MODELING AND OPTIMIZATION OF PROCESS PARAMETERS USING NEURAL NETWORKS AND SIMULATED ANNEALING ALGORITHM FOR ELECTRICAL DISCHARGE MACHINING OF AISI2312 HOT WORKED STEEL

M. A. Moghaddam[1*], F. Kolahan[2]

[1'2]Department of Mechanical Engineering,
Ferdowsi University of Mashhad, P.O. Box 91775-111, Mashhad, Iran

ABSTRACT

The present study addresses the multi-criteria modeling and optimization of Electrical Discharge Machining (EDM) for AISI2312 hot worked steel parts via optimized back propagation neural networks (OBPNN) and Simulated Annealing (SA) algorithm. The process response characteristics considered are surface roughness, tool wear rate and material removal rate. The process input parameters include voltage, peak current, pulse off time, pulse on time and duty factor. The weighted normalized grades, obtained from Taguchi design of experiments, are used to develop the arteficial neural network (ANN) model. In order to enhance the prediction capability of the proosed model, its architecture is tuned by simulated annealing algorithm. Next, the developed model is embaded into the SA algorithm to determine the best set of process parameters values for a desired set of outputs. Validation of the results has been carried out through a test run under the optimal machining conditions. Experimental results indicate that the proposed modeling and optimization procedures are quite efficient in modeling and optimization of EDM process parameters.

***KEYWORDS**: Electrical discharge machining (EDM); AISI2312 steel; Optimized back propagation neural network (OBPNN); Simulated annealing algorithm*

1.0 INTRODUCTION

In recent years various machining processes have been developerd or modified to cope with high alloy materials. Among these alloyes, hotworked AISI2312 steel is one of the most difficult-to-cut materials. The need for producing complex shapes along with reasonable speed and surface finish are very difficult to achieve by traditional machining processes. Electric Discharge Machining (EDM) is a non-conventional

* Corresponding author email: masoud_azadi88@yahoo.com

machining process suitable for shaping this alloy. In EDM process, the material erosion mechanism primarily makes use of electrical energy and turns it into thermal energy through a series of discrete electrical discharges occurring between the electrode and workpiece immersed in a dielectric fluid (Figure 1). This unique feature of using thermal energy to machine electrically conductive parts has been its distinctive advantage in the manufacture of moulds, dies, aerospace and surgical components (Pushpendra et al. (2012)).

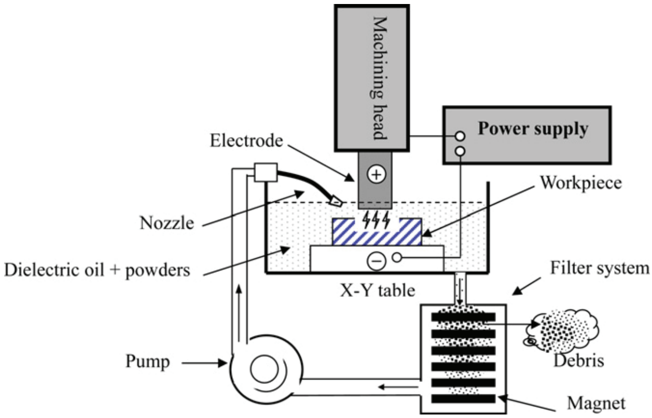

Figure 1. Schematic illustration of Electrical Discharge Machining

The thermal energy generates a channel of plasma between the work piece electrode (cathode) and the tool electrode (anode) at a temperature in the range of 8000 to 12,000°C. initialising a substantial amount of heating and melting of material at the surface of each pole. When the pulsating direct current supply occurring at the rate of approximately 20,000–30,000 Hz is turned off, the plasma channel breaks down. This causes a sudden reduction in the temperature allowing the circulating dielectric fluid to implore the plasma channel and flush the molten material from the pole surfaces in the form of microscopic debris. This process of melting and evaporating material from the workpiece surface is in complete contrast to the conventional machining processes, as chips are not mechanically produced (Ho & Newman, 2003).

The most infulential process parameters of EDM process are dischrge voltage, peak current, pulse duration (pulse on time and pulse off time), duty factor, polarity, type of dielectric flushing, spark gap, pulse frequency and corresponding performance measures are Material Removal Rate (MRR), Tool Wear Rate (TWR), Surface Roughness (SR), total machining time and etc, (Figure 2). However, optimizing any of these meaures alone have a limited value in real practice, due to the complex nature of the process where several different and sometimes contradictory objectives must be simultaneously considered (Chowdary & Yuvaraj, 2012). For this, in recent years multi-criteria process modeling and optimization has recived more attention by researchers and practitioners (Asoka & Kumar, 2008).

