

Artificial Neural Network to Support Thermohydraulic Design Optimization for An Advanced Nuclear Heat Removal System

Artit Ridluan^a, Ondrej Linda^b, Milos Manic^b, and Akira Tokuhiro^a

^aNuclear Engineering Program, Department of Mechanical Engineering, University of Idaho

^bDepartment of Computer Science, University of Idaho

1776 Science Center Dr, Suite 306, Idaho Falls, Idaho, 83402

Tel: +1(208) 533-8125, Fax: +1(208) 282-7950, Email aridluan@vandals.uidaho.edu.

Abstract – The U.S. Department of Energy (DOE) is leading a number of initiatives, including one known as the Next Generation Nuclear Plant (NGNP) project. One of the NGNP nuclear system concepts is the Very High Temperature (gas-cooled) Reactor (VHTR) that may be coupled to a hydrogen generating plant to support the anticipated hydrogen economy.

For the NGNP, an efficient power conversion system using an Intermediate Heat Exchanger (IHX) is key to electricity and/or process heat generation (hydrogen production). Ideally, it's desirable for the IHX to be compact and thermally efficient. However, traditional heat exchanger design practices do not assure that the design parameters are optimized.

As part of NGNP heat exchanger design and optimization project, this research paper thus proposes developing a recurrent-type Artificial Neural Network (ANN), the Hopfield Network (HN) model, in which the activation function is modified, as a design optimization approach to support a NGNP thermal system candidate, the Printed Circuit Heat Exchanger (PCHE). Four quadratic functions, available in literature, were used to test the presented methodology. The results computed by an artificially intelligent approach were compared to another approach, the Genetic Algorithm (GA). The results show that the HN results are close to GA in optimization of multi-variable second-order equations.

I. INTRODUCTION

More and more engineering and scientific research relies on use of complex computational codes and performing computational experiments (simulations) to investigate a design point of an system of interest. In recent years, Experimental Fluid Dynamics (EFD) and Computational Fluid Dynamics (CFD) have been used to study complex flow systems and phenomena. The corresponding structural performance can then be analyzed and investigated using a Finite Element Method (FEM). The design space is studied by changing the system parameters during iterative computations; typically, based on the designer's experience or via a trial-and-error basis. Thus for a given thermal component the optimal design point and best efficiency is not guaranteed. For complex convective phenomena, the design exercise (and optimization) can be computationally expensive.

To address these concerns, a combination statistical-mathematical technique is introduced. The approximate model or so-called surrogate model (also metamodel), characterized by input factors and corresponding responses, is constructed. Specifically, the "Response Surface Methodology (RSM)" has been successfully used to construct these surrogate models [1,2]. The RSM facilitates characterization of system behavior through functional relationships between input and output as shown by Fig. 1.

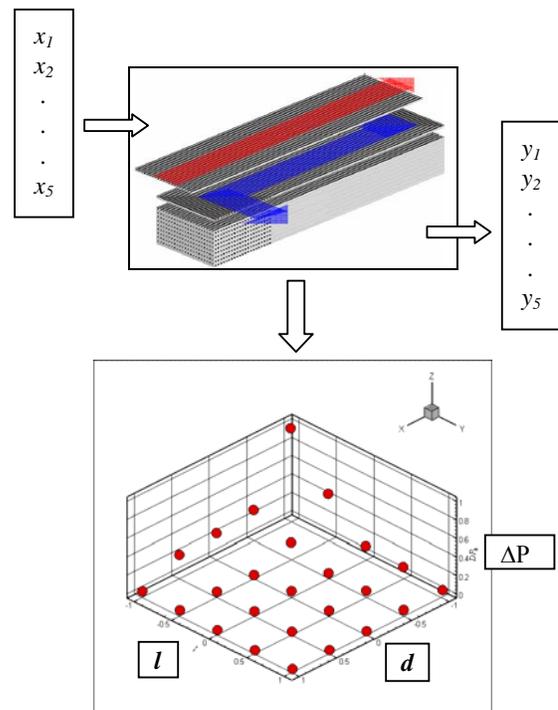


Fig 1. NGNP PCHE design and functional requirements and response surface plot of hot- and cold-sided pressure drop with diameter and zigzag length

