Modeling for a small-hole drilling process of engineering plastic PEEK by Taguchi-based neural network method

Dar-Yuan Chang · Chien-Hung Lin · Xing-Yao Wu

Received: 16 June 2021 / Accepted: 23 November 2021 / Published online: 16 January 2022 © The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract
Engineering plastics have specific properties in strength, hardness, impact resistance, and aging persistence, often used for structural plates and electronic components. However, the holes made by the drilling process always shrink after the cutting heat dispersion due to their high thermal expansion coefficient. Drilling parameters must be discussed thoughtfully especially in the small-hole fabrication to acquire a stable hole quality. This study developed parameter models by the Taguchi-based neural network method to save the experimental resources on the drilling of engineering plastic, polyetheretherketone (PEEK). A three-level full-factorial orthogonal array experiment, L27, was first conducted for minimizing the thrust force, hole shrinkage in diameter, and roundness error. In terms of the network modeling, four variables were designated to the input layer neurons included the three drilling parameters (spindle speed, depth of peck-drilling, feed rate) and the thrust force detected, and that of the output layer neurons were two hole characteristics of diameter shrinkage and roundness. The models were trained by a stepped-learning procedure to expand the network’s field information stage by stage. After three stages of training, the models developed can provide precise simulations for the network’s training sets. For the non-trained cases, the prediction accuracy of the hole’s characteristics discussed was below 1 μm in the drilling of a 1-mm-diameter hole.

Keywords Neural network · Stepped-learning procedure · Taguchi’s method · Thrust force · Diameter shrinkage

1 Introduction
For a mechanical part made by cutting processes, usually, the first operation is the main body fabrication by means of turning or milling process, and then uses the drilling process to finish the last hole features. Drilling is one of the most frequent machining operations, accounting for over 30% of all cutting operations in the industry [1]. Esim and Yildirim [2] indicated that many undesired effects occur during the drilling process due to the complexity of the drill and cutting geometry. Roth et al. [3] believed the use of multiple cutting edges in the drilling process results in variable cutting speeds along these edges, which increase interactions between the chips and cutting tool during the chip evacuation and altered the heat transfer. As the product trends in modernization, lightweight, space reduction, and sustainability, fabrication of miniature parts and components is crucial in the industry [4]. Correlations between cutting parameters and hole’s characteristics are worth investigating especially for a small-hole process.

Artificial neural network (ANN) has excellent abilities in learning complex non-linear and multivariable relationships among variables/parameters and makes them very useful in many engineering applications. For instance, Zerti et al. [5] investigated ANN models to predict the surface roughness and thrust force in the turning of stainless steel AISI 420 based on the experimental data of cutting speed, feed rate, and depth of cut. Özden et al. [6] demonstrated the modeling of cutting parameters in the turning process of PEEK composite using ANN and adaptive-neural fuzzy inference (ANFIS) systems to predict the cutting forces. Daniel et al. [7] developed an ANN model for milling the aluminum metal matrix composites using the back-propagation neural network (BPNN) algorithm to predict the surface roughness, temperature, material removal rate (MRR), and thrust force components. Yadav et al. [8] built an ANN model for the electrical discharge diamond grinding process to predict the output responses of MRR and surface roughness. Biswas et al. [9] presented an ANN model for the Nd:YAG laser
microdrilling of TiN-Al2O3 composite by the BPNN for the predictions of the hole’s circularity and taper. Sarıkaya and Yılmaz [10] presented a two-hidden-layer architecture network for micro-electrical discharge drilling process to access the drillability of machining the stainless steel AISI 304.

For the mechanical drilling process, drilling parameters of cutting velocity, spindle speed, feed rate, depth of drilling, and point angle are mostly analyzed; cutting characteristics discussed include thrust force, torque, vibration, cutting temperature, driving power, and tool wear; and hole’s quality characteristics measured are displacement of hole center, surface roughness of drilled wall, roundness, hole oversize, and burrs [4, 11]. In general, the cutting force is the most sensitive indicator of machining performance for observing the tool condition and chip formation [12]. Neural networks can be used to model drilling dynamics, and thereafter, a neural controller can be used to control feed rate to get desired thrust force [13]. Several ANN models on the thrust force and spindle power in the twist drilling have been presented. Patra et al. [14] developed an ANN model based on thrust force signals and drilling parameters to predict the drilled hole number in the micro peck-drilling process of the tool steel of AISI P20. Efklidis et al. [15] presented ANN models on drilling of aluminum Al 7075 for predicting the thrust force and torque based on the data of full-factorial experiments. Kharwar et al. [16] developed ANN models to predict the thrust force, torque, and surface roughness in the drilling of reinforced polymer nanocomposites. Corne et al. [17] built an ANN model based on the data of spindle power to predict tool wear and to detect drill breakage during the drilling of Inconel steel. Soepangkat et al. [18] applied the BPNN network to model the drilling process of carbon fiber reinforced polymer, and to predict the optimum responses of thrust force, torque, and delamination.

Vibration induced during the machining process has adverse effects on the life span of the cutting tool, measurement accuracy, and surface quality of the workpiece. Ulase et al. [19] presented an ANN model by vibration signals to save the resources in experiments and determined the optimum cutting parameters in the drilling of tool steel AISI D3. Eism et al. [2] presented two neural networks (NN) predictors modeling by BPNN and radial basis neural network to predict the vibration variation in the drilling of steel and aluminum. Caggiano et al. [20] developed a BPNN model taking the specific hole number as the input layer neurons and the diagnosis of flank wear can be accurately simulated based on the multiple sensor signals of thrust force, torque, vibrations, and acoustic emission.

In terms of hole’s characteristics, Vrabel et al. [21] presented an ANN adaptive control system for the drilling of nickel-based superalloy of Udimet 720. Six input parameters were fed to the network for predicting the roughness of the hole drilled. Akınçoğlu et al. [22] proposed an ANN model for predicting the surface roughness and roundness in the drilling of the tool steel AISI D2. Cruz et al. [23] installed a multiple sensors system to monitor cutting conditions to develop the ANN model for predicting the hole diameter and roundness in the drilling of a metal sandwich plate. Mondal et al. [24] investigated burrs generated in aluminum alloy drilling. An ANN model was built to derive the optimal process parameters by the flower pollination algorithm for diminishing the burr height. Ahn and Lee [25] discussed the burr formation in the drilling of ductile metals of copper and brass. A BPNN model was developed to predict the burr size and type.

