University College of Takestan

Available online at **http://UCTjournals.com**

UCT Journal of Research in Science ,Engineering and Technology

UCT . J. Resea. Scien. Engineer.Techno.(UJRSET)

2(2): 65-74(2014)

Volume 2,Issue 2

ISSN 2382-977X

Integrating Genetic Algorithm and Neural Network to Optimize Statistical Problems

Mahsa Mesgaran, M.Sc.[1], Seyed Hamid Reza Pasandideh[*], Ph.D.[2]

[1]. *Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Theran, Iran*

[2].*Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Theran, Iran*

[*] Corresponding Author Email: *m_mesgaran@yahoo.com*

Phone: +98 (21) 88830891, Fax: +98 (21) 88329213

ABSTRACT

Many real word product or process design problems involve multi response problems which have stochastic nature. This paper proposes a hybrid approach involved genetic algorithm and artificial neural network methodology to solve these problems. Usually, in these problems the relationship between responses and independent variables is indeterminate; therefore to generate required input data we are interested to use a method to approximate this relationship. Artificial Neural Network (ANN) is a methodology employed in this research to evaluate linear and nonlinear relationship between variables. We model the statistical multi response problem by three different multi objective decision making (MODM) techniques. Moreover, four different genetic algorithms are proposed in which four pairwise multiple comparisons statistical tests are used to control the random nature of the problem. Finally, the performance of the proposed methodology is demonstrated using a tow way analysis of variance (ANOVA) for a numerical example and the results are compared statistically..

Original Article:

Received 29 Jan. 2014
Accepted 15 Mar. 2014
Published 30 jun . 2014

Keywords: Statistical multi-response optimization problems Genetic algorithm; Neural networks; Multi objective decision making; Statistical pairwise multiple comparison.

1.Introduction

A common problem in product or process design is the selection of parameter levels for optimizing multiple responses simultaneously (Myers and Montgomery, 1995). As an example, in product systems with multiple quality characteristics (response variables) for productions, determination of the optimal control factors (independent variables) value such that the quality characteristics reach to desirable levels is an important proposition.

Note that in some cases, the problem has the stochastic nature. That means, if the input or independent variables are fixed, in each of the system or process execution, the outputs may be different (Pasandideh and Akhavan Niaki, 2006).

Generally in optimization process, the first step is estimating the relationship between independent and response variables. In this regard, experimental design techniques are conducted to demonstrate a cause and effect relation between one or more control factors and response variables (Kutner, Nachtsheim, Neter, & Li, 2005). One of these techniques to model and analyze problems is Response Surface Methodology (RSM) (Montgomery, 2001). It should be note that in the problems that the independent assumptions of the input variables are rejected or a complex relationship between response and control factor variables are observed, the relationship evaluated with this method has poor quality (Kim, & Lin, 2001).

Compared with classical modeling techniques, such as RSM, artificial neural network (ANN) is superior modeling technique for data sets showing nonlinear and complex relationships (Bourquin, Schmidli, Hoogevest, & Leuenberger, 1998). Thus, ANN has been recommended in the literature to approximate the relationship between variables to generate required input data. For example, Hsieh (2010) proposed a procedure based on ANNs technique to model a logical analysis to achieve parameter optimization of a multiple responses problem. Bashiri and Hosseininezhad (2009), presented a method based on neural network to estimate the relationship between responses and several independent variables for multiple response optimization and uses desirability of each response

[*] Corresponding Author

for forecasting. Šibalija and Majstorović (2010) proposed a novel and general approach to multiple response process optimizations for correlated responses. This approach is based on Taguchi quality loss function and multivariate statistical methods. The process modeling is performed using ANN, presenting an input for GA.