Fig. 1 shows a NGNP PCHE 3D distribution of data points on coordinates corresponding to targeted design parameters (as example): diameter (d), zigzag length (l) versus the hot- (DP_h) and cold- (DP_c) side pressure drops. The data points were each generated obtained by performing CFD simulation for a set of design parameters. Here one can imagine the ‘surface’ produce by such data. In general Fig. 1 shows that smaller diameter and shorter the zigzag length, higher pressure drops are observed. To mathematically describe the surface, a multi-variable polynomial model was applied. The generated equation is then used to obtain the optimal solution of the system.

In the process of the RSM-based optimization and analysis, the Multi-Variable Quadratic Polynomial (MVQP) equation or the Quadratic Response Surface (QRS) are often used in regions where an optima (or optimum) exists. For example, Timothy et al. [3] constructed QRS models for weight, thrust, and GLOW in aerospike design. Chiang and Chang [4] applied RSM to construct MVQP equations for parametric optimization of a pin-fin heat sink. In their work, the QRS models for thermal resistance and pressure drop were constructed. Lepadatu [5] optimized a springback in a bending process by applying the QRS model; the springback was related to the die corner radius and the tool clearance.

In the current work, the multi-variable quadratic polynomial equations are written in the conventional summation form as,

$$y = \beta_0 + \sum_{i=1}^n \beta_{1i} x_i + \sum_{i=1}^n \beta_{2i} x_i^2 + \sum_{i<j}^n \sum_{j=1}^n \beta_{ij} x_i x_j \quad (1)$$

Here y is the response or output of the system, x_i and x_j are the input factors/design variables, β_0 is the intercept and $\beta_{1,2}$ and β_{ij} are the coefficients of the MVQP of the first and second orders, respectively.

When constructed, the MVQP equation is applied to the design optimization task. Various techniques have been used to obtain the optimal response point. Chiang [6] minimized QRS models of thermal resistance and pressure drop by a Sequential Approximation Optimization (SAO). Adinarayana [7] solved the QRS optimization by means of multi-stage, Monte-Carlo for the critical medium components, of maximum alkaline protease¹ production. Gangadharan [8] considered the optimum level of the alpha amylase production² by bacillus amyloliquefaction using the Box-Behnken design³. Shyy [9] used artificial neural network techniques to enhance the QRS models for rocket engine trajectory predictions. Here the rocket injector optimization problem was solved using the Generalized

Reduced Gradient (GRG) method.

The weakness of the existing optimization methods for solving the MVQP equations are their stochastic nature and high complexity. To alleviate these drawbacks, the present work proposes a design optimization methodology for multi-variable quadratic polynomial equation by means of artificial neural network, namely via the Hopfield network. The developed Hopfield Network (HN)-based design optimization constitutes a novel method that is relatively easy to implement, fast, and does not require a specific complex algorithm for solving the MVQP equation. Moreover, we demonstrate the ability of the HN method to solve MVQP-based design optimization problems; thus encompassing many types of engineering problems.

The paper is organized as follows. Section II discusses the general optimization problem. We then describe the artificial neuron and the Hopfield neural network model in Sections III and IV respectively. Results from the HN-based design optimization and comparison against the Genetic Algorithm (GA) approach developed by Schreyer [10] are presented in Section V. Conclusion to date are given in Section VI.

II. GENERAL OPTIMIZATION PROBLEM

In the design optimization exercise, an optimization model is composed principally of three components: design variables or parameters, objective function, and constraints. Fundamentally, the optimal design problem is mathematically formulated as follows. For a single objective optimization:

$$\begin{aligned} &\text{Optimize,} && y(X) \\ &\text{subject to,} && h(X) = 0 \\ & && g(X) \leq 0 \\ & && X^l \leq X \leq X^u \end{aligned}$$

Here $y(X)$ is cast in the form of the MVQP equation.