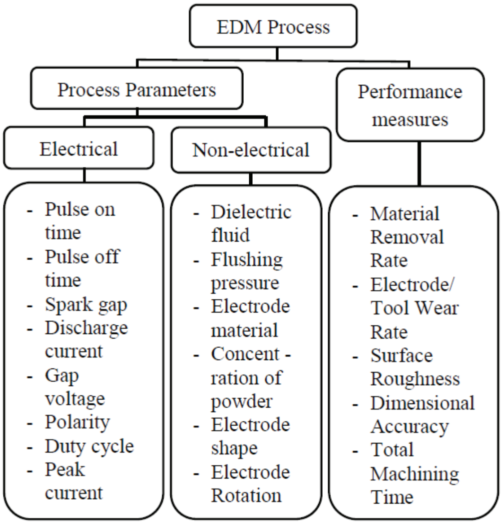

Figure 2. Process parameters and performance measures of EDM (Sanghani & Acharya, 2014)

To improve the process performance of EDM, it is essential to understand how performance measures depend on input parameters. The success of optimization techniques depend on the establishment of proper relationships between input parameters and performance measures. But the stochastic and complex nature of the process makes it too difficult to establish such relationship. During earlier days, based on the actual mechanism physical models of EDM process were developed. But, the physical models could not establish the relationships between input parameters and performance measures accurately as the process involves thermal, electrical and metallurgical variables. Further, inevitable assumptions made while physical modeling of the process, induce large deviations from actual process. Thus, inability of physical models has led the researchers to develop empirical or data based model of EDM process. Many empirical, statistical and regression techniques have been used for modeling EDM process (Zorepour, et al. (2004)). Fitting curves to non-linear data becomes difficult when number of inputs is high. Therefore statistical techniques find limited application in modeling EDM process. Regression techniques also do not provide satisfactory results because of the presence of noise in the system variables of the EDM process.

In recent years, Arteficial Neural Networks (ANNs) have demonstrated great potential in modeling of the input–output relationships of complicated systems.There are many types of artificial neural networks which differ in architecture, implementation of transfer functions and strategy of learning. In view of their universal approximation property, Back Propagation Neural Network (BPNN) has received considerable attention. The feature subsets, the number of hidden layers, and the number of processing elements in hidden layers are the architectural factors of BPNN to be determined in advance for the modeling process (Jung & Kwon, 2008).

For ANN, the heuristic algorithms are popularly applied to select best architecture including the optimal number of processing elements (Assarzadeh & Ghoreishi, 2008).

Several researchers have shown the applicability and superiority of ANN in modeling and optimization of machining processes.

Pushpendra et al. (2012) used controlled elitist non-dominated sorting genetic algorithm to optimize the EDM process for Inconel 718. ANN with back propagation algorithm has been used to model EDM process. ANN has been trained with the experimental data set. Controlled elitist non-dominated sorting genetic algorithm has been employed in the trained network and a set of pareto-optimal solutions was obtained. Next, the prediction ability of the trained network has been verified experimentally. The average percentage difference between experimental and ANN's predicted value was 4 and 4.67 for MRR and SR respectively.

Mohana and Hanumantha (2010) used BPNN models to determine the settings of pulse on time, pulse off time, peak current and resistanse for the estimation of MRR and SR. Based on their results, peak current has the most influence on the two machining responses.

Mahdavinejad et al. (2011) used ANN with back propagation algorithm to model the EDM process for machining of Silicon Carbide (SiC). An ANN model has been trained within the experimental data. Various ANN architectures have been studied, and 3-5-5-2 is selected. Material removal rate and surface roughness have been optimized as objectives by using a multi-objective optimization method. Testing results demonestrated that the model is suitable for predicting the response parameters.

Bhavesh et al. (2013) proposed an artificial neural network for the prediction of MRR, SR and TWR in EDM of AISI H13 Steel. The neural network based on process model has been generated to establish relationship between input process conditions (gap voltage, peak current, pulse on time, pulse off time and electrode material) and process responses (MRR, SR and TWR). The ANN model has been trained and tested using the data generated from a series of experiments on EDM. The trained neural network system has been used to predict MRR, SR and TWR for different input conditions. The ANN model has been found efficient to predict EDM process responses for selected process conditions.

Previous studies proved the efficiency of ANN techniques and heuristic algorithms to model and optimize process parameters setting of EDM. In most of the cases the number of neurons in the hidden layers for the training algorithm is being selected through tial and error. But in this study the problem was tackled using simulated annealing (SA) algorithm. This study proposes a hybrid approach composed of ANN and SA algorithm to undertake the multi-criteria modeling and optimization for EDM of AISI2312 hot worked steel parts. To the best of our knowledge, there is no published work to study the EDM process of AISI2312 steel through the proposed method. First, the experimental data are gathered based on L36 Taguchi design matrix. Then, the process is modeled using an optimized ANN. To enhance the prediction accuracy of the proposed model, the architecture (number of hidden layer and neurons in each layer) of the network has been optimized using a SA algorithm. Finally, the optimized BPNN model has been embedded into a SA optimization procedure, to determine the best set of process parameter values.