One of the techniques to optimize the multi responses problems is selected a strategy used to converts a multi response problem into single one. These methods are popular approaches to multi response optimization. Some methods to combine response variables to single one are multiple criteria decision making (MCDM) techniques such as goal programming, goal attainment, Global Criteria. Moreover, Derringer and Suich (1980), employed a technique called desirability function, first defined by Harrington (1965) to solve multi objective optimization problems. This method turns the multiple response problems into a single one and solves it by maximizing the combined desirability.

many, or even most, real engineering problems have multiple objectives, i.e., minimize cost, maximize performance, maximize reliability, etc. Genetic algorithm (GA) is a popular meta-heuristic approach that is particularly suitable for this class of problems (Konaka, Coitb & Smithc, 2006). Generally, GA is an appealing tool to solve optimization problems (Bäck, 1996). For instance, Baseler and Sepulveda (2000), integrated goal programming and Genetic Algorithm (GA) methods to solve the problem. Moreover, they used some statistical tests to control the random nature of the problem. Pasandideh and Akhavan Niaki (2006) presented two methods to solve multi response statistical problems. In these methods they used desirability function to model the problem, genetic algorithm for problem optimization and simulation methodology to generate required inputs. Amiri, Karimi, and Jamshidi (2008) proposed a hybrid method for optimizing statistical multi response problems. This methodology integrates simulation, fuzzy goal programming and genetic local search algorithm. C.B. Cheng, C.J. Cheng, and Lee (2002) presented a neuro-fuzzy and GA method for optimizing the multiple response problems. Another approach to multiple objective optimizations is lexicographic ordering, in which the designer ranks the objectives in order of importance. The optimum solution is then obtained by optimizing the objective function, starting with the most important and proceeding according to the assigned order of importance. For instance, Fourman (1985) suggested use of a GA based method on the lexicographic ordering problem. Kim and Rhee (2004) proposed a method based on the desirability function and GA and applied this method to optimize a welding process.

In this study, we propose a methodology based on a neural network, a modeling technique, and four different genetic algorithms for multiple response statistical optimization problems when the problem have stochastic nature and the relationship between control factors and the responses is unknown and so complicated.

The structure of the paper is as follow. The definition of the investigated problem comes in Section 2. Section 3 contains the modeling approaches and Section 4 defines the neural network approach and its application in the proposed problem. Genetic algorithm procedures, to solve the problem proposed in section 5. To evaluate the performance of the proposed methodology, numerical illustrations are given in Section 6. Conclusion presents in section 7.

2. Problem definition

In this study we supposed to have a production system in which events occur stochastically so that there are the statistical relationships between input and output variables. It means, if the input variables, $x_1, x_2, \dots, x_p$, are fixed, in each of the system or process execution, the outputs, $y_1, y_2, \dots, y_k$, may be different. We want to determine the input variable levels such that the outputs maximized or minimized or set to a target value. As an example for these problems, suppose that we are interested in improving the yield of a chemical process. We know from the result of the experiment that the two most important variables that influence the quality characteristics of the yield are temperature and reaction time. Also the quality characteristics can be the adhesiveness, percent conversion and the corrosion.

In order to generate required input data for the system, we used artificial neural network to describe relationship between input and output and then MCDM methods are employed to model the problem. Finally four different structures of genetic algorithm are employed to solve the model and find the adjusted levels of the control factors such that the responses are optimized simultaneously.

3. problem modeling

Generally, the mathematical model of the multi objective problems is defined as follows:

$$\text{Max } f_i(\boldsymbol{x}), \qquad i = 1, \dots, k$$

s.t.

$$\boldsymbol{x} \in X \qquad (1)$$

Here $\boldsymbol{x}$ is the vector of input variables, $x_1, \dots, x_p$, and f_i denotes the relationship between responses, y_i, and input variables. Note that in practice f_i is unknown and this function may differ for each response.

The simplest idea to model a multi response statistical optimization problem is to combine all of the responses into

a single one within the MCDM methods framework. For this, we used three commonly methods which discussed in following:

3.1. Global Criteria method (LP-metric method)

Using the Global Criterion method the optimal solution is found by minimizing a global criterion, *F* (Gomes et al., 2012). This global criterion is as follow:

$$F = \left(\sum_{i=1}^{k} w_i \left\{ \frac{T_i - y_i}{T_i} \right\}^r \right)^{1/r}, \quad i = 1, \dots, k \qquad (2)$$

The value of r relies on the choice of the decision maker (DM), where the usual practice has been $r = 1$, $r = 2$ and $r = \infty$. T_i is the target defined for the objective y_i and $w_i$representing importance of each objective.

