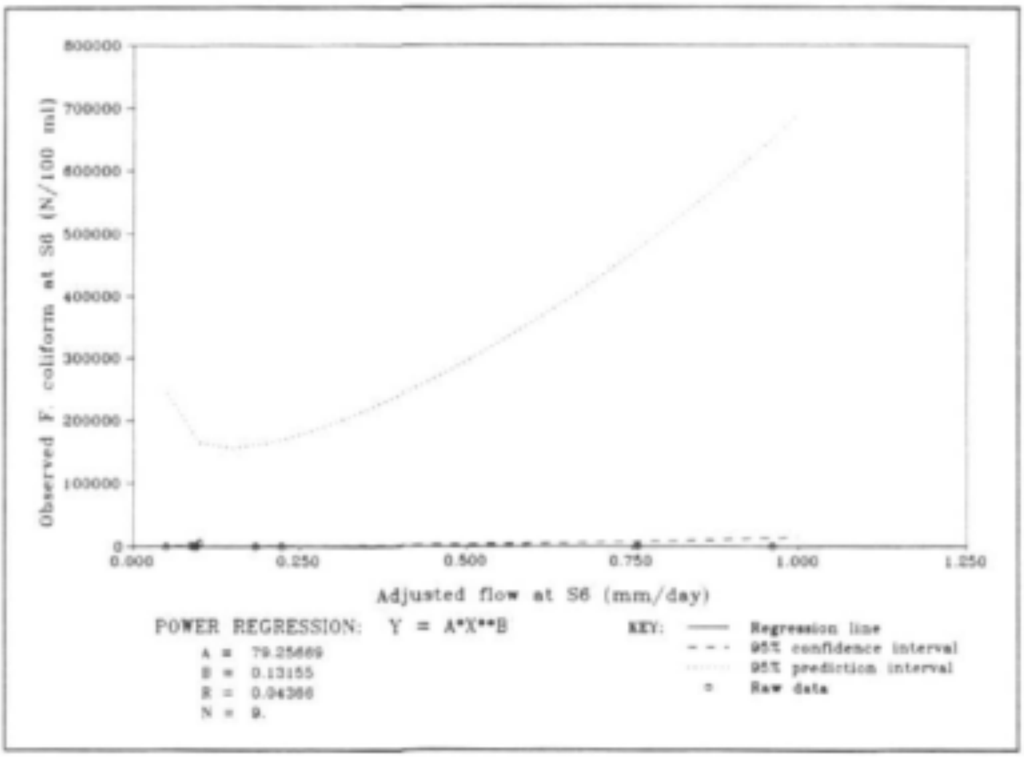


Figure B.6.12 : Power regression between flow and faecal coliform at station S6

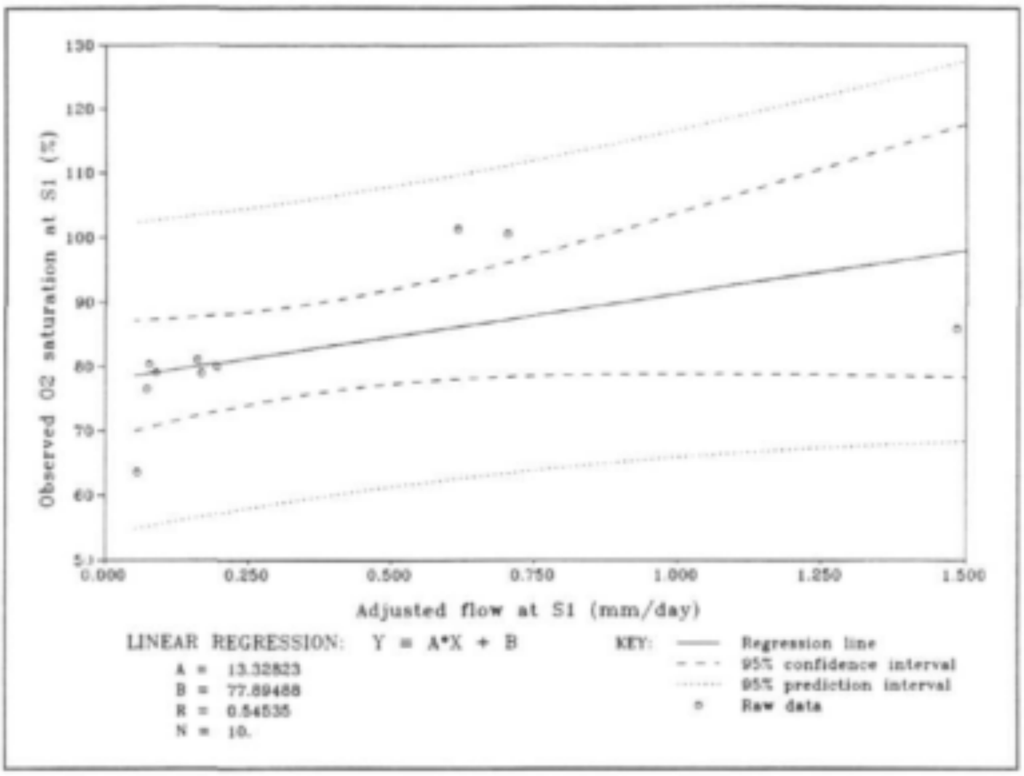


Figure B.7.1 : Linear regression between flow and oxygen saturation at station S1

