Journal of the Brazilian Society of Mechanical Sciences and Engineering (2020) 42:73
https://doi.org/10.1007/s40430-019-2149-1

TECHNICAL PAPER

Optimization of backpropagation neural network-based models in EDM process using particle swarm optimization and simulated annealing algorithms

Ali Saffaran[1] · **Masoud Azadi Moghaddam**[2] · **Farhad Kolahan**[2]

Received: 12 July 2019 / Accepted: 9 December 2019

Abstract

In the present study, artificial neural network (ANN) along with heuristic algorithms, namely particle swarm optimization (PSO) and simulated annealing (SA), has been employed to carry out the modeling and optimization procedure of electrical discharge machining (EDM) process on AISI2312 hot worked steel parts. Surface roughness (SR), tool wear rate (TWR) and material removal rate (MRR) are the process quality measures considered as process output characteristics. Determination of a process variables (pulse on and off time, current, voltage and duty factor) combination to minimize TWR and SR and maximize MRR independently (as single objective) and also simultaneously (as multi-criteria) optimization is the main objective of this study. The experimental data are gathered using Taguchi L_{36} orthogonal array based on design of experiments approach. Next, the output measures are used to develop the ANN model. Furthermore, the architecture of the ANN has been modified using PSO algorithm. At the last step, in order to determine the best set of process output variables values for a desired set of process quality measures, the developed ANN model is embedded into proposed heuristic algorithms (SA and PSO) with which their derived results have been compared. It is evident that the proposed optimization procedure is quite efficient in modeling (with less than 1% error) and optimization (less than 4 and 7 percent error for single- and multi-objective optimizations, respectively) of EDM process variables.

Keywords Electrical discharge machining · Taguchi technique · Design of experiments · Artificial neural network · Simulated annealing algorithm · Particle swarm optimization algorithm

1 Introduction

Electrical discharge machining (EDM) is the most extensively and successfully applied process for machining of difficult-to-cut alloys (such as hot worked and super alloys) among the several non-conventional ones. In EDM process, electrical energy in the form of a series of discrete electrical discharges occurring between the tool electrode and workpiece electrode (both submerged in a dielectric fluid) through which a channel of plasma (Fig. 1) is generated. Due to the electrical energy discharges, a considerable amount of heat melts and evaporates the material at the surface of the workpiece electrode. This exclusive feature of using thermal energy to machine electrically conductive parts is considered as the main merits of EDM in the machining of molds, dies, aerospace and surgical components [1].

Like other manufacturing processes, in EDM process, the proper process variable settings are a crucial feature to reduce production cost and improve product quality. There are several process input variables in EDM process, out of which discharge voltage (V), peak current (I), pulse on time (T_{on}), pulse off time (T_{off}) and duty factor (η) are the most influential ones. Material removal rate (MRR), tool wear rate (TWR) and surface roughness (SR) are their corresponding quality measures considered as the process responses. Due to the complex nature of the process where several different and sometimes contradictory objectives must be simultaneously considered, optimizing any of these measures alone has a limited value in real practice. As a result,

Technical Editor: Lincoln Cardoso Brandao.

✉ Masoud Azadi Moghaddam
masoudazadi888@gmail.com

1 Department of Mechanical Engineering, Khayyam University of Mashhad, Mashhad, Iran

2 Department of Mechanical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

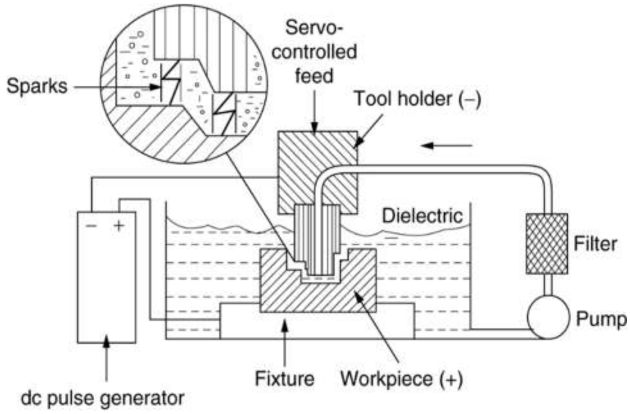

Fig. 1 Graphic illustration of electrical discharge machining process [2]

multi-criteria process optimization has received more attention by researchers in this field of study [2]. The successful implementations of optimization methods depend on the proper establishment of relationships between process input variables and the corresponding performance characteristics. Nonetheless, establishing such relationships is difficult due to the stochastic nature of EDM process which has got the attention of scholars [3, 4].

Effect of wire-cut electric discharge machining (WEDM) process variables (pulse off time, pulse on time and servo voltage) on the most important process responses (MRR and SR) on Inconel 625 alloy has been investigated using Taguchi method by Subrahmanyam and Nancharaiah [5]. To gather the required data, orthogonal array Taguchi method has been used. Furthermore, for optimizing purpose, ANOVA has been used so that maximum MRR and minimum SR were obtained. To achieve optimum MRR, reduced SR and TWR for LM6-alumina stir casted metal matrix composites (MMC) and gray relation analysis (GRA) have been applied to design and optimize the EDM multiple performance characteristics by Palanisamy et al. [6]. The process input variables such as discharge current, pulse on time and off time have been considered to be optimized. The study revealed that discharge current was the most influential variable that affects the SR and MRR.

GRA has been used to deal with the multi-objective optimization of EDM process of Al7075 alloy considering MRR and SR by Tharian et al. [7]. Moreover, single-objective optimization has been performed using Taguchi technique. For performing multi-objective optimization purpose, the problem has been solved by GRA. Experimental design matrix required for conducting tests has been made according to the L_9 orthogonal array Taguchi method. To find out the significant process variables affecting the single-objective and multi-objective problems, ANOVA has been performed. To establish the relationships between process input variables and output responses, mathematical modeling has also been performed. The effect of EDM process input variables (workpiece electrical conductivity, gap voltage, gap current, pulse on time and off time) on the process responses (MRR and TWR) has been investigated. Experiments have been designed and performed as per the orthogonal array Taguchi method (L_{18} $(6^1 \times 3^4)$). For optimizing the conflicting responses, Taguchi's approach along with utility concept have been employed. Based on the obtained results, the overall utility was significantly affected by gap voltage, gap current, pulse off and on time. Therefore, the optimal values of MRR and TWR were 9.157 mm^3/min and 0.128 mm^3/min, respectively [8].

For optimization of multiple responses of EDM process, fuzzy method coupled with Taguchi has been used by Nagaraju et al. [9]. AISI 304 stainless steel has been considered as the specimen material. The effect of process variables (discharge current and voltage, pulse on time, and inter electrode gap) on the most important responses (MRR, TWR and Ra) has been investigated. To design the experimental matrix, L_9 orthogonal array has been used. Fuzzy logic known as multi-performance characteristic index (MPCI) has been used to convert the multiple responses into a single characteristic index. Finally, Taguchi has been used to optimize the MPCIs. Multi-objective optimization of EDM on Mg–RE–Zn–Zr alloy has been carried out using the novel meta-heuristic algorithm (passing vehicle search (PVS)) by Parsana et al. [10]. Pulse on and off time and peak current have been considered as the process input variables. To formulate a mathematical model for MRR, TWR and SR, response surface method (RSM) has been used. Multi-objective PVS calculated optimal solutions for different weights to generate 2D and surface Pareto fronts have been performed using the weighted sum method. These Pareto fronts were evaluated for performance determination of PVS using novel and established metrics such as spacing, spreading, hyper volume and pure diversity.