3.2. Goal Programming method

Goal programming is a multi-objective decision making introduced by Charnes and Cooper (1961). Goal programming method requires the decision maker to set definite aspiration values or goals for each objective that he wishes to achieve. With this approach the mathematical model of the problem becomes (Chang, 2011):

$$\text{Min } F = \sum_{i=1}^{k} w_i (d_i^+ + d_i^-), \quad i = 1, \dots, k$$

s.t.

$$\begin{aligned} & y_i - d_i^+ + d_i^- = g_i, && i = 1, \dots, k \\ & d_i^+, d_i^- \geq 0, && i = 1, \dots, k \\ & \underline{x} \epsilon X \end{aligned} \qquad (3)$$

where d_i^+ and d_i^- are over- and under- achievement of the *i*th goal, respectively, g_i is the goal of ith objective. It is also requirement that $d_i^+ . d_i^- = 0$ must always hold true.

3.3. Desirability function

The desirability function approach is based on the idea that the quality of a product that has multiple quality characteristics is completely unacceptable if one of the characteristics lies outside the desired limits (Pasandideh and Akhavan Niaki, 2006).

A desirability function, $d_i(y_i)$, assigns numbers between 0 and 1 to the each response and increases as the desirability of the corresponding response increases. The overall desirability, *D*, is defined as follow:

$$D = (d_1(y_1) \times d_2(y_2) \times \dots \times d_k(y_k))^{1/k} \qquad (4)$$

where *k* denotes the number of the responses. If a response y_i is completely undesirable, i.e. $d_i(y_i) = 0$, then the overall desirability value is equal to 0. Therefore, maximization of the overall desirability is the purpose of this method.

Depending on whether a particular response y_i is to be maximized, minimized, or assigned a target value, we can use different desirability functions.

4. Artificial neural network

In order to generate required data in this research, we should approximate the statistical relationship between input and output variables. The most important attractive of the artificial neural networks is their ability to approximate any complex relationship between input and response variables.

ANNs can be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges (with weights) are connections between neuron outputs and neuron inputs. Based on the connection pattern (architecture), ANNs can be grouped into two categories: feed-forward and feed-back networks (Jain, Mao, & Mohiuddin, 1996).The most common family of feed forward networks is called Multi-Layer Perceptron (MLP).

The set of test data defines as input vectors and their corresponding target vector employed to train network. Training is updating network architecture and connection weights so that the variation of the actual output network and the expected output called network error, minimized. There are three main type of training: supervised, unsupervised, and hybrid. In supervised learning, the network weights are adjusted with a correct answer (target) for every input vectors to produce outputs as close as possible to the targets. Unsupervised learning is a type of learning in which the network is provided with only input vectors and does not require a correct answer associated with each input vector in the training data set. Hybrid learning combines supervised and unsupervised learning. Part of the weights is usually determined through supervised learning, while the others are obtained through unsupervised learning (Jain et al., 1996).

The commonly used supervised training networks are Perceptron, back propagation neural network (BPNN), and learning vector quantization (LVQ). Back propagation minimizes the mean square error (MSE) between the target data and the output of ANN.

The number of neurons per layer and the number of layers greatly influences the performance of the MLP. Too few of them prevent the learning process, and too many of them over fit the training data set (Ghaffari et al., 2006).

Here, a back propagation feed forward multilayer neural network is designed to estimate the relationships between input variables and their corresponding response. Next, given the test and the training data set, different back propagation networks are evaluated to select an appropriate

network. Finally, a network with the lowest MSE is chosen to be the optimal network.

5. The methodology

In this study, because of the stochastic nature of the problem, we apply a heuristic search algorithm to solve it. Genetic algorithm is a stochastic search technique based on the mechanism of natural selection and natural genetics developed initially by Holland (1975), and Goldberg (1989) described the usual form of it. GA process starts with generating an initial set of solutions called population. Each solution in the population is called a chromosome. The chromosomes expand through successive repeats, called generations. During each generation, some measures of fitness are used to appraise individuals. Fitness value used to select chromosome from the current generation to advance into the next generation. In the next generation, new chromosomes, called offspring, are created by crossover and mutation operator. After several generations, the algorithm converges to the best chromosome.
In this paper we present four GA methods, which are different in controlling the stochastic nature of the problem.

5.1. Initial condition

Required initial information to start the GA method is as following:

- Population size: The number of chromosomes or scenarios in each generation denoted by *N*.
- Number of iteration: The number of simulated iteration of each chromosome denoted by *n*.
- Crossover rate: The probability of performing crossover denoted by p_c.
- Mutation rate: The probability of performing mutation denoted by p_m.