Several scholars used the BPNN to develop their network models for the input patterns simulation or prediction in machining processes [2, 3, 5–10, 14–17, 19–25]. Data of the experimental trials are divided randomly, approximately 60–70% for network training and the remaining for validating and testing. During the network learning, when the procedure of network training has answers to one of the converging conditions, then the validation and testing procedures can be started. The performance of the NN model can be evaluated by the mean square error (MSE) and the correlation coefficient (R). For instance, Cruz et al. [23] implemented 162 drilling trials, and the experimental data were divided, with 60% for network training, 20% for validation, and 20% for testing. Ahn and Lee [25] conducted 60 data sets for NN modeling and sliced them into a training set of forty and a testing set of twenty. Yadav et al. [8] prepared 31 sets of experimental data used as the inputs in network training, and a separate 8 experimental trials were arranged for network testing. Mondal et al. [24] specified the training ratio of 80%, testing ratio of 10%, and validation ratio of 10% used in NN modeling. At the same time, the prediction was validated with another L9 OA experimental data.

In the references cited, for the input patterns (training sets) used in network’s training, usually, the models can derive precise simulations after some NN’s experiments with different hidden layer architectures or training parameters. However, for the non-trained cases (beyond the training sets), the models might predict the output characteristics inaccurately and even occur overfitting. It is owing to the lack of field knowledge. This study presents a “stepped-learning procedure” to advance the network’s training process. Through “differential analyses” of the network’s training sets, some additional drilling trials and training sets are decided for the next-stage training. Accordingly, the field information of the network model can be expanded in a planned way, and the prediction reliability would be improved as well.

The remainder of this paper is presented as follows. Section 2 focuses on the proposed stepped-learning procedure, followed by the machining setup and experimentation in Sect. 3. Section 4 presents the results of Taguchi’s
experiments. Section 5 describes the modeling of the ANN models with data, followed by the conclusions in Sect. 6.

2 Taguchi-based neural network model by a stepped-learning procedure

2.1 Back-propagation neural network

ANN is a mathematical or computational model that is inspired by the structure or functional aspects of biological neural networks. If a network is able to compute some functional relationship between its input and its output, this neural network is called a mapping network [26]. The BPNN is a layered, feedforward network, and fully interconnected by layers. There are no feedback connections and no connections that bypass one layer to go directly to a later layer in this network. It can be used to solve complex pattern-matching problems. Figure 1 shows the architecture of a BPNN network with a double-hidden-layer used in this study.

Several NN training algorithms have been proposed such as Levenberg–Marquardt (LM), Bayesian regularization, and scaled conjugate gradient. LM algorithm is a combination of two training processes: steepest descent algorithm and Gauss–Newton algorithm, the network’s convergence speed can be improved significantly [14]. Sarıkaya and Yılmaz [10] also indicated that LM algorithm operates quicker once it trains a medium-sized feed forward neural network. This study built the drilling parameters model by BPNN network trained with LM algorithm. For a neural network, “learning” means setting a proper architecture with sufficient neurons to find appropriate connected weights for modeling. In this study, four types of double-hidden-layer neurons, $5 \times 5$, $10 \times 10$, $15 \times 15$, and $20 \times 20$, were experimented to determine the network architecture (see Fig. 1). The network was developed by the MATLAB neural network toolbox (v. 2019a).

2.2 Input pattern transformation

The experimental method proposed by Taguchi is based on an orthogonal array (OA) table. The trials in the OA table are all equipped with the orthogonal property on each factorial level among the control factors analyze; therefore, the Taguchi’s analysis can be completed by the least trials. These trials are very suitable to be assigned as the initial training sets for the network learning.

Before network training, experimental data obtained from the OA experiment must be transformed into the input patterns by means of a normalization process. Values of the input and output variables must be normalized between 0 and 1. Their ranges in normalization are schemed by the experimental parameter values or the hole’s characteristics measured. However, to expand the field information beyond the training sets, the normalized range $(x_{\text{max}} - x_{\text{min}})$ of each variable is specified at a slightly larger than the actual experimental values. The normalized value $(x_{\text{nor}})$ is calculated by

$$x_{\text{nor}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \in [0, 1],$$

(1)

where $x$ is the value of the variable to be normalized.

2.3 Training by a stepped-learning procedure

This study trained the NN parameter model by a stepped-learning procedure as shown in Fig. 2. In the first-stage training, the sampling distribution is selected as the defaulted ratio of MATLAB: i.e., 70% for training, 15% for validation, and the remaining 15% for testing, randomly sampled from the input patterns. The transfer function of the hidden layer is tan-sigmoid (Tansig), and that of the output layer is pure linear (Purelin). The performance of sample regression can be evaluated in terms of the MSE and R values. The MSE is given as

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{exp}})^2,$$

(2)

where $n$ is the number of the training sets, $y_i$ is the simulated outputs of the training set examined, and $y_{\text{exp}}$ is its expected value in training. The $R$ of the model is calculated by

Fig. 1 Architecture of the BPNN model with double hidden layers
Experiment data from the L_{27} OA experiment

Normalization
(drilling parameter & hole characteristics)

Input pattern transformation
(network training sets)

Network architecture
(4×5×5×2) (4×10×10×2)
(4×15×15×2) (4×20×20×2)

Network training

Drilling experiments
(for additional/extra drilling trials)

Scheme the model testing sets
(8 drilling trials)

Scheme the additional/extra training sets
(20 drilling trials)

Decide the datum sets
(for additional/extra drilling trials)

Sorting of the differentials of the network training sets

Adjust the network architecture or training parameters

Converge?

Yes

Network validation & testing

Differential analysis
(simulated vs. experimental)

Taguchi-based neural network parameter model by stepped-learning procedure

Fig. 2 Taguchi-based neural network model by a stepped-learning method
\[ R = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}. \quad (3) \]

where \( x \) and \( y \) are the experimental values and the model’s simulated values, respectively, and \( \bar{x} \) and \( \bar{y} \) are their averages. A lower MSE and a larger \( R \) indicate a better performance.

Three converging conditions are set here to terminate the NN training as follows:

1. The maximum number of epochs in training is reached in 1000.
2. The minimum gradient descent in training is below \( 1 \times 10^{-7} \).
3. The MSE of the validation sets cannot be descended during 6 continuous epochs.

When the NN training is terminated, differential analyses are conducted for checking the deviations of training sets in network simulations. Afterward, to select the training sets that underperformed in network training to be the “datum sets” for planning “additional training sets” and “model’s testing sets.” The former is used to expand the field information that the model lacked, and the latter is prepared for estimating the prediction accuracy only, not involved in the training process.