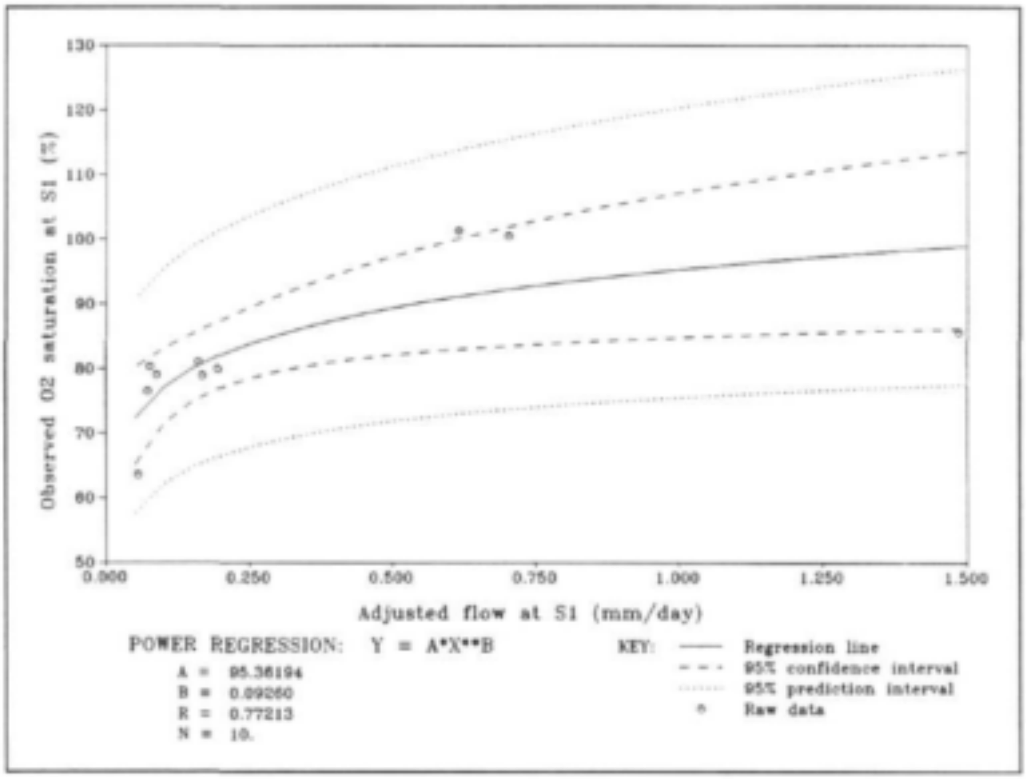


Figure B.7.2 : Power regression between flow and oxygen saturation at station S1

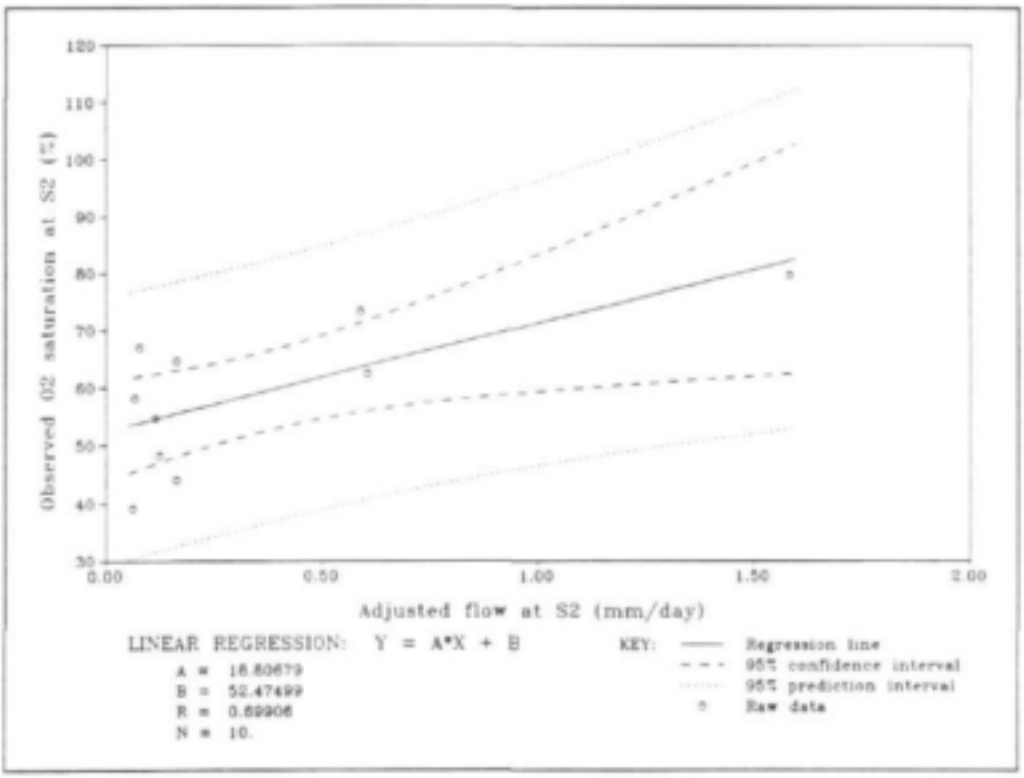


Figure B.7.3 : Linear regression between flow and oxygen saturation at station S2