5.2. Chromosome

Chromosome is a possible solution to the problem which is defined as a set of input variable values, $x_1, \dots, x_p$.

5.3. Initial population

In GA method, generating an initial population of *N* solution is the first step to start optimization problem. The population is randomly initialized. To do this, after preparation of chromosome *j*, $x_{1j}, \dots, x_{pj}$, we generate the chromosome in *n* replications and determine the fitness value of responses in each replication. We define the following parameters (Pasandideh and Akhavan Niaki, 2006):

x_{ij} is the input variable *i* in chromosome *j*, $i = 1, \dots, p$, $j = 1, \dots, N$,

y_{ijr} is the response variable *i* in chromosome *j* and in replication *r*, $i = 1, \dots, k$, $j = 1, \dots, N$, $r = 1, \dots, n$,

F_{jr} is the output value of chromosome *j* in replication *r*, $j = 1, \dots, N$, $r = 1, \dots, n$,

, and we obtain its value based on the function values described in Section 3,

$\overline{F}_j$ is the mean of the function values in chromosome *j*, $j = 1, \dots, N$ and we calculate it by $\overline{F}_j = \sum_{r=1}^{n} F_{jr}/n$.

5.4. Crossover and Mutation operators

Crossover is the main genetic algorithm operator. GA employs the crossover operator to generate new offspring. Using this operator a pair of chromosome is selected randomly with the probability p_c from the population. In this study we employ the two point crossover.
Mutation is the second operation in the GA methods for exploring new solutions. Mutation is an operator which produces random changes in various chromosomes. In this paper, we replace a gene with a randomly selected number within the boundaries of the parameter (Gen, & Cheng, 2000).

5.5. Objective function evaluation

After producing the new chromosomes by crossover and mutation operators and approximating the responses using suitable artificial neural network, we evaluate each chromosome by fitness functions described in section 3 and use the result to reproduce a new generation.

5.6. Chromosomes grouping

To control the stochastic nature of the problem we apply the statistical tests to compare the chromosomes and grouping them based on their fitness values. For this we apply four different pairwise multiple comparison statistical tests. The chromosomes are grouped such that there is no statistical difference between chromosomes in each the groups but there exist differences among different groups (Pasandideh and Akhavan Niaki, 2006). The statistical tests used in this study are described in the following:

5.6.1. *LSD procedure*

Using this statistical test the least significant difference (LSD) critical value is determined as (Montgomery, 2001):

$$LSD = t(\alpha/2, f)\sqrt{2MSE/n} \qquad (5)$$

where the probability of making a type I error is α and $t(\alpha/2, f)$ is the tabular value of Student's t for the selected significance level, $\alpha/2$. *N* is the population size, *MSE* is the mean squared error of $\overline{F}_j$'s and *f* is degrees of freedom associated with MSE which is equal to $N(n-1)$. First we calculate the LSD value and then based on the fitness function value, $\overline{F}_j$, we rank the chromosomes in

ascending order. For each pairs of $\overline{F}_j$'s,we consider the following equation:

$$\left|\overline{F}_i - \overline{F}_j\right| \leq LSD \qquad (6)$$

If the equation is true, we can said that the chromosomes *i* ,*j* and all of the chromosomes between this range are statistically equal and we can put them in the same group.

5.6.2. *Tukey procedure*

The second method we used in this paper is the Tukey's test (Montgomery, 2001). This method is exactly the same as the first method, except that the critical value is computed as:

$$T_\alpha = q_\alpha(N,f)\sqrt{MSE/n} \qquad (7)$$

Where $q_\alpha(N,f)$ is the upper percentile point of the studentized range statistic that obtained from a table of studentized ranges for the selected significance level, α and the degrees of freedom, *f*, which is equal to $N(n-1)$.

