An optimised back propagation neural network approach and simulated annealing algorithm towards optimisation of EDM process parameters

Masoud Azadi Moghaddam* and
Farhad Kolahan

Department of Mechanical Engineering,
Ferdowsi University of Mashhad,
P.O. Box 91775-1111, Mashhad, Iran
Email: masoud_azadi88@yahoo.com
*Corresponding author

Abstract: The present research addresses the multi-criteria modelling and optimisation of electrical discharge machining (EDM) process, via optimised back propagation neural networks (OBPNN) and simulated annealing (SA) algorithm. The process response characteristics considered are material removal rate, surface roughness, and tool wear rate. The process input parameters include voltage, peak current, pulse off time, and pulse on time and duty factor. The three performance characteristics are combined into a single objective using weighted normalised grades (WNG) obtained from experimental study based on Taguchi method, to develop the artificial neural network (ANN) model. In order to enhance the prediction capability of the proposed model, its architecture is tuned by SA algorithm. Next, the developed model is embedded into SA algorithm to determine the best set of process parameters values for an optimal set of outputs. Experimental results indicate that the proposed optimisation procedure is quite efficient in modelling and optimisation of EDM process parameters.

[Received 25 January 2015; Revised 12 April 2015; Accepted 3 May 2015]

Keywords: electrical discharge machining; EDM; Taguchi method; optimisation; artificial neural network; ANN; optimised back propagation neural network; OBPNN; simulated annealing algorithm; SA.

Reference to this paper should be made as follows: Moghaddam, M.A. and Kolahan, F. (2015) 'An optimised back propagation neural network approach and simulated annealing algorithm towards optimisation of EDM process parameters', *Int. J. Manufacturing Research*, Vol. 10, No. 3, pp.215–236.

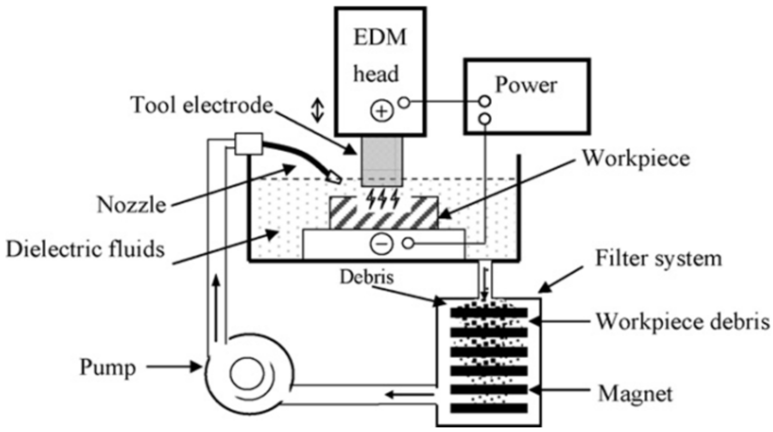
Biographical notes: Masoud Azadi Moghaddam holds a BSc degree in Production and Manufacturing Engineering. He has also obtained his MSc degree in Mechanical Engineering (Manufacturing Engineering) from Ferdowsi University of Mashhad, I.R. Iran. He is now a PhD candidate in the same university, under the supervision of Dr. Kolahan. The research he has undertaken includes modelling and optimisation of advanced manufacturing processes. He has published several journal and conference papers in his field of research during his postgraduate studies.

Farhad Kolahan is an Associate Professor at the Department of Mechanical Engineering, Ferdowsi University of Mashhad, I.R. Iran. He received his BSc degree in Production and Manufacturing Engineering from Tabriz University, Iran. He then continued his postgraduate studies abroad and graduated with a PhD degree in Industrial and Manufacturing Engineering from Ottawa University, Canada in 1999. His research interests include production planning and scheduling, manufacturing processes optimisation and applications of heuristic algorithms in industrial optimisation.

1 Introduction

In recent years various machining processes have been developed or modified to cope with high alloy materials. Among these alloys, hot worked 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. Electrical discharge machining (EDM) is a non-conventional 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 work piece 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 manufacturing of moulds, dies, aerospace and surgical components (Pushpendra et al., 2012).

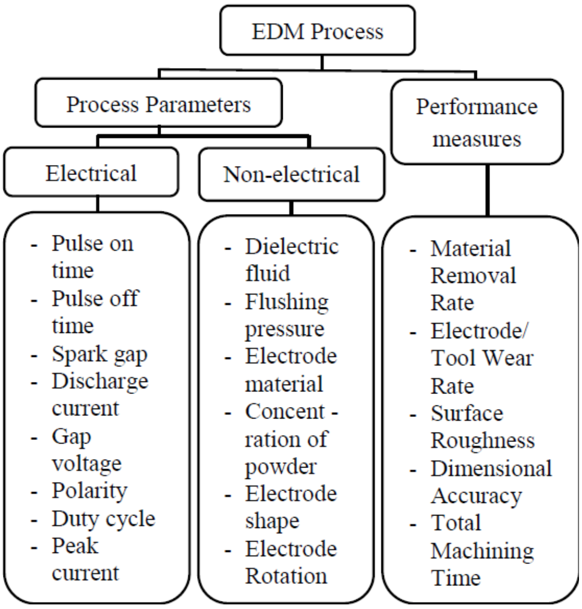
Figure 1 Schematic illustration of EDM



Source: Ho and Newman (2003)

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 8,000 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 work piece surface is in complete contrast to the conventional machining processes, as chips are not mechanically produced (Ho and Newman, 2003).

