Neural Network Design by Taguchi Method

Akash Saxena[1],Deepak Saini[1], Shalini Shekhawat[2]
[1]Department of Electrical Engineering, [2]Department of Mathematics
Swami Keshvanand Institute of Technology Management and Gramothan,Jaipur.
Email- akash@skit.ac.in
Received 02 September 2016, received in revised form 07 September 2016, accepted 09 September 2016

Abstract: **The development of new learning algorithms for design of the neural networks is a potential research area. The design of neural network is based on selection of the optimal parameters to achieve higher learning speed and accuracy during recall. These design parameters are weights, no. of hidden layers and no. of neurons. With this motivation, this paper presents an application of Taguchi method to meet high accuracy requirements. Three design factors are considered as control factors which are namely input representation scheme, no. of hidden layers and no. of input nodes to design the neural network. This design has been tested over three benchmark optimization functions. It is observed that the proposed network shows better accuracy.**

Keywords: **Artificial Neural Network (ANN),Design of Experiment (DOE), Feed Forward Neural Networks (FFNN),Signal to Noise Ratios (SNRs),Taguchi Method.**

1. INTRODUCTION

The Feed Forward Neural Networks (FFNNs) are suitable for many real problems like pattern recognition, optimization, regression, forecasting, signal filtering and many more. Despite the capability of neural networks in data matching, the black box characteristic of the network is an Achilles heel. The main problem associated with the network design is the time taking process associated with the training of the network. Along with this, accuracy of the network is highly influenced by its structure. Broadly, the neural design can be classified as a design problem of micro and macro structural level. At micro structural level, the modification in the transfer function, input representation scheme and modification in learning rules are considered. On the other hand, at macro structural level the design parameters are no. of hidden layers and neuron and their connections schemes. The design of neural network can be classified in two categories. In first, researchers optimize the size of the network by choosing hidden layers and neurons as variables. In second category, the development of fast and accurate learning algorithms is explored. In reference [1] the Multi Layer Perceptron (MLP) has been trained with the help of biography based optimization method. In reference [2] a newly introduced metaheuristic algorithm Grey wolf Optimizer (GWO) has been employed to train the MLP. It was observed by the author that the results obtained from GWO were quite competitive as compared with conventional evolutionary trainers. An evolutionary programming based approach was applied John R McDonnell et.al. [3] for determination of hidden layers in neural networks. Application of metaheuristic techniques is explored in this area to optimize size of network and to enhance the efficiency of the network. On the basis of this discussion, following objectives are framed for this work.

a. To develop FFNN models based on different input representation schemes, hidden layers and no of nodes.
b. To develop L_8 orthogonal array design to obtain optimal architecture of neural network and to obtain and analyzed signal to noise ratios.
c. To test the architecture for prediction of three benchmark functions namely Exponential, Beale's and Matyas function.

Remaining part of the manuscript is organized as follows: in section 2 details of experimental design is explained, in section 3 construction of orthogonal arrays is explained; in section 4 and 5 simulation results and conclusions are presented.

2. EXPERIMENTAL DESIGN

The traditional nondeterministic optimization algorithms have been focusing on continuous variable problems. In different approaches involvement of multiple variables not only makes the process time consuming but also less efficient. For the different type of design problem, the Taguchi method has been widely used to define the optimal level of each parameter to condense deviations in responses as well as increase the system performance [4-5].

Design of experiment is an important tool to analyze the influence of control factors on performance output. The most important stage of design lays in the choice of control factors. Therefore many designs reported in literature extensively included large no of variables so that the non-significant variables can be identified at earlier stage [6]. Amar et.al. [7] presented solid particle erosion wear performance of a multi component hybrid composite consisting of polyester. In this design experiment it is concluded by the authors that glass–polyester composite without any filler suffers greater erosion loss than the hybrid composite with alumina filling. Significant control factors and their interactions that influence the wear rate are identified. Several design of experiments are performed on the taguchi implementations few of them are namely dry sliding wear characterization of epoxy TiO_2 particulate filled functionally graded composite materials. Mechanical and erosion behavior of hybrid composites were discussed in approach [8].

Further in electromagnetic field studies and antenna design, taguchi method is augmented. The usage of Taguchi method for the electromagnetic problems is explained in various approaches [4-8]. Rectangular micro strip patch antenna design, Antenna optimization and many other approaches in literature verified the significance of the taguchi implementation in design stage [9]. Similarly in Electrical networks, Power System Stabilizer design problem is discussed by S.R. Karnik et. al[10]. In this approach estimation of PSS parameters are handled by non-iterative taguchi method. The process is performed with simultaneously reduction in the influence of variations in system operating conditions [10].