5.6.3. *Duncan procedure*

In the next method we employed the multiple range test developed by Duncan (1995). In this method we ranked the chromosomes based on the fitness function value from largest to smallest. This method begins with obtain the critical value of the least significant range as follow:

$$R_p = r_\alpha(p,df)\sqrt{MSE/n} \qquad (8)$$

for each $p = 2,\dots,N$. p is the number of chromosomes included in the range of a comparison. If the two chromosomes have consecutive ranking, then $p = 2$ and for the chromosomes with the highest and lowest ranking it is equal to N. The values of $r_\alpha(p,df)$ can be obtained from Duncan's table of significant ranges. The observed differences between $\overline{F}_j$'s, beginning with largest versus smallest, compared with the least significant range R_N. Next, the difference of the largest and the second smallest is computed and compared with the least significant range R_{N-1}. This process is continued until the differences of all possible $N(N-1)/2$ pairs of chromosomes have been considered. If an observed difference is less than the corresponding least significant range, then we conclude that the pair of chromosomes is significantly equal. Therefore these pairs of chromosomes and those included in the range of the comparison are in the same group.

5.6.4. *Newman-Keuls procedure*

This method is same as the Duncan procedure except that the critical value is computed as (Montgomery, 2001):

$$k_p = q_\alpha(N,f)\sqrt{MSE/n} \quad p = 2,\dots,N \qquad (9)$$

5.7. Chromosome selection

In order to generate chromosomes for the next generation we apply the most common selection technique which is roulette wheel selection. In this technique, a probability of selection assigned to each group based on its fitness value. The probability for each group is calculated by:

$$p_k = \frac{\sum_{\forall j \epsilon Group\ k} F_j}{\sum_{j=1}^{N} F_j} \qquad (10)$$

Then, a group is chosen randomly based on p_k values and its best chromosome is selected based on its fitness function value. We done this procedure *N* times, making a generation with *N* chromosomes.

After making the next generation, the crossover and mutation operators will operate on the new generation again and the selection phase will be repeated until the stopping criteria are met.

In order to demonstrate the application of the proposed methodology, in the next section a numerical illustration is given.

6. Illustrative example

The numerical illustration of this research is about a chemical process that was used by Pasandideh and Akhavan Niaki (2006). The problem used in their research is a numerical example used by Cheng et al. (2002). In this problem, there are three design variables and two responses. The design variables are reaction time (x1), temperature (x2) and percent catalyst (x3) and the responses are percent conversion (y1) and thermal activity (y2). The objective is to maximize y1 while keeping y2 between 55 and 60, with a target value of 57.5. The experimental observations are shown in Table 1.

Table (1) is here

After statistical normalization of the observation, the architecture of the neural network for each response is determined to generate the required input data based on the number of hidden layers, activation function and the number of neurons in each layer.

To select the number of neurons, we started with 2 neurons in each layer and we gradually increased the number of neurons until 10 neurons. The tangent hyperbolic function and sigmoid function were examined as the transfer (activation) function for the hidden layer and linear function for output layer. The network with minimum mean square

error (MSE) and maximum correlation (R) between the actual and the expected output have been selected.

In order to code and train the network, special routine of the feed forward back-propagation neural network in MATLAB 2009 is employed for each response. Moreover, the training algorithm is Levenberg-Marquardt (Lee, Ho, Lin, & Kang, 2011) and the ratio of the test data to the whole data is 20%.

Following tables show the suitable number of neurons according to respective MSE and R values, for two types of transfer function. We can see the network with two hidden layers and 7-10-1 neuron respectively in both hidden layers and output layer and tangent hyperbolic function for response1and the network with 4-6-1 neuron and sigmoid function for response2 are best architectures.

Table 2. Optimal parameters of the artificial neural network according to response1

Activation function	Number of neurons	R	MSE
Log-sigmoid	3-1	0.8841	0.2118
Tan-sigmoid	2-1	0.9142	0.1695
Log-sigmoid	2-9-1	0.9691	0.0714
Tan-sigmoid	7-10-1	0.9733	0.0546

Table 3. Optimal parameters of the artificial neural network according to response2

Activation function	Number of neurons	R	MSE
Log-sigmoid	6-1	0.8732	0.2892
Tan-sigmoid	6-1	0.8902	0.2325
Log-sigmoid	4-6-1	0.9959	0.0079
Tan-sigmoid	3-4-1	0.9933	0.0138

Then, the MSE values obtained using both the selected artificial neural network and the regression approach compared and the result (Table 4) indicate better performance of the neural network method.