Figure 2 Process parameters and performance measures of EDM



Source: Yanamandala et al. (2012)

The most influential process parameters of EDM process are discharge 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, optimising any of these measures alone has 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 (Yanamandala et al., 2012). For this, in recent year's multi-criteria process modelling and optimisation has received more attention by researchers (Asoka et al., 2008).

To improve the process performance of EDM, it is essential to understand how performance measures depend on input parameters. The success of optimisation 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 modelling of the process, induce large deviations from actual process. Thus, inability of physical models has led the researchers to develop empirical or databased model of EDM process. Many empirical, statistical and regression techniques have been used for modelling EDM process (Zorepour et al., 2007; Petropoulos et al., 2004). Fitting curves to nonlinear data become difficult when number of inputs is high. Therefore statistical techniques find limited application in modelling EDM process. Regression techniques also do not provide satisfactory results because of the presence of noise in the system variables of the EDM process (Zorepour et al., 2007; Petropoulos et al., 2004; Jung and Kwon, 2008).

In recent years, artificial neural networks (ANNs) have demonstrated great potential in modelling of the input–output relationships of complicated systems. There are many types of ANNs 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 modelling process (Jung and Kwon, 2008; Assarzadeh and Ghoreishi, 2008).

Several researchers have shown the applicability and superiority of ANN in modelling and optimisation of machining processes.

Pushpendra et al. (2012) used controlled elitist non-dominated sorting genetic algorithm to optimise 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 dataset. 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 (Pushpendra et al., 2012).

Mohana et al. (2009) used BPNN models to determine the settings of pulse on time, pulse off time, peak current and resistance for the estimation of MRR and SR. Based on their results, peak current has the most influence on the two machining responses.

Mahdavi Nejad (2011) used ANN with back propagation 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. MRR and SR have been optimised as objectives by using a multi-objective optimisation method. Testing results demonstrated that the model is suitable for predicting the response parameters (Mahdavi Nejad, 2011).

Bhavesh et al. (2013) proposed an ANN 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 (Bhavesh et al., 2013).

Previous studies proved the efficiency of ANN techniques and heuristic algorithms to model and optimise process parameter setting of EDM. In these cases the number of neurons in the hidden layers for the training algorithm is being selected through trial and error. But in the proposed 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 modelling and optimisation 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. The purpose of this paper is to present an efficient and integrated approach for the determination of appropriate machining parameters yielding the objective of maximum MRR and minimum SR and TWR simultaneously. First, the experimental data are gathered based on L_{36} Taguchi design matrix. Then, the process is modelled using an optimised back propagation neural network (OBPNN). 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 optimised using SA algorithm. Finally, the OBPNN model has been embedded into a SA optimisation procedure, to determine the best set of process parameter values.

2 Experimental details

2.1 Material properties

The first step necessary to initiate the study was to select the material and size of the specimens for tests. 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 mechanical properties of this alloy are provided in Table 1.

Table 1 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 work pieces having 10 mm thickness and 40×20 mm^2 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 order to obtain different machining process parameters and output features for the training and testing of the neural network, a series of experiments were performed on an Azerakhsh-304H die-sinking machine. Die-sinking machine used is shown in Figure 3. Table 2 illustrates the technical specification of the machine tool used.

Figure 3 The Azerakhsh-304H EDM machine used for experiments (see online version for colours)

Table 2 The detailed technical specifications of the machine tool used

Specification	*Size*
Work table size	500 × 300 mm
Cross travel Y	250 mm
Spindle travel and head stock travel	180 + 200 mm
Maximum electrode weight	50 kg
Loading capacity of table	500 kg
Work table size	500 × 300 mm

2.3 Electrode material

The materials which are 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 possesses 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 materials 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 structural integrity. More so, the structural 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 on the basis of these facts and literature survey, copper has been used as electrode material in this work.

3 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 fulfil 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. 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 [12].

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 method

Taguchi method 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 OA, 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, 2010; Gopalsamy et al., 2009; Payal et al., 2008).

At first, some preliminary tests were carried out to determine the stable domain of the machine parameters and also the different ranges of process variables. Based on literature survey, preliminary test results and working characteristics of the EDM machine, peak current (I), voltage (V), pulse off time (T_{off}), pulse on time (T_{on}), and duty factor (η) were chosen as the independent input parameters. During these experiments, by altering the values of the input parameters to different levels, stable states of the machining conditions have also been specified. 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, gap voltage and duty factor. Therefore, L_{36} ($2^1 \times 3^4$) has been used to carry out experiments. Out of five, one factor has two levels and the rest of the factors have three levels each. Therefore 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 MRR TWR and SR.

Table 3 Machining parameters and levels

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

Table 3 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).

4 Machining performance evaluation

MRR, SR, and 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 WRW which is obtained using equation (2) (Sanghani and 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 parameter is measured by equation (3) (Sanghani and Acharya, 2014).

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

where *Ra* is the arithmetic average deviation from the mean line, *L* is the sampling length, and *Z*(*x*) is the ordinate of the profile curve (Sanghani and Acharya, 2014).