In recent years, artificial neural networks (ANNs) have demonstrated sufficient potential in modeling of complicated nonlinear systems such as the EDM process. There are many types of ANNs which vary in architecture, implementation of transfer functions and learning strategy. In view of their universal approximation property, backpropagation neural network (BPNN) has received considerable attention. The architectural factors of BPNN to be determined in advance for the modeling of the process under consideration are the feature subsets, the number of hidden layers and processing elements number in hidden layers [7].

Along this line, Naveen et al. [11] have employed orthogonal array Taguchi method to get the optimized MRR and SR of WEDM on Inconel 750 considering process input variables (pulse on, pulse off, voltage and current). To model the MRR and SR, an ANN model has been developed. Finally, the model was optimized using PSO algorithm. NN models have been established to predict cutting speed, SR and wire consumption for effective modeling of WEDM by Sen et al. [12]. Fuzzy logic has been incorporated to convert the multi-objective problem into a single-objective one. Teaching learning-based optimization and genetic algorithm (GA) technique have been used for optimizations. Validation tests verified that teaching learning-based optimization is the more appropriate method. To reach maximum MRR, better SR and minimum dimensional error during WEDM of Al7075-TiB_2 in situ composite, prediction and comparison of machining performances have been focused. Pulse on and off time, current and bed speed have been considered as the process input variables. Orthogonal array Taguchi technique has been used for designing the experimental matrix required. An ANN model has been developed in order to predict the process responses. Based on the results predicted, machining responses are in good agreement with experimental values [13].

To estimate SR, accuracy, volumetric MRR and TWR based on the WEDM variables including pulse on off time, current and bed speed, ANN has been employed. Orthogonal array Taguchi method has been used to determine the experimental matrix required. To establish the relations between process input and output characteristics, ANN has been used. Based on the results, the proposed model is quite efficient in predicting the process responses [14]. The application of assisted electrical discharge machining (AAEDM) of D3 steel has been investigated by Singh et al. [15]. A mathematical model was actuated to realize the SR by utilizing dimensional analysis hypothesis. The experimental and foreseen assessments of SR during the procedure, acquired by Buckingham pi theorem, ANN and ANFIS, were observed to be as per one another. Be that as it may, the ANFIS strategy demonstrated to be all the more fitting to the EDM output when contrasted with the ANN and the semi-empirical model. An ANN model for predicting the SR has been proposed by Khan et al. [16]. DOE approach has been used for determining the experimental matrix required for training and testing. The result showed that the proposed ANN model can predict the SR effectively. Low discharge energy level ended in smaller craters and micro-cracks producing a suitable structure of the surface. This approach helps in economic EDM machining. A review has been carried out by Venkata and Kalyankar [17], in which various modern machining processes (including electric discharge machining, ultrasonic machining, abrasive jet machining, electro chemical machining, micro-machining and laser beam machining) have been optimized using different approaches.

There is a great deal of papers in which modeling and optimization of EDM process have been considered. In this paper, ANN has been used to model the EDM process responses (MRR, SR and TWR). Moreover, a heuristic algorithm (PSO) has been used for determination of ANN architecture (including number of neurons/nodes and hidden layers) and optimization purposes. In different papers, the architecture of the ANNs has been determined based on trial and error method. In the proposed manuscript, PSO algorithm has been used twice (for ANN architecture determination and single- and multi-objective optimization purposes). Furthermore, the performance of PSO algorithm has been checked using SA algorithm. Thus, this study recommends a hybrid method composed of BPNN and heuristic algorithms (SA and PSO) to undertake the single-objective and multi-objective modeling and optimization in EDM process. The purpose of this study is to present an efficient and integrated approach for the determination of appropriate variables setting yielding the objective of maximum MRR and minimum SR and TWR independently and also simultaneously. Then, the results derived from the heuristic algorithms have been compared with. To validate the results, some confirmation tests have been conducted. To the best of our knowledge, there is no published work in which study of EDM process of AISI2312 hot worked steel parts through the proposed method has been considered. First, to gather the experimental data required, L_{36} orthogonal array Taguchi (OA-Taguchi) design matrix based on which empirical tests have been carried out, has been used. Then, the process was modeled using a BPNN. Besides, the architecture of the PBNN has been determined and modified using PSO algorithm. Finally, the BPNN models have been embedded into SA and PSO optimization algorithms to determine the best set of process variables in order to achieve maximum MRR and minimum SR and TWR as single- and multi-criteria optimization. Finally, the performance of the proposed integrated PSO-BPNN has been validated through experimental tests. Moreover, the article concludes with the confirmation of the computational and experimental results and a summary of the major findings.

2 Experimental setup and material used

AISI2312 hot worked steel specimens were used on which the experiments have been conducted. This alloy is widely used for such parts as plastic molds due to its high erosion and heat resistance. However, because of its controlled sulfur content, AISI2312 is one of the most difficult-to-cut steel alloys. This calls for more research on employing non-traditional machining. Specimens have been cut out of a plate with 10 mm thickness into 40×20 mm dimension.

Fig. 2 The Azarakhsh-304H EDM machine used for conducting experiments

Table 1 Detailed technical specifications of the die sinking machine 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

In EDM process, a wide variety of materials such as brass, copper and tungsten alloys as well as graphite may be used as tool electrode. The applications of brass and tungsten are limited to certain materials. Due to extreme high melting point, graphite rate of erosion is less in comparison with copper. On the contrary, very fine surfaces could be achieved using copper electrode. Furthermore, the machinability of copper is much better than that of graphite [4]. Therefore, based on these facts and the literature survey in this regard, copper electrodes, with 99% purity and 8.98 g/cm^3 density, were opted as electrode tools in our experiments.

An Azarakhsh-304H die sinking machine, shown in Fig. 2, has been employed to carry out the experiments. In this machine, the *X* and *Y* axis are manually controlled and the *Z*-axis is servo controlled.

The technical specification of the EDM machine used to carry out the experiments is shown in Table 1. It should be noted that tests have been done in random to increase the accuracy of the results. Moreover, tool electrodes were replaced after each test run. The machining time for all test runs was set at 45 min. Besides, pure kerosene has been used as dielectric.

3 Design of experiments

Design of experiments (DOE) is extensively used for gaining knowledge of the existing processes and/or optimizing the processes quality characteristics. In carrying out DOE, to observe changes in the output characteristics, changes are made to the input variables of the system. Modeling and optimization of the process could be used by the information gathered from correctly planned experiments. Full factorial (FF) designs are most popular strategies such as response surface methodology (RSM), center composite design (CCD) and orthogonal array (OA) Taguchi design of experiment. Detailed information about DOE approach and its various applications may readily be found in the related literatures [18, 19].

Taguchi procedure has been extensively used in various engineering requests among various DOE strategies due to its distinct advantages. With fewer number of experiments (and hence lesser cost), Taguchi can provide much useful information which, in turn, can be used for process modeling and analysis.

In this study, attempt to find optimum variables of EDM process on AISI2312 hot worked steel parts in order to minimize TWR and SR and maximize MRR using Taguchi matrix and BPNN integrated with heuristic algorithms (PSO and SA) has been made.

