Application of Artificial Neural Networks in Prediction Tool Life of PVD Coated Carbide When End Milling of Ti6Al4V Alloy

Salah Al-Zubaidi, Jaharah A. Ghani, and Che Hassan Che Haron

Abstract—Nowadays, the application of artificial neural networks (ANN) is often utilized in solving numerous problems in machining processes. There has been evidence of the significance of a tool life prediction of coated and uncoated cutting tools. The current study aims at applying ANN in the prediction of the tool life of PVD cutting tools using low experimental data sets. It used a feed forward back propagation neural network with a Levenberg-Marquard (L-M) training algorithm is used in modeling the tool life of a PVD insert cutting tool when end milling of Ti6Al4V under dry cutting conditions. One hundred and ten (110) models were designed, trained and tested using Matlab neural network tool box. Based on the same experimental data, a regression model (RM) has been constructed employing SPSS software, and based on the mean square error of ANN and RM models, the two models were compared. The findings revealed that the ANN model resulted into minimum mean square error compared with RM model.

Key-words: artificial neural network, coated carbide, end milling, prediction, Ti6Al4V alloy, tool life

I. INTRODUCTION

The importance of machining cost and product quality as two factors is realized in predicting tool life in the end milling process. In evaluating the performance of a machining process, tool wear/tool life is regarded as an important aspect. Moreover, it is stated that tool wear/tool life predictions and the corresponding economic analysis are two of the most important topics in process planning and machining optimization [1]. Factors such as cutting conditions, tool material, type and mode of milling, tool geometry...etc., as displayed in Fig. 1, have effect on the tool life in end milling.

Developing models for the execution is possible to be achieved through using analytical, statistical & mathematical, or artificial intelligence techniques.

Fig. 1 Factors affecting tool life of cutting tool [2].

The process of machining has been proved to be a complex or sophisticated process in previous studies. This indicates that it is characterized as a non linear problem. It is stated that any approach employed or applied for the purpose of modeling such processes shows a relation between inputs parameters and response. In general, although the techniques being applied maybe different, each model relates to the input parameters with the output responses. However, the problem in using the conventional approaches is that they cannot be representatives of the nonlinearity of these kinds of problems. Here, at this particular point, the importance of
artificial intelligence (A.I) methods is emphasized as they are more capable of doing this specific function, thus, meeting the need and achieving this purpose. Generally pointing out the features of all A.I methods, it is stated that they are considered as a mimic and simulation of nervous cell [3-5]. Moreover, the literature has proved that the artificial neural network (ANN) method is regarded as the most well-known method among all A.I methods. This is because it is the most frequently used technique which has been used by researchers for modelling linear and non-linear problems. Furthermore, A.I methods are more capable of mapping non-linear input-output relation than statistical techniques. Due to the evidence of the importance of artificial neural network (ANN) as an artificial intelligence approach, the current utilized A.I. in predicting tool life as one of the uncoated carbide cutting tools and its role in ending milling of Ti6Al4V alloy under dry cutting conditions.

Concerning the tool wear, it is pointed out that this tool wear leads to degradation of the quality of machined part being cut, so it can be modeled employing or making a use of the ANN. the majority of the previous studies and investigations have been conducted on surface roughness in comparison to those studies on tool wear and life especially in milling process [6,7]. For instance, R.K. Dutta et al. [8] conducted a study examining the convergence speed and prediction accuracy of modified back propagation in comparison to normal back propagation in monitoring the tool conditions. The findings of the study showed that the momentum factor significantly affects the convergence speed more than the learning rate does. In a study by Chen & Chen [9], an online tool wear prediction system was presented and the ANN was used. Moreover, 100 and 9 experimental data sets were applied in training and testing the feed forward back propagation. In this study, input parameters were average peak force in Y axis, feed rate and depth cut. The investigation obtained good results by using this system. Another study by Palanisamy et al. [10] investigated the application of both of regression model (RM) and the ANN model for tool wear prediction in end milling AISI 1020 steel with carbide inserts. The design of the experiments was carried out based on three full factorials which have five levels namely; cutting speed, feed rate, and depth of cut which constituted the input parameters whereas the output was flank wear. The findings obtained from the experiments revealed that the ANN model generated accurate results than those results obtained through the RM. In addition, C. Bruni et al. [11] enhanced ANN and multiple regression models to be used or adopted by studies on surface roughness and tool wear prediction in end milling AISI 420B stainless steel under different cutting and lubricant conditions ((dry, wet, and minimum lubricant quantity MQL). It was found that under high cutting speed with MQL, minimum surface roughness and wear can be obtained.