2.0 EXPERIMENTAL DETAILS

2.1. Material properties

Hot worked alloys are among the hardest materials to shape because of their strength and chemical reactivity with tool materials. AISI2312 hot worked steel is a popular alloy used for plastic moulds, mould frames, pressure casting moulds and recipient sleeves. Because of its controlled sulphur content this material has poor toughness. Despite its unique properties, the usage of this alloy is limited due to the high processing costs. In this study AISI2312 hot worked steel parts have been applied since only a few researchers have done the studies regarding this material using EDM process. The chemical composition and mechanical properties of this alloy is provided in Table 1 and 2 respectively.

Table 1. Chemical composition of AISI2312

Composition						
C	Si	Mn	P	S	Cr	Mo
0.38	0.30	1.50	0.02	0.07	2.00	0.20

Table.2 Mechanical properties of AISI2312

Property	Unit	Value
Hardness	HRC	55-60 HRC
Average Coefficient of Thermal Expansion	µm/m · K	11.6
Young's modulus	GPa	212
Thermal Conductivity	W/m · K	34.0

The EDM operation is performed on workpeices having 10mm thickness and 40×20 mm2 dimension. The machining time for each test was 45 minutes. Furthermore, the experiments have been done in random order to increase accuracy.

2.2. Machine tool

In the present study, an Azerakhsh-304H die-sinking machine has been used to perform the experiments. Die-sinking machine used is shown in figure 3. Table.3 illustrates the mechanical specification of the machine tool used.

Figure 3. The Azerakhsh-304H EDM machine used for experiments

Table. 3 The detailed technical specifications of the machine tool used

SPECIFICATION	Size
Work Table Size	500×300 mm
Cross Travel Y	250 mm
Spindle Travel & Head Stock Travel	180+200 mm
Maximum Electrode weight	50 kg
Loading Capacity of Table	500 kg

2.3. Electrode material

The materials which are been used as tool electrode in EDM are brass, copper, tungsten and graphite. Brass is seldom used as electrode material in modern EDM because of its high wear rate. Use of tungsten as tool electrode is also limited to certain applications only. However, tungsten possess some qualities such as high density, high tensile strength and high melting point, its cutting speed is much slower than that of brass and copper due to its relatively poor electrical conductivity. In addition, high cost and low machinability are also disadvantages of tungsten to be used as EDM tool electrode. Copper and graphite are most commonly used electrode material in EDM. The wear rate of graphite is much less than that of copper due to its extremely high melting point. Copper can produce very fine surface due to its structure integrity. More so, the structure integrity makes copper electrode highly resistant to DC arcing in case of poor flushing conditions. It is difficult to machine graphite electrode being it polycrystalline and brittle. On the contrary, the machinability of copper is better than that of graphite.

Therefore based on the basis of these facts and literature survey, copper has been used as electrode material in this work. Table 4 shows the electrode used properties.

Table.4 Properties of the electrode used

Electrode Material	Density (g/cm3)	Thermal Conductivity (W/mK)	Electrode Resistivity (µ-ohms)	Hardness (BHN)
Copper	8.9	399	1.69	48

3.0 DESIGN OF EXPERIMENTS

Experimentation is an integral part of any engineering investigation. The word 'design' in the expression Design of Experiments (DOE), is used in a general sense to convey planning of experiments to fulfill intended objectives. To design the experiment is to develop a scheme or layout of the different conditions to be studied. In practice, 'design' refers to some form of engineering communication, such as a set of specifications, drawings or physical models that describe the concept. Since an experiment design should satisfy primarily the conditions for each experimental run. Therefore, before designing an experiment, knowledge of the product / process under investigation is of the prime importance for identifying the factors that influence the outcome. The general scenario in an experiment is that there is an output variable (generally quantitative in nature), which depends on several input variables, called factors. Each factor has at least two settings, called levels. A combination of the levels of all the factors involved in the experiment is called a treatment combination. DOE is a statistical technique used to study the effects of multiple variables on performance measures simultaneously. It provides an efficient experimental schedule and statistical analysis of the experimental results (Roy et al., 1990). DOE strategy involves, a) Selection of process parameters and their levels. b) Selection of performance measures. c) Selection of the matrix of experiments.

3.1. Taguchi technique

Taguchi technique constructed a special set of general designs for factorial experiments that overcomes the drawbacks of partial factorial experiment. The method is popularly known as Taguchi's method. The special set of designs consists of Orthogonal Arrays (OA). The OA is a method of setting up experiments that only requires a fraction of full factorial combinations. The treatment combinations are chosen to provide sufficient information to determine the factor effects using the analysis of means. Orthogonal refers to the balance of the various combinations of factors so that no one factor is given more or less weight in the experiment than the other factors. Orthogonal also refers to the fact that effect of each factor can be mathematically assessed independent of the effect of the other factors. Taguchi's method, firstly, clearly defines orthogonal arrays, each of which can be used for many experimental situations. Secondly, Taguchi's method provides a standard method for analysis of results. Taguchi's method provides consistency and reproducibility that is generally not found in other statistical methods. (Roy et al., 1990).