Firstly, to determine the stable domain of the input variables and also the feasible levels of them some preliminary tests were carried out [14]. Peak current (I), voltage (V), pulse off time (T_{off}), pulse on time (T_{on}), and duty factor (η), based on literature surveys, preliminary test results (based on the screening method using DOE) and working characteristics of the EDM process were selected as the independent input variables. During these experiments, by altering the values of the input variables to different levels, stable states of the machining conditions have also been identified. Initial experiments were conducted for the wide range of pulse on time, discharge current and gap voltage. Reasonable range of peak current was attained for 6-30A. Below 6A, MRR was very low and beyond 18A, MRR was good but SR was very poor. Similar observations were made for range of pulse on and off time, gap voltage and duty factor. Thus, L_{36} ($2^1 \times 3^4$) design of experiments matrix has been used to conduct the experiments required. A certain number of levels for some of the process variables may also be dictated such as limitations of test equipment. The die sinking EDM machine used for experiments in this study had only two settings for pulse off time $- T_{\text{off}}$ (10 and 75 µs). Out of five, one factor has 2 levels and the rest of the factors have 3 levels each (Table 2). Subsequently, this study has been carried out 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 2 EDM Machining variables and their viable intervals and levels

variables	Symbol	Range	Level 1	Level 2	Level 3
Peak current (A)	I	6–30	6	18	30
Voltage (V)	V	50–60	50	55	60
Pulse on time (μs)	T_{on}	25–200	25	100	200
Pulse off time (μs)	T_{off}	10–75	10	75	–
Duty factor (S)	η	0.4–1.6	0.4	1.0	1.6

Fig. 3 The electronic balance and surface roughness tester used

4 Evaluation of the process quality characteristics

In this survey, MRR, SR and TWR are used to evaluate EDM process of AISI2312 hot worked steel parts [20].

MRR is a measure of machining speed and is expressed as the specimen removal weight in a machining time in minute. In this regard, the specimens have been weighed twice, before the each test runs begin and after the test completed. The difference value between them has been considered as the material removed from the specimen during the process. The value of ratio of the material removed from the specimen to machining time has been considered as MRR. TWR, usually expressed as a percentage, is defined by the ratio of the tool wear weight to the specimen removal weight. Surface quality is usually measured in terms of surface roughness (SR). The average roughness (Ra) is the area between the roughness profile and its mean line. After machining process, the surface finish of each sample has been measured using an automatic digital surface roughness tester (Mitutoyo Model). Also, to measure the MRR and TWR, an A&D electronic balance (with 0.01gr accuracy) has been used (Fig. 3).

Table 3 shows the L_{36} experimental design matrix based on the Taguchi procedure and their corresponding measured results. The first five columns are the process input variables (including voltage, duty factor, current, pulse on and off time, respectively). The second three columns represent the process responses (material removal rate, tool wear rate and surface roughness).

5 Backpropagation neural network (BPNN)

For establishing relations between process input variables and output responses, different procedures have been proposed among which ANNs are extensively used in this regard. ANNs are reminiscent of the creature's nervous system, which is a highly nonlinear, complex and processing system. Learning, generalization and parallel processing are significant merits of ANNs that make them appropriate for modeling different processes such as EDM process [21].

ANNs are built by connecting processing units, named neurons or nodes. Each of the input (X_i) is associated with some weight (W_i) which takes a share of the input to the neuron for processing. The neuron combines the inputs ($X_i \times W_i$) and produces net input which in turn is transformed into output with the help of transfer function/activation function [21].

Many scholars have proposed that multi-layered networks are capable of computing a wider ranges of nonlinear functions than the single-layered networks [15–17]. However, the computational effort required for modeling purpose increases substantially. The backpropagation neural networks (BPNNs) are found most suitable for dealing with such large learning problems. This type of neural network is known as a supervised network due to a desired process quality measures in order to learn. A BPNN consists of multiple layers of neuron in a directed scheme, with each layer connected to the next one. Except for the input nodes, each neuron is a processing element with a nonlinear activation function defined in Eq. (1) [22]:

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

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

$$P\left(W_{i,j-1}, O_{i,j-1}\right) = \sum_{j=1}^{m} \sum_{i=1}^{n} \left(W_{i,j-1} \cdot O_{i,j-1}\right) \tag{2}$$

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

In this research, for modeling of the EDM process, the total number of input nodes is five (pulse on and off time, current, voltage and duty factor). The best architecture for modeling (the number of hidden layers and the number of nodes in each hidden layer) has been chosen using PSO algorithm. Besides, the transfer function of each processing element is identified and the next network is trained to interrelate the process variables to responses. The outputs of trained model are MRR, TWR and SR. As the number of responses for the trained model is three, linear transfer functions have dealt with outputs of nodes in the last hidden layer to calculate the network outcomes ($y_k^{(\mathrm{net})}$) as Eq. (3):

Table 3 The L_{36} orthogonal array experimental design matrix and results

No	V (V)	η (s)	I (A)	T_{on} (μs)	T_{off} (μs)	MRR (gr/min)	TWR (%)	SR (μm)
1	1	1	1	1	1	0.0078	11.4	3.9
2	2	2	2	2	1	0.0676	2.6	7.1
3	3	3	3	3	1	0.1487	0.6	13.5
4	1	1	1	1	1	0.0073	9.0	3.2
5	2	2	2	2	1	0.0462	3.3	6.9
6	3	3	3	3	1	0.1520	0.4	12.7
7	3	2	1	1	1	0.0100	6.7	3.8
8	1	3	2	2	1	0.1227	2.7	8.4
9	2	1	3	3	1	0.0629	0.7	12.5
10	2	3	1	1	1	0.0124	5.3	4.4
11	3	1	2	2	1	0.0347	3.8	7.6
12	1	2	3	3	1	0.2364	0.7	13.7
13	1	3	2	1	1	0.0378	35.9	4.8
14	2	1	3	2	1	0.0562	7.5	8.1
15	3	2	1	3	1	0.0196	1.1	6.2
16	2	3	2	1	1	0.0284	36.0	4.6
17	3	1	3	2	1	0.0498	6.6	10.0
18	1	2	1	3	1	0.0253	0.8	5.8
19	3	1	2	1	2	0.0127	35.1	4.9
20	1	2	3	2	2	0.0664	11.0	7.0
21	2	3	1	3	2	0.0189	1.2	6.5
22	3	2	2	1	2	0.0267	39.2	4.8
23	1	3	3	2	2	0.0984	7.9	8.7
24	2	1	1	3	2	0.0082	2.7	6.1
25	1	2	3	1	2	0.0444	46.5	5.5
26	2	3	1	2	2	0.0171	1.3	5.8
27	3	1	2	3	2	0.0387	0.6	11.1
28	2	2	3	1	2	0.0409	44.6	4.9
29	3	3	1	2	2	0.0149	1.5	4.6
30	1	1	2	3	2	0.0424	0.5	11.6
31	3	3	3	1	2	0.0349	42.0	4.9
32	1	1	1	2	2	0.0098	2.3	6.3
33	2	2	2	3	2	0.0947	0.7	8.8
34	2	1	3	1	2	0.0189	47.0	4.9
35	3	2	1	2	2	0.0142	1.6	5.5
36	1	3	2	3	2	0.1140	0.2	9.8

$$y_k^{(\mathrm{net})} = c_{ik} \times P\left(W_{i,j-1}, O_{i,j-1}\right) = b_{ik} \quad k = 1-3, \tag{3}$$

where c_{ik} and b_{ik} are constant real numbers (b_{ik} = biases).