5 Experimental results and weighted normalised grade

5.1 Experimental results

Table 4 shows the L_{36} experimental matrix for five input variables and three output parameters of the EDM process.

Table 4 The L_{36} experimental design and results

No	$T_{off}(\mu s)$	$T_{on}(\mu s)$	*I(A)*	*η (Sec)*	*V(V)*	*SR (μm)*	*MRR (gr/min)*	*TWR (%)*	*WNG*
1	1	1	1	1	1	3.6	0.35	11.4	0.64
2	1	2	2	2	2	7.2	3.04	2.6	0.60
3	1	3	3	3	3	3.2	0.33	0.6	0.66
4	1	1	1	1	1	7.2	2.08	9.0	0.57
5	1	2	2	2	2	13.0	6.84	3.3	0.55
6	1	3	3	3	3	3.8	0.45	0. 4	0.64
7	1	1	1	2	3	8.8	5.52	6.7	0.60
8	1	2	2	3	1	13.0	2.83	2.7	0.42
9	1	3	3	1	2	4.6	0.56	0.7	0.62
10	1	1	1	3	2	7.6	1.56	5.3	0.54
11	1	2	2	1	3	13.4	10.64	3.8	0.64
12	1	3	3	2	1	5.0	1.70	0.7	0.43
13	1	1	2	3	1	8.4	2.53	35.9	0.50
14	1	2	3	1	2	6.4	0.88	7.5	0.58
15	1	3	1	2	3	4.8	1.28	1.1	0.48
16	1	1	2	3	2	10.2	2.24	36	0.45
17	1	2	3	1	3	6.0	1.14	6.6	0.60
18	1	3	1	2	1	4.4	0.57	0.8	0.56
19	2	1	2	1	3	7.0	2.99	35.1	0.51
20	2	2	3	2	1	6.4	0.85	11.0	0.58
21	2	3	1	3	2	4.6	1.20	1.2	0.48
22	2	1	2	2	3	8.4	4.43	39.2	0.50
23	2	2	3	3	1	5.8	0.37	7.9	0.58
24	2	3	1	1	2	5.8	2.00	2.7	0.30
25	2	1	3	2	1	5.8	0.77	46.5	0.59
26	2	2	1	3	2	11.2	1.74	1.3	0.45
27	2	3	2	1	3	4.6	1.84	0.6	0.38
28	2	1	3	2	2	4.4	0.67	44.6	0.63
29	2	2	1	3	3	11.6	1.91	1.5	0.44

Table 4 The L_{36} experimental design and results (continued)

No	$T_{off}(\mu s)$	$T_{on}(\mu s)$	*I(A)*	*η (Sec)*	*V(V)*	*SR (μm)*	*MRR (gr/min)*	*TWR (%)*	*WNG*
30	2	3	2	1	1	5.2	1.57	0.5	0.40
31	2	1	3	3	3	6.6	0.44	42.0	0.56
32	2	2	1	1	1	8.8	4.26	2.3	0.60
33	2	3	2	2	2	5.0	0.85	0.7	0.48
34	2	1	3	1	2	5.4	0.64	47.0	0.60
35	2	2	1	2	3	9.2	5.13	1.6	0.62
36	2	3	2	3	1	3.2	0.91	0.2	0.68

Figure 4 Interaction plot for TWR (see online version for colours)

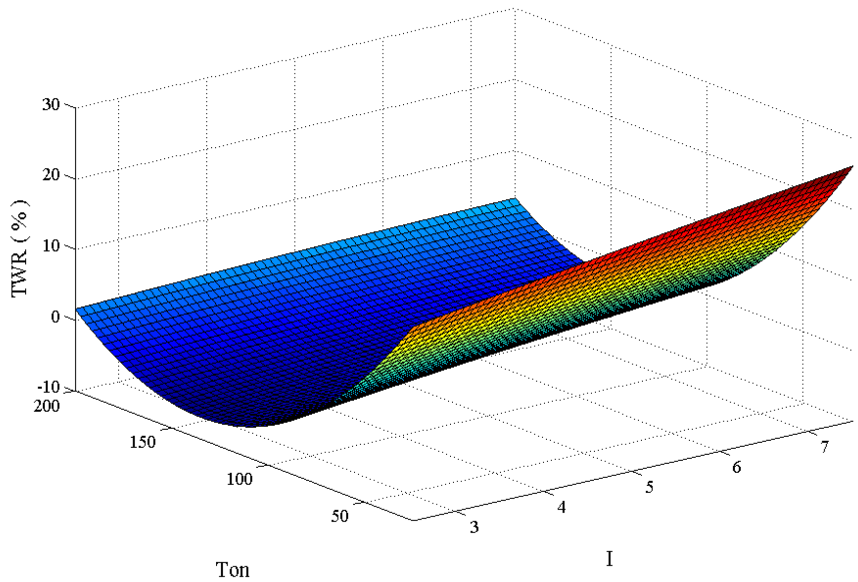


Figure 4, demonstrates the interaction effect of peak current and pulse on time on TWR (three out of the five parameters remained constant). As illustrated, within the range of 25–200 μs, by increasing the pulse on time up to about 120 μs the TWR decreases afterwards increases slightly. Similarly by increasing the peak current, within the range of 2.5–7.5A, the TWR increases.