3. CONSTRUCTION OF ORTHOGONAL ARRAYS

In the Taguchi method, the term 'signal' represents the desirable value (mean) for the output characteristic and the term 'noise' represents the undesirable value (standard deviation) for the out-put characteristic. The SNR is the ratio of the mean to the standard deviation. Taguchi method employs the SNR to measure the quality characteristic deviating from the desired value [6]. The reason of quality vacillation is uncontrollable factors, named as noise factors, which can be classified into external factors namely temperatures and human errors, manufacturing limitations and product corrosion. Depending on design objective, there are three quality characteristics namely "the-nominal-the-better", "the-smaller-the-better", and "the-larger-the-better" [11]. Their mathematical expressions are formulated as follows:

Case 1: "Smaller the better": aiming to minimize the performance

$$SNR = -10\log_{10}\left(\frac{\sum_{i=1}^{N} y_i^2}{N}\right)$$

Where the y denotes the performance indicator, subscript i experiment number, N is number of replicates of experiment'i'.[20]

Case 2. "Larger the better" aiming to maximize the performance

$$SNR = -10\log_{10}\left(\frac{\sum_{i=1}^{N}(1/y_i^2)}{N}\right)$$

Case 3: "Nominal is best" aiming to target the predetermined nominal value

$$\begin{cases} SNR = 10\log_{10}\left[\left(\frac{\bar{y}}{s}\right)^2\right] \\ \bar{y} = \frac{y_1 + y_2 + y_3 + \ldots + y_N}{N} \\ S = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^2}{N-1} \end{cases}$$

To develop different experiments we consider three factors and two level designs. Out of three available designs, we chose L_8 orthogonal array due to seven columns to match the needs of experiments. Following are the major steps for creating the taguchi design:

a.Identify the design factors and decide the objective to achieve

b.Define the matrix experiments and data analysis procedure.

c.Conduct the experiments and compute the statistics.

4. RESULTS

In this work we consider three factors namely input representation scheme, hidden layers and no. of neurons. Two levels of input representation scheme are random data and normalized data pattern. Similarly no. of hidden layers 5 and 10 and no. of input nodes are 10 and 12. The objective of the neural network is defined as prediction of the shape of three bench mark functions (Beale's, Matyas and Exponential, function) [12] the detail properties of the functions are given below.

$$f(x,y) = (1.5 - x + xy)^2 + (2.25 - x + xy^2)^2 + (2.625 - x + xy^3)^2 \quad (1)$$

Range $-4.5 \le x, y \le 4.5$

$$f(x,y) = 0.26(x^2 + y^2) - 0.48xy \quad (2)$$

Range $-10 \le x, y \le 10$

$$f(x,y) = x.e^{x^2 - y^2} \quad (3)$$

Design of experiment is shown in table I. the objective is to achieve highest adjusted R square values. R values for testing, validation and training with overall values of R is incorporated in design to take care of the efficiency of supervised learning engine at the stage of testing, training and validation. Figures 1-2 shows the values of SNRs and the interaction of the factors associated with the designs. The simulation studies are performed over Intel ®core ™, i7, 2.9 GHz 4.00 GB RAM processor unit. Design of Experiment based analysis is performed by using Minitab ® 17.1.0.