The next is the second-stage refining procedure that aims to improve the network’s simulation and prediction accuracy by adding some additional training sets that underperformed in the first-stage training. In this stage, all the trials obtained from the experiments are designated as the network’s training sets, with no validation sets and testing sets again. Prediction accuracy of the non-trained cases is estimated by a specific model’s testing sets. Termination of the network training is determined by the specific MSE directly. However, the model’s prediction accuracy achieved can be estimated in terms of root-mean-square error, RMSE, calculated by the simulated value of the output variable (\( y_j \)) and its corresponding measured data (\( y_{j,\text{mea}} \)).

\[ \text{RMSE} = \sqrt{\frac{1}{m} \sum_{j=1}^{m} (y_j - y_{j,\text{mea}})^2} \quad (4) \]

\( \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \)

3 Machining setup and experimentation

3.1 Material and Tools

Thermoplastic plastics have a specific property of high thermal expansion coefficient whose cutting reaction is different from the metal drilling. The holes drilled will be shrunk, while the cutting heat is dispersed. Therefore, the final hole diameter would be smaller than the drill diameter. To acquire a stable quality of the hole drilled, the selection of drilling parameters is very important [4]. This study conducted small-hole drilling experiments of the polyetheretherketone (PEEK), MDS 100, which was produced by Quadrant Systems Ltd, UK. Table 1 lists its important properties.

To avoid the tool “walking” on the workpiece surface at the beginning of the operation, a center piloting hole was first made by a short-rigid drill of 0.5-mm diameter with a drill body of 0.8-mm length and a point angle of 90°. The depth of the center hole was 0.2 mm. After that, the peck-drilling cycle of the target hole of 1.0-mm diameter was started as shown in Fig. 3. The retreating point was set at 0.2 mm above the workpiece surface. A tungsten carbide twist drill with a point angle of 118 was used made by Sphinx Tools Ltd, Swiss.

3.2 Apparatus and machining setup

A vertical machine center (NXV-560A made by YCM Ltd, Taiwan) was used as shown in Fig. 4. The force system consists of a tri-axial force sensor (261A01, made by PCB Ltd, USA) and a data acquisition card (NI 9234, made by National Instrument, USA). The sensitivity of the force sensor on the Z axis is 0.56 mV/N, broadband resolution is 0.027 N-rms, and its low frequency response is 0.01 Hz. After the analog signals are transformed into the digital, import them to the computer and execute the signal analysis by a dynamic acquisition software (m+p analyzer, v. 5.2.1, made by m+p international, Germany).

This study designed a bakelite working base installed on the machine’s platform for mounting the force sensor as shown in Fig. 5. An acrylic workpiece fixture was screwed on the top of the force sensor, and then, the workpiece can be fastened by the four clamps. Figure 6 shows a PEEK workpiece plate of 80 mm × 30 mm × 4.5 mm. Twenty-one position holes of 2-mm diameter are distributed over the plate. Every plate could arrange 20 drilling zones at the most. Because burrs are easy to be formed during the
thermoplastic machining, a shallow square of 7 mm × 7 mm × 0.25 mm was first milled for each drilling zone on both surfaces to prevent the measuring errors. Thus, the total drilling length of the peck-drilling cycle was actually 4 mm. To make sure that the drilling zone of each experimental trial can be positioned correctly, two pins with 2-mm diameter were inserted in the acrylic fixture for coupling with the position holes of the drilling zone. Each zone had five experimental holes, but only the third was the hole drilled directly above the force sensor (see the drilling sequence marked in the right-enlarged drawing). Therefore, the force signals of the third hole were collected and used in the following thrust force analysis.

In terms of the hole’s characteristics, an image measurement instrument (Baty-6490 Venture plus, UK) was used here. Its resolution is 0.5 μm, and measurement accuracy is 2.5 μm + L/150. L is the measuring length in millimeter. Both hole’s characteristics of the hole diameter and roundness error were measured by an automatic optical inspection (AOI) procedure.

### 3.3 Experimental plan

Most research believes that cutting speed and feed rate are two important parameters affecting the performance in drilling processes, and they are usually selected as control factors in experiments. In order to avoid the adverse effects of chip piling in the small-hole fabrication, the holes drilled in this study were machined by the peck-drilling process and cooling cyclically with an oil lubricant (C-1170A-1, made by Peisun Chemical Co., LTD., Taiwan).

As a result, three factors of spindle speed, depth of peck-drilling, and feed rate was considered in this work, and the experiments were schemed by Taguchi’s three-level L27 OA for carrying out a full-factorial experiment. On the basis of the machining information provided by the material vendor (www.mcam.com) and some preliminary trials, the levels of the control factors are presented in Table 2. Each parameter
set in the L27 OA had three replicate trials, and then data of the 81 drilling trials were obtained, included the thrust force and the two hole’s characteristics measured.

A smaller thrust force indicates that the drilling resistance of the tool is low and the machining is stable. A smaller shrinkage in diameter or a lower roundness points out that the hole’s quality is good. Therefore, the three characteristics mentioned in this experiment all belong to the smaller-the-better (STB) problem. According to Taguchi’s methods [27], the definition of single-to-noise (S/N) ratio is

\[
S/N_{\text{STB}} = -10 \cdot \log\left(\bar{y}^2 + S^2\right),
\]

(5)

where \(\bar{y}\) is the average value, \(S\) is the standard deviation, and \(n\) is the number of replications of each parameter set. As to the hole shrinkage, this study quantified the diameter shrinkage, \(\Delta D\), by the difference between the diameter measured \((D_{\text{mea}})\) and the tool diameter \((D_{\text{tool}})\) expressed as

\[
\Delta D = D_{\text{tool}} - D_{\text{mea}}
\]

(8)

4 Results of Taguchi’s drilling experiment

According to the factorial levels planned in Table 2, the L27 OA experiment was performed and obtained the data of 81 drilling trials. The results were reported below.

4.1 Cutting characteristic-thrust force

Three levels of peck-drilling depth (Factor B) were analyzed here including 0.3, 0.4, and 0.5 mm, whereas the numbers of drilling cycles to finish a hole were 14, 10, and 8, respectively. Figure 7 shows time-domain diagrams of the thrust force in the trials of Trial-1–1, Trial-4–1, and Trial-7–1 that they were operated by an identical spindle speed of 3000 rpm and a feed rate of 200 mm/min but with different peck-drilling depths. Because the thrust force peak during the first drilling cycle was higher than that of the following cycles and decreased abruptly when the drill pierced the bottom surface, hence the force signals of the first and last cycles were excluded in analyses. Table 3 presents the thrust force data and S/N ratio of the L27 OA experiment. The averages of thrust force peaks of the three trials above mentioned were 4.50, 4.26, and 4.07 N, respectively. The high-level drilling depth (B3) presented a lower thrust force (Trial-7–1).