Choosing the best network architecture (number of hidden layers and the number of neurons in each layer) is one of the most significant tasks in ANN modeling. In ANN models, the number of neurons in the first layer and last layer correspond to the number of input variables and output characteristics, respectively. Thus, the ANN model for EDM process has **5** neurons in the first layer (input) and **3** neurons (for the single objective) or **1** neuron (for the multi-criteria modeling) in the last (output) layer (Fig. 3). In the past, the adequacy of an ANN model would be checked by mean square error (MSE) between desired outputs (Y_k) and predicted outputs (y_k). The best network would then be selected based on MSE criterion. The general form of MSE function is expressed in Eq. (4):

$$\mathrm{MSE} = \frac{1}{p} \sum_{k=1}^{p} (Y_k - y_k)^2 \tag{4}$$

Nowadays, however, a more comprehensive criterion called authority of the net ($M_{(\mathrm{net})}$) has been proposed to account for both training and testing errors in evaluating a given ANN model [23]. The training of an ANN implies finding desired net's architecture and weights that minimize error between the desired process quality characteristics and the predicted ones. Forward phase is the first step in training, which occurs when an input vector X is presented and propagated through the network to compute an output characteristics. Therefore, an error between the desired output (Y_k) and predicted output (y_k) of the neural network is calculated. So, the modeling authority of the net ($M_{(\mathrm{net})}$) is given in Eq. (5) [23]:

$$M_{(\mathrm{net})} = \alpha \cdot \frac{1}{p_0} \sum_{r=1}^{p_0} (Y_r - y_r)^2 + \beta \cdot \frac{1}{q_0} \sum_{s=1}^{q_0} (Y_s - y_s)^2 \quad (5)$$

The recent relation corresponds to fitness function for developing the BPNN construction, where α and β are the coefficients that determine the relative importance of learning and generalization capability of ANN. Also, p_0 and q_0 are the numbers of training and testing data, respectively. Furthermore, Y_r and y_r are preventative variables for desired and predicted values used for training purpose. In the same token, variables, Y_s and y_s are desired and predicted test values.

The best network may then be selected based on $M_{(\mathrm{net})}$ criterion. Likewise, in this study, number of hidden layers (No = 2) and neuron in each layer (No = 5) have been determined using PSO algorithm and assessed base on the $M_{(\mathrm{net})}$ criterion (Fig. 4).

6 Analysis of variance (ANOVA) interaction analysis

To determine how well a model fits the experimental values and represent the authentic process under study, analysis of variance (ANOVA) is performed [19].

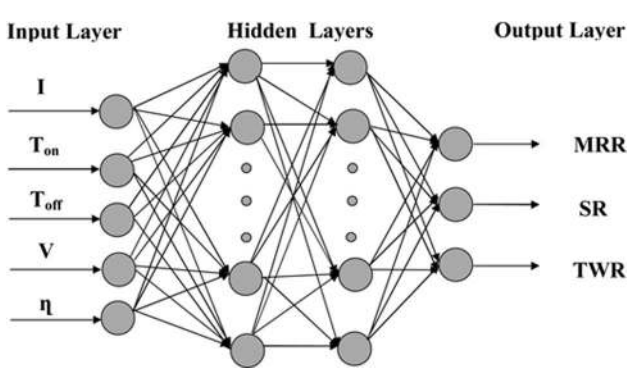

Fig. 4 Architecture of proposed artificial neural network used for single-objective modeling

6.1 Interaction plot

An interactions plot is a powerful graphical tool which plots the mean response of two factors at all possible combinations of their settings. If the lines are parallel, this indicates that there is an interaction between the factors. Non-parallel lines are an indication of the presence of interaction between the factors. The interaction plots for the process variable are illustrated in Fig. 5.

It could be noticed that the current is the most important variable affecting MRR. By the same token, the most important variable affecting SR and TWR is pulse on time. Furthermore, pulse off time is the least important parameter affecting MRR, TWR and SR.

7 Heuristic algorithms

There is a plenty of different heuristic algorithms (including GA, SA, PSO, Ant and Bee colony, and etc.) among which SA and PSO based on their merits are employed extensively for optimization of different problems and processes. Easy to program and converge fast are the significant advantages of using PSO algorithm. While, falling into local optimum traps in high-dimensional space could be considered as disadvantage of using this algorithm. GA is another widely used algorithm coding which is a time consuming process due to setting its large number of parameters. SA's major merit over other algorithms (such as PSO) is its ability to avoid getting trapped in local minima. As with GA, a major advantage of SA is its flexibility and robustness as a global search method. SA algorithm does not use gradient information and makes relatively few assumptions about the problem being solved. It can deal with highly nonlinear problems and no differentiable functions as well as functions with multiple local optima. SA is a very powerful and important tool in a variety of disciplines.

Based on the above-mentioned reasons, in this study, PSO and SA algorithms have been considered as the heuristic algorithms to optimize the process responses. At the first step, PSO algorithm has been considered for determining BPNN architecture. Then, the process responses have been optimized using PSO algorithm. Next, the performance of PSO has been checked using SA algorithm. At the last step, the optimization results have been confirmed using experimental tests. The details of these algorithms procedures are well documented in Ref. [24].

7.1 Simulated annealing algorithm

Simulated annealing (SA) algorithm, first proposed by Kirkpatrick in 1983, is originally inspired by the process