It has been proved that artificial intelligence (AI) models are more capable of modeling nonlinear and complex problems like metal cutting than other approaches. One of the more important techniques of AI is stated to be artificial neural network (ANN) is. It is defined as a computing method which is employed for solving many problems, and it mimics the nervous cell. Most of the previous literature focuses on modeling surface roughness by employing the ANN compared to other responses like cutting forces and tool life, especially for super alloys [4]. Therefore, the aim of the present study is to apply ANN in predicting the tool life of PVD cutting tools when end milling of Ti6Al4V alloy under dry cutting conditions using low experimental data sets.

II. FEED FORWARD BACK PROPAGATION NETWORK

The development of a feed forward back propagation network was dated back to 1970 by Rumelhart, Hinton, and Williams [12]. Currently, this network is regarded as the most well known one among other networks because it is more efficient and accurate than others [6, 13]. As depicted in Figure 2, the architecture of the back propagation neural network consists of input, hidden and output layers where each layer has a number of processing elements called neurons. The number of neurons at the input and output are determined by the problem parameters. However, it is pointed out that researchers can be free to choose the number of hidden layers and their neurons because there is no clear-cut method in selecting network parameters [14]. Feed forward back propagation neural networks are regarded as supervised networks providing both the inputs and outputs. Fig. 3 displays the flowchart of training the supervised network.
As described in the Neural Network Tool Box User Guide [15], the following points are done in a sequence for designing the neural networks:

1. collecting the data
2. creating the neural network
3. configuring the neural network
4. initializing the weights and biases
5. training the network
6. validating the network
7. using and applying the network.

In carrying out training BPNN, the following four stages need to be followed by a researcher [16]:

1. Initialization of weights
2. Feed forward phase
3. Back propagation phase for error calculations
4. Tuning of weights and biases

Selecting low and random weight values needs to be done by a researcher in order to avoid saturation during training.

After setting the various weights, it is noticed that each neuron receives an input signal $X_i$, which sends it directly to all hidden neurons in the hidden layer, and which sum up their weighted input signals as expressed in Equation (1) below:

$$h_k = \sum_{j=1}^{I} W_{j,k} \times X_j + b_{hk}$$  \hspace{1cm} (1)$$

Selecting low and random weight values needs to be done by a researcher in order to avoid saturation during training.

An activation function must be applied for this net:

$$H_k = f(h_k)$$  \hspace{1cm} (2)$$

Consequently the output layer receives this signal and its neurons sum up their weighted inputs, as in the hidden layer

$$y_z = \sum_{k=1}^{K} W_{k,z} \times h_k + b_{yz}$$  \hspace{1cm} (3)$$
The activation function must be applied to \( y_z \) in order to obtain a network output for Equation (4)

\[
y_z = f(yz) \tag{4}
\]

At the end of the feed forward phase this output is compared with the target to calculate the errors in the back propagation phase as below:

\[
\delta_z = (yz - y) f'(y) \tag{5}
\]

Then delta inputs (\( \delta_k \)) for each hidden neuron can be calculated from the output layer neurons.

\[
\delta_k = \sum_{z=1}^{K} \delta_i W_{i z} \tag{6}
\]

After that, the error can be determined:

\[
\delta_k = \delta f(h) \tag{7}
\]

Consequently, both the weights and biases of the output and hidden layers are updated, respectively.

The correction of weights and biases between hidden and output layers is calculated by:

\[
\Delta W_{i z} = \alpha \delta_k h_i \tag{8}
\]

\[
\Delta b_{i z} = \alpha \delta_z \tag{9}
\]

Finally, new weights and biases can be updated:

\[
W_{i z}(\text{new}) = W_{i z}(\text{old}) + \Delta W_{i z} \tag{10}
\]

\[
 b_{i z}(\text{new}) = b_{i z}(\text{old}) + \Delta b_{i z} \tag{11}
\]

Terminating the training takes place in one of the two cases, either having training epochs reached, or having the calculated error being equal to the goal error, after which testing the network with new data pairs to check its generalization has to be conducted.