In the metal drilling, when a lower feed rate is applied, as the increase of the number of holes drilled, tool wear increases rapidly due to ploughing and rapid heating that causes the increasing of the thrust force [14]. However, in our drilling experiments of thermoplastics-PEEK, the cutting heat results from a lower feed rate would soften the material and decrease the thrust force. Figure 8 a shows the factorial level effects analyzed by the S/N ratio. The sequence affecting the thrust force was feed rate (Factor C), spindle speed (Factor A), and depth of peck-drilling (Factor B). A high-level spindle speed (A3)
with a low-level feed rate (C1) obtained the best performance in the thrust force. From the statistics of the full-factorial experiment, L27 OA (Table 3), the optimal factorial level set was A3, B2, and C1 (Trial-22) that had the smallest thrust force of 3.01 N. On the contrary, the worst one was A1, B2, and C3 (Trial-6) that it enlarged to 8.55 N.

### 4.2 Hole characteristics-diameter shrinkage and roundness

Five holes were drilled in each drilling zone (see Fig. 6); therefore, the factorial level effect was evaluated by the average of the hole characteristics measured. Three replicas were done for every parameter set in the L27 OA experiment. Table 4 presents the measurement data of the two hole’s characteristics, diameter shrinkage (D) and roundness, and the S/N ratio.

For the diameter shrinkage, the factorial level effects analyzed by the S/N ratio are shown in Fig. 8b. The factor of spindle speed (Factor A) presented the largest effect, and the high-level spindle speed (A3) still obtained the smallest hole shrinkage in diameter. It is because that the operation was preceded by peck-drilling and the drill lifted above the workpiece surface during the drilling cycle (see Fig. 3), the chips can be smoothly removed from the machining hole. In addition, the liquid lubricant was sufficiently applied; consequently, the adverse effect of the high-level spindle speed on the hole shrinkage was unobvious. From Table 4, the parameter sets of Trial-23 and Trial-20 had the smallest means of 8.22 and 8.24 μm, respectively. The optimal factorial level set was A3, B3, and C2 (Trial-23).

In terms of the roundness, Fig. 8c shows the factorial level effects plotted by the S/N ratio. The level effects of the three factors were approximate. In principle, the use of a high-level spindle speed (A3) with a high-level feed rate (C3) could obtain a better performance. According to the statistics, Trial-23 was also the optimal parameter set with the smallest mean of 3.48 μm.
5 Parameter models trained by stepped-learning procedure

In the network modeling, four variables were designated to the input layer neurons, included the spindle speed, depth of peck-drilling, feed rate, and thrust force measured, and the variables of the output layer were the diameter shrinkage and roundness, as Fig. 1 shows. After Taguchi’s experiment, 81 sets of experimental data were provided for the following network training.

| Trial | A  | B  | C  | Avg. of thrust force peaks (N) | Mean (N) | Std. dev. (N) | S/N (dB) |
|-------|----|----|----|-------------------------------|----------|---------------|---------|
| L27 OA | A1 | A2 | A3 | N1 | N2 | N3 | | |
| Trial-1 | 1  | 1  | 1  | 4.50 | 4.47 | 4.38 | 4.45 | 0.049 | -12.97 |
| Trial-2 | 1  | 1  | 2  | 6.61 | 6.55 | 6.55 | 6.57 | 0.031 | -16.35 |
| Trial-3 | 1  | 1  | 3  | 8.21 | 8.45 | 8.52 | 8.39 | 0.133 | -18.48 |
| Trial-4 | 1  | 2  | 1  | 4.26 | 4.17 | 4.16 | 4.20 | 0.046 | -12.46 |
| Trial-5 | 1  | 2  | 2  | 6.50 | 6.42 | 6.50 | 6.47 | 0.037 | -16.22 |
| Trial-6 | 1  | 2  | 3  | 8.56 | 8.58 | 8.50 | 8.55 | 0.035 | -8.64  |
| Trial-7 | 1  | 3  | 1  | 4.07 | 4.18 | 4.03 | 4.09 | 0.062 | -12.24 |
| Trial-8 | 1  | 3  | 2  | 6.34 | 6.28 | 6.23 | 6.29 | 0.044 | -15.97 |
| Trial-9 | 1  | 3  | 3  | 8.49 | 8.50 | 8.54 | 8.51 | 0.020 | -18.60 |
| Trial-10| 2  | 1  | 1  | 3.65 | 3.59 | 3.62 | 3.62 | 0.025 | -11.18 |
| Trial-11| 2  | 1  | 2  | 5.07 | 5.11 | 5.00 | 5.06 | 0.044 | -14.08 |
| Trial-12| 2  | 1  | 3  | 6.84 | 6.81 | 6.83 | 6.83 | 0.015 | -16.69 |
| Trial-13| 2  | 2  | 1  | 3.51 | 3.39 | 3.40 | 3.43 | 0.052 | -10.71 |
| Trial-14| 2  | 2  | 2  | 4.96 | 4.99 | 4.93 | 4.96 | 0.025 | -13.91 |
| Trial-15| 2  | 2  | 3  | 6.78 | 6.82 | 6.78 | 6.79 | 0.016 | -16.64 |
| Trial-16| 2  | 3  | 1  | 3.33 | 3.34 | 3.35 | 3.34 | 0.008 | -10.47 |
| Trial-17| 2  | 3  | 2  | 4.60 | 4.78 | 4.79 | 4.72 | 0.089 | -13.48 |
| Trial-18| 2  | 3  | 3  | 6.63 | 6.59 | 6.68 | 6.63 | 0.037 | -16.43 |
| Trial-19| 3  | 1  | 1  | 3.18 | 3.13 | 3.17 | 3.16 | 0.021 | -9.99  |
| Trial-20| 3  | 1  | 2  | 4.21 | 4.27 | 4.24 | 4.24 | 0.026 | -12.55 |
| Trial-21| 3  | 1  | 3  | 5.55 | 5.56 | 5.42 | 5.51 | 0.062 | -14.82 |
| Trial-22| 3  | 2  | 1  | 2.99 | 2.99 | 3.05 | 3.01 | 0.027 | -9.57  |
| Trial-23| 3  | 2  | 2  | 4.15 | 4.07 | 4.02 | 4.08 | 0.052 | -12.22 |
| Trial-24| 3  | 2  | 3  | 5.42 | 5.44 | 5.43 | 5.43 | 0.007 | -14.69 |
| Trial-25| 3  | 3  | 1  | 2.99 | 3.04 | 3.03 | 3.02 | 0.022 | -9.61  |
| Trial-26| 3  | 3  | 2  | 4.03 | 4.09 | 4.01 | 4.04 | 0.034 | -12.14 |
| Trial-27| 3  | 3  | 3  | 5.34 | 5.36 | 5.35 | 5.35 | 0.005 | -14.57 |

Fig. 8 Factorial level effect plots: (a) thrust force, (b) diameter shrinkage, and (c) roundness
5.1 The first-stage training

5.1.1 Network training experiments

First, the experimental data were transformed into the input patterns required in NN training by the normalization procedure. Then, the input patterns were randomly divided into three groups of training sets, validation sets, and testing sets by the ratio of 57:12:12 as the MATLAB NN toolbox defaulted. Four types of hidden-layer architecture were designated in the first-stage training including 5 × 5, 10 × 10, 15 × 15, and 20 × 20. The experimental results are listed in Table 5. The terminal moment was determined by the performance of the validation sets. The training of the four networks stopped at the 18th, 33rd, 11th, and 16th epochs, respectively, when their MSEs did not improve during 6 continuous epochs. From the MSE of the training sets, both models of 10 × 10 and 15 × 15 had better performance whose MSEs were below 0.0065.