For many engineering and science problems, the process of design optimization begins by defining the input factors that are relevant for the system of interest. In the design optimization, the input factors are associated with the design variables (X). For specific design variables, system outputs are generated. Based on the relationship between design variables and system output, an objective function $y(X)$ can be constructed. The function $y(X)$ is either maximized or minimized (optimized), while satisfying applicable equality and inequality constraint functions $h(x)$ and $g(x)$; further, X is bounde by lower and upper limit values, X^l and X^u .

¹ Protease is any enzyme that conducts proteolysis; that is, initiates protein catabolism by hydrolysis of the peptide bonds linking amino acids together in a polypeptide chain. Cf. en.wikipedia.org/wiki/Protease

² A-amylase is a major form of amylase found in humans; it can cut alpha-bonds in large sugar molecules.

³ These biological examples are given to substantiate its general applicability in science and technology.

III. ARTIFICIAL NEURON AND HOPFIELD MODELS

A biological neuron is idealized as a computational unit, as in Fig. 1, and connected to other neurons (synapses) and receives input via its input connections (dendrites). A weighted sum of input signals (net value) is then compared to the threshold or a certain activation function. An output signal is transmitted via other synapses to other neurons.

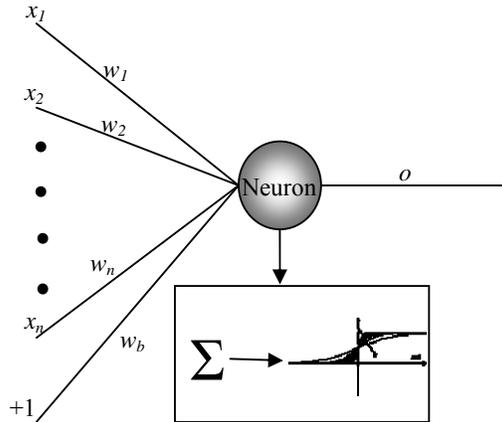


Fig. 2. Single artificial neuron with input and output connections.

To mimic a biological neuron, its artificial counterpart performs similar functions; that is, it calculates a weighted sum of input signals (represented by Σ) and compares it against the activation function or a threshold (indicated by graphical interpretation), as shown in Fig 2.

If the weighted sum of input signals (*net*) is above the threshold, a neuron will generate an output signal. The general neuron function can be expressed in conventional form as,

$$net = w_i x_i + w_b \quad (2)$$

Here, x_i is the neuron's input, while w_i and w_b are the neural and bias weights.

An output of neuron (o) is computed as,

$$o = f(net) \quad (3)$$

where f represents the activation function.

The applicability of artificial neural networks (ANN) are not only limited to data classification and prediction but equally, it can be used to solve optimization problems. For example, Lee and Chen [11] applied the back-propagation neural network to solve a generalized design optimization problem. In this section, the recurrent type of ANN, the Hopfield Network (HN), is described and used to solve specific problems frequently found in RSM-based design optimization.

The HN is illustrated in Fig. 3. The HN consists of a

single layer neural network. The output of each neuron is fed back to the input, which is propagated to each neuron except itself. Therefore, no self feedback (dash line) is presented in the HN. Further, the HN constitutes a specific network with delayed dynamic flow of input data. The HN can be considered dynamic since the inputs are fed to the network and the HN will 'run' until it stabilizes. The HN can also be considered a network with delay because the input data are propagated through the recurrent feedback connections and given time is required for the network to stabilize. The stabilization of the HN 'energy' represents the optimal solution.