Taguchi's method has been used as a DOE technique for the present work. Taguchi's method provides the special set of design that requires a fraction of full factorial combinations. This study has been undertaken to investigate the effects of peak current (I), voltage (V), pulse off time (T_{off}), pulse on time (T_{on}), and duty factor (η) on material removal rate (MRR) tool wear rate (TWR) and surface roughness (SR). Therefore, L_{36} ($2^1 \times 3^4$) has been used to carry out experiments. Five process (input) parameters have been selected on the basis of literature survey and preliminary investigations. Out of five, one factor has 2 levels and the rest of the factors have 3 levels each (Roy et al., 1990). Preliminary experiments were conducted for the wide range of pulse-on-time, discharge current and gap voltage. Satisfactory results were obtained for 2.5-7.5A, range of peak current. Below 2.5 A, MRR was very low and beyond 7.5 A, MRR was good but SR was vey poor. Similar observations were made for specified range of pulse on and off time and gap voltage. The range of the other parameters was fixed based on the ranges as specified in machine manufacturer's manual.

Table 5 lists the machining parameters and their levels. The limitations of test equipment may also dictate a certain number of levels for some of the process parameters. In our case, the die-sinking EDM machine used for experiments had only two settings for pulse off time - T_{off} (10 and 75 μs).

Table 5. Machining Parameters and Levels

Parameters	Range	Level 1	Level 2	Level 3
Peak current (A)	2.5-7.5	2.5	5	7
Voltage (V)	50-60	50	55	60
Pulse on time (μs)	25-200	25	100	200
Pulse off time (μs)	10-75	10	75	-
Duty factor (S)	0.4-1.6	0.4	1	1.6

4.0 MACHINING PERFORMANCE EVALUATION

Material removal rate(MRR), surface roughness (SR), and tool wear ratio (TWR) are used to evaluate machining performance. The MRR is expressed as the workpiece removal weight (WRW) under a period time of machining (T) in minute (45 minutes).

$$MRR = \frac{\text{WRW}}{\text{time of machining}} \tag{1}$$

The TWR, usually expressed as a percentage, and is defined by the ratio of the tool wear weight (TWW) to the work piece removal weight (WRW) which is obtained using Equation (2) (Sanghani & Acharya, 2014).

$$TWR(\%) = \frac{\text{TWW}}{\text{WRW}} \times 100 \tag{2}$$

The SR value of the machined product is also one of the most important quality characteristics. The parameter Ra is used in this study. The average roughness (Ra) is the area between the roughness profile and its mean line, or the integral of the absolute value of the roughness profile height over the evaluation. This partameter is measured by Equation (3) (Sanghani & Acharya, 2014)

$$Ra = \frac{1}{L} \int_0^L |\ Z(x)\ |\ dx \tag{3}$$

where Ra is the arithmetic average deviation from the mean line, L the sampling length, and Z(x) is the ordinate of the profile curve (Pushpendra et al., 2012).

Figure 3. Digital surface roughness tester and electronic balance used

5.0 EXPERIMENTAL RESULTS AND THEIR NORMALIZED GRADES

In EDM, multi-criteria modeling is needed to simultaneously achieve high MRR, good SR and low TWR. Multiple output responses can be transformed into a single measure by using Weighted Normalized Grades (WNG) method (Jung & Kwon, 2008). This is done by normalizing the values of process responses and then averaging these normalized values. In this way the effects of adopting different units may be eliminated (Gao et al., 2008). The MRR corresponds to the higher-the-better quality characteristics and hence the experimental measurements are normalized using the following formula:

$$Z_i = \frac{(y_i - \min(y_i, i = 1,2\ldots, n))}{(\max(y_i, i = 1,2\ldots, n) - \min(y_i, i = 1,2\ldots, n))} \tag{4}$$

Likewise, for lower-the-better quality characteristics such as SR and TWR normalized data are calculated by:

$$Z_i = \frac{(\max(y_i, i = 1,2\ldots, n) - y_i)}{(\max(y_i, i = 1,2\ldots, n) - \min(y_i, i = 1,2\ldots, n))} \tag{5}$$

In the above equations, n is the number of trials (in our case 36), y_i is the value of the observed response in the i^{th} trial and Z_i is the corresponding normalized value for y_i. Now the multiple output responses can be transformed into a single WNG through the following equation:

$$G_i = \sum_{k=1}^{p} \beta_k . Z_i \tag{6}$$

where p is the number of performance measures, β_k is the normalized relative weight of the k^{th} response. In this study, p is 3 which correspond to MRR, SR and TWR. The weighting coefficient, β_k, is assumed to be the same (β_k = 0.333) for all three process outputs. This single measure may now be used for model development. The 36 experimental settings and their corresponding measured outputs are recorded in Table 3. The last column of Table 3 shows the calculated WNG for each test.