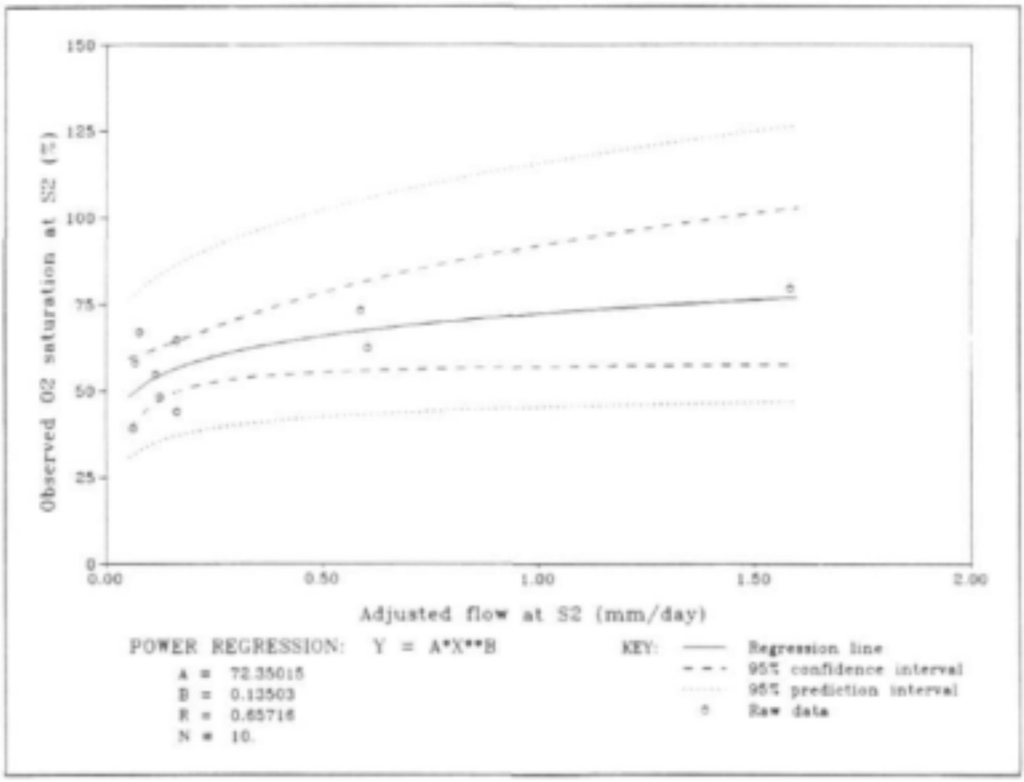


Figure B.7.4 : Power regression between flow and oxygen saturation at station S2

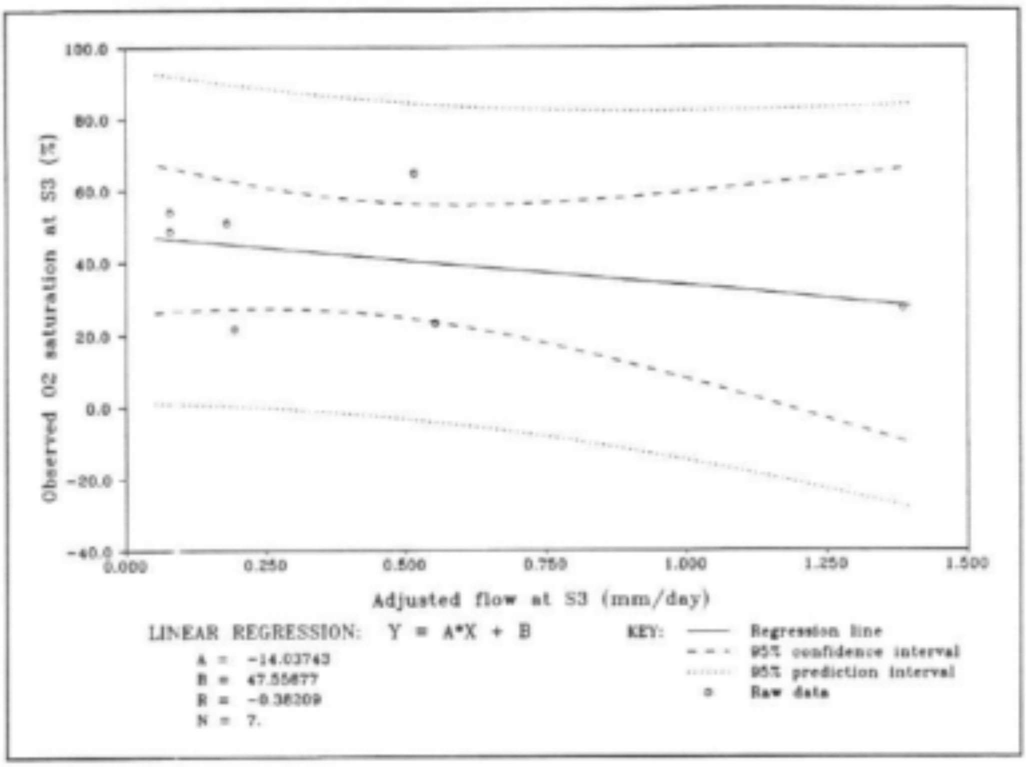


Figure B.7.5 : Linear regression between flow and oxygen saturation at station S3

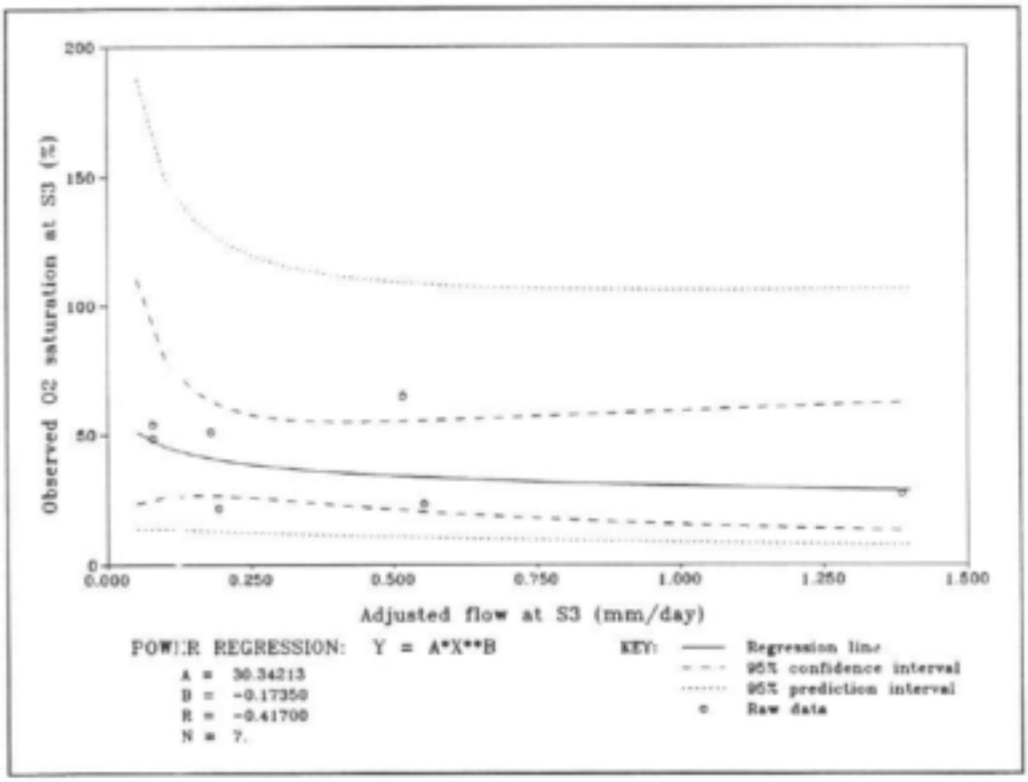


Figure B.7.6 : Power regression between flow and oxygen saturation at station S3

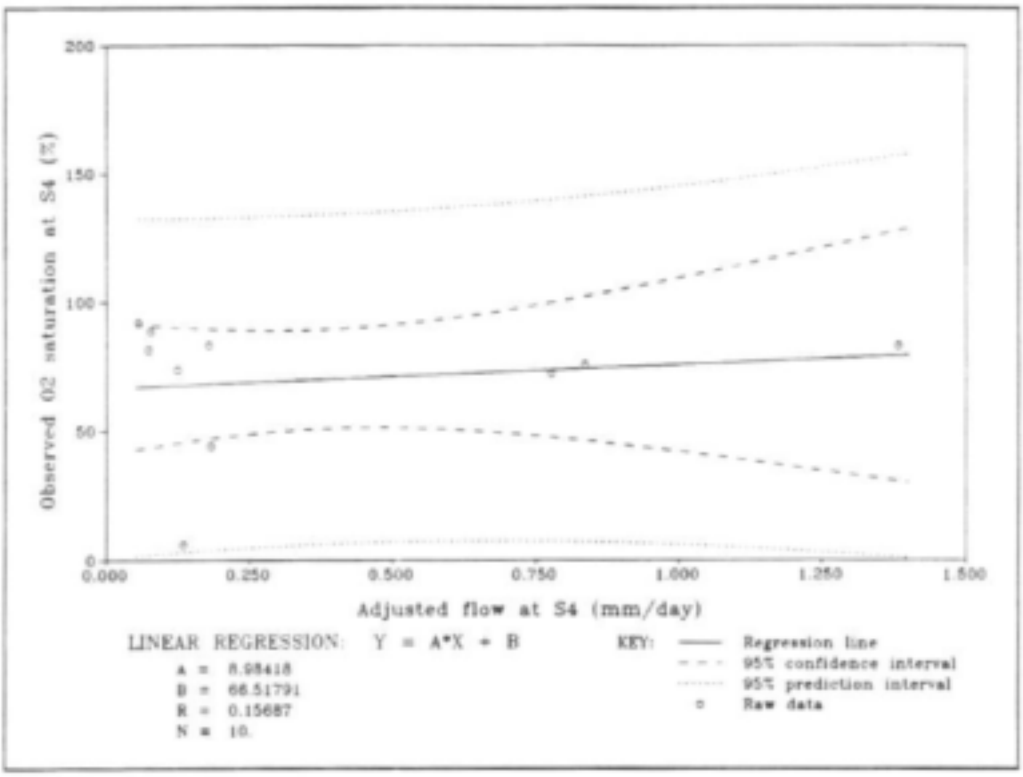


Figure B.7.7 : Linear regression between flow and oxygen saturation at station S4