Model’s performance in the output simulations was evaluated by the correlation coefficient (R), discussed by training sets, validation sets, testing sets, and all input patterns, respectively. The R value of the hidden-layer neurons of 5 × 5 of the training sets achieved 0.79039, but that of the all patterns was 0.66139 only. The simulation precision was insufficient. While the hidden-layer neurons increased to 10 × 10, 15 × 15, and 20 × 20, the R values of the all patterns were all above 0.7. Regression plots of the four NN models are shown in Appendix 1.

5.1.2 Differential analyses for the additional training sets

When the network training is terminated, the simulated values of the output-layer variables can be derived through a simple NN’s calculation. Because that the network is trained by the training sets sampled from the drilling trials of the L27 OA experiment only, learning information of the model is limited. In order to enhance the simulation precision, the differential analysis of each training set is executed individually for calculating the deviation between the simulated output and its expected value. Moreover, these data can be used to compare the learning outcomes of the models trained by different hidden-layer architectures.
Based on the results of differential analysis, the worse five parameter sets in L27 OA were determined as the datum sets for planning the additional drilling trials. They were Trial-22, Trial-2, Trial-24, Trial-4, and Trial-14 as listed in Table 6. Since the spindle speed (Factor A) and feed rate (Factor C) were two significant factors in factorial level analysis (refer Fig. 8), four additional drilling trials were designed for each datum set, with a spindle speed interval of ±500 rpm and a feed rate interval of ±50 mm/min. Accordingly, twenty drilling trials were planned. After actual drilling works, the thrust force and hole characteristics were measured as the variable data of the additional training sets for the second-stage training.

### 5.1.3 Trials for the model’s testing sets

For a NN model, a higher $R$ value indicates that the model’s simulation for the training sets is good, but it cannot ensure the prediction reliability of the non-trained cases. To improve the accuracy in the model’s prediction, in addition to the five datum sets decided for the additional drilling trials, two datum sets that learning underperformed in the first-stage are more selected for planning the model’s testing sets. In this case, Trial-20 and Trial-13 were chosen, and both designated four drilling trials as well, as Table 7 presents. They are specified for the model prediction test only, not involved in the network training. Figure 9 shows the results of RMSE calculated by the model’s testing sets with four different models. The predicted errors of the diameter shrinkage were larger than the roundness. The model trained by the hidden-layer of 5×5 had a better performance, but only 57 sets of training data were used in training; the information for the network was insufficient. It still needed a network refining procedure to improve the model’s prediction accuracy.

| No  | Datum sets | Input variables | Output variables |
|-----|------------|-----------------|-----------------|
| Add-1 | Trial-22 | 4500 0.4 150 2.50 | 8.1 6.3 |
| Add-2 | Trial-22 | 4500 0.4 250 3.51 | 8.2 5.1 |
| Add-3 | Trial-22 | 4500 0.4 150 2.23 | 7.7 5.8 |
| Add-4 | Trial-22 | 4500 0.4 250 3.06 | 7.2 4.8 |
| Add-5 | Trial-2 | 2500 0.3 250 6.10 | 9.0 5.8 |
| Add-6 | Trial-2 | 2500 0.3 350 8.25 | 11.2 6.3 |
| Add-7 | Trial-2 | 3500 0.3 250 4.47 | 9.9 6.9 |
| Add-8 | Trial-2 | 3500 0.3 350 6.27 | 10.2 5.1 |
| Add-9 | Trial-24 | 4500 0.4 350 4.80 | 8.8 6.8 |
| Add-10 | Trial-24 | 4500 0.4 450 6.21 | 8.8 6.9 |
| Add-11 | Trial-24 | 5500 0.4 350 3.92 | 9.4 7.5 |
| Add-12 | Trial-24 | 5500 0.4 450 5.19 | 8.0 5.6 |
| Add-13 | Trial-4 | 2500 0.4 150 3.54 | 11.2 6.9 |
| Add-14 | Trial-4 | 2500 0.4 250 5.76 | 9.8 6.7 |
| Add-15 | Trial-4 | 3500 0.4 150 2.96 | 8.7 5.2 |
| Add-16 | Trial-4 | 3500 0.4 250 4.28 | 8.9 8.3 |
| Add-17 | Trial-14 | 3500 0.4 250 4.21 | 7.8 6.7 |
| Add-18 | Trial-14 | 3500 0.4 350 6.15 | 8.9 6.2 |
| Add-19 | Trial-14 | 4500 0.4 250 3.50 | 8.2 5.5 |
| Add-20 | Trial-14 | 4500 0.4 350 4.76 | 11.0 5.7 |


Table 7 Variable data of the model’s testing sets

| No  | Datum sets | Input variables | Output variables |
|-----|------------|-----------------|------------------|
|     |            | Spindle speed (rpm) | Depth of peck-drilling (mm) | Feed rate (mm/min) | Avg. thrust force peaks (N) | \(\Delta D\) (μm) | Roundness (μm) |
| Test-1 | Trial-20 | 4500 | 0.3 | 250 | 3.60 | 8.9 | 4.8 |
| Test-2 | Trial-20 | 4500 | 0.3 | 350 | 4.92 | 8.0 | 5.2 |
| Test-3 | Trial-20 | 5500 | 0.3 | 250 | 3.19 | 10.0 | 5.7 |
| Test-4 | Trial-20 | 5500 | 0.3 | 350 | 4.24 | 9.4 | 5.1 |
| Test-5 | Trial-13 | 3500 | 0.4 | 150 | 2.81 | 9.3 | 5.5 |
| Test-6 | Trial-13 | 3500 | 0.4 | 250 | 4.26 | 9.3 | 4.9 |
| Test-7 | Trial-13 | 4500 | 0.4 | 150 | 2.53 | 8.4 | 3.6 |
| Test-8 | Trial-13 | 4500 | 0.4 | 250 | 3.59 | 8.9 | 4.6 |