Table 3. The L_{36} experimental design and results

No	T_{off}(μs)	T_{on}(μs)	I(A)	η(sec)	V(V)	SR (μm)	MRR(gr/min)	TWR (%)	WNG
1	1	1	1	1	1	3.6	0.35	0.04	0.64
2	1	2	2	2	2	7.2	3.04	0.08	0.60
3	1	3	3	3	3	3.2	0.33	0.03	0.66
4	1	1	1	1	1	7.2	2.08	0.07	0.57
5	1	2	2	2	2	13.0	6.84	0.03	0.55
6	1	3	3	3	3	3.8	0.45	0.03	0.64
7	1	1	1	2	3	8.8	5.52	0.15	0.60
8	1	2	2	3	1	13.0	2.83	0.02	0.42
9	1	3	3	1	2	4.6	0.56	0.03	0.62
10	1	1	1	3	2	7.6	1.56	0.06	0.54
11	1	2	2	1	3	13.4	10.64	0.07	0.64
12	1	3	3	2	1	5.0	1.70	0.61	0.43
13	1	1	2	3	1	8.4	2.53	0.19	0.50
14	1	2	3	1	2	6.4	0.88	0.01	0.58
15	1	3	1	2	3	4.8	1.28	0.46	0.48
16	1	1	2	3	2	10.2	2.24	0.15	0.45
17	1	2	3	1	3	6.0	1.14	0.01	0.60
18	1	3	1	2	1	4.4	0.57	0.20	0.56
19	2	1	2	1	3	7.0	2.99	0.33	0.51
20	2	2	3	2	1	6.4	0.85	0.01	0.58
21	2	3	1	3	2	4.6	1.20	0.47	0.48
22	2	1	2	2	3	8.4	4.43	0.35	0.50
23	2	2	3	3	1	5.8	0.37	0.01	0.58
24	2	3	1	1	2	5.8	2.00	0.93	0.30
25	2	1	3	2	1	5.8	0.77	0.01	0.59
26	2	2	1	3	2	11.2	1.74	0.01	0.45
27	2	3	2	1	3	4.6	1.84	0.82	0.38
28	2	1	3	2	2	4.4	0.67	0.01	0.63
29	2	2	1	3	3	11.6	1.91	0.01	0.44
30	2	3	2	1	1	5.2	1.57	0.66	0.40
31	2	1	3	3	3	6.6	0.44	0.01	0.56
32	2	2	1	1	1	8.8	4.26	0.03	0.60
33	2	3	2	2	2	5.0	0.85	0.40	0.48
34	2	1	3	1	2	5.4	0.64	0.01	0.60
35	2	2	1	2	3	9.2	5.13	0.01	0.62
36	2	3	2	3	1	3.2	0.91	0.01	0.68

6.0 MODEL DEVELOPMENT - THE OPTIMIZED BACK PROPAGATION NEURAL NETWORK

6.1 Simulated Annealing (SA) Algorithm

Simulated annealing (SA) algorithm is an optimization process whose operation is strongly reminiscent of the physical annealing of crystalline compounds such as metals and metallic alloys (Kirkpatric et al., 1983). In condensed matter physics, annealing is a physical process that is used to reconstruct the crystal structure of a solid with a low energy state. A solid in a state bath is first heated up to a temperature above the melting point of the solid. At this temperature, all particles of the solid are in violent random motion. The temperature of the heat bath is then slowly cooled down. All particles of the solid rearrange themselves and tend toward a low energy state. As the cooling of the particle is carried out sufficiently slowly, lower and lower energy states are obtained until the lowest energy state is reached. Similarly, in EDM an energy function is created which is minimized. While minimizing efforts are made to avoid local minima and to achieve global minima. The lowest energy level gives the optimized value of EDM parameters. In recent years, the simulated annealing algorithm has emerged as a leading tool for large-scale combinational optimization problems.

A standard SA procedure begins by generating an initial solution at random. At initial stages, a small random change is made in the current solution. Then the objective function value of new solution is calculated and compared with that of current solution. A move is made to the new solution if it has better value or if the probability function implemented in SA has a higher value than a randomly generated number. The probability of accepting a new solution is given as follows:

$$p = \begin{cases} 1 & if \;\; \Delta < 0 \\ e^{-\Delta/t} & if \;\; \Delta \geq 0 \end{cases} \tag{7}$$

The calculation of this probability relies on a temperature parameter, *T*, which is referred to as temperature, since it plays a similar role as the temperature in the physical annealing process. To avoid getting trapped at a local minimum point, the rate of reduction should be slow (Yang & Srinivas, 2009). In our problem the following method to reduce the temperature has been used:

$$T_{i+1} = cT_i \qquad i = 0,1,... \;\; and \qquad 0.9 \leq c < 1 \tag{8}$$

Thus, at the start of SA most worsening moves may be accepted, but at the end only improving ones are likely to be allowed. This can help the procedure jump out of a local minimum. The algorithm may be terminated after a certain volume fraction for the structure has been reached or after a pre-specified run time.