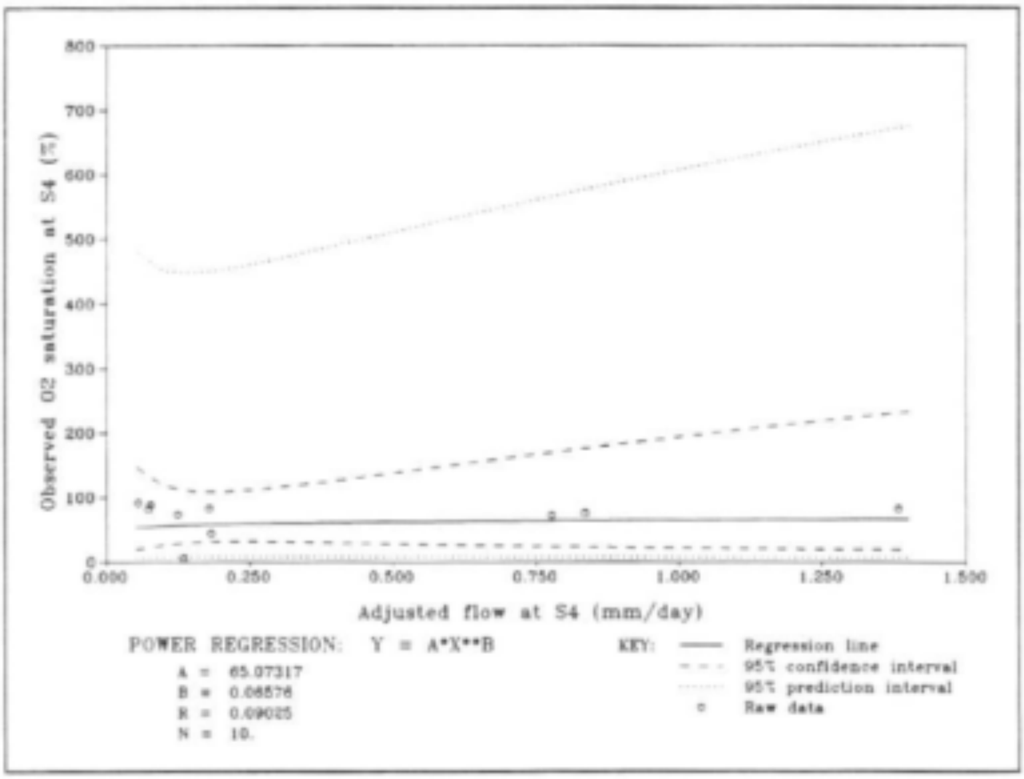


Figure B.7.8 : Power regression between flow and oxygen saturation at station S4

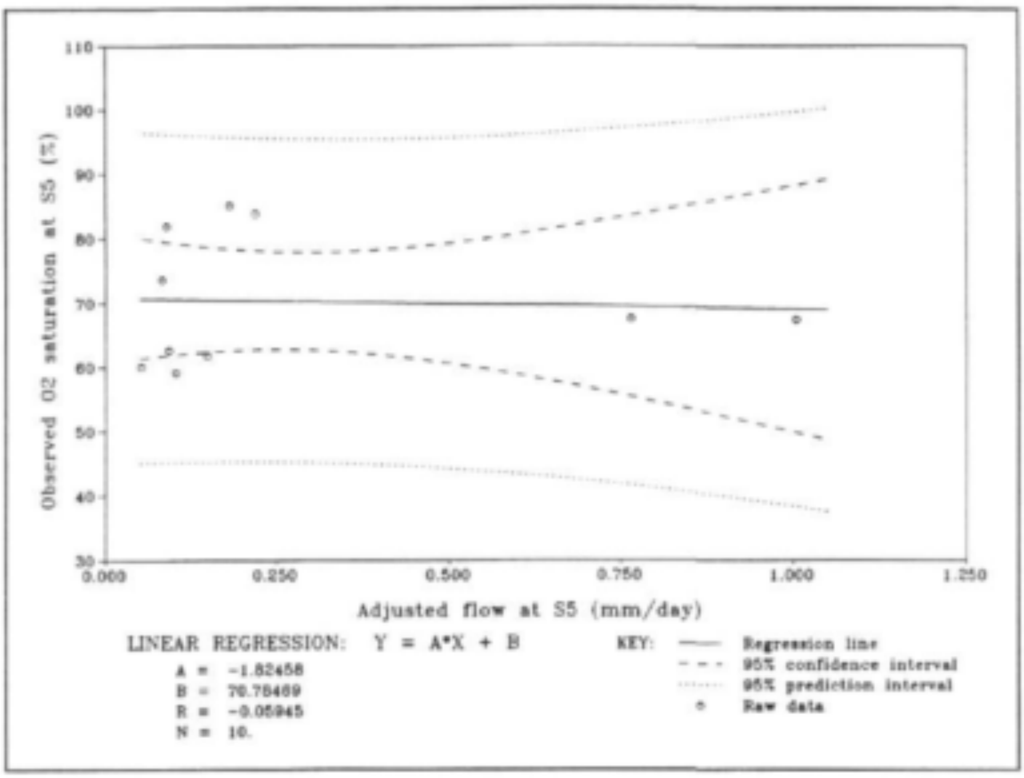


Figure B.7.9 : Linear regression between flow and oxygen saturation at station S5