Figure 5, demonstrates the interaction effect of peak current and pulse on time on SR. As illustrated, by increasing the peak current and pulse on time the SR increases.

Figure 6, demonstrates the interaction effect of peak current and pulse on time on MRR. As illustrated, by increasing the peak current and pulse on time the MRR increases.

Figure 5 Interaction plot for SR (see online version for colours)

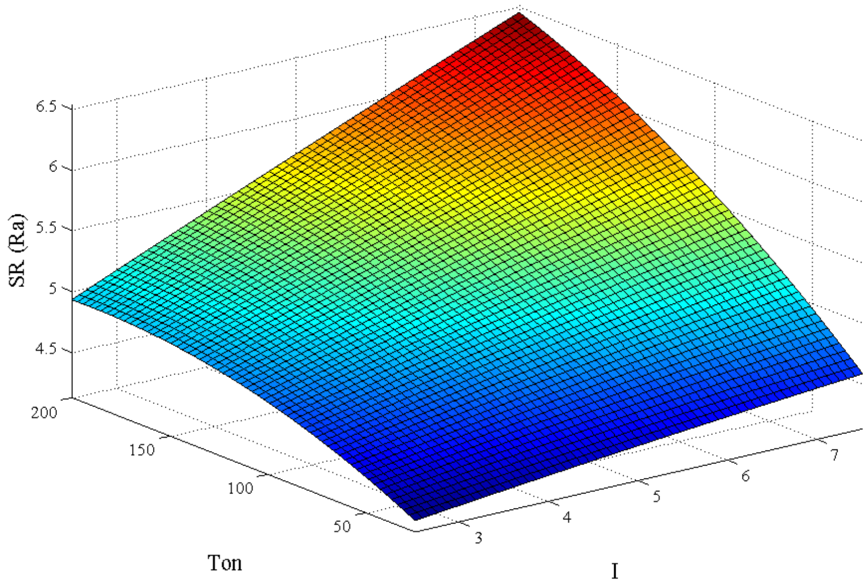
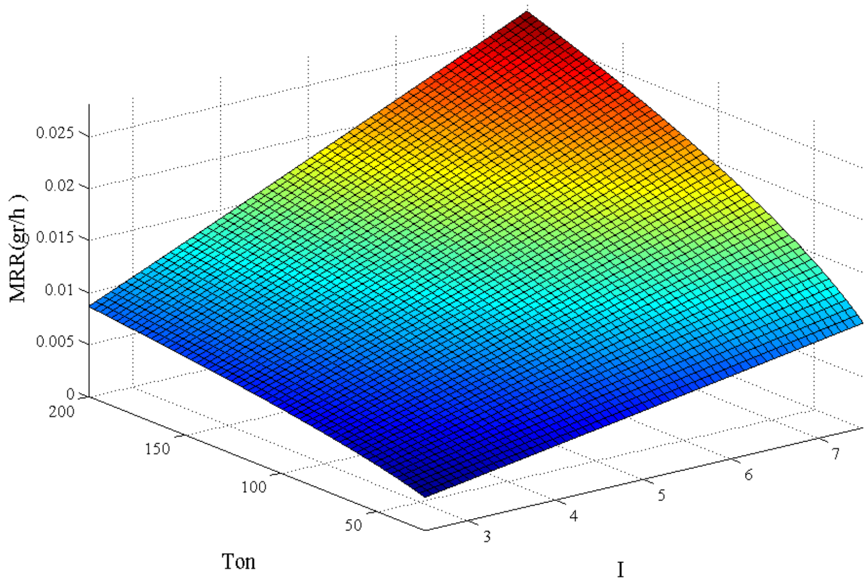


Figure 6 Interaction plot for MRR (see online version for colours)



5.2 Analysis of variance (ANOVA)

The ANOVA is used to investigate the most influential parameters to the process factor-level response. In this investigation, the experimental data are analysed using the F-test and the contribution percentage (Kolahan et al., 2012). ANOVA has been performed on the above data within the confidence limit of 95%.

Table 5 Result of ANOVA for MRR

Machining parameters	*Degree of freedom (Dof)*	*Sum of square (SS_j)*	*Adjusted (MS_j)*	*F-value*	*Contribution percentage (%)*
A	1	56.76	16.810	6.895*	2.34
B	1	428.92	35.195	14.435*	18.80
C	1	1227.77	203.516	83.471*	54.00
D	1	240.79	52.439	21.508*	10.50
E	1	35.83	36.870	15.122*	1.45
BB	1	33.42	33.415	13.705*	1.36
CC	1	123.50	123.498	50.652*	5.33
DD	1	35.15	46.875	19.225*	1.44
BC	1	26.88	26.880	11.025*	1.81
Error	26	63.39	2.438	-	-
Total	35	2272.39	-	-	-

Note: *Significant parameters, $F_{0.05,1,26} = 4.23$

Table 6 Result of ANOVA for TWR

Machining parameters	*Degree of freedom (Dof)*	*Sum of square (SS_j)*	*Adjusted (MS_j)*	*F-value*	*Contribution percentage (%)*
B	1	5479.63	146.49	9.537*	80.00
C	1	329.12	853.36	55.557*	4.54
BB	1	102.73	102.73	6.688*	1.27
BC	1	530.11	530.11	34.512*	7.45
Error	27	476.16	15.36	-	-
Total	35	6917.75	-	-	-