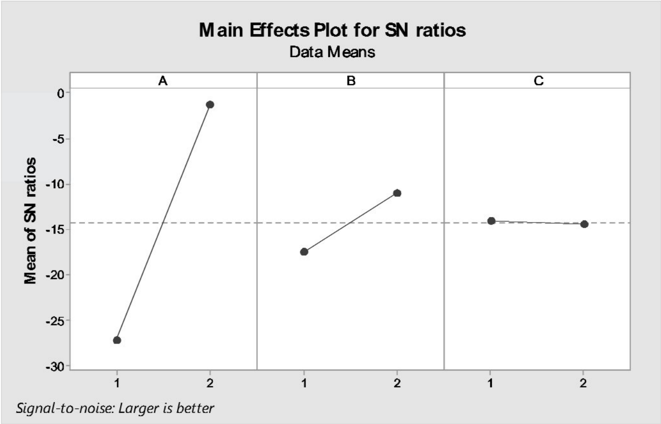


Figure 1. SNR for Matyas function

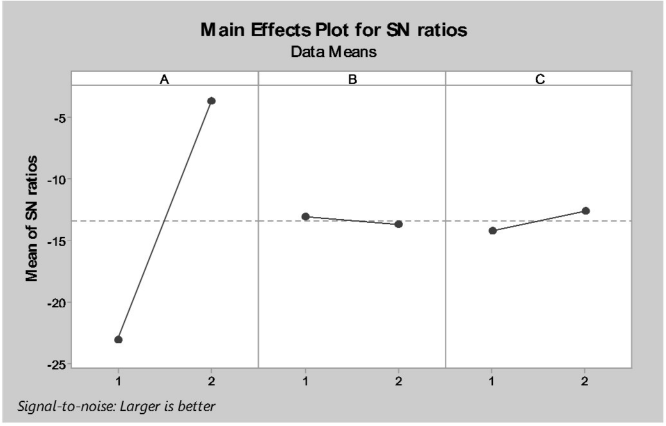


Figure 2. SNR for exponential function

TABLE I Design of Experiment L_8 orthogonal Arrays

Benchmark functions	A	B	C	$R_{training}$	$R_{validation}$	$R_{testing}$	R_{all}	SNR
Matyas Function	Random	5_h	10_n	0.715	0.525	0.024	0.475	-26.4002
	Random	5_h	12_n	0.779	0.006	0.09	0.588	-38.4363
	Random	10_h	10_n	0.034	0.537	0.055	0.322	-24.8027
	Random	10_h	12_n	0.081	0.119	0.192	0.103	-19.3503
	Normalized	5_h	10_n	0.558	0.666	0.477	0.562	-5.12796
	Normalized	5_h	12_n	0.999	0.998	0.998	0.999	-0.01304
	Normalized	10_h	10_n	0.999	0.994	0.995	0.995	-0.03704
	Normalized	10_h	12_n	0.998	0.997	0.999	0.998	-0.0174
Exponential Function	Random	5_h	10_n	0.257	0.115	0.121	0.192	-16.6788
	Random	5_h	12_n	0.364	0.006	0.164	0.193	-38.4276
	Random	10_h	10_n	0.815	0.007	0.039	0.492	-37.2163
	Random	10_h	12_n	0.135	0.409	0.108	0.156	-16.7134
	Normalized	5_h	10_n	0.995	0.966	0.994	0.994	-0.11351
	Normalized	5_h	12_n	0.986	0.99	0.984	0.98	-0.13145
	Normalized	10_h	10_n	0.996	0.997	0.945	0.983	-0.17208
	Normalized	10_h	12_n	0.9943	0.9882	0.984	0.99	-0.09516
Beales function	Random	5_h	10_n	0.59	0.058	0.14	0.372	-19.5229
	Random	5_h	12_n	0.336	0.102	0.037	0.241	-23.2871
	Random	10_h	10_n	0.069	0.025	0.068	0.052	-27.7745
	Random	10_h	12_n	0.149	0.106	0.052	0.104	-21.7342
	Normalized	5_h	10_n	0.328	0.467	0.292	0.348	-9.27704
	Normalized	5_h	12_n	0.988	0.988	0.977	0.985	-0.13596
	Normalized	10_h	10_n	0.999	0.999	0.999	0.999	-0.00869
	Normalized	10_h	12_n	0.459	0.716	0.652	0.522	-5.02408