III. THE MULTIPLE LINEAR REGRESSION MODELS

Regression analysis is defined as a statistical technique which is usually used for modeling and examining the relationship between two or more variables. It is pointed out that in a simple linear regression model, there is only one independent variable. However, sometimes, many diverse applications may have many different independent factors affecting the outcome of a process [17]. In such a case, the model used for analyzing is called a multiple regression model which is expressed as follows:

\[
y = \beta_0 + \sum_{j=0}^{K} \beta_j x_j \tag{12}
\]

\( K \) is used to symbolize the independent variables and the parameters \( \beta_j, j=0, 1, 2, \) and \( k \) are referred to as the regression coefficients. Thus, the analysis of variance (ANOVA) is considered as a well-known commonly used statistical technique for determining the percent contribution of each parameter for:

1. The total sum of squares,

\[
\sum_{i=1}^{n} (y_i - y)^2 \tag{13}
\]

is a chi-squared random variable with \( n-1 \) degrees of freedom

2. The sum of squares due to regression,

\[
\sum_{i=1}^{n} (\hat{y}_i - y)^2 \tag{14}
\]

is a chi-squared random variable with 1 degree of freedom

3. The error sum of squares,

\[
\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{15}
\]

is a chi-squared random variable with \( n-2 \) degrees of freedom

4. The sums squares and are independently distributed.

5. The ratio between the mean square for regression and the mean square for error

\[
\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2/(n-2)} \tag{16}
\]

follows the F (1, \( n-2 \)) distribution, [18].

IV. ANN MODELLING

The current study adopted the work by Nagi[19] as a case study for modeling the ANN in which Ti6Al4V alloy was machined with PVD-coated carbide employing a CNC Milling machine. Based on this work conducted by this researcher, a factorial design of the experiment which is integrated with response surface methodology to build mathematical models for surface roughness, cutting forces, and tool life prediction was used. Table (1) presents the mechanical features of Ti6Al4V alloy whereas Table (2) gives the geometry of the cutting tool insert.
### Table (1) Mechanical properties of Ti6Al4V alloy and uncoated carbide cutting tool

<table>
<thead>
<tr>
<th>Ultimate tensile strength MPa</th>
<th>Yield strength MPa</th>
<th>Rockwell Hardness HRC</th>
<th>Modulus of elasticity GPa</th>
<th>Poisson's ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>950</td>
<td>880</td>
<td>36</td>
<td>113.8</td>
<td>0.342</td>
</tr>
</tbody>
</table>

### Table (2) Tool geometry of uncoated carbide tool

<table>
<thead>
<tr>
<th>ISO grade K20</th>
<th>Insert cutting rake angle</th>
<th>Insert side clearance angle</th>
<th>Insert helix angle</th>
<th>Insert radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S20-S30)-XOMX090308TR ME06, F40 (PVD- Coated Carbide Ti-Al-N) with chamfer of 0.06 width at 4°</td>
<td>24°</td>
<td>11°</td>
<td>15°</td>
<td>8mm</td>
</tr>
</tbody>
</table>

It is obvious that, when compared to the data used in other previous studies, the data sets used in the present study are considered relatively low. Since titanium alloys are regarded difficult to machine and very expensive [19-21], it turns out to be almost difficult to conduct experiments with such material.

It is also stated that selecting the network parameters significantly affects the network performance. For the number of hidden layers and their neurons, they are stated to be related to the complexity and degree of non-linearity of the problem. So far, no obviously precise rules or guides for the selection of network parameters have been found in previous studies except some hints and notes referred to by some researchers which cannot be applied to all cases and problems [13]. Therefore, selecting the parameters required in any study mainly depends on trial and error.