5.2 The second-stage training

5.2.1 Network refining experiments

In the second-stage refining experiment, in addition to the 81 trials from the \(L_{27}\) OA experiment, 20 additional drilling trials were included. After the input pattern transformation, 101 training sets were prepared for NN refining experiments. In terms of the network architecture, because the \(R\) value of the model of hidden-layer 5×5 evaluated by all patterns was only 0.66139 in the first-stage training (see Table 5), the simulation precision was unsatisfied due to the insufficiently connected weights among the neurons. Hence, only three hidden-layer architectures of 10×10, 15×15, and 20×20 were experimented here. The outcome of network learning can be estimated by the MSE of the network’s training sets. A larger MSE would result in a worse precision; however, a smaller MSE may incur the model to be overfitting. In this stage, three converging conditions for training termination were designated, included MSE < 0.001, MSE < 0.005, and MSE < 0.010. As a result, 9 models were built after the network experiments. The training data and models’ performance are presented in Table 8. The \(R\) values of the three models trained by the condition of MSE < 0.001 were all above 0.98. This indicated that the simulated outputs were very close to the expected value having a good performance in the training sets simulation. When the converging condition enlarged to MSE < 0.005, the \(R\) values were slightly decreased near 0.92 (0.92601, 0.92612, 0.92650). However, when the MSE more increased to MSE < 0.010, although the terminal epochs of the three models were only 24, 9, and 5, respectively, their \(R\) values were dropped below 0.85 (0.85114, 0.84994, 0.85385). The learning performance was fair. As to the training efficiency, for the hidden-layer neurons of 15×15 and 20×20, the maximum training epochs to satisfy the converging conditions were about 300 which were lower than the network of 10×10 obviously (see Table 8). The more neurons used in the hidden layer can acquire higher training efficiency. Regression plots of the 101 training sets by the 9 NN models are shown in Appendix 2.

5.2.2 Analyses of the model’s testing sets

After training, import the 8 model’s testing sets into the 9 models obtained from the NN experiments to derive their predictions of the output variables. The prediction accuracy of each model was evaluated by the RMSE as shown in Fig. 10. The models trained by the converging condition of MSE < 0.001 presented high RMSEs. This indicated that these models could not derive accurate predictions for the non-trained cases (the model’s testing sets), and they were in an overfitting state. However, for the three models whose converging condition loosened to MSE < 0.010, their RMSEs were descended obviously. These models had a better performance in predictions, but their accuracy was still above 1 μm. More training sets were required to provide field information for network learning.

5.2.3 Differential analyses for the extra training sets

From the previous subsection, when the converging condition was specified at MSE < 0.010, the RMSEs were the smallest, and no overfitting occurred in the predictions of the non-trained cases. Therefore, the differential analyses were executed based on the simulated outputs derived by the models of MSE < 0.010. From the 101 training set, five training
sets were decided as the datum sets for planning the extra drilling sets required for the next stage. They were Trial-18, Add-20, Add-17, Add-19, and Add-18. Since some datum sets had been experimented in the second-stage (coded in Add-), but their learning outcomes were still underperformed. More information was required in these cases. Hence, the intervals setting for the extra drilling trials were only half of that in the previous stage, that is, 250 rpm for the spindle speed and 25 mm/min for the feed rate. Similarly, four trials were designated for each datum set, and the extra drilling trials are presented in Table 9. After actual experiments, characteristic values of the thrust force, diameter shrinkage, and roundness were measured and prepared as the extra training sets used in the third-stage training.

5.3 The third-stage training

5.3.1 Network refining experiments

In the third-stage training, there were 121 training sets including 81 sets obtained from Taguchi’s L$_{27}$ OA experiment, 20 sets from the additional drilling trials, and 20 sets from the extra drilling trials. Because the models trained by the converging condition of the MSE $< 0.001$ had obvious overfitting in the second-stage training, only two conditions of MSE $< 0.005$ and MSE $< 0.010$ remained. In addition, a condition of MSE $< 0.0125$ was added for observing the effect of enlarging the MSE in training. In terms of network architecture, the setting was the same as the second-stage. Table 10 presents the training data of nine models in the third-stage refining experiments. From the $R$, the converging condition of MSE $< 0.005$ presented the best learning results for the training sets as well. The $R$ values were all above 0.92 for the three architectures. However, when the converging condition specified to MSE $< 0.0125$, only 20, 11, and 8 epochs were required to terminate the training procedure, respectively. But the $R$ values were dropped near 0.81 (0.81293, 0.81700, 0.81269). Compared with the second-stage training, the number of training sets was increased from 101 to 121, the field information of the network model was more integrated, while the training cycles needed for the models of 15$\times$15 and 20$\times$20 were around 40 epochs to meet the converging condition. The training efficiency was improved.

![Fig. 10 RMSEs of the model’s testing sets evaluated by the second-stage models](image)
5.3.2 Analyses of the model’s testing sets

After the network refining experiments, the predicted outputs of the model’s testing sets can be derived by the 9 NN models. Figure 11 shows the RMSEs of the models estimated by the model’s testing sets. Although the prediction accuracy of the models trained by the MSE < 0.005 was still large, the RMSEs of the others models derived by MSE < 0.010 and MSE < 0.0125 were all below 2 μm and no overfitting occurred. The optimal model for both characteristics was trained by the hidden-layer architecture of 20 × 20 with the converging condition of MSE < 0.0125. The RMSEs for the diameter shrinkage and roundness were 0.821 and 0.635 μm, respectively.

5.4 Discussion

In the network training, the model’s precision can be estimated by the $R$ value based on the simulated outputs of the training sets. However, the prediction accuracy for the non-trained sets (model’s testing sets) may be estimated by the average RMSE of the output variables examined. Two hole’s characteristics discussed in this study; thus, we have

$$\text{RMSE}_{\text{avg}} = \frac{\text{RMSE}_{\text{DS}} + \text{RMSE}_{\text{R}}}{2},$$

(9)

where RMSE_{DS} is the RMSE of the hole shrinkage in diameter and RMSE_{R} is that of the roundness.