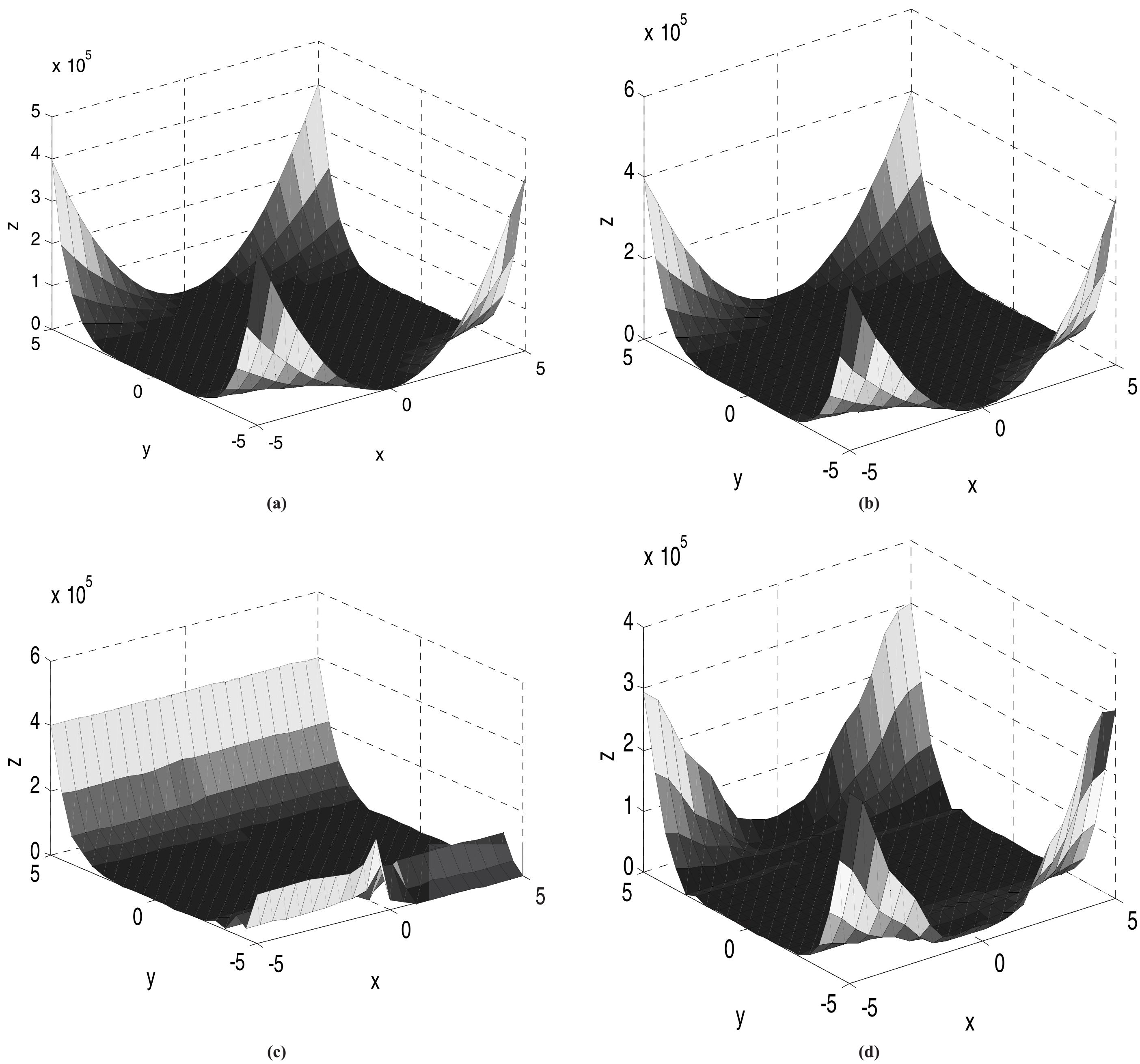


Figure 3 (a) Original Beale's function (b) Beale's function by design 7 (Normalized -10h10n) (c) Beale's function by design 5 (Normalized -5h10n) (d) Beale's function by design 6 (Normalized -5h12n)

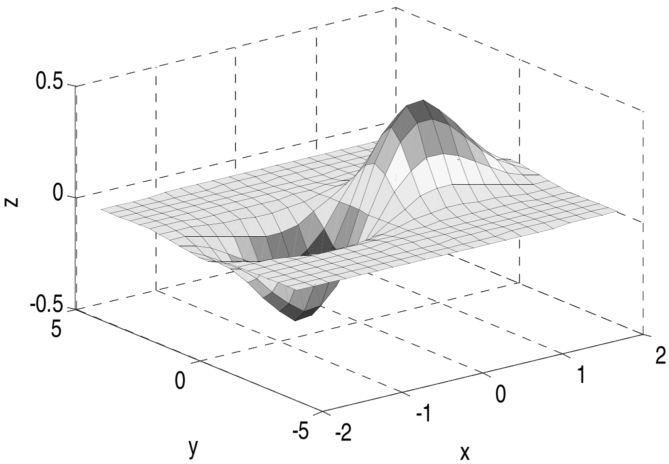
Different 8 experiment runs are exhibited in table 1 for prediction of the shape of Matyas function, Beale's function and exponential functions. Different designs obtained from the neural topologies are exhibited in figure3 4 and 5. Following are the observations from the design of the experiments.

a. It is observed for figure 1 and 2 that input representation scheme (factor A) of level2 i.e. normalized input representation scheme has a significant impact on the SNR. Hence, the input representation scheme should contain normalized data.