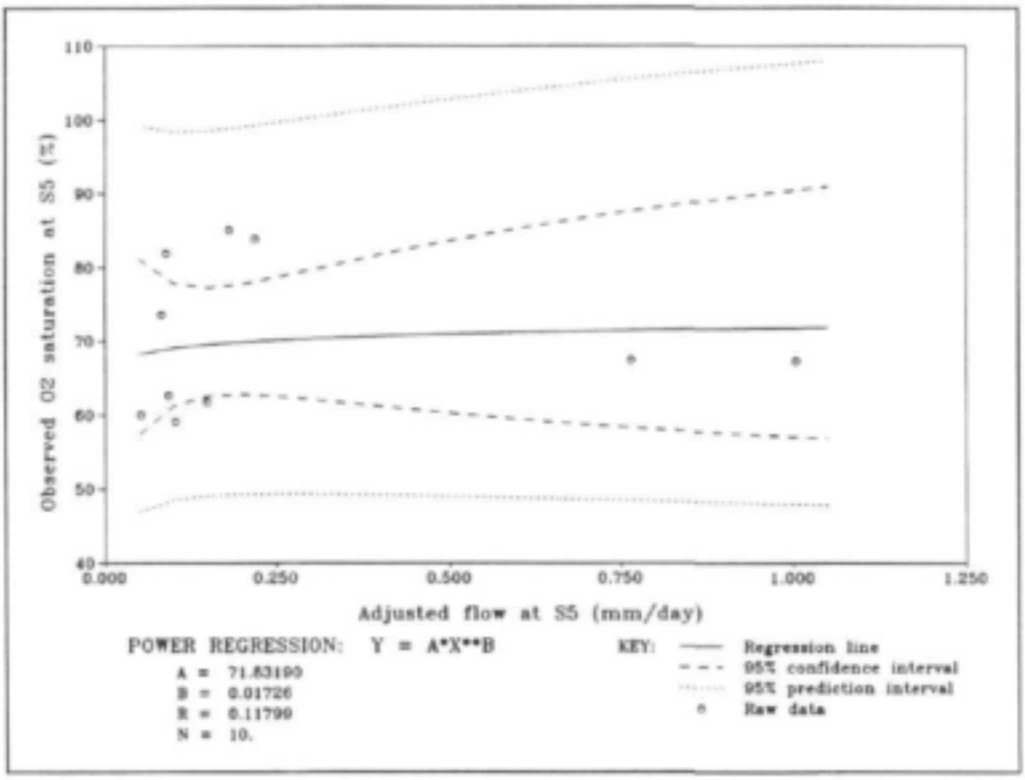


Figure B.7.10 : Power regression between flow and oxygen saturation at station S5

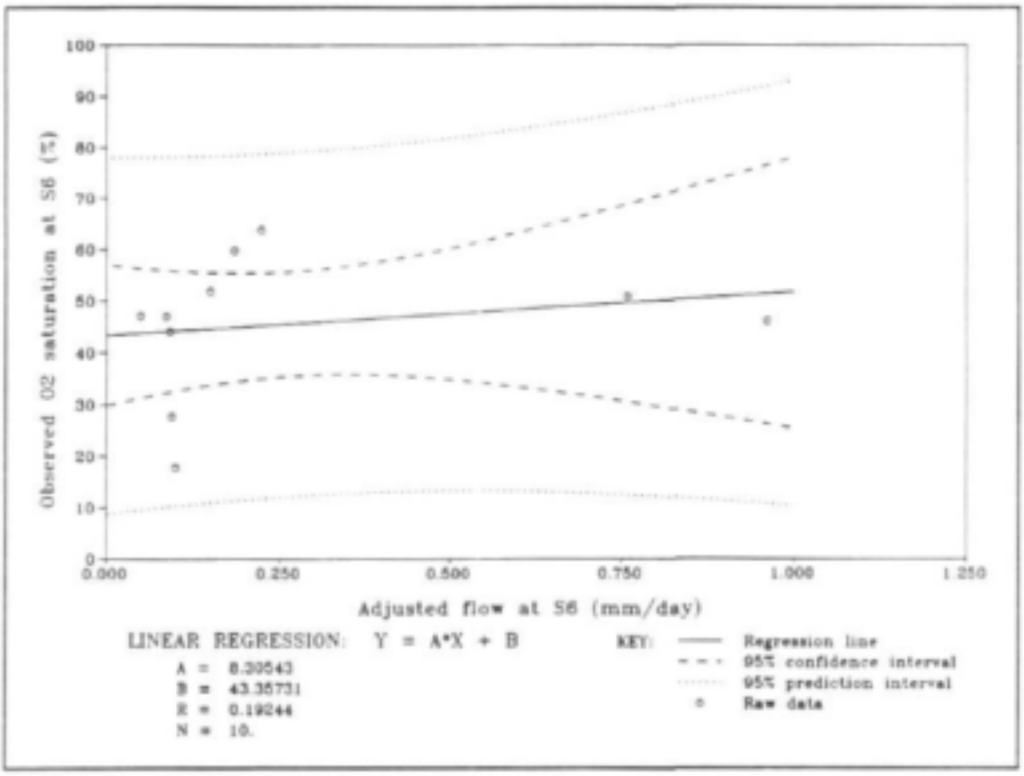


Figure B.7.11 : Linear regression between flow and oxygen saturation at station S6

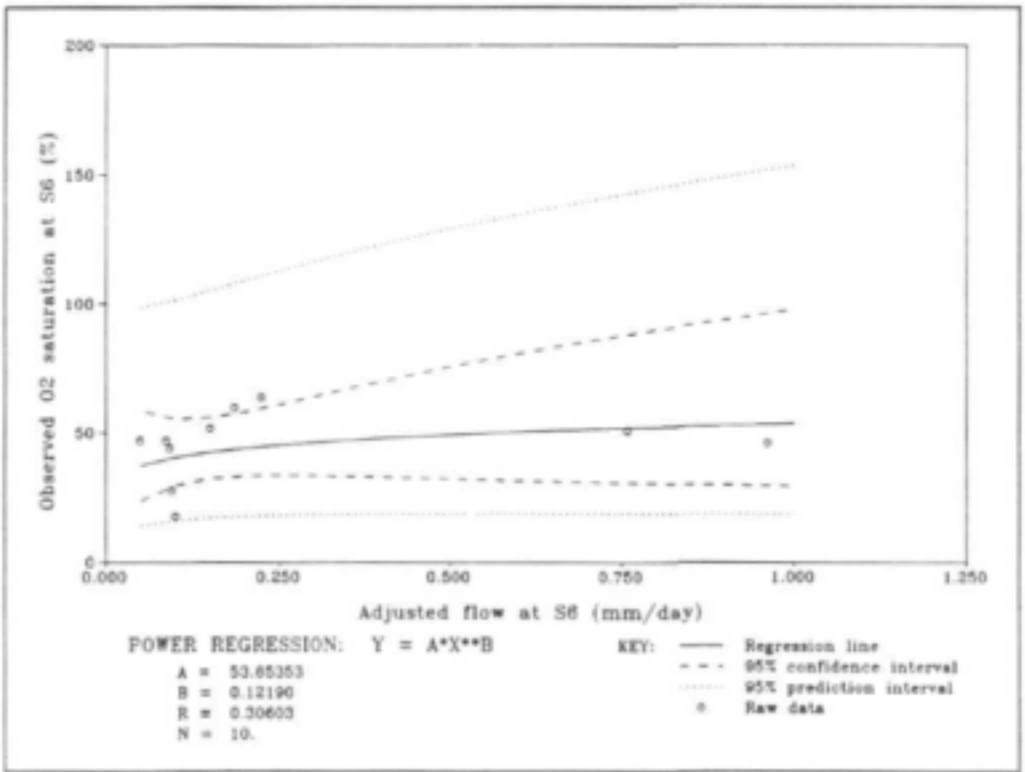


Figure B.7.12 : Power regression between flow and oxygen saturation at station S6