Note: *Significant parameters, $F_{0.05,1,26} = 4.23$

Table 7 Result of ANOVA for SR

Machining parameters	*Degree of freedom (Dof)*	*Sum of square (SS_j)*	*Adjusted (MS_j)*	*F-value*	*Contribution percentage (%)*
B	1	236.984	12.3943	17.3869*	59.25
C	1	106.886	9.1071	12.7756*	26.65
BB	1	9.925	9.9252	13.9232*	2.32
CC	1	6.735	6.7348	9.4477*	2.50
BC	1	14.633	14.6330	20.5274*	3.50
Error	26	38.242	1.4708	-	-
Total	35	398.771	-	-	-

Note: *Significant parameters, $F_{0.05,1,26} = 4.23$

According to ANOVA procedure, large F-value indicates that the variation of the process parameter makes a big change on the performance characteristics. In this study, a confidence level of 95% is selected to evaluate parameters significances. Therefore, F-values of machining parameters are compared with the appropriate values from confidence table, $F_{\alpha,v1,v2}$; where α is risk, v_1 and v_2 are degrees of freedom associated with numerator and denominator which illustrated in Tables 5, 6 and 7 (Kolahan et al., 2012; Vishwakarma et al., 2012; Vishnu et al., 2013). As the F-value of each parameter is greater than the $F_{\alpha,v1,v2}$ observed from the table means the corresponding parameter is influential in the process characteristic.

ANOVA results may provide the percent contributions of each parameter (Roy, 2010).

$$P_i(\%) = \frac{SS_i - (DOF_i \times MS_{error})}{Total\ Sum\ of\ Squre} \tag{4}$$

In the above formula according to the ANOVA results, P_i is contribution percentage, SS_i is sum of square, DOF_i is degree of freedom of i^{th} factor, and MS_{error} is mean sum of square of error (Roy, 2010).

Figure 7 The effect of machining parameters on the MRR (see online version for colours)

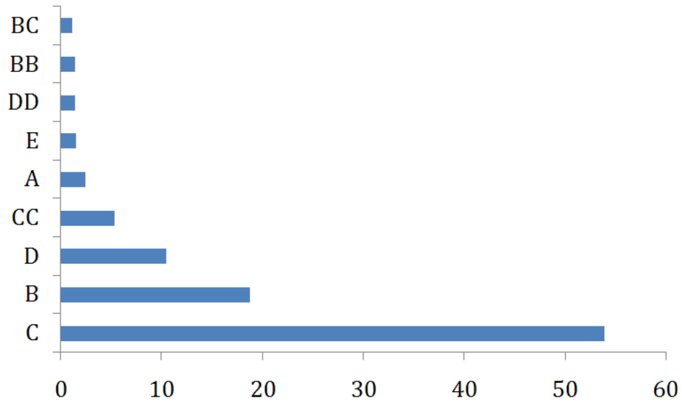


Figure 8 The effect of machining parameters on the SR (see online version for colours)

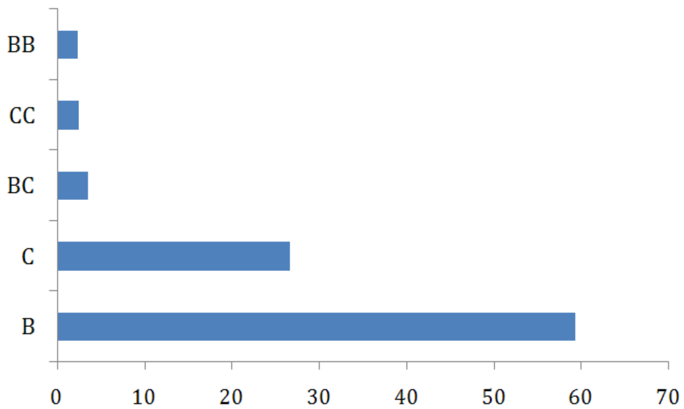
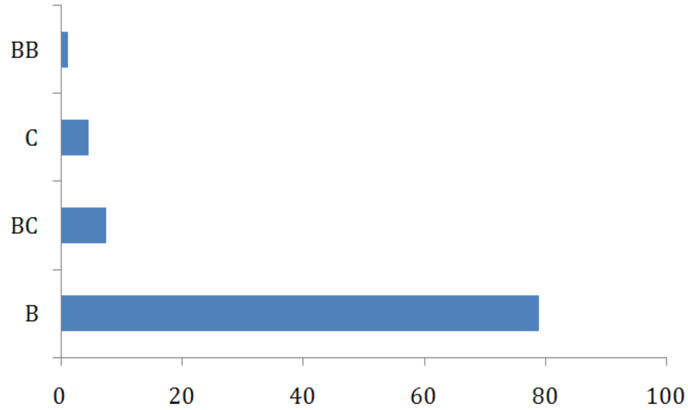


Figure 9 The effect of machining parameters on the TWR (see online version for colours)



The percent contributions of the EDM parameters on MRR, SR and TWR are shown in Figures 7, 8 and 9 respectively.

According to Figure 7, peak current is the major factor affecting the MRR with 54% contribution. It is followed by pulse on time and voltage with 18.8% and 10.5% respectively. The remaining (4%) effects are due to noise factors or uncontrollable parameters.