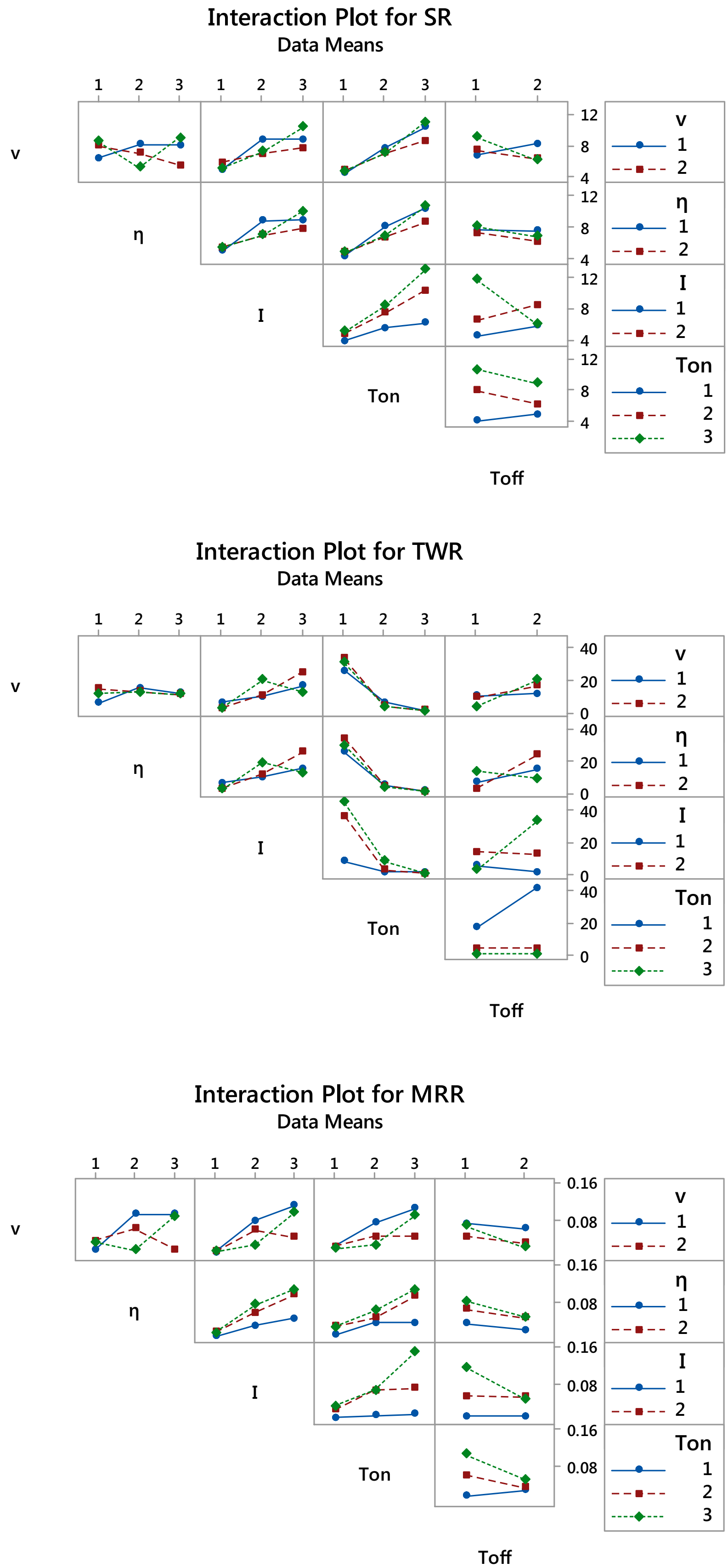

Fig. 5 Interaction plots for EDM process variables

of physical annealing in metal work and widely used in optimization problems.

The first step in a standard SA algorithm approach, is generating an initial solution randomly. A small random change is made in the existing solution at initial stages. Then, the objective function value of new solution is calculated (E_{i+1}) and compared to that of the current solution (E_i). A move is made to the new solution if it has better value or if the probability function, implemented in SA algorithm, has a higher value than a randomly generated number between 0 and 1. The probability of accepting a new non-improving solution is given in Eq. (6): [25].

$$P_i = \exp -(\Delta E / T_i) \tag{6}$$

where T_i is the current temperature, along with the difference between current solution and the new solution (ΔE). For reduction of the temperature, Eq. (7) is used:

$$T_{i+1} = \alpha \times T_i \quad i = 0, 1, \ldots \text{ and } \quad 0.9 \leq \alpha < 1 \tag{7}$$

SA algorithm has varied applications in various engineering problems [26, 27]. In this study, SA algorithm has been used in the optimization of EDM process variables. In this stage, the proposed BPNN model is implanted into a SA procedure to find the optimal set of EDM process variables in order to maximize the MRR and minimize the TWR and SR simultaneously.

7.2 Particle swarm optimization algorithm

Particle swarm optimization (PSO) algorithm is a population-based stochastic optimization procedure, inspired by social behavior of birds flocking or fish schooling, developed by Eberhart and Kennedy in 1995 [28]. The optimization process is initialized with a population of random solutions and searches for optima by updating generations. The potential solutions named particles fly through the problem space by following the current optimal particles. PSO algorithm is implemented easily and has few variables to adjust. The algorithm can be illustrated based on the following scenario: A group of birds are randomly searching food in an area. There is only one piece of food in the area being searched. All the birds do not know where the food is. But they know how far the food is in their search [29]. So, the best approach to achieve the food is to simply follow the bird, which is nearest to the food. In optimization problems, each bird in the search space is referred to as 'particle.' All the particles are evaluated by the fitness function to be optimized and have velocities for the particles. The particles fly through the problem space by following the current optimum particles. The problem is initialized with a group of random particles and then searches for optima by updating generations. In all the iterations, each particle is updated by following two 'best' values. The best solution achieved so far among the particle is called as 'particle best' termed as $\mathrm{p_{best}}$, and the best solution obtained so far in the population is called as 'global best' termed as $\mathrm{g_{best}}$. A particle takes the entire particle toward its $\mathrm{p_{best}}$ and $\mathrm{g_{best}}$ locations. After finding the two best values, the particles are updated with its velocity and positions using Eqs. (8) and (9) [28].

$$\begin{aligned} V[] = {} & (C_1 \times \mathrm{rand}()) \times \ (\mathrm{pbest}[] - \mathrm{present}[]) \\ & + (C_2 \times \mathrm{rand}()) \times (\mathrm{gbest}[] - \mathrm{present}[]) \end{aligned} \tag{8}$$

$$p[] = \ V[] + \mathrm{present}[] \tag{9}$$

V[] is the particle velocity, present is the current particle, p_{best} and g_{best} are defined as stated before, rand() is the random number between 0 and 1, c_1, c_2 are learning factors usually varies from 1 to 4, and p[] is new particle position. Compared to other optimization techniques, the information sharing mechanism in PSO is significantly different. Only g_{best} gives out information to others, which is a one-way information sharing mechanism. The evolution looks only for the best solution, and hence, all the particles tend to converge to the best solution quickly in most cases. The advantages of using PSO are that it takes real numbers as particles and there are few variables to adjust [28].

The searching is a repeat process, and the stop criteria are that the maximum iteration is reached or the minimum error condition is satisfied. The various variables in PSO are number of particles, dimension of particles, and range of particles, learning factor, stop condition and global versus local version [28].

7.3 Algorithm procedures

The PSO algorithm procedure is as the following sentences

Step 1: Initialization: The position and velocity of all particles are set at random within pre-specified or legal range.
Step 2: Calculate fitness function value for each particle. If the fitness function value is better than the best fitness function value (p_{best}) in history, set current value as the new p_{best}.
Step 3: Choose particle with the best fitness function value of all the particles considered so far as the g_{best}.
Step 4: Calculate particle velocity and position for each particle using Eqs. (7) and (8).
Step 5: Particle velocities on each dimension are closed to a maximum velocity $v_{\max}$. If the sum of acceleration would cause the velocity on that dimension to exceed $v_{\max}$ (Specified by the user), the velocity on the dimension is limited to $v_{\max}$.

Step 6: Dismiss if maximum number of iterations is reached.
Else, go to Step 2.
Step 7: End.

The SA algorithm procedure is as the following sentences

Step 1. Initialize the temperature parameter T_0; cooling schedule; r $(0<r<1)$ and the termination criterion (e.g., number of iterations $k=1\ldots K$); Generate and evaluate an initial candidate solution (perhaps at random); call this the current solution, c.
Step 2. Generate a new neighboring solution, m, by making a small change in the current permutation of jobs and evaluate this new solution
Step 3. Consider this new solution as the current solution if the following sentences are satisfied:

(a) The objective value of new solution, $E(m)$, is better than of the current solution, $E(c)$.
(b) The value of acceptance probability function given by $(\exp^{(f(m)-f(c))/Tk})$ is greater than a uniformly generated random number "rand"; where $0<\text{rand}<1$.