Table 4. MSE comparison of the proposed neural networks and regression

Response Variable	MSE	
	ANN	Regression
y_1	0.055	0.061
y_2	0.008	0.069

6.1. *Parameter tuning*

Tuning the parameters of an evolutionary algorithm is essential for good algorithm performance. In this Section, the parameters of the proposed GA consisting of the crossover rate (P_C), the mutation rate (P_M), and the population size (N) are tuned using response surface methodology (RSM).

First in terms of four test procedures and three modeling method used in this research, using MATLAB computer software, the algorithms run 30 times, each time changing its parameters in their corresponding ranges and obtained the responses values for each algorithm. The crossover and mutation operations rates vary in the range of 0.6–0.9 and 0–0.05, respectively. Furthermore, different integer population sizes between 10 and 100 are considered in this experiment.

Then, a first or second order polynomial mathematical model is fitted using regression based on various combinations of algorithm parameters and corresponding fitness values. This model describes the relationship between fitness and algorithm parameters. Then we selected the optimal level of these parameter such that response variable optimize. The optimal values for algorithm parameters are obtained as:

Table 5. Optimal values of algorithms parameters

MODM methods	Tests	$npop^*$	$p_c{}^*$	$p_m{}^*$
LP metric method	Duncan	20	0.7	0
	LSD	30	0.8	0.03
	Newman	15	0.8	0.01
	Tukey	50	0.9	0.05
Desirability method	Duncan	80	0.8	0.03
	LSD	70	0.8	0.03
	Newman	10	0.6	0.05
	Tukey	10	0.6	0.04
Goal Programming method	Duncan	50	0.9	0.02
	LSD	10	0.8	0.04
	Newman	70	0.8	0
	Tukey	80	0.6	0.02

After determination of optimal algorithm parameters the levels of the control factors and responses obtained as follow:

Table (6) is here

6.2. *Statistical comparison*

In some problems it is important to study the effects of two or more factors on the responses. In general, factorial designs are most efficient by which in each complete trial or replication of the experiment all possible combination of the levels of the factors are investigated (Montgomery, 2001).
In this research two factors are effective on the responses. The four tests procedure performed in the structure of the algorithms are considered as the first factor and the second factor being the three MODM methods used in this research to model the problem. We were interested in testing hypotheses about the equality of the performance means of the algorithms for the first factor or in the other words, equality of the treatment effects on the responses and also the equality of second treatment effects. Also we were interested in determining whether two treatments interact.

We tested the hypothesizes, at an overall 95% significant level, based on the experimental results obtained from employing each algorithm several times in terms of their desirability. For this we employed the SPSS statistical software and the results are summarized in a tow way analysis of variance table, as shown in table 7.

Table (7): Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3.827[a]	11	.348	4.294	.000
Intercept	103.780	1	103.780	1280.824	.000
Tests	.071	3	.024	.293	.830
MODM methods	3.413	2	1.707	21.063	.000
VAR00002 * VAR00003	.343	6	.057	.705	.646
Error	28.197	348	.081		
Total	135.804	360			
Corrected Total	32.024	359			

a. R Squared = .120 (Adjusted R Squared = .092)

We can see from the analysis of variance table that there is not a significant interaction between tests type and MCDM methods used in the algorithms. Furthermore, the main effect of tests type is not significant. Also the analysis of variance indicates that the performances mean based on the MODM factor differ. So it was interested to make comparisons between the individual means for this factor to discover the specific differences. The multiple comparison methods are useful in this case. For this, the Tukey test was employed in SPSS software.

Table (8) is here

Note that the performance means and p-value are displayed in homogeneous subsets tables.

Table (9): Homogenous subset
Tukey HSD

MODM methods	N	Subset		
		1	2	3
GP	120	.4208		
LP metric	120		.5309	
Desirability function	120			.6591
Sig.		1.000	1.000	1.000

The results that are concluded from this table show that the desirability function method significantly has the better desirability than the others.

7. Conclusions

In processes with stochastic nature and data sets showing nonlinear and complex relationships between control and response variables and complicated optimization models, to overcome the limitation and weakness of the methods which are used before in this cases, in this paper a new approach based on a neural network, three modeling techniques, and four different genetic algorithms was proposed to solve multiple response statistical optimization problems. The neural network approach generated the required input data and the MODM techniques modeled the problem and the four different structures of genetic algorithm optimized the model to find the levels of the control factors. At the end, the performance of the proposed methodology was evaluated using a numerical example. The MSE values obtained using both the selected artificial neural network and the regression approach compared and the result indicate better performance of the proposed methodology. Also the tests procedure employed in the structure of the algorithms were statistically equal and the desirability function method have better performance than the other MODM methods used in this study.