APPENDIX C

GENERATED CONCENTRATION PLOTS

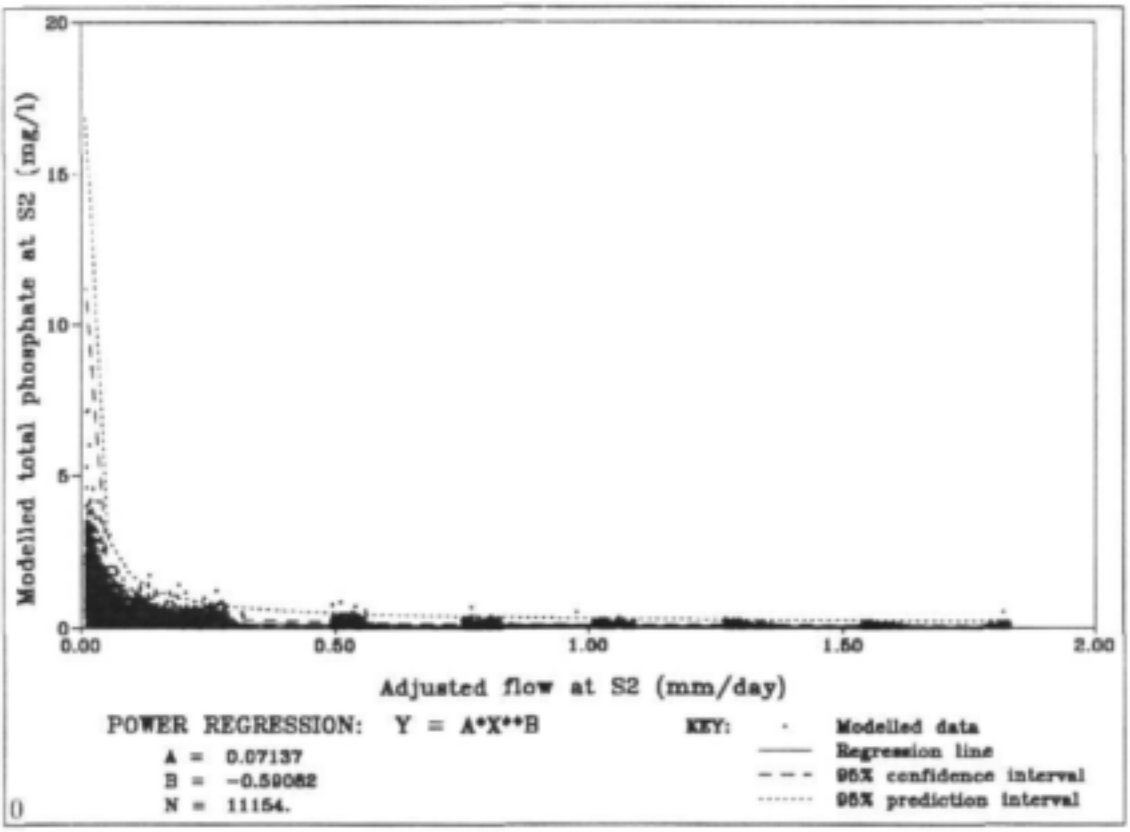


Figure C.1 : Estimated total phosphate concentrations at station S2

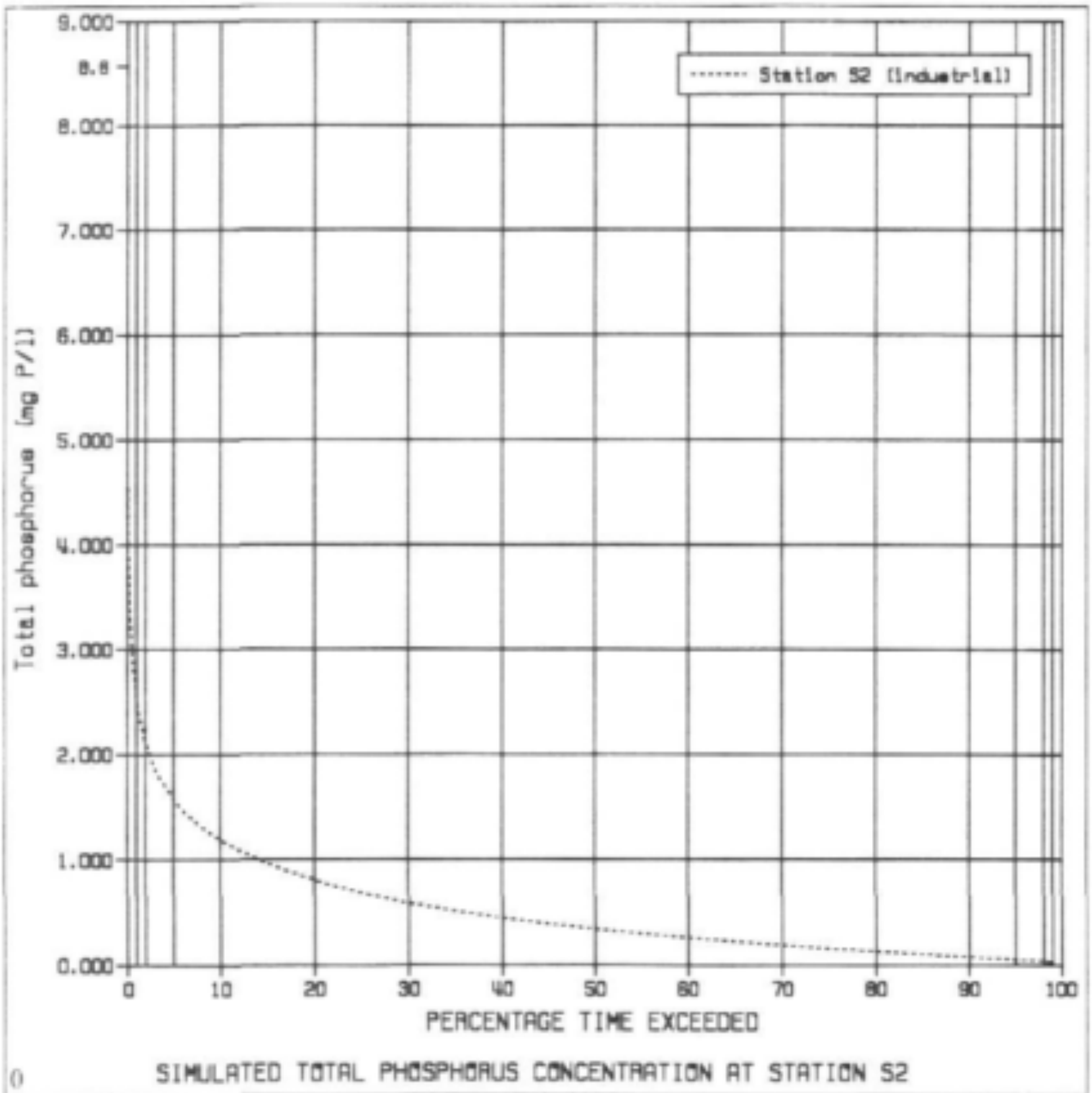


Figure C.2 : Duration curve of total phosphate concentration at station S2

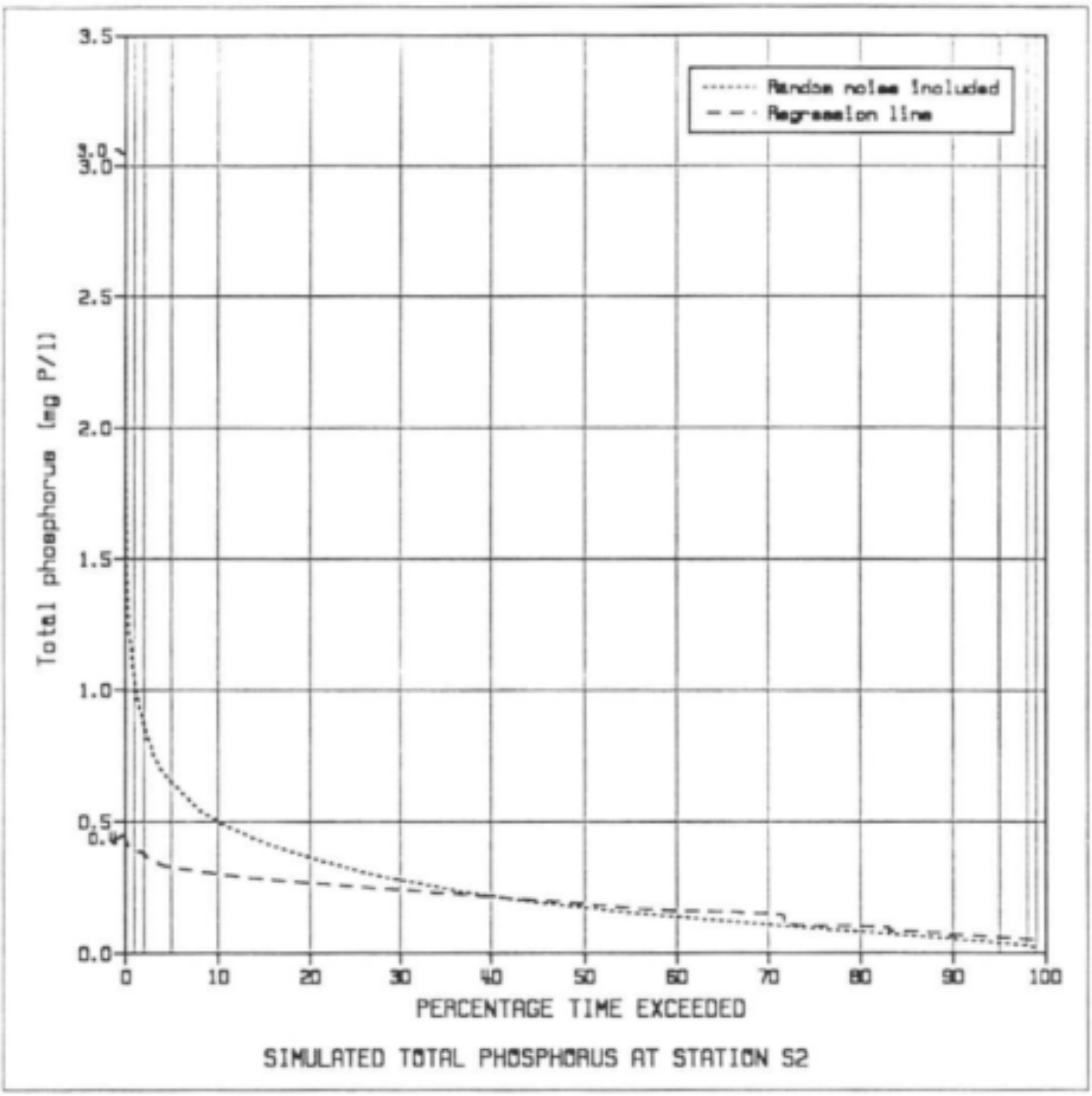