| Table 9 | Variable data of the extra training sets for the third-stage training |
|---|---|---|---|---|
| No | Datum sets | Input variables | Output variables |
| | | Spindle speed (rpm) | Depth of peck drilling (mm) | Feed rate (mm/min) | Avg. thrust force peaks (N) | ΔD (μm) | Roundness (μm) |
| Extra-1 | Trial-18 | 3500 | 0.5 | 350 | 6.06 | 13.5 | 4.6 |
| Extra-2 | Trial-18 | 3500 | 0.5 | 450 | 7.81 | 8.2 | 5.1 |
| Extra-3 | Trial-18 | 4500 | 0.5 | 350 | 4.79 | 6.5 | 4.6 |
| Extra-4 | Trial-18 | 4500 | 0.5 | 450 | 6.35 | 7.7 | 4.4 |
| Extra-5 | Add-20 | 4250 | 0.4 | 325 | 4.72 | 7.2 | 4.1 |
| Extra-6 | Add-20 | 4250 | 0.4 | 375 | 5.54 | 8.9 | 3.8 |
| Extra-7 | Add-20 | 4750 | 0.4 | 325 | 4.30 | 8.4 | 4.4 |
| Extra-8 | Add-20 | 4750 | 0.4 | 375 | 5.01 | 8.7 | 4.1 |
| Extra-9 | Add-17 | 3250 | 0.4 | 225 | 4.11 | 8.4 | 4.6 |
| Extra-10 | Add-17 | 3250 | 0.4 | 275 | 4.93 | 6.4 | 3.0 |
| Extra-11 | Add-17 | 3750 | 0.4 | 225 | 3.64 | 6.7 | 3.4 |
| Extra-12 | Add-17 | 3750 | 0.4 | 275 | 4.23 | 7.4 | 4.2 |
| Extra-13 | Add-19 | 4250 | 0.4 | 225 | 3.33 | 7.2 | 5.0 |
| Extra-14 | Add-19 | 4250 | 0.4 | 275 | 3.82 | 13.7 | 5.2 |
| Extra-15 | Add-19 | 4750 | 0.4 | 225 | 3.21 | 6.9 | 4.4 |
| Extra-16 | Add-19 | 4750 | 0.4 | 275 | 3.60 | 6.9 | 4.1 |
| Extra-17 | Add-18 | 3250 | 0.4 | 325 | 5.98 | 8.7 | 3.9 |
| Extra-18 | Add-18 | 3250 | 0.4 | 375 | 6.92 | 7.4 | 3.8 |
| Extra-19 | Add-18 | 3750 | 0.4 | 325 | 5.24 | 9.4 | 4.4 |
| Extra-20 | Add-18 | 3750 | 0.4 | 375 | 5.80 | 14.1 | 5.8 |

| Table 10 | Training data in the third-stage training |
|---|---|---|---|---|
| Neurons of hidden layer$^*$ | Converging condition (MSE) | Number of training cycles (epochs) | MSE of network’s training sets | Correlation coefficient ($R$) |
| 10 × 10 | <0.005 | 141 | 0.005 | 0.92897 |
| | <0.010 | 28 | 0.00993 | 0.85313 |
| | <0.0125 | 20 | 0.0124 | 0.81299 |
| 15 × 15 | <0.005 | 41 | 0.005 | 0.92906 |
| | <0.010 | 26 | 0.00991 | 0.85372 |
| | <0.0125 | 11 | 0.0122 | 0.81700 |
| 20 × 20 | <0.005 | 39 | 0.005 | 0.92906 |
| | <0.010 | 9 | 0.00999 | 0.85299 |
| | <0.0125 | 8 | 0.0125 | 0.81700 |
Table 11 summarizes the models’ performance during the three training stages. Several points are found:

1. In the first-stage training, only 57 training sets were used. The $R$ values of all models were below 0.75. As to the second-stage training, the number of training sets increased to 101. The $R$ values of the three models trained by the converging condition of MSE $< 0.001$ were high to 0.98. This indicates that the model’s simulation precision for the training sets was excellent; nevertheless, their predicted errors (RMSE$_{avg}$) were so large for the non-trained cases (model’s testing sets). These models were in an overfitting state.

2. For the models terminated at MSE $< 0.005$, in the second- and third-stage training, their $R$ values were near 0.92. Although their simulation precisions ($R$ value) were slightly decreased to the models trained by MSE $< 0.001$, their predicted errors (RMSE$_{avg}$) were improved. When the converging condition enlarged to MSE $< 0.010$, the $R$ values were dropped again, near 0.85; however, the RMSE$_{avg}$ were more improved to 1.2794 and 1.3797 μm, respectively. As to the three models trained by the condition of MSE $< 0.0125$, their $R$ values were near 0.81 in the third-stage training.

3. In this case analyzed, the optimal model for output variables prediction was trained by the hidden-layer neurons of 20×20 with the converging condition of MSE $< 0.0125$. The average RMSE of the model’s testing sets (non-trained cases) was 0.7281 μm.

6 Conclusions

Hole shrinkage always produces in the drilling of engineering plastics; as a result, the dimensions and profile of the hole drilled are hard to control. This study conducted the Taguchi L$_{27}$ OA experiment of the PEEK-MDS 100 to derive the optimal parameter conditions for minimizing the thrust force and lowering the hole’s shrinkage and roundness in the drilling of a 1-mm-diameter hole. In addition, a stepped-learning procedure is presented for developing a Taguchi-based NN model that can save the experimental data, simulate the hole’s characteristics of the network’s training sets, and predict the non-trained cases. The field information of the network model was expanded by three stages of training, and the model’s prediction accuracy was improved gradually. The ANN models for a small-hole drilling process of the thermoplastic PEEK were successfully built.
Appendix 1

Fig. 12 Models’ performance presented by correlation coefficient in the first-stage training: (a) $4 \times 5 \times 5 \times 2$, (b) $4 \times 10 \times 10 \times 2$, (c) $4 \times 15 \times 15 \times 2$, and (d) $4 \times 20 \times 20 \times 2$
Appendix 2

Fig. 13 Models’ performance presented by correlation coefficient in the second-stage training: (a) $4 \times 10 \times 10 \times 2$, (b) $4 \times 15 \times 15 \times 2$, and (c) $4 \times 20 \times 20 \times 2$
Author contribution All authors participated in the writing, read and approved the final manuscript.

Data availability All the data have been presented in the manuscript.

Declarations

Ethical approval Not applicable.

Consent to participate The authors declare that they all consent to participate this research.

Consent for publication The authors declare that they all consent to publish the manuscript.

Competing interests The authors declare no competing interests.