References

[1] Amiri, M. Karimi, N. Jamshidi, S. F. (2008). A methodology for optimizing statistical multi-response problems using genetic

local search algorithm through fuzzy goal programming. *Journal of Applied Sciences*, *8*, 3199-3206.

[2] Baseler, F. F. & Sepulveda, J. A. (2000). *Multi response simulation optimization using stochastic genetic search within a goal programming framework*. Proceedings of the 2000 Winter Simulation Conference.

[3] Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice*. Oxford University Press.

[4] Bashiri, M. & Hosseininezhad, S. J. (2009). Optimization of Multiple Response Process by Neural Networks Based on Desirability Concept. *International Journal of Industrial Engineering and Production Management*, *2*(4), 63-53.

[5] Bourquin, J., Schmidli, H., Hoogevest, P. & Leuenberger, H., (1998). Advantages of artificial neural networks (ANNs) as alternative modeling technique for data sets showing non-linear relationship using data from a galenical study on a solid dosage form. *Eur. J. Pharm. Sci*, *7*, 5-16.

[6] Chang, C. T. (2011). Multi-choice goal programming with utility function. *European Journal of Operational Research*, *215*, 439–445.

[7] Charnes, A. & Cooper, W.W. (1961). *Management Models and Industrial Applications of Linear Programming*, John Wiley and Sons, New York.

[8] Cheng, C. B., Cheng, C. J. & Lee. E. S. (2002). Neuro Fuzzy and Genetic Algorithm in Multiple Response Optimization. *Computers and Mathematics with Applications*, *44*, 1503-1514.

[9] Derringer, G. & Suich, R. (1980). Simultaneous optimization of several response variables. *Journal of Quality Technology*, *12*, 214-219.

[10] Duncan, D. B. (1995). Multiple range and multiple F-Test. *Biometrics, 11*, 1-42.

[11] Fourman, M. P. (1985). Comparison of symbolic layout using genetic algorithms. *Proceedings of the First International Conference on Genetic Algorithm and Their Applications*, 141-153.

[12] Ghaffari, A., Abdollahi, H., Khoshayand, M. R., Soltani Bozchalooi, I., Dadgar, A. & Rafiee-Tehrani, M., (2006). Performance comparison of neural network training algorithms in modeling of bimodal drug delivery. *International Journal of Pharmaceutics*, *327*, 126–138.

[13] Goldberg, D. (1989) *Genetic Algorithms in Search, optimization and Machine Learning, Addison-Wesley*, Reading, MA, USA.

[14] Gomes, J. H., Salgado, A. R., Paiva, A. P., Ferreira, J. R., Costa, S. C. & Balestrassi, P. P. (2012). Global Criterion Method Based on Principal Components to the Optimization of Manufacturing Processes with Multiple Responses. *Journal of Mechanical Engineering*, *5*(85), 345-353.

[15] Harrington, J. E. (1965). The desirability function. *Industrial Quality Control*, *21*, 494-498.

[16] Holland, J. H. (1975) *Adaptation in Natural and Artificial Systems,* University of Michigan Press, Michigan.

[17] Hsieh, K. L. (2010). Employing Artificial Neural Network into Achieving Parameter Optimization of Multi Response Problem with Different Importance Degree Consideration. *Information Technology Journal*, *9*(5), 918-926.

[18] Jain, A.K., Mao, J., & Mohiuddin, K.M. (1996). Artificial Neural Network: a tutorial computer. *IEEE*, 31-44.

[19] Kim, D. & Rhee, S. (2004). Optimization of a gas metal arc welding process using the desirability function and the genetic algorithm. *Proceedings of the Institution of Mechanical Engineers*, *Part B: Journal of Engineering Manufacture*, *218(1)*, 35-41.

[20] Kim, K. J. & Lin, D. K. J. (2001). Simultaneous Optimization of mechanical properties of steel by polynomial regression functions. *Applied Statistics*, *49*, 311-325.