Figure C.3 : Effect on duration curve of not adding random noise to regression

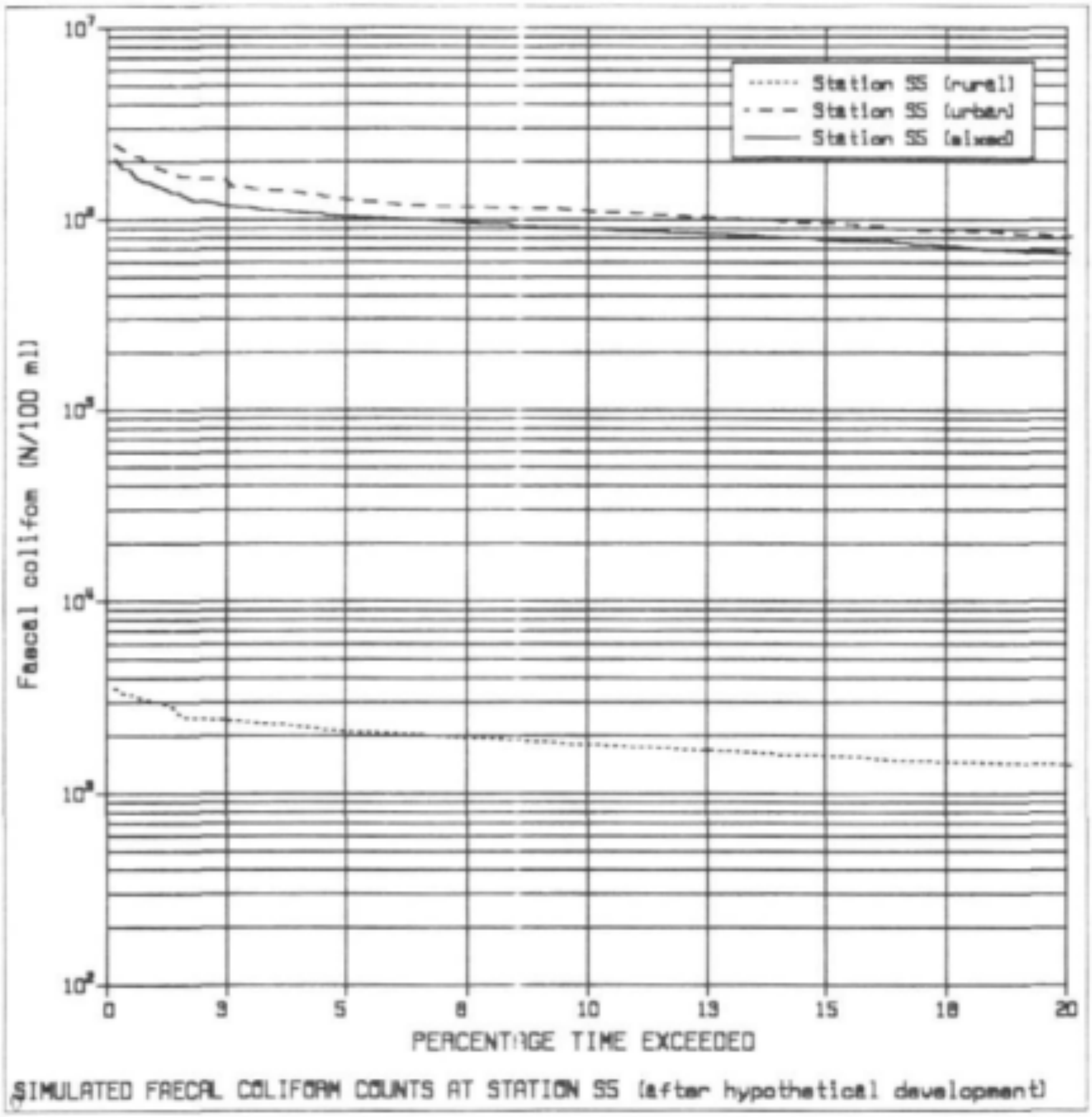
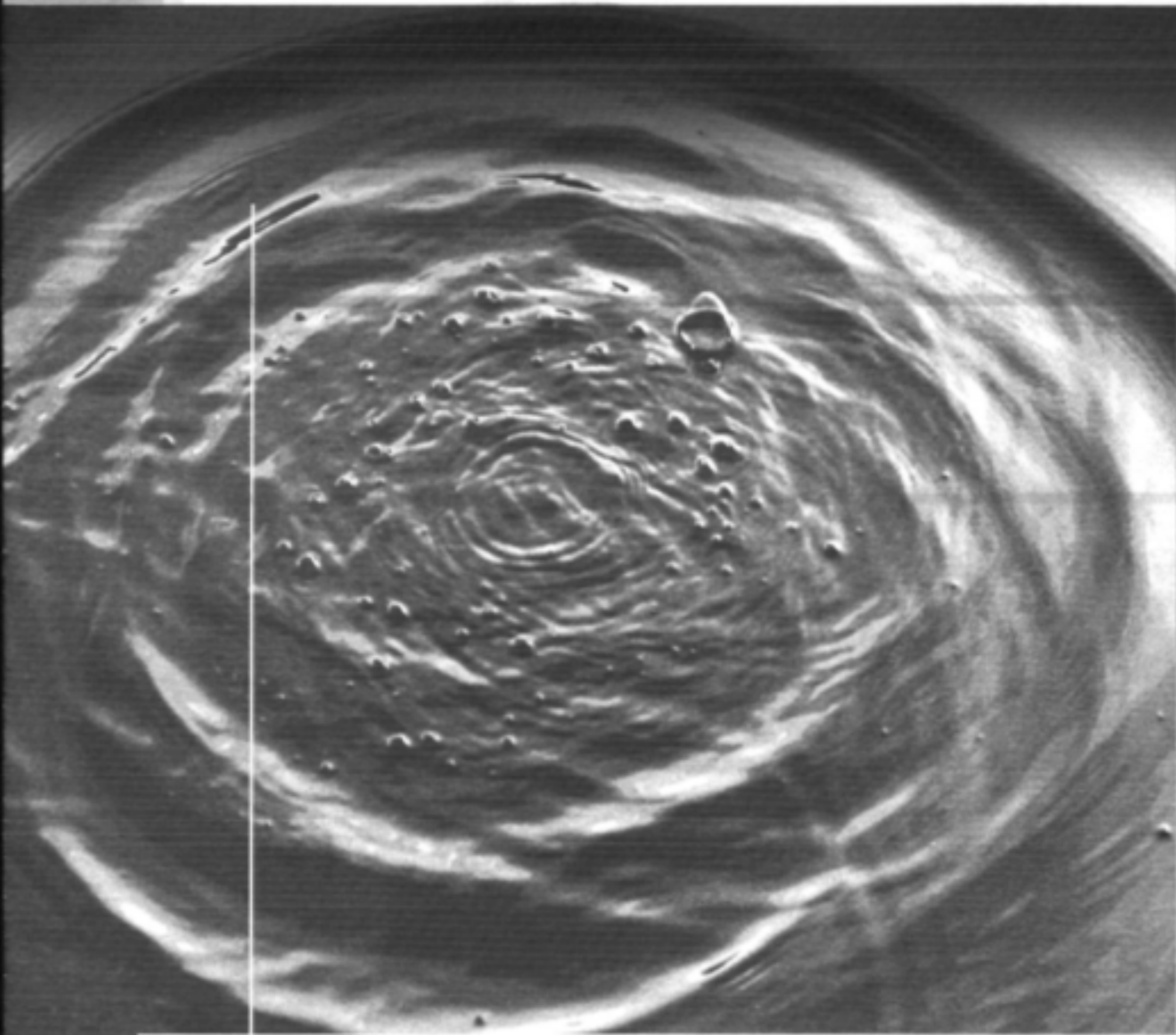
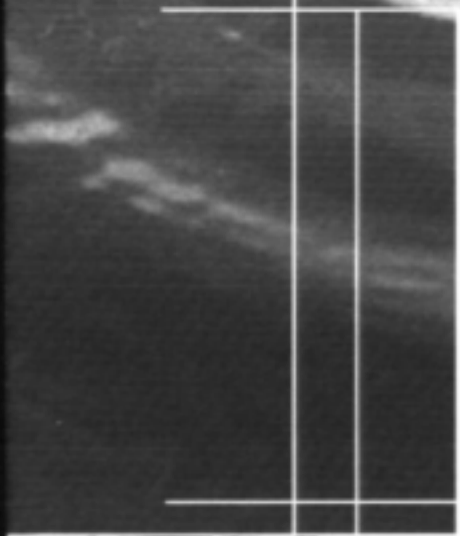
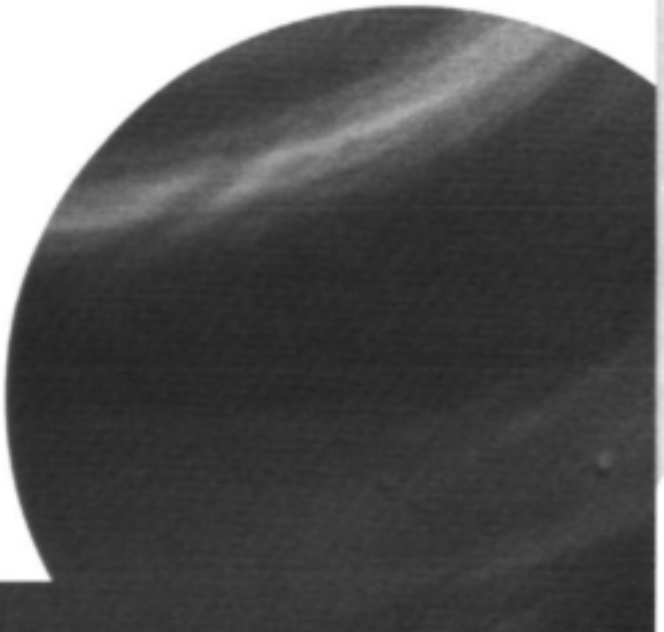
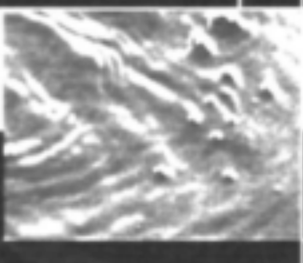


Figure C.4 : Hypothetical duration curves of faecal coliform at station S5 assuming high density residential development of half the catchment









Water Research Commission
PO Box 824, Pretoria, 0001, South Africa
Tel: +27 12 330 0340, Fax: +27 12 331 2565
Web: http://www.wrc.org.za

1868457532