References

1. Choi YJ, Park MS, Chu CN (2008) Prediction of drill failure using features extraction in time and frequency domains of feed motor current. Int J Mach Tools Manuf 48(1):29–39. https://doi.org/10.1016/j.ijamttools.2007.08.009
2. Esim E, Yildirim Ş (2017) Drilling performance analysis of drill column machine using proposed neural networks. Neural Comput & Appl 28:79–90. https://doi.org/10.1007/s00521-016-2322-8
3. Roth JT, Djurdjanovic D, Yang X, Mears L, Kurfess T (2010) Quality and Inspection of Machining Operations: Tool Condition Monitoring. ASME J Manuf Sci Eng 132(4):041015–41016. https://doi.org/10.1115/1.4002022
4. Hyacinth Sanguthi X, Natarajan U, Ramasubbu N (2015) A review of accuracy enhancement in microdrilling operations. Int J Adv Manuf Technol 81:199–217. https://doi.org/10.1007/s00170-015-6900-1
5. Zerti A, Yallese MA, Meddour I, Belhadi S, Haddad A, Mabrouki T (2019) Modeling and multi-objective optimization for minimizing surface roughness, cutting force, and power, and maximizing productivity for tempered stainless steel AISI 420 in turning operations. Int J Adv Manuf Technol 102:135–157. https://doi.org/10.1007/s00170-018-2984-8
6. Özden G, Öteyaka MO, Cabrera FM (2021) Modeling of cutting parameters in turning of PEEK composite using artificial neural networks and adaptive-neural fuzzy inference systems. J Thermoplast Compos Mater Online First. https://doi.org/10.1177/0892705721103070
7. Daniel SAA, Pugazhenthri R, Kumar R, Vijayananth S (2019) Multi objective prediction and optimization of control parameters in the milling of aluminium hybrid metal matrix composites using ANN and Taguchi-grey relational analysis. Def Technol 15(4):545–556. https://doi.org/10.1016/j.dt.2019.01.001
8. Yadav RN, Yadava V, Singh GK (2013) Multi-objective optimization of process parameters in Electro-Discharge Diamond Face Grinding based on ANN-NSGA-II hybrid technique. Front Mech Eng 8:319–332. https://doi.org/10.1007/s11465-013-0269-3
9. Biswas R, Kuar AS, Biswas SK, Mitra S (2009) Artificial neural network modeling of Nd:YAG laser microdrilling on titanium nitride alumina composite. Proc Inst Mech Eng B J Eng Manuf 224(3):473–482. https://doi.org/10.1243/09544054JEM1576
10. Sarkanay M, Yilmaz V (2018) Optimization and predictive modeling using S/N, RSM, RA and ANNs for micro-electrical discharge drilling of AISI 304 stainless steel. Neural Comput & Applic 30:1503–1517. https://doi.org/10.1007/s00521-016-2775-9
11. Kumar SPL, Jerald J, Kumanan S, Prabakaran RA (2014) Review on current research aspects in tool-based micromachining processes. Mater Manuf Process 29(11–12):1291–1337. https://doi.org/10.1080/10426914.2014.952037
12. Harun MHS, Ghazali MF, Yusoff AR (2017) Analysis of tri-axial force vibration sensors for detection of failure criterion in deep twist drilling process. Int J Adv Manuf Technol 89:3535–3545. https://doi.org/10.1007/j.composites.2012.10.038
13. Singh AP, Sharma M, Singh I (2013) A review of modeling and control during drilling of fiber reinforced plastic composites. Composites: Part B 47:118–125. https://doi.org/10.1016/j.compositesb.2012.10.038
14. Patra K, Jha AK, Szalay T, Ranjan J, Monostori L (2017) Artificial neural network based tool condition monitoring in micro mechanical peck drilling using thrust force signals. Precis Eng 48:279–291. https://doi.org/10.1016/j.precisioneng.2016.12.011
15. Eftholidis N, Hernández CG, Talón JH, Kyратис P (2018) Modeling and prediction of thrust force and torque in drilling operations of Al7075 using ANN and RSM methodologies. J Mech Eng Sci 64(6):351–361. https://doi.org/10.15545/sjme.2017.5188
16. Kharwar PK, Verma RK, Singh A (2020) Neural network modeling and combined compromise solution (CoCoSo) method for optimization of drilling performances in polymer nanocomposites. J Thermoplastic Composite Materials, Online first. https://doi.org/10.1177/0892705720939165
17. Corne R, Nath C, Mansori MEI, Kurfess T (2017) Study of spindle power data with neural network for predicting real-time tool wear/ breakage during inconel drilling. J Manuf Syst 43(2):287–295. https://doi.org/10.1016/j.jmsy.2017.01.004
18. Soepangkat BOP, Norcahyo R, Khoirul Effendi M, Pramujati B (2020) Multi-response optimization of carbon fiber reinforced polymer (CFRP) drilling using back propagation neural network-particulate swarm optimization (BPNN-PSO). Engineering Science and Technology, an International Journal 23(3):700–713. https://doi.org/10.1016/j.jestch.2019.10.002
19. Ulas HB, Ozkan MT, Malkoc Y (2019) Vibration prediction in drilling processes with HSS and carbide drill bit means by means of artificial neural networks. Neural Comput & Applic 31:5547–5562. https://doi.org/10.1007/s00521-018-3379-3
20. Caggiano A, Angelone R, Napolitano F, Nele L., Teti R (2018) Dimensionality reduction of sensorial features by principal component analysis for ANN machine learning in tool condition monitoring of CFRP drilling. Procedia CIRP 78:307–312. https://doi.org/10.1016/j.procir.2018.09.072
21. Vrabel M, Maňková I, Beňo J (2016) Monitoring and control of manufacturing process to assist the surface workpiece quality when drilling. Procedia CIRP 41:735–739. https://doi.org/10.1016/j.procir.2015.12.092
22. Akınçoğlu S, Mendi F, Çiçek A, Akınçoğlu G (2013) ANN-based prediction of surface and hole quality in drilling of AISI D2 cold work tool steel. Int J Adv Manuf Technol 68:197–207. https://doi.org/10.1007/s00170-012-4719-6
23. Cruz CED, Aguiar PR, Machado AR, Bianchi EC, Contrucci JG, Neto FC (2013) Monitoring in precision metal drilling process using multi-sensors and neural network. Int J Adv Manuf Technol 66:151–158. https://doi.org/10.1007/s00170-012-4314-x
24. Mondal N, Mandal S, Mandal MC (2020) FPA based optimization of drilling burr using regression analysis and ANN model. Measurement 152:107327. https://doi.org/10.1016/j.measurement.2019.107327
25. Ahn Y, Lee SH (2017) Classification and prediction of burr formation in micro drilling of ductile metals. Int J Prod Res 55(17):4833–4846. https://doi.org/10.1080/00207543.2016.1254355
26. Freeman JA, Skapura DM (1991) Neural networks: algorithms, application, and programming techniques. Addison-Wesley, Massachusetts (chapter 3)

27. Fowlkes WY, Creveling CM (1984) Engineering methods for robust product design: using taguchi methods in technology and product development. Addison-Wesley, Massachusetts, chapter 4 and chapter 5

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.