Simulated annealing algorithm has diverse applications including improving the performance of other artificial intelligence techniques and determining the optimal set of process parameters (Yang & Srinivas, 2009 and Markopoulos, 2008). In this research, SA has been used twice. First it is employed to determine the best architecture (number of layers and number of neurons in each layer) of the ANN to model the EDM process. Once the best architecture of the ANN is determined, the proposed model is implanted into a SA procedure to find the optimal set of EDM process parameters.

6.2 The optimized Back Propagation Neural Network (BPNN)

The first model of the artificial neural network (ANN) was given by McCulloch and Pitts in 1943. ANNs are simplified models of biological nervous system inspired by the computing performed by a human brain. Kohonen defined neural network as "massively parallel interconnected networks of simple usually adaptive elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous system do." ANNs have the capability to learn and thereby acquire knowledge and make it available for use (McCulloch & Pitts, 1943).

ANNs are built by connecting processing units, called nodes or neurons. Each of the input (Xi) is associated with some weight (Wi) which takes a portion of the input to the node for processing. The node combines the inputs (XiWi) and produces net input which in turn is transformed into output with the help of transfer function/activation function (McCulloch & Pitts, 1943).

Traditional modeling methods are mostly relied on assumptions for model simplifications, and consequently may lead to inaccurate results. Recently, ANN has become a powerful and practical method to model complex non-linear systems. The basis of NN modeling is to capture the underlying trend of the data set presented to it, in the form of a complex nonlinear relationship between the input parameters and the output variable (Debabrata et al., 2007). Learning, generalization, and parallel processing are important advantages of ANN. These characteristics of the ANNs make them suitable for EDM process modeling.

Many researchers have proposed that multilayered networks are capable of computing a wide range of Boolean functions than networks with a single layer of computing units (Huang & Huang 1991). However, the computational effort needed for modeling a system increases substantially when more parameters and more complicated architecters are considered. The Back Propagation Neural Networks (BPNN) is found most suitable for handling such large learning problems. This type of neural network is known as a supervised network because it requires a desired output in order to learn. A BPNN consists of multiple layers of nodes in a directed scheme, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function (Sexton & Allidae, 1998).

The architecture of a neural network specifies the number of layers, the number of neurones, each layer contains and how the neurons are interconnected. The design of an BPNN architecture is "more of an art than a science", in the sense that optimal design is based more on metha heuristic or experience rather than on proven methods (Sexton & Allidae, 1998). Metha heuristic methods such as SA algorithm can be used to construct proper architecture design for NN model development. Also in that generates a single output value from all of the input values that are applied to the neuron. Every connection has a transfer function that is applied to the input value associated with the connection (Debabrata et al., 2007). In this study log-Sigmoid transfer function is used for network training, defined as follows:

$$\mathrm{F}_{\mathrm{i,j}} = \frac{1}{1+\exp^{-P(W_{i,j-1},O_{i,j-1})}} \tag{9}$$

where, for ith neuron in the j^{th} layer, $P\ (W_{i,j-1}, O_{i,j-1})$ is given by:

$$P(W_{i,j-1}, O_{i,j-1}) = \sum_{j=1}^{m}\sum_{i=1}^{n} W_{i,j-1}.O_{i,j-1} \tag{10}$$

where, n and m are number of hidden layers and neurons in each layer respectively. $W_{i,j-1}$ is the weight of the ith neuron in $(j\text{-}1)^{th}$.

One of the most important tasks in ANN modeling is to choose the best network architecture, namely the number of hidden layers and the number of neurons in each layer. Since the number of possible combinations may be very large, the trial-and-error approach is inefficient. In this study, to specify the best ANN architecture SA is employed. Usually the performance of the network will be checked by Mean Square Error (MSE) between desired outputs (Y_k) and predicted outputs (y_k) which is expressed as:

$$MSE = \frac{1}{\mathrm{p}}\sum_{\mathrm{k}=1}^{\mathrm{p}}(\mathrm{Y_k} - \mathrm{y_k})^2 \tag{11}$$

Learning MSE and the generalization MSE, detect the two main characteristics of "learning" and "generalization" of ANN. The effectiveness of developed net dependence on this features (Krishna & Hanumantha, 2009).

8.3 Proper structure derivation of the model

The appropriate neural network architecture for model development was tuned via SA. Number of hidden layers was varied from 1 to 4; hence a $5–n_1–n_2–n_3–n_4–1$ structure was constructed; where n_1, n_2, n_3 and n_4 are the number of nods in the 1st to the 4th hidden layers. The training of a neural network implies finding desired net's architecture and weights that minimize error between the desired output and the predicted outputs. The first step in training is the forward phase which occurs when an input vector X is presented and propagated through the network to compute an output (Krishna & Hanumantha, 2009). Hence, an error between the desired output (Y_k) and predicted output (y_k) of the neural network is computed. So the modeling authority of the net (M(net))canbegiven by:

$$M(\text{net}) = \alpha.\frac{1}{p_0}\sum_{r=1}^{p_0}(Y_r - y_r)^2 + \beta.\frac{1}{q_0}\sum_{s=1}^{q_0}(Y_s - y_s)^2 \qquad (12)$$

where α and β are coefficients which determine the relative importance of learning and generalization capability of respect NN. Also p_0 and q_0 are number of training and testing data respectively. The recent relation corresponds to fitness function for developing the optimized BPNN construction. In backward phase of BP training, to minimize the error between the desired and actual outputs, the gradient descent method with a momentum coefficient, is used (Markopoulos & Manolakos, 2008). The weights are updated using the following rule:

$$\Delta w_{i,j}{}^{(n+1)} = -\Gamma.\frac{\partial(M(\text{net}))}{\partial w_{i,j}{}^{n}} + \omega.w_{i,j}{}^{n} \qquad (13)$$

where Γ is learning rate and Δwi,jn is change of neuron's weight at previous step and ω is momentum.