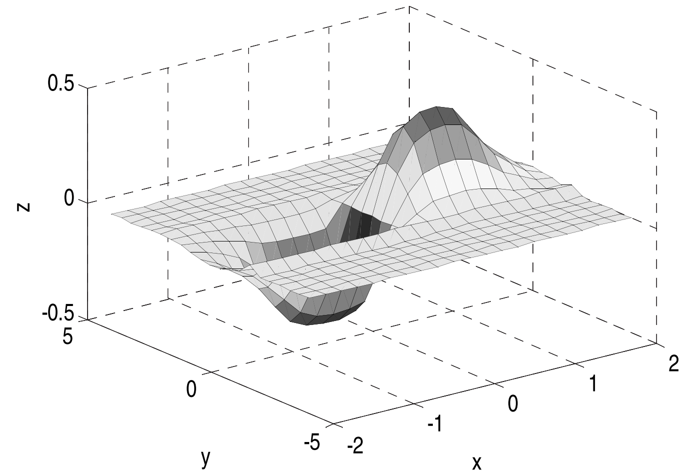
b. It is observed that for factor B (hidden layers) level 1 is suitable for obtaining the exponential function. On the other hand, level 2 i.e. 10 hidden layers are suitable for prediction of function Matyas and Beale's.

c. Similar observations are carried out for no. of input nodes it is observed that optimal values for obtaining high values of R for exponential function is 12, for Matyas function is 10 and for Beale's function is 12.

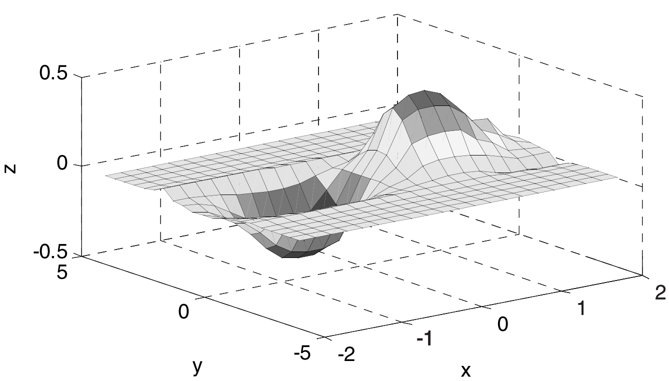
d. It is interesting to judge that with the change in the prediction task the optimal values of these designs parameters are different. For prediction of the Beale's function normalized input representation scheme Factor A level 2, with 5 hidden layers Factor B level1 and 12 Factor C level 2 input nodes are suitable. Similarly for the prediction of Matyas and exponential functions. The shapes obtained from input representation schemes are shown. It is observed that neural network designed from taguchi is able to judge the relationships of variables effectively. Figure 6 shows regression results for exponential function by design 7.



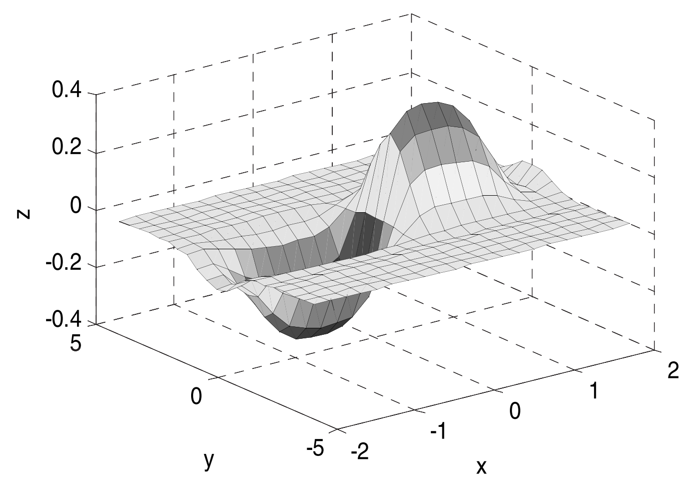
(a)



(b)