Moreover, pulse on time is the major factor affecting the SR (Figure 8) with 59.25% contribution, followed by peak current with 26.65% effect. The remaining parameters have little effects on this output. The main process parameter affecting TWR is pulse on time with 80% contribution (Figure 9).

In EDM, multi-criteria modelling is needed to simultaneously achieve high MRR, low TWR and SR. Multiple output responses can be transformed into a single measure by using weighted normalised grades (WNGs) method (Jung and Kwon, 2008). This is done by normalising the values of process responses and then averaging these normalised 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 normalised using the following formula:

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

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

$$Z_i = \frac{\left(\max\left(y_i, i = 1, 2 \ldots, n\right) - y_i\right)}{\left(\max(y_i, i = 1, 2 \ldots, n) - \min\left(y_i, i = 1, 2 \ldots, n\right)\right)} \tag{6}$$

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 normalised value for y_i. Now the multiple output responses can be transformed into a single WNG through the following equation:

$$G_i = \sum_{k=1}^{p} \beta_k . Z_i \tag{7}$$

where p is the number of performance measures, β_k is the normalised 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. Corresponding measured outputs are recorded in Table 4. The last column of Table 4 shows the calculated WNG for each test.

6 Model development

6.1 SA algorithm

SA algorithm is an optimisation process whose operation is strongly reminiscent of the physical annealing of crystalline compounds such as metals and metallic alloys (Kirkpatrick and Vecchi, 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 minimised. While minimising efforts are made to avoid local minima and to achieve global minima. The lowest energy level gives the optimised value of EDM parameters. In recent years, the SA algorithm has emerged as a leading tool for large-scale combinational optimisation 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:

$$p = \begin{cases} 1 & \text{if} \quad \Delta < 0 \\ e^{-\Delta/t} & \text{if} \quad \Delta \geq 0 \end{cases} \tag{8}$$

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 [21]. In our problem the following method to reduce the temperature has been used:

$$T_{i+1} = cT_i \quad i = 0, 1, \ldots \quad \text{and} \quad 0.9 \leq c < 1 \tag{9}$$

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.

SA algorithm has diverse applications including improving the performance of other artificial intelligence techniques and determining the optimal set of process parameters (Yang et al., 2009; Markopoulos et al., 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 in order to maximise the MRR and minimise the TWR and SR simultaneously via increasing WNG.

6.2 The OBPNN

The first model of the ANN was given by McCulloch and Pitts (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 (Kohonen, 1987).

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

Traditional modelling 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 nonlinear systems. The basis of NN modelling is to capture the underlying trend of the dataset presented to it, in the form of a complex nonlinear relationship between the input parameters and the output variable (Debabrata et al., 2007). Learning, generalisation, and parallel processing are important advantages of ANN. These characteristics of the ANNs make them suitable for EDM process modelling.

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 (Debabrata et al., 2007). However, the computational effort needed for modelling a system increases substantially when more parameters and more complicated architectures are considered. The BPNN are 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 et al., 1998).

$$F_{i,j} = \frac{1}{1+\exp^{-P\left(W_{i,j-1},O_{i,j-1}\right)}} \tag{10}$$

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

$$P\left(W_{i,j-1}, O_{i,j-1}\right) = \sum_{j=1}^{m}\sum_{i=1}^{n} W_{i,j-1}.O_{i,j-1} \tag{11}$$

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

One of the most important tasks in ANN modelling 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, in order 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:

$$MSE = \frac{1}{p}\sum_{k=1}^{p}\left(Y_k - y_k\right)^2 \tag{12}$$

Learning MSE and the generalisation MSE, detect the two main characteristics of 'learning' and 'generalisation' of ANN. The effectiveness of developed net depends on these features.

6.3 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 first to the fourth hidden layers. The training of a neural network implies finding desired net's architecture and weights that minimise 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 and Hanumantha, 2009). Hence, an error between the desired output (Y_k) and predicted output (y_k) of the neural network is computed. So the modelling authority of the net ($M(net)$) can be given by:

$$M(net) = \alpha.\frac{1}{p_0}\sum_{r=1}^{p_0}\left(Y_r - y_r\right)^2 + \beta.\frac{1}{q_0}\sum_{s=1}^{q_0}\left(Y_s - y_s\right)^2 \tag{13}$$

where α and β are coefficients which determine the relative importance of learning and generalisation 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 optimised BPNN construction. In backward phase of BP training, to minimise the error between the desired and actual outputs, the gradient descent method with a momentum coefficient, is used (Gao et al., 2008). The weights are updated using the following rule:

$$\Delta w_{i,j}{}^{(n+1)} = -\Gamma.\frac{\partial(M(net))}{\partial w_{i,j}{}^{n}} + \omega.w_{i,j}{}^{n} \tag{14}$$

where Γ is learning rate and $\Delta w_{i,j}^{n}$ is change of neuron's weight at previous step and ω is momentum.

In first step initial parameters and an initial net structure is configured. Based on the fitness 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*(*ne*t)) are 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 ones has been derivate. This iterative step is continued until the algorithm has been converged after the predetermined number of iterations.