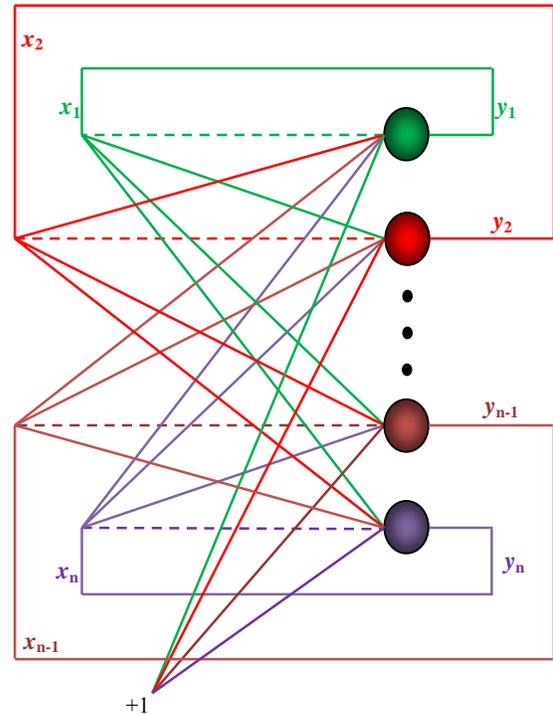


Fig. 3. A representation of the Hopfield network.

The general form of the HN is written in the indexed notation as follows,

$$net_i = \sum_{\substack{j=1 \\ i \neq j}}^n w_{ij} x_j + x_i^{ini} + b_i \quad (4)$$

Here net_i is the weighted sum of each neuron, w_{ij} is the neural weight associated with an output I_j of neuron j , b_j is the bias weight of each neuron, and x_i^{ini} is the initial input.

For the HN, the so-called computational energy function is used to evaluate the stability of the system. The principle is that a dynamic system should stabilize at some point and that this state of stability is identified by no further change in its energy state. The Lyapunov Energy Function is written as,

$$E = -\frac{1}{2} \sum_{\substack{j=1 \\ i \neq j}}^n w_{ij} y_i y_j + b_i y_i \quad (5)$$

or in matrix form,

$$E = -\frac{1}{2} Y' W Y + B Y \quad (6)$$

Here y_i and Y are the output and the output matrix respectively.

The gradient of the energy function is then given as:

$$\nabla E = -\frac{1}{2} (W' + W) Y + B Y \quad (7)$$

If the weight matrix, W , is symmetrical and contains zeros along its diagonal, the energy equation reduces to,

$$\nabla E = -W Y + B Y \quad (8)$$

The change in energy is proportional to the gradient and the output is,

$$\Delta E = (\nabla E)' \Delta Y \quad (9)$$

Substituting eq. (8) into eq. (9) gives:

$$\Delta E = (-W' Y + B Y) \Delta Y = -(W' Y - B Y) \Delta Y \quad (10)$$

or in the tensor form,

$$\Delta E = -\sum_{j=1}^n (w'_{ij} y_j - b_i y_i) \Delta y_i = net_i \Delta y_i \quad (11)$$

We note here that for HN, the optimal point is equivalent to no change in the energy state ($\Delta E=0$). To facilitate considerations of constraints, a linear activation function is modified. A comparison of a simple linear to a modified linear activation functions is shown in Figs . 4 (a) and (b).

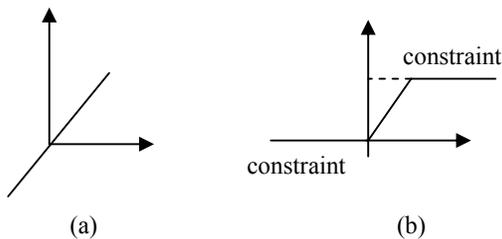


Fig. 4. The plot of (a) the linear activation function and (b) the modified linear activation function.

To apply the HN for the optimization problem, the quadratic polynomial equation is first differentiated with respect to x_i and then rearranged into a form, expressed by

equation (6). In this way, the weight and bias sets of the HN are specified. An example and also an algorithm used for training the HN is given in section IV.