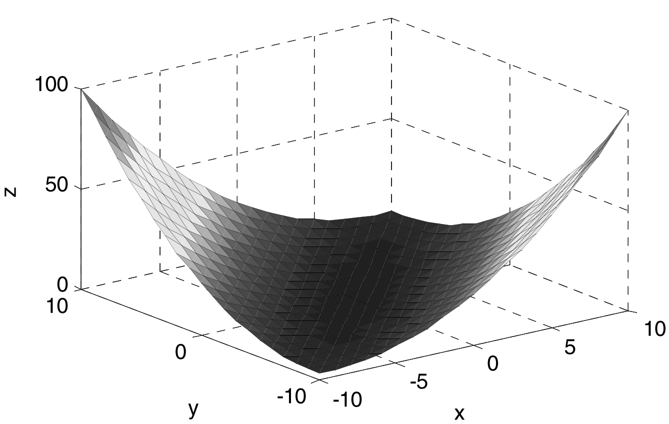


(c)

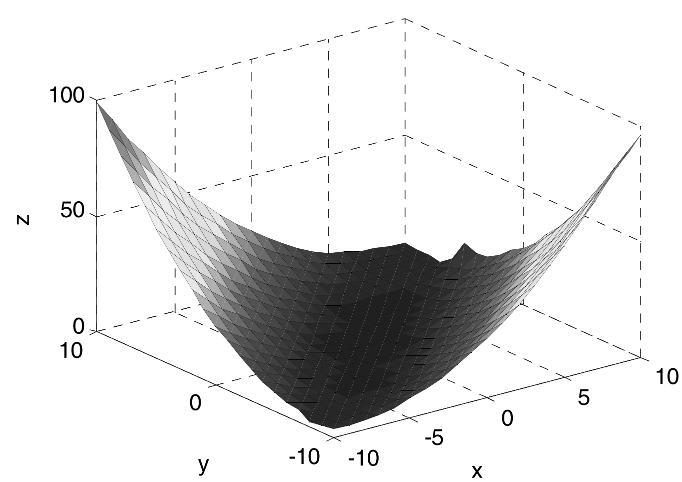


(d)

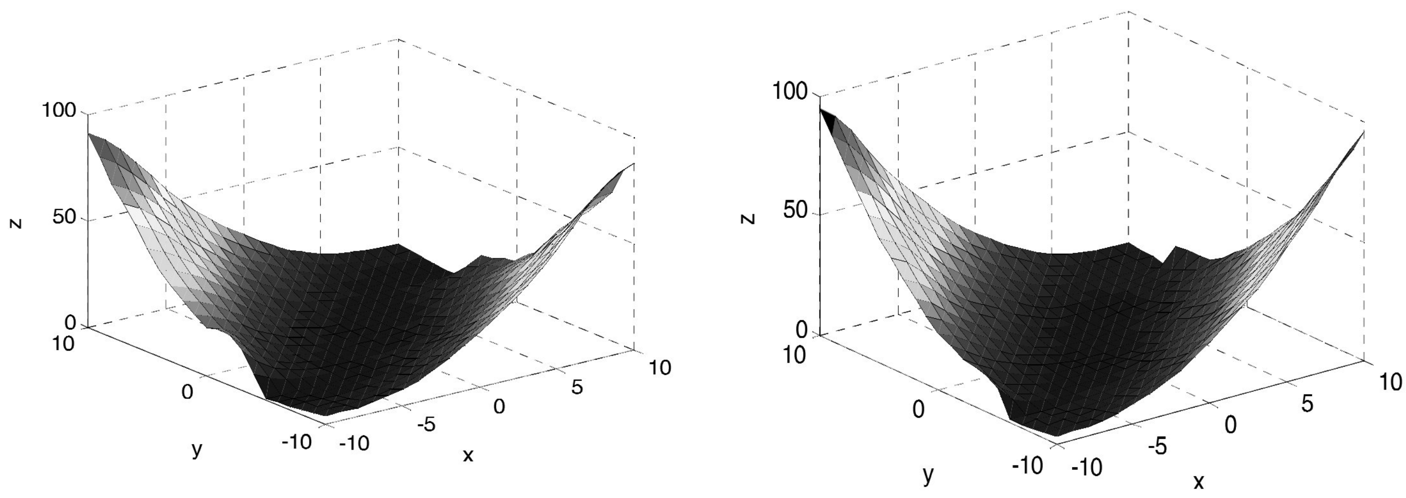
Figure 4 (a) original Exponential function (b) Exponential function by design 7 (Normalized -10h10n) (c) Exponential function by design 8 (Normalized -10h12n) (d) Exponential function by design 6 (Normalized -5h12n)