In first step initial parameters and an initial net structure is been configured. Based on the fittness function M(net) the approximation aptitude of developed model is been evaluated. At each iteration a new architecture based on the current structure is generated and evaluated. This new model's structure is then accepted if the objective functions (M(net)) is lower than the current one or if the value of the probability function implemented in SA has a higher value than a randomly generated number between zero and one. otherwise the algorithm data have been updated and a new structure based on the current structure have been derivated. This iterative steps is continued until the algorithm has been converged after the predetermined number of iterations.

7.0 RESULTS AND DISCUSSION

7.1 OBPNN prediction results

Figure 4 shows convergence of SA for objective function (M(net)). The architecture of the OBPNN is shown in Figure 5. The optimum number of hidden layers is 2 with 7 and 3 hidden layer nodes in first and second layer respectively (architecture of OBPNN is: 5–7–4–1). The linear regression analysis is conducted to compute the correlation correlation (R2adj) between actual experimental and predicted WNGs. The correlation coefficients at the train and testing stage is 0.998 and 0.987 respectively. The related fitness function for trained model is 1.53 (M(OBPNN)= 1.53), mean square error between actual and trained data is 0.14 while maximum and minimum of absolute errors are 0.01% and 1.33% respectively. It is clear that the proposed model predictions follow the experimental results very closely therefore can accurately predict the actual WNGs.

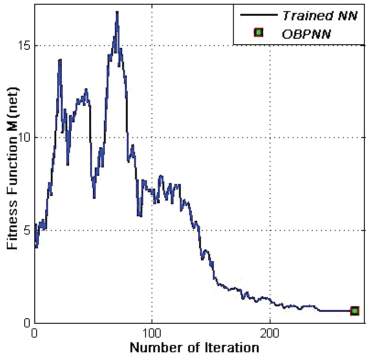

Figure 4. Convergence of SA in OBPNN training

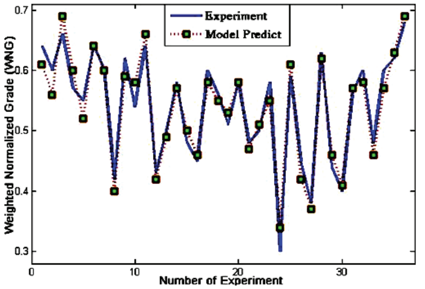

Figure 5. Comparison between experimental and verification of WNGs

It can be deducted that developed model is adequate and based on it, optimum set of machining can be found. In Figure 5 results of comparison between experimental and verification of OBPNN for WNGs is seen. Predicted WNGs by OBPNN pursuit the experimental results closely.

7.2 EDM process input parameters optimization

After successfully process modeling, another algorithm based on SA has been developed and optimal value of machining outputs have been determined. The developed single model of OBPNN was considered as objective function of this algorithm; where maximum WNG is desirable. In complementary section, in order to evaluate the accuracy of the predicted values, another actual experiment was carried out based on the optimized process parameters and the obtained experimental responses was compared with the initial parameters design. The results which are presented in Table 4, show that the hybrid model can improve machining performances. As observed in Table 4, optimal machining set is: T_{off}= 15μs, T_{on}= 25 μs, I=3 A, η= 0.8 sec V= 65v. For optimal machining set rather than initial machining design, material removal rate increases form 0.35 g/min to 3.43 g/min; tool wear rate is reduced from 0.04 to 0.02 mm; and surface roughness value was reduced from 3.6 μm to 3.16 μm. It is evident that quality characteristics can be greatly improved through proposed method.

Table 4. Optimal EDM parameters and related responses

Set	Process parameter	MRR (g/min)	TWR (%)	SR (μm)	WNG
Initial process design	T_{off} (μs) = 10,T_{on} (μs) = 25 I (A) =2.5, η (sec) = 0.40, V (v) = 50	0.35	0.04	3.60	0.64
Optimal process design	T_{off} (μs) = 15 , T_{on} (μs) = 25 I (A) = 3 , η (sec) = 0.80, V (v) = 65	3.43	0.02	3.16	0.73

7.0 CONCLUSION

Hybrid modeling and optimization of process parameters and responses in EDM process of AISI2312 hot worked steel have been implemented. Experimental data for process modeling obtained from conducted experiments by Taguchi methodology, a systematic tool for design of experiments. Multiple process output measure transformed to the single measure namely weighted normalized grade (WNG).