IV. OPTIMIZATION RESULTS

The HN optimization exercise for a quadratic model is performed by testing the following functions; first,

$$f(x_1, x_2, x_3) = 3x_1^2 + 4x_2^2 + 3x_3^2 + 2x_1x_2 - x_2x_3 - 2x_1x_3 + 2x_1 - 3x_2 + 4x_3 \quad (12)$$

The second test function is,

$$f(R, C) = 5.2559 + 0.1378R + 0.1613C - 0.0345R^2 - 0.1102C^2 + 0.0285RC \quad (13)$$

subject to, $2 \leq R \leq 6$, $2\% \leq C \leq 15\%$

The third test function is,

$$f(x_1, x_2, x_3) = 2.13 - 0.15x_1^2 - 0.13x_2^2 - 0.052x_3^2 + 0.04x_1x_2 + 0.02x_2x_3 - 0.047x_1x_3 + 0.094x_1 + 0.17x_2 + 6.25 \times 10^{-3} x_3 \quad (14)$$

subject to, $-1 \leq x_1, x_2, x_3 \leq 1$

The fourth test function is,

$$f(x_1, x_2, x_3, x_4, x_5) = -3.59538 + 2.70570x_1 + 0.249231x_2 + 3.51131x_3 + 0.540555x_4 + 1.00495x_5 - 6.450564x_1^2 - 0.928410x_2^2 - 30.8026x_3^2 - 0.03330256x_4^2 - 0.0318385x_5^2 - 0.290625x_1x_2 + 4.03906x_1x_3 + 0.357031x_1x_4 + 0.108218x_1x_5 + 0.693750x_2x_3 + 0.484875x_2x_4 - 0.0828704x_2x_5 - 0.0467188x_3x_4 + 0.730324x_3x_5 - 0.0589120x_4x_5 \quad (15)$$

subject to, $0.3 \leq x_1 \leq 0.7$, $0.3 \leq x_2 \leq 0.8$, $0.15 \leq x_3 \leq 0.35$, $0.5 \leq x_4 \leq 2.5$, $10.20 \leq x_5 \leq 12.90$

Note that the quadratic models available in literature were obtained by the technique, so called Response Surface Methodology (RSM) [1,2].

The function expressed by eq. (12) is a quadratic polynomial test function with three independent variables (x_1 , x_2 , and x_3). Lepadatu et al [5] proposed a relationship between springback to a die's corner radius (R) and tool clearance (C) in the quadratic form as written in eq. (13). The third equation originates from the bio-energy field. It presents the MVQP equation for the optimization of the hydrogen production [12]. Eq. (14) was constructed to account for process control conditions for hydrogen production. The rate of hydrogen production,

$f(x_1, x_2, x_3)$, is written as the function of three factors: glucose (x_1), Fe^{2+} (x_2), and Mg^{2+} (x_3). Finally, the natural length of a 'cyclone' [13] is described in terms of the MVQP equation and shown as eq. (15). The cyclone here is a fluid power system that moves and separates airborne solid particulate through flow channels. The natural length, $f(x_1, x_2, x_3, x_4, x_5)$, is written as the function of five factors: diameter (x_1), height of the inlet (x_2), width of the inlet (x_3), difference of the cylinder and depth of the vortex finder (x_4), and the logarithm of the Reynold number (x_5). This problem is substantially more complex due to the increased number of input variables and constraints.

To demonstrate the HN in the search for an optimal point, equation (12) is selected. As noted, we then first differentiate the second-order polynomial equation (12) with respect to x_1, x_2 , and x_3 , respectively;

$$\left. \begin{aligned} \frac{\partial f(x_1, x_2, x_3)}{\partial x_1} &= 6x_1 + 2x_2 - 2x_3 + 2 = 0 \\ \frac{\partial f(x_1, x_2, x_3)}{\partial x_2} &= 2x_1 + 8x_2 - x_3 - 3 = 0 \\ \frac{\partial f(x_1, x_2, x_3)}{\partial x_3} &= -2x_1 - x_2 + 4x_3 + 4 = 0 \end{aligned} \right\} \quad (16)$$

By rearranging eq. (16), the weight (W) and bias (B) matrices of the HN are specified, respectively,

$$W = \begin{bmatrix} 0 & -\frac{1}{3} & +\frac{1}{3} \\ -\frac{1}{4} & 0 & +\frac{1}{8} \\ +\frac{1}{2} & +\frac{1}{4} & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -\frac{1}{3} \\ +\frac{3}{8} \\ -1 \end{bmatrix}$$

Note that the first row of matrix W and B is the weight and bias of the first neuron and so on. The HN algorithm for optimization is schematically shown in Fig. 5. A set of the initialized inputs is used to compute the 'energy function' and the net (weighted sum) of each neuron in the HN. The output associated with each neuron is then fed back to the input. According to the use of modified linear activation function (see Fig. 4), the outputs defined by eq. (3) is directly proportional to the net. The HN will 'run' until a stable condition has reached ($\Delta E=0$).