(a)



(b)



(c) **(d)**

Figure 5 (a) original Matyas function (b) Matyas function by design 6 (Normalized -5h12n) (c) Matyas function by design 7 (Normalized -10h10n) (d) Matyas function by design 8 (Normalized -10h12n)

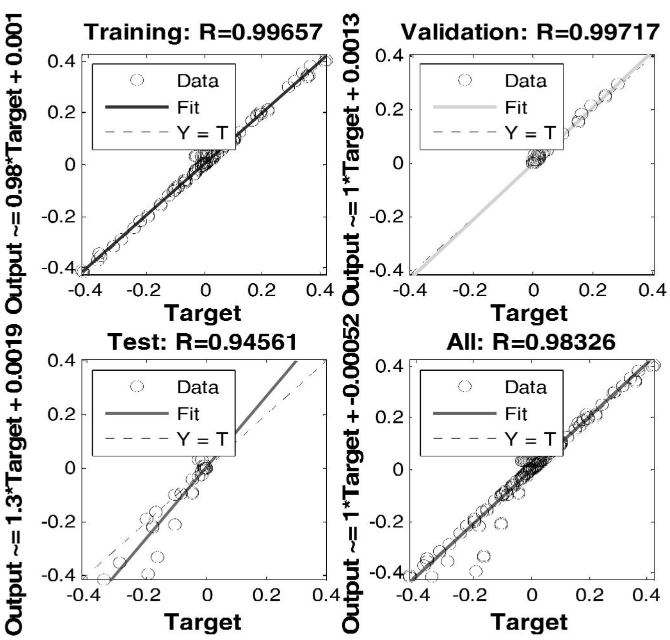


Figure 6 Regression results for prediction of Exponential function by design 7

5. CONCLUSION

Taguchi design of experiment method is useful for product development process. This method is applied to know the effect of different control factors and noise factors on the process. This manuscript has presented an application of this method to determine the optimal architecture of ANN. L_8 orthogonal array design has been developed and tested for the prediction of the shapes of three different bench mark functions. It is observed that normalized input representation scheme gives better results and other parameters like hidden layers and nodes are problem specific. Application of taguchi method for designing neural network for analyzes the behavior of special functions lies in the future scope.

REFERENCES

[1] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrew Lewis, “Let a biogeography-based optimizer train your Multi-Layer Perceptron”, Information Sciences, Volume 269, Pages 188-209, June 2014.

[2] Seyedali Mirjalili. “How effective is the Grey Wolf optimizer in training multi-layer perceptrons”Applied Intelligence Volume 43, Pages 150-161,July 2015.

[3] John R Mcdonnell and Don Waggen” Determining Neural Network Hidden Layer Size using Evolutionary Programming”proceddings of world congress on neural network, Volume 3, Pages 564-567, November 1983.

[4] Badkar, D. S., Pandey, K. S., & Buvanashekaran, G. Parameter optimization of laser transformation hardening by using Taguchi method and utility concept. The International Journal of Advanced Manufacturing Technology, 52, 1067–1077, 2011.

[5] Sibalija, T. V., & Majstorovic, V. D. . An integrated approach to optimise parameter design of multi-response processes based on Taguchi method and artificial intelligence. Journal of Intelligent Manufacturing, 23, 1511–1528, 2012.

[6] Sun, Guangyong, et al. "Discrete robust optimization algorithm based on Taguchi method for structural crashworthiness design." *Expert Systems with Applications* 42.9 ,4482-4492,2015.

[7] Patnaik, Amar, et al. "Tribo-performance of polyester hybrid composites: damage assessment and parameter optimization using Taguchi design."*Materials & Design* 30.1,57-67, 2009.

[8] Mahapatra, S. S., and Amar Patnaik. "Study on mechanical and erosion wear behavior of hybrid composites using Taguchi experimental design." *Materials & Design* 30.8 2791-2801, 2009.

[9] Taguchi, Genichi. *Introduction to quality engineering: designing quality into products and processes*. 1986.

[10] Karnik, S. R., A. B. Raju, and M. S. Raviprakasha. "Genetic algorithm based robust power system stabilizer design using Taguchi principle." Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on. IEEE, 2008.

[11] Nalbant, M., Gokkaya, H., & Sur, G. Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning. Materials & Design, 28, 1379–1385, 2007.

[12] http://www.sfu.ca/~ssurjano/optimization.html.

◆ ◆ ◆