Step 4. Check the termination criterion and update the temperature parameter (i.e., $T_k=r\times T_{k-1}$) and return to Step 2.

7.3.1 Variables of the algorithms

Variables of SA and PSO algorithm used in the proposed model are given below, respectively.

SA variables:

Initial temperature: 700, Temperature reduction rate: 0.91, Processing time: 30 s

PSO variables:

Number of iteration performed: 30, Population: 50, Learning factor c_1: 2, Learning factor c_2: 2

7.3.2 Calculation of optimum machining variables

Peak current (I) is calculated randomly within the limits using Eq. (10) [3]

$$I = I_{\min} + (I_{\max} - I_{\min}) \times \text{rand}() \tag{10}$$

Similarly, voltage (V) is also calculated randomly within the limits using Eq. (11).

$$V = V_{\min} + (V_{\max} - V_{\min}) \times \text{rand}() \tag{11}$$

Similarly, pulse on time (T_{on}), pulse off time (T_{off}) and duty factor (η) are also calculated randomly within the limits using Eqs. (12)–(14), respectively.

$$T_{\text{on}} = T_{\text{on(min)}} + (T_{\text{on(max)}} - T_{\text{on(min)}}) \times \text{rand}() \tag{12}$$

$$T_{\text{off}} = T_{\text{off(min)}} + (T_{\text{off(max)}} - T_{\text{off(min)}} \times \text{rand}() \tag{13}$$

$$\eta = \eta_{\min} + (\eta_{\max} - \eta_{\min}) \times \text{rand}() \tag{14}$$

The proposed integrated method and related algorithm are presented in Fig. 6.

The 36 data set gathered using Taguchi approach were organized in input/output pairs and has been divided into two subsets randomly. With 26 input/target pairs (experimental sets), training has been done. Furthermore, 10 input/target pairs for evading overfitting and attaining good generalization by means of cross-validation. Some test runs (2, 4, 7, 11, 14, 19, 22, 23, 27, 31 and 34) has been Collected for validating (testing) of developed model.

8 EDM process optimization results

8.1 Single-objective optimization

In this section, the developed BPNN model has been embedded into SA and PSO algorithms to maximize MRR and minimize TWR and SR independently. The results of optimization are shown in Table 4.

8.2 Multi-criteria optimization

As these objectives (MRR, TWR and SR) are conflicting, they have been converted into a single measure for multi-criteria optimization. To maximize MRR and minimize TWR and SR, the process variables values should be found in such a way that minimize the following Equation.

$$\begin{aligned} &\text{Minimize} \quad f\ (I,\ T_{\text{on}},\ T_{\text{off}}, \eta,\ V\ ) = W_1\text{TWR} + W_2\text{SR} - W_3\text{MRR} \\ &\text{Subjected to} \\ &6 \le I \le 30 \\ &25 \le T_{\text{on}} \le 200 \\ &10 \le T_{\text{off}} \le 75 \\ &0.4 \le \eta \le 1.6 \\ &50 \le V \le 60 \end{aligned} \tag{15}$$

where W_1, W_2 and W_3 are the weights considered for TWR, SR and MRR, respectively.

The algorithms were run from different starting points and with various variables settings. The best results obtained from the optimization procedure are reported in Tables 5 and 6.

In complementary section, in order to evaluate the accuracy of the predicted values, a set of experimental experiment was carried out based on the optimized process variables. Moreover, the obtained experimental responses derived

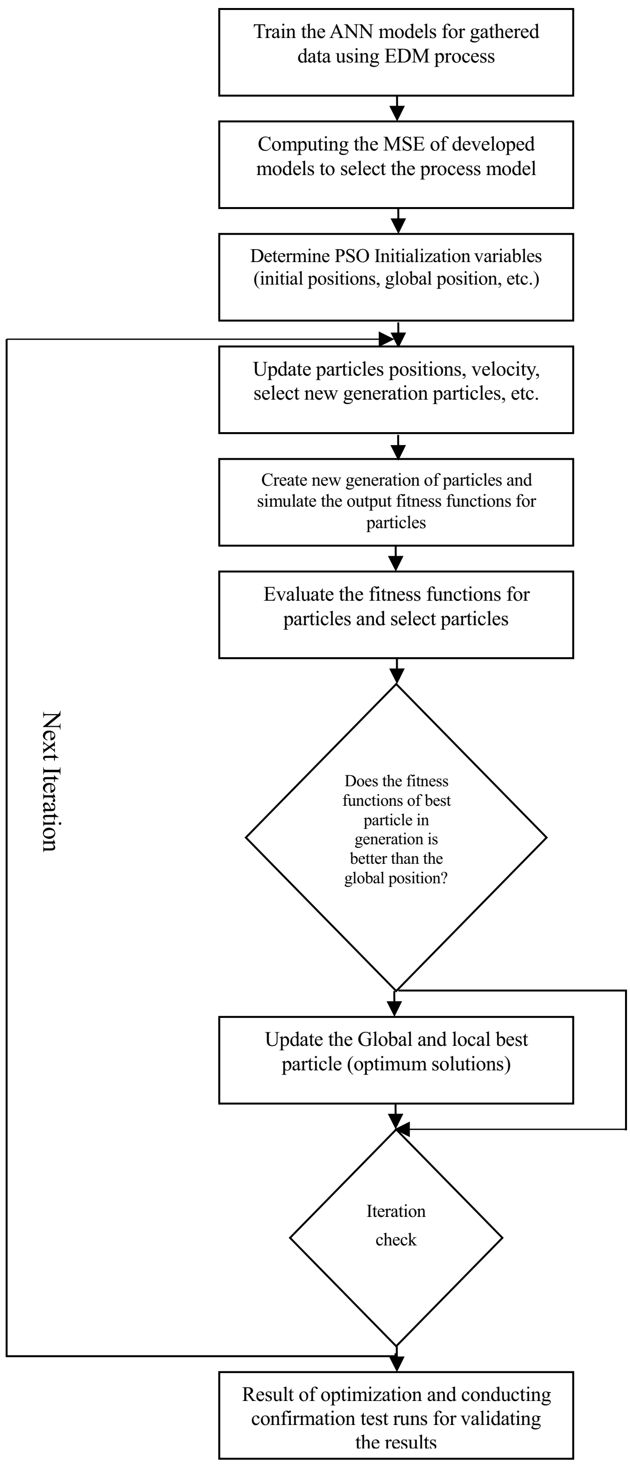

Fig. 6 Flowchart illustration of proposed method used for EDM process optimization

from SA and PSO algorithms were compared. The results which are presented in Tables 5 and 6 for equal (0.333) and different (0.750) weighing of each process characteristics show that the hybrid model can improve quality characteristics of the process.

Figure 7 illustrates the convergence trends for the heuristic algorithms used for TWR ($W_1 = 0.750$ and $W_2 = W_3 = 0.125$). By the same token, Figs. 8, 9 and 10 show the convergences of the proposed algorithms for SR ($W_2 = 0.750$ and $W_1 = W_3 = 0.125$), MRR ($W_3 = 0.750$ and $W_1 = W_2 = 0.125$) and the corresponding equal weighing ($W_1 = W_2 = W_3 = 0.333$). As shown, PSO converges quicker than SA. However, the optimized values are approximately the same. The termination factor has been considered time (30 s).