7 Results and discussion

7.1 OBPNN prediction results

Figure 10 shows convergence of SA for objective function (*M*(*net*)). 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: 5–2–7–3–1). The linear regression analysis is conducted to compute the correlation coefficient (R^2_{adj}) between actual experimental and predicted WNGs. The correlation coefficients at the train and testing stage are 0.998 and 0.987 respectively. The related fitness function for trained model is 1.53 (*M*(OBPNN)= 1.53), MSE 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. It can be seen that developed model is adequate and optimum set of machining can be found based on it. In Figure 11 comparison between experimental and verification of OBPNN for WNGs is shown. Predicted WNGs by OBPNN pursuit the experimental results closely.

Figure 10 Convergence of SA in OBPNN training (see online version for colours)

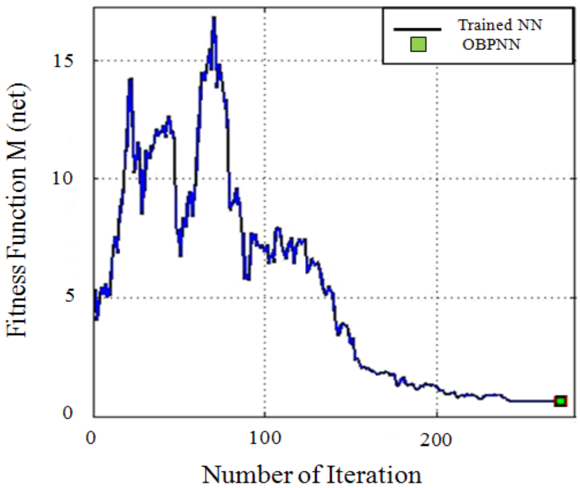
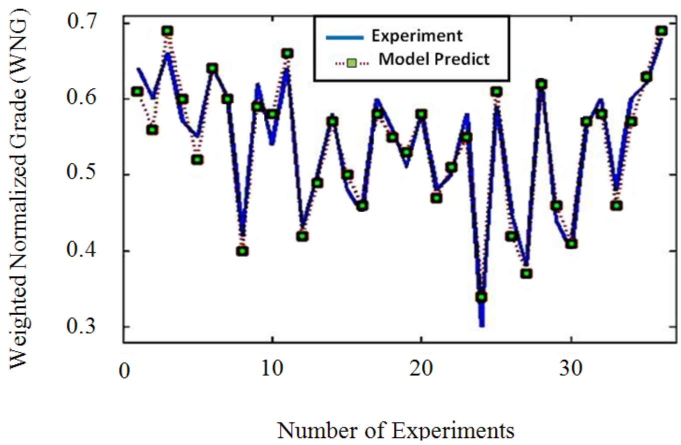


Figure 11 Comparison between experimental and verification of WNGS (see online version for colours)



7.2 EDM process input parameters optimisation

After successful process modelling, another algorithm based on SA has been developed and optimal values 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 optimised process parameters and the obtained experimental responses were compared with the initial parameters design. The results which are presented in Table 8 show that the hybrid model can improve machining performances. As observed in Table 8, optimal machining set is: Toff = 15 µs, Ton = 25 µs, I=3 A, η = 0.8 sec and V = 65 v. For optimal machining set rather than initial machining design, MRR increases form 0.35 g/min to 3.43 g/min; TWR is reduced from 0.04 to 0.02 mm; and SR 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 8 Optimal EDM parameters and related responses

Set	*WNG*	*SR (µm)*	*TWR (%)*	*MRR (g/min)*	*Process parameter*
Initial process design	0.64	3.60	0.04	0.35	T_{off} (µs) = 10, T_{on} (µs) = 25 I (A) =2.5, η (sec) = 0.4, V (v) = 50
Optimal process design	0.73	3.16	0.02	3.43	T_{off} (µs) = 15, T_{on} (µs) = 25 I (A) = 3, η (sec) = 0.8, V (v) = 65

8 Conclusions

In this study, hybrid modelling and optimisation of EDM process on AISI2312 hot worked steel have been implemented. Experimental data for process modeling obtained from conducted experiments by Taguchi method, a systematic tool for DOE. The tests are conducted under varying peak current (I), voltage (V), pulse on time (T_{on}), pulse off time (T_{off}) and duty factor (η). The process output characteristics include SR, TWR and MRR. The significance of the input parameters on output characteristics of the EDM process was evaluated using analysis of variance (ANOVA) method. According to the findings in this regard, peak current is the major factor affecting the MRR followed by pulse on time and voltage respectively. Similarly, pulse on time is the most effective factor affecting the SR followed by peak current and the main process parameter affecting TWR is pulse on time. The objective of this research is to find a combination of process parameters to minimise TWR and SR and maximise MRR simultaneously. Multiple process output measure was transformed to the single measure namely WNG to achieve the objective. The OBPNN was developed to establish accurate model of process multiple performance characteristics. The merit of this study was that the optimal net's architecture (number of neurons and hidden layers) of OBPNN has been specified using SA algorithm, whereas in the most studies they have been selected through trial and error. Then, the SA procedure has been used to find the optimal set of EDM process parameters in order to increase WNG. Correlation coefficient (R_{adj}^2) and 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 realised comparatively with the other commonly used modelling methods. From the present analysis it is evident that the proposed hybrid model will be very beneficial in multi-objective process modelling and optimisation.