A comparison of optimal values calculated by the HN and GA for each problem, are given in Tables I, II, III, and IV for eq. (12), (13), (14), and (15) respectively.

TABLE I
 Optimal Results for eq. (12)

Variables	HN	GA
x_1	-0.93150	0.010000
x_2	0.43840	0.373812
x_3	-1.35620	0.010000

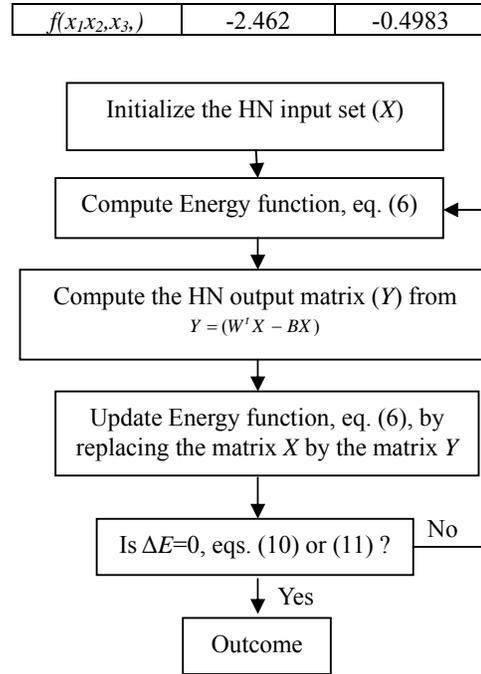


Fig. 5. The HN training algorithm for the optimization.

TABLE II
 Optimal Results for eq. (13)

Variables	HN	GA
R	2.823200	2.835000
C	2.000000	2.000000
$f(R, C)$	5.406	5.406

TABLE III
 Optimal Results for eq. (14)

Variables	HN	GA
x_1	0.4068	1
x_2	0.7175	0.597513
x_3	0.0143	2.21×10^{-8}
$f(x_1, x_2, x_3)$	2.210	2.153

TABLE IV
 Optimal Results for eq. (15)

Variables	HN	GA
x_1	0.33900	0.7000
x_2	0.30000	0.3000
x_3	0.23510	0.2590
x_4	0.54180	2.5000
x_5	12.9000	12.900
$f(x_1, x_2, x_3, x_4, x_5)$	6.078	5.661

Three-dimensional surface plots of each equation, together with the optimal results are illustrated in Fig. 5 for the both the GA and HN methods. In the figures shown, it is evident that the results calculated by the HN provide a comparatively optimal solution in contrast to application of the GA. In the research both GA and HN were initialized at the same initial location.