9 Conclusions

Selection of process variables levels significantly affects the quality of final product in electrical discharge machining (EDM) process. On the other hand, the interactions of these variables call for simultaneous selection of their optimal values. In this research, the problem of modeling and optimization (both single and multi-criteria) of EDM process for AISI2312 hot worked steel alloy has been addressed. The process modeling has been carried out using experimental data gathered as per L_{36} orthogonal array Taguchi (OA-Taguchi) method. First, three important process characteristics including material removal rate (MRR), surface roughness (SR) and tool wear rate (TWR) have been combine into an equally weighed single measure called weighted normalized grade (WNG). Next, the backpropagation neural network (BPNN) was developed to establish accurate relationships between input process variables (current, voltage, pulse off and on time and duty factor) and multiple performance characteristics. Furthermore, in most of the studies, the ANN architecture (number of hidden layers and nudes/neurons) has been determined based on trial and error method. In this study, to tackle the problem of selection of ANN architecture, PSO algorithm has been used. In the next stage, the developed BPNN model has been implanted into heuristic algorithms (SA and PSO) to find the optimal set of EDM process variables in order to maximize MRR and minimize SR and TWR independently and simultaneously. Then, the algorithms performances have been compared. The derived results manifests that the performance for both algorithms in single-objective optimization are approximately the same. Furthermore, the results of the optimization illustrate that the PSO algorithm converges quicker than the SA algorithm. The proposed modeling and optimization approach, with minor changes, can be applied to other manufacturing process.

Table 4 Result of single-objective optimization using SA and PSO algorithms

Output	Algorithm	Process variables					Predicted	Experiment	Error (%)
		I (A)	V (V)	T_{on} (μs)	T_{off} (μs)	η (S)			
MRR (gr/min)	SA	27	50	200	12	1.6	0.270	0.260	3.7
MRR (gr/min)	PSO	30	50	200	10	1.6	0.270	0.260	3.7
SR (μm)	SA	8	60	25	71	0.4	2.820	2.710	3.9
SR (μm)	PSO	6	60	25	75	0.4	2.800	2.750	1.7
TWR (%)	SA	6	60	200	32	1.4	0.123	0.126	2.4
TWR (%)	PSO	6	60	200	29	1.5	0.120	0.116	3.3

Table 5 Optimal EDM variables and corresponding process quality measures for equal weighing in multi-criteria optimization

Output	Algorithm	Process variables					Predicted	Experiment	Error (%)
		I (A)	V (V)	T_{on} (μs)	T_{off} (μs)	η (S)			
MRR (gr/min)	SA	19	53	118	33	1.2	0.243	0.236	2.90
SR (μm)	SA						3.700	3.570	3.51
TWR (%)	SA						0.190	0.200	5.26
MRR (gr/min)	PSO	22	54	129	39	1.2	0.249	0.240	3.61
SR (μm)	PSO						3.800	3.600	5.26
TWR (%)	PSO						0.200	0.190	5.00

Table 6 Optimal EDM variables and corresponding process quality measures for 75% weighing for each out put in multi-criteria optimization

Output	Algorithm	Process variables					Predicted	Experiment	Error (%)
		I (A)	V (V)	T_{on} (μs)	T_{off} (μs)	η (S)			
MRR (gr/min)	SA	22	50	200	43	1.30	0.256	0.242	5.46
MRR (gr/min)	PSO	22	50	200	41	1.30	0.256	0.248	3.12
SR (μm)	SA	9	58	95	53	0.70	2.900	2.800	3.44
SR (μm)	PSO	8	57	95	56	0.70	2.900	2.800	3.44
TWR (%)	SA	18	56	176	51	0.96	0.170	0.180	5.88
TWR (%)	PSO	18	57	182	50	1.00	0.160	0.150	6.25

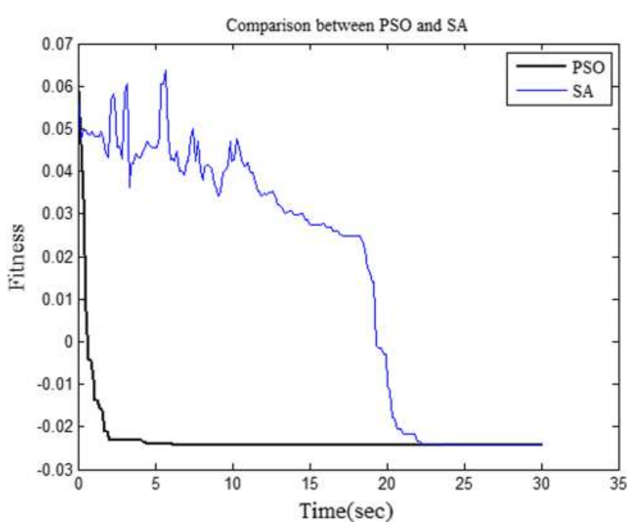

Fig. 7 Convergence of the heuristic algorithms for TWR ($W_1=0.75$)

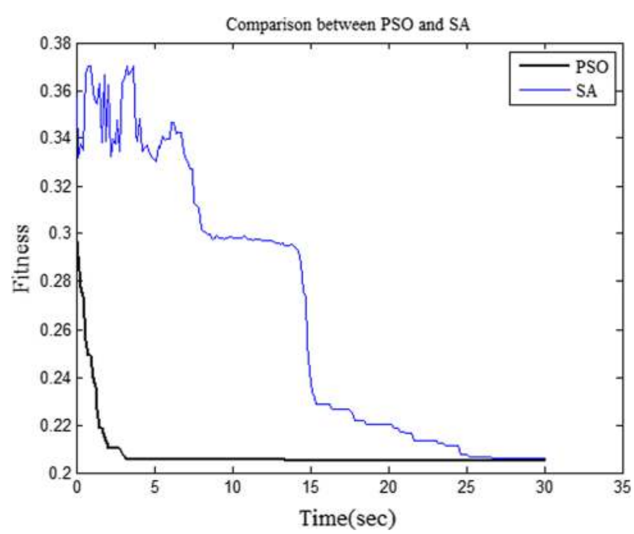

Fig. 8 Convergence of the heuristic algorithms for SR ($W_2=0.75$)

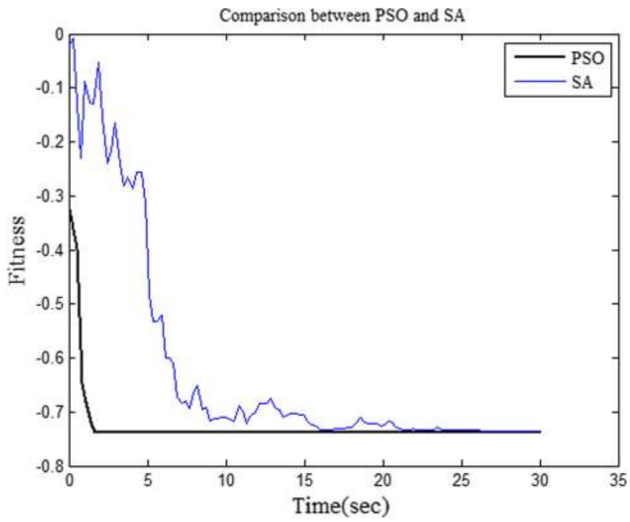

Fig. 9 Convergence of the heuristic algorithms for MRR ($W_3 = 0.75$)