References

Asoka, P., Kumar, R.R., Jeyapaul, R. and Santhi, M. (2008) 'Development of multi-objective optimization models for electrochemical machining process', *Journal of Materials Manufacturing Technology*, Vol. 39, No. 2, pp.55–63.

Assarzadeh, S. and Ghoreishi, M. (2008) 'Neural-network-based modeling and optimization of the electro-discharge machining process', *Journal of Advance Manufacturing Technology*, Vol. 39, No. 5, pp.488–500.

Bhavesh, A., Patel, D.S. and 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*, Vol. 1, No. 12, pp.858–869.

Debabrata, M., Surjya, K. and 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*, Vol. 186, No. 1, pp.54–162.

Gao, Q., Zhang, Q., Su, S. and Zhang, J. (2008) 'Parameter optimization model in electrical discharge machining process', *Journal of Zhejiang University Science*, Vol. 9, No. 1, pp.104–108.

Gopalsamy, B.M., Mondal, B. and Ghosh, S. (2009) 'Taguchi method and ANOVA: an approach for process parameters optimization of hard machining while machining hardened steel', *Journal of Scientific and Industrial Research*, Vol. 68, No. 8, pp.686–695.

Ho, K.H. and Newman, S.T. (2003) 'State of the art electrical discharge machining (EDM)', *International Journal of Machine Tools and Manufacture*, Vol. 43, No. 3, pp.1287–1300.

Jung, H. and Kwon, W.T. (2008) 'Optimization of EDM process for multiple performance characteristics using Taguchi method and Grey relational analysis', *Journal of Advance Manufacturing Technology*, Vol. 37, No. 5, pp.523–533.

Kirkpatrick, G.J. and Vecchi, D.D. (1983) 'Optimization by simulated annealing', *Science*, Vol. 220, No. 4598, pp.671–680.

Kohonen, T. (1987) 'Adaptive and associative, and self-organization functions in neural computing', *Applied Optics*, Vol. 26, pp.4910–4918.

Kolahan, F., Golmezerji, R. and Azadi Moghaddam, M. (2012) 'Multi objective optimization of turning process using grey relational analysis and simulated annealing algorithm', *Applied Mechanics and Materials*, Vol. 110, No. 5, pp.2926–2932.

Krishna, M.R. and 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*, Vol. 209, No. 3, pp.1512–1520.

Mahdavi Nejad, 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*, Vol. 2, No. 1, pp.669–675.

Markopoulos, A.P., Manolakos, D.E. and Vaxevanidis, N.M. (2008) 'Artificial neural network models for the prediction of surface roughness in electrical discharge machining', *Journal of Intelligence Manufacturing*, Vol. 19, No. 3, pp.283–292.

McCulloch, W. and Pitts, W. (1943) 'A logical calculus of the ideas immanent in nervous activity', *Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp.115–133.

Mohana, G.K., Hanumantha, D. and Rajurkar, K. (2009) 'Hybrid modeling and optimization of surface produced by electric discharge machining using artificial neural network and genetic algorithm', *Journal of Engineering and Applied Sciences*, Vol. 5, No. 3, pp.72–81.

Payal, H.S., Choudhary, R. and Singh, S. (2008) 'Analysis of electro discharge machined surface of EN-31 tool steel', *Journal of Scientific and Industrial Research*, Vol. 67, No. 12, pp.1072–1077.

Petropoulos, G., Vaxevanidis, N.M. and Pandazaras, C. (2004) 'Modeling of surface finish in electro-discharge machining based on statistical multi parameter analysis', *Journal of Material processing Technology*, Vol. 155, No. 5, pp.1247–1251.

Pushpendra, S., Bharti, S., Maheshwari, C. and Sharma, C. (2012) 'Multi-objective optimization of electric-discharge machining process using controlled elitist NSGA-II', *Journal of Mechanical Science and Technology*, Vol. 26, No. 6, pp.1875–1883.

Roy, R.K. (2010) 'A primer on the Taguchi method', *Society of Manufacturing Engineers*, Vol. 5, No. 2, pp.25–41.

Sanghani, C.R. and 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*, Vol. 4, No. 1, pp.433–450.

Sexton, R., Allidae, B., Dorsey, R.E. and Johnson, J.D. (1998) 'Global optimization for artificial neural networks: a tabu search application', *Journal of Operational Research*, Vol. 106, No. 3, pp.570–584.

Vishnu, D., Asal, R.I., Patel, A. and Choudhary, B. (2013) 'Optimization of process parameters of EDM using ANOVA method', *International Journal of Engineering Research and Applications*, Vol. 3, No. 2, pp.1119–1125.

Vishwakarma, M., Parashar, V. and Khare, V.K. (2012) 'Regression analysis and optimization of material removal rate on electric discharge machine for EN-19 alloy steel', *International Journal of Scientific and Research Publications*, Vol. 2, No. 11, pp.145–153.

Yanamandala, R.C., Yuvaraj, K. and 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*, Vol. 12, No. 2, pp.730–739.

Yang, S.H. and Srinivas, J., Mohan, S., Lee, D.M. and Balaji, S. (2009) 'Optimization of electric discharge machining using simulated annealing', *Journal of Materials Processing Technology*, Vol. 209, No. 9, pp.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*, Vol. 187, No. 12, pp.711–714.