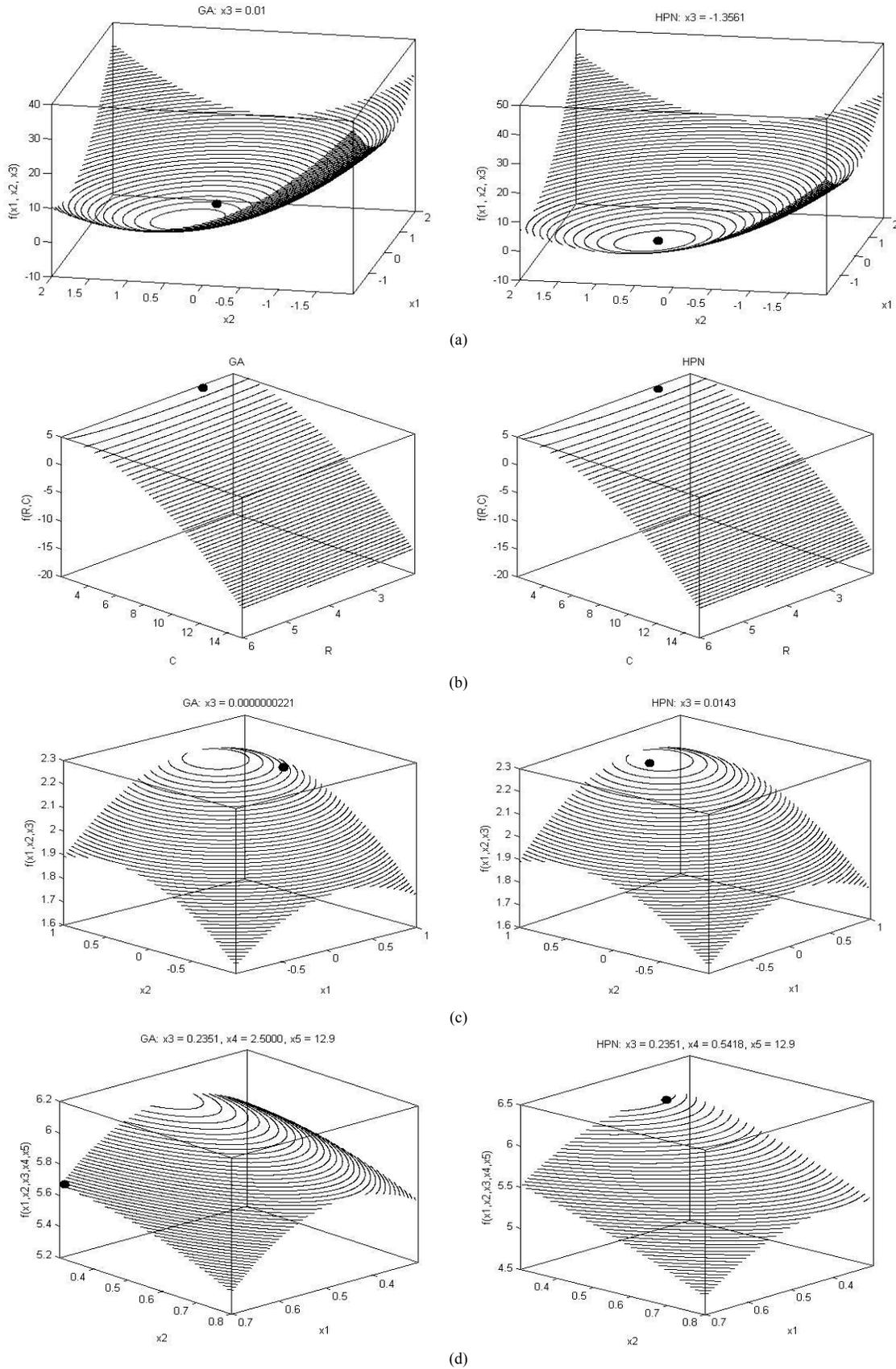


Fig. 5. The optimal surface plots for (a) eq. (12), (b) eq. (13), (c) eq. (14), and (d) eq. (15) computed by GA (left) and HN (right).

IV. CONCLUSIONS

To support the anticipated application of compact heat exchangers for future advanced nuclear power plants, we presented the Hopfield Network (HN) method for optimization of Multi-Variable Quadratic Polynomial (MVQP) equation. This equation characterized design parameters and functional requirements of engineered components and have broadly been applied in science and technology. A modified linear activation function was proposed to facilitate modeling of the constraints of the system under design.

The HN architecture was tested on four different reference quadratic functions, selected from available literature. The equations defined problems with two, three and five input variables.

Further, a comparison was made between the HN and the GA. The optimal results revealed that the present approach appeared to perform better than the previous approach for all test functions. It is apparent that the HN-based design optimization, is a method that is easier to implement, relatively faster, and one that does not require development of a complex algorithm for solving the MVQP optimization problem. Further, we found GA to be sensitive to the input initialization (different initial set of inputs give different optimal solutions), while this sensitivity was not encountered when using HN. While GA generally can find optima in a defined region (see Schreyers [10] for example) HN seems better suited to investigate a design parameter space. Obviously, this requires further work. Computing times for both GA and HN are on the order of minutes; each CFD simulation took many more hours.