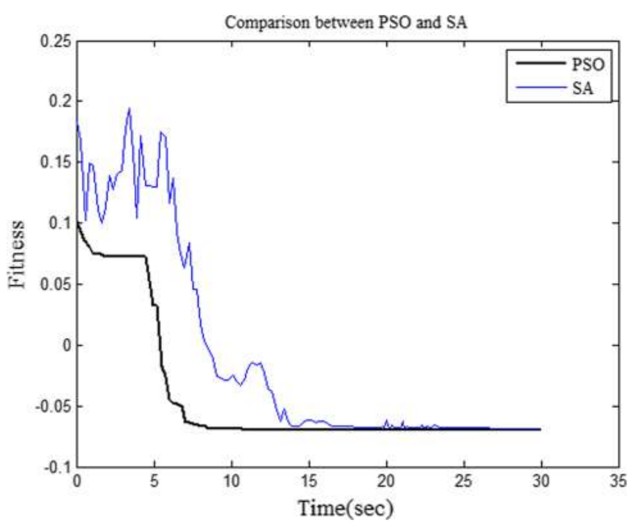

Fig. 10 Convergence of the heuristic algorithms for the equal weighing ($W_1 = W_2 = W_3 = 0.333$)

References

1. Ho KH, Newman ST (2003) State of the art electrical discharge machining (EDM). Int J Mach Tools Manuf 43:1287–1300
2. Muthuramalingam T, Mohan B (2014) A review on influence of electrical process variables in EDM process. Arch Civil Mech Eng 15:87–94
3. Zarepour H, Tehrani AF, Karimi D, Amini S (2007) Statistical analysis on electrode wear in EDM of tool steel DIN 1.2714 used in forging dies. J Mater Process Technol 188:711–714
4. Asokan P, Kumar RR, Jeyapaul R, Santhi M (2008) Development of multi-objective optimization models for electrochemical machining process. Int J Adv Manuf Technol 39:55–63
5. Subrahmanyam M, Nancharaiah T (2019) Optimization of process variables in wire-cut EDM of Inconel 625 using Taguchi's approach. Materialstoday Proc. https://doi.org/10.1016/j.matpr.2019.05.449
6. Bodukuri AK, Chandramouli S, Eswaraiah K, Laxman J (2019) Experimental investigation and optimization of EDM process parameters on Aluminum metal matrix composite. Materialstoday Proc. https://doi.org/10.1016/j.matpr.2018.10.271
7. Tharian BK, Jacob E, Johnson J, Hari V (2019) Multi-objective parametric optimization in EDM using grey relational analysis. Materialstoday Proc 16:1013–1019
8. Kumar V, Jatti S (2018) Multi-characteristics optimization in EDM of NiTi alloy, NiCu alloy and BeCu alloy using Taguchi's approach and utility concept. Alex Eng J 57:2807–2817
9. Nagaraju N, Venkatesu S, Ujwalaa NG (2018) Optimization of process variables of EDM process using fuzzy logic and Taguchi methods for improving material removal rate and surface finish. Materialstoday Proc 5:7420–7428
10. Parsana S, Radadia N, Sheth M, Sheth N, Savsani V, Eswara N, Prasad B, Ramprabhu T (2018) Machining parameter optimization for EDM machining of Mg–RE–Zn–Zr alloy using multi-objective Passing Vehicle Search algorithm. Archiv Civil Mech Eng 18:799–817
11. Babu KN, Karthikeyan R, Punitha A (2019) An integrated ANN–PSO approach to optimize the material removal rate and surface roughness of wire cut EDM on INCONEL 750. Materialstoday Proc 19:501–505. https://doi.org/10.1016/j.matpr.2019.07.643
12. Sen R, Choudhuri B, Barma JD, Chakraborti P (2018) Optimization of wire EDM variables using teaching learning based algorithm during machining of maraging steel 300. Materialstoday 5:7541–7551
13. Vishwa R, Surya VR, Kumar KMV, Keshavamurthy R, Ugrasen G, Ravindra HV (2017) Prediction of machining characteristics using artificial neural network in wire EDM of Al7075 based In-situ composite. Materialstoday 4:203–212
14. Gurupavan HR, Devegowda TM, Ravindra HV, Ugrasen G (2017) Estimation of machining performances in WEDM of aluminium based metal matrix composite material using ANN. Materialstoday 4:10035–10038
15. Singh NK, Singh Y, Kumar S, Sharma A (2019) Predictive analysis of surface roughness in EDM using semi-empirical, ANN and ANFIS techniques: a comparative study. Materialstoday Proc. https://doi.org/10.1016/j.matpr.2019.08.234
16. Ashikur Rahman Khan M, Rahman MM, Kadirgama K (2014) Neural network modeling and analysis for surface characteristics in electrical discharge machining. Procedia Eng 90:631–636
17. Venkata RR, Kalyankar VD (2017) Optimization of modern machining processes using advanced optimization techniques: a review. Int J Adv Manuf Technol 73:1159-118
18. Gul T, Bischoff R, Permentier HP (2015) Optimization of reaction variables for the electrochemical oxidation of lidocaine with a design of experiments approach. Electrochim Acta 171:23–28
19. Kim Y, Jeon ES (2015) Establishment of regression model for estimating shape variables for vacuum-sealed glass panel using design of experiments. Vacuum 121:113–119
20. Kumar NM, Kumaran SS, Kumaraswamidhas LA (2015) An investigation of mechanical properties and material removal rate, tool wear rate in EDM machining process of AL2618 alloy reinforced with Si3N4, AlN and ZrB2. J Alloy Compd 650:318–327
21. Korkut I, Acır A, Boy M (2011) Application of regression and artificial neural network analysis in modelling of tool–chip interface temperature in machining. Expert Syst Appl 38:11651–11656
22. Niaki STA, Hoseinzade S (2013) Forecasting S&P 500 index using artificial neural networks and design of experiments. J Ind Eng Int 9:1–9
23. Kapusuz M, Ozcan H, Yamin JA (2015) Research of performance on a spark ignition engine fueled by alcohol–gasoline blends using artificial neural networks. Appl Therm Eng 91:525–534

24. Lee KH, Kim KW (2015) Performance comparison of particle swarm optimization and genetic algorithm for inverse surface radiation problem. Int J Heat Mass Transf 88:330–337
25. Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Am Assoc Adv Sci 220:671–680
26. Bashiri M, Karimi H (2012) Effective heuristics and meta-heuristics for the quadratic assignment problem with tuned variables and analytical comparisons. J Ind Eng Int 8:1–9
27. Jahromi MHMA, Tavakkoli-Moghaddam R, Makui A, Shamsi A (2012) Solving an one-dimensional cutting stock problem by simulated annealing and tabu search. J Ind Eng Int 8:1–8
28. Raja B, Baskar S (2012) Application of particle swarm optimization technique for achieving desired milled surface roughness in minimum machining time. Expert Syst Appl 39:5982–5989
29. Cheng R, Jin Y (2015) A social learning particle swarm optimization algorithm for scalable optimization. Inf Sci 291:43–60