[21] Konaka, A., Coitb, D. W. & Smithc, A. E. (2006). Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering and System Safety*, *91*, 992–1007.

[22] Kutner, M. H., Nachtsheim, C. J., Neter, J. & Li, W. (2005). *Applied Linear Statistical Models*, 5^{th} Ed., McGraw-Hill, Irwin.

[23] Lee, T. E., Ho, C. Y., Lin, C. H. & Kang, M.S. (2011) 'Design of artificial neural networks for distribution feeder loss analysis', *Expert SystAppl*, Vol.38, pp.14838-14845.

[24] Montgomery, D.C. (2001) *Design and analysis of experiments*, 5^{th} ed., John Wiley & Sons, New York, NY, USA.

[25] Myers, R. H. & Montgomery, D. C. (1995). *Response Surface Methodology: Process and Product Improvement with Designed Experiments*, John Wiley & Sons, New York, NY.

[26] Pasandideh, S. H. R. & Akhavan Niaki, S. T. (2006). Optimizing Multi-Response Statistical

[27] Problems Using a Genetic Algorithm. *Scientia Iranica*, *13*(1), 50-59.

[28] Šibalija, T. V. & Majstorović, V. D. (2010). Novel Approach to Multi Response Optimization for Correlated Responses. *FME Transactions*, *38*, pp.39-48.

[29] Yang, J., Rivard, H. & Zmeureanu, R. (2005, August). *Building Energy Predication with Adaptive Artificial Neural Networks*. Ninth International IBPSA Conference Montréal.

Table 1. Sampling data

observation	x_1	x_2	x_3	x_4	y_1	y_2
1	2	18	330	25	52	50
2	7	18	360	25	50	45
3	7	23	330	25	120	117
4	2	23	360	25	170	159
5	7	18	330	30	120	110
6	2	18	360	30	94	90
7	2	23	330	30	186	178
8	7	23	360	30	180	176
9	4.5	20.5	345	27.5	166	160
10	4.5	20.5	345	27.5	165	163
11	4.5	20.5	345	27.5	167	165
12	4.5	20.5	345	27.5	161	166
13	4.5	20.5	345	31.04	172	169
14	4.5	20.5	345	23.96	160	157
15	4.5	24.04	345	27.5	173	174
16	4.5	16.96	345	27.5	155	150
17	4.5	20.5	366.2	27.5	171	167
18	4.5	20.5	323.8	27.5	157	159
19	8.035	20.5	345	27.5	169	161
20	0.965	20.5	345	27.5	162	159

Table (6): Levels of the control factors and responses

MODM methods	Tests	x_1	x_2	x_3	y_1	y_2
LP metric method	Duncan	-0.8662	0.2833	-0.2581	89.0699	57.6309
	LSD	-0.6968	0.7058	-0.2016	90.4415	57.9670
	Newman	-0.6371	0.6061	-0.1178	88.5262	58.2596
	Tukey	-0.8507	0.2660	-0.3044	85.9853	57.9781
Desirability method	Duncan	-0.6952	0.0130	-0.0883	81.5769	58.289
	LSD	-0.6915	0.2207	-0.1224	85.0801	58.4304
	Newman	-0.8717	0.0588	-0.0099	82.2812	58.7057
	Tukey	-0.7393	0.1403	-0.0700	83.1862	58.331
Goal Programming method	Duncan	-0.6389	0.6169	-0.2290	94.2729	58.4349
	LSD	-0.5861	1.0385	-0.0851	94.8479	58.4471
	Newman	-0.4424	0.7037	-0.1025	93.5833	58.7305
	Tukey	-0.4930	0.6925	-0.1926	90.2800	58.6231

Table (8): Multiple Comparisons
Tukey HSD

(I) MODM methods	(J) MODM methods	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Desirability function	GP	.2383*	.0367	.000	.1518	.3248
	LP metric	.1282*	.0367	.002	.0417	.2147
GP	Desirability function	-.2383*	.0367	.000	-.3248	-.1518
	LP metric	-.1101*	.0367	.008	-.1966	-.0236
LP metric	Desirability function	-.1282*	.0367	.002	-.2147	-.0417
	GP	.1101*	.0367	.008	.0236	.1966

*. The mean difference is significant at the 0.05 level.