Finally, we note that the results presented here is part of a larger NGNP PCHE design optimization study and thus, still on-going. Further, we plan to extend HN-based optimization to higher order polynomial equations and to develop a multilayer HN to serve as multi-objective optimizer. These results will be reported in the future.

ACKNOWLEDGMENTS

The first author would like to acknowledge the financial and technical support from Mr. Michael Patterson, Hydrogen Process & Heat Transport Systems NGNP Engineering, Idaho National Laboratory. The first and second authors also acknowledge the support of Dr. Fred Gunnerson, Nuclear Engineering Program, Mechanical Engineering Department, University of Idaho.

NOMENCLATURE

net_i	weighted sum of inputs each neuron
w_{ij}	neural weights
b_j	bias weights of each neuron
E	Lyapunov Energy Function
f	Activation function

REFERENCES

1. R. H. MYERS, *Response Surface Methodology*, Published by Author, Blacksburg, VA (1976).
2. R. H. MYERS and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization using Designed Experiment*, John Wiley & Son, New York (2002).
3. T. W. SIMPSON, T. M. Mauery, J. J. Korte, and F. Mistree "Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization," 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, **AIAA 98-4758**, 1-11 (1998).
4. K. T. CHIANG and F. P. Chang, "Application of Response Surface Methodology in the Parametric Optimization of a Pin-Fin Type Heat Sink," *International Communication in Heat and Mass Transfer*, **33**, 836-845 (2006).
5. D. LEPADATU, R. Hambli, A. Kobi, and A. Barreau, "Optimization of Springback in Bending Process using FEM Simulation and Response Surface Model," *International Journal of Advanced Manufacturing Technology*, **27**, 40-47 (2005).
6. K. T. CHIANG, "Modeling and Optimization of Designing Parameters for A Parallel-Plain Fin Heat Sink with Confined Imping Jet using the Response Surface Methodology," *Applied Thermal Engineering*, **27**, 2475-2482 (2007).
7. K. ADINARAYANA and P. Ellaiah "Response Surface Optimization of the Critical Medium Components for Production of Alkaline Protease by a Newly Isolated Bacillus sp," *J Pharm Pharmaceut Sci*, **5**, 3, 272-278 (2002).
8. D. GANGADHARAN, S. Sivaramkrishnan, K. M. Nampoothiri, R. K. Sulumaran, and A. Pandey "Response Surface Methodology for the Optimization of Alpha Amylase Production by Bacillus Amyloliquefaciens," *Bioresource Technology*, **99**, 11, 4597-4602 (2008).
9. W. SHYY, P. K. Tucker, and R. Vaidyanathan "Response Surface and Neural Network Techniques for Rocket Engine Injector Optimization," 35th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibition, **AIAA 99-2455**, 1-15 (1999).
10. A. C. SCHREYERS "Thermal and Structural Stud Wall Design Optimization in Excel using Genetic Algorithms," *Optimization in Engineering Design*, **MIE 616**, 1-8 (2005).
11. S-J. LEE and H. Chen, "Design Optimization with

- Back-Propagation Neural Network,”. *Journal of Intelligent Manufacturing*, **2**, 293-303 (1991).
12. W.-Q. GUO, N.-Q. Ren, X.-J. Wang, W.-S. Xiang, J. Ding, Y. You, and B.-F. Liu, “Optimizatin of Culture Conditions for Hydrogen Production by Ethanoligenens Harbinense B49 using Response Surface Methodology,” *Bioresource Technology*, **100**, 1192-1196 (2009).
 13. F. QIAN and M. Zhang, “Study of the Natural Vortex Length of a Cyclone with Response Surface Methodology,” *Computer and Chemical Engineering*, **29**, 2155-2162 (2005).