The optimized back propagation neural network (OBPNN) was developed to establish accurate model of process multiple performance characteristics. The optimal net's architecture (number of neurons and hidden layers) of OBPNN has been specified using simulated annealing algorithm. Correlation coefficient (R2adj) and mean square error (MSE) between the experimental and predicted values have been calculated. Results demonstrates that proposed model of OBPNN models the EDM process efficiently; so the proper machining input parameters determined via SA based on the developed model.

The validation of proposed method was evaluated based on a confirmation test; which the actual experiment outputs for optimal design compared to initial machining set. Using this approach, substantial improvements of the prediction capability of the ANNs could be realized comparatively with the other commonly used modeling methods. From the present analysis it is evident that the proposed hybrid model will be very beneficial in multi-objective process modeling and quality performance optimization.

REFERENCES

Assarzadeh, S. & Ghoreishi, M. (2008). Neural-network-based modeling and optimization of the electro-discharge machining process. *Journal of Advance Manufacturing Technology*, *39*, 488-500.

Asoka, P., Kumar, R. R., Jeyapaul, R. & Santhi, M. (2008). Development of multi-objective optimization models for electrochemical machining process. *Journal of Materials Manufacturing Technology*, *39*, 55-63.

Bhavesh, A., Patel, D. S., & Haresh, A. (2013). Influence of electrode material and process parameters on surface quality and MRR in EDM of AISI H13 using ANN. *International Journal on Recent and Innovation Trends in Computing and Communication, 1*(12), 858 – 869.

Debabrata, M., Surjya, K. & Partha, S. (2007). Modeling of electrical discharge machining process using back propagation neural network and multi-objective optimization using non dominating sorting genetic algorithm-II. *Journal of Materials Processing Technology*, *186*, 154-162.

Gao, Q., Zhang, Q., Su, S., & Zhang, J. (2008). Parameter optimization model in electrical discharge machining process. *Journal of Zhejiang University Science, 9*, 104-108.

Ho, K. H. & Newman, S. T. (2003). State of the art electrical discharge machining (EDM). *International Journal of Machine Tools and Manufacture*, *43*, 1287–1300.

Jung, H. & Kwon, W. T. (2008). Optimization of EDM process for multiple performance characteristics using Taguchi method and Grey relational analysis. *Journal of Advance Manufacturing Technology, 37*, 523-533.

Krishna. M. R. & Hanumantha. D. (2009). Development of hybrid model and optimization of surface roughness in electric discharge machining using artificial neural networks and genetic algorithm. *Journal of Materials Processing Technology, 209*, 1512-1520.

MahdaviNejad, R. A. (2011). Modeling and optimization of electrical discharge machining of SiC parameters, using neural network and non-dominating sorting genetic algorithm (NSGA II). *Materials Sciences and Applications, 2*(1), 669-675.

Markopoulos, A. P., Manolakos, D. E., & Vaxevanidis, N. M. (2008). Artificial neural network models for the prediction of surface roughness in electrical discharge machining. *Journal of Intelligence Manufacturing*, 283-292.

Mohana G. K., Hanumantha. D., & Rajurkar. K. (2010). Hybrid modeling and optimization of surface produced by electric discharge machining using artificial neural network and genetic algorithm. *Journal of Engineering and Applied Sciences, 5*, 72-81.

Petropoulos, G., Vaxevanidis, N. M. & Pandazaras, C. (2004). Modeling of surface finish in electro-discharge machining based on statistical multiparameter analysis. *Journal of Material Processing Technology*, 1247-1251.

Pushpendra S. Bharti, S. Maheshwari, & Sharma, C. (2012). Multi-objective optimization of electric-discharge machining process using controlled elitist NSGA-II. *Journal of Mechanical Science and Technology, 26*(6), 1875-1883.

Sanghani, C. R. & Acharya, G. D. (2014). A review of research on improvement and optimization of performance measures for electrical discharge machining. *Journal of Engineering Research and Applications, 4*(1), 433-450.

Sexton, R., Allidae, B., Dorsey. R. E., & Johnson. J. D. (1998). Global optimization for artificial neural networks: A tabo search application. *Journal of Operational Research, 106*, 570-584.

Yanamandala, R. C., Yuvaraj, K., & Prahlada R. B. (2012). Neural network for prediction of EDM of Al/Sic-graphite particulate reinforced hybrid composites. *International Journal of Emerging Technology and Advanced Engineering, 12*(2), 730-739.

Yang S. H. and Srinivas, J., Mohan, S., Lee D. M., & Balaji, S. (2009). Optimization of electric discharge machining using simulated annealing, *Journal of Materials Processing Technology, 209*, 4471–4475.

Zorepour, H., Tehrani, A.F., Karimi, D. and Amini, S. (2007). Statistical analysis on electrode wear in EDM of tool steel DIN 1.2714 used in forging dies. *Journal of Material Processing Technology*, 711-714.