For the present study, one hundred and ten (110) different topologies were created and used for later evaluation. Single and two hidden layers were used with the number of neurons ranging from 1 to 10. Tansigmoid and Purline are recognized as transfer functions which were employed in the hidden and output layers, respectively. The first function was commonly used and it's differentiated from the second one which was used in function approximation [15, 16, 22] problems. The neural network was trained by using the Levenberg-Marquard algorithm. Furthermore, maximum epochs were 10000. Normalization included all the input and output data sets so that computing problems could be avoided as much as possible. Cutting speed, feed rate, and depth of cut have been chosen as input parameters, while the output was tool life. The first fourteen data sets were used for training and the last three were used for testing. Thus, the neural networks were designed, configured, trained and tested using the Matlab neural network tool box.
After training and testing the 110 configurations were carried out, only four of them were chosen for the evaluation while the other remaining ones were discarded. Table (4) presents the actual and predicted values of the four ANN models. Based on these findings, some of the predicted values were found to be identical with the experimental values, thus, indicating that there was a kind of some agreement or consistency. For the memorization (accuracy) of the networks, it is gained or obtained in the training phase. To provide evidence of the effectiveness of these models in prediction, it was important to test them with the new data sets that were not used before in the training phase. Therefore, in the present study, the last three rows of data sets (See Table 3) were used in the testing phase to check the generalization of the developed models. After conducting testing the models with new pairs of input-output data, the evaluation of both models was carried out. For this purpose, the mean square error (MSE) was chosen as the evaluation criteria for selecting the best ANN model. As proved by the results presented in Table (5), the 3-9-1 obtained the minimum average mean square error (MSE). Moreover, the MSE of other networks (3-9-5-1, 3-9-6-1, and 3-10-3-1) were relatively close to each other. As a result of this, 3-9-1 (single hidden layer with nine neurons) was proved to be the best network. As displayed in Figure 4, the real and predicted data of the best model show that some of the values are close to each other. The current study also used the mathematical model given in Eq. (12) for the purpose of developing the RM as to evaluate the tool life values, and it is expressed as in the following equation:

\[ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \]  

(V)

Table (5) MSE of ANN models

<table>
<thead>
<tr>
<th>Structures</th>
<th>3-9-1</th>
<th>3-9-5-1</th>
<th>3-9-6-1</th>
<th>3-10-3-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.22812</td>
<td>2.26092</td>
<td>2.01012</td>
<td>2.07772</td>
</tr>
</tbody>
</table>

Where \( y \) symbolizes the predicted value of tool life, and the symbols \( x_1, x_2, x_3 \) reflect the cutting speed, feed rate, and depth of cut respectively whereas \( \beta_0, \beta_1, \beta_2 \) and \( \beta_3 \) represent the coefficients of these cutting parameters. Thus, the development of the RMs for PVD cutting tool presented in Eq. (17) was carried out based on the data for the real machining experiments as presented in Table 3 through using the SPSS software. Like ANN modeling, these data sets were divided into two subsets: fourteen (14) data sets for network training and the other three (3) sets for network testing.

Table (6) below shows the results in relation to the values of coefficients for the model parameters of PVD coated cutting tools.

By transferring the values of coefficients for PVD coated cutting tool from Table 6 into Eq. (17), the following the equations of the RM were generated:

\[ Y = 18.270 - 0.08x_1 - 42.791x_2 - 1.127x_3 \]  

(VI)

As a consequence, a comparison between the scores of tool life values of the experimental data (given in Table 3) and the tool life predicted values of the RM (eq. 18) as shown in Table (7) below was carried out. Table (8) shows the findings obtained from comparing the 3-9-1 ANN and RM models. As previously shown in the above, it was found that the MSE of the ANN model was proved to be higher than that of the RM in and testing phases. The MSE of the ANN model reached 0.22812 in testing. However, the MSE of the RM model was 1.374958. This implies that results generated from using the ANN model are much better those obtained through the RM. This is also an indication of the capability and effectiveness of this model in modeling the nonlinear problems which
were found to be higher and better than those of other conventional techniques.

Table 6 Coefficients values for PVD Tool

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>18.267</td>
<td>2.066</td>
</tr>
<tr>
<td>Cutting speed</td>
<td>-.080</td>
<td>.014</td>
</tr>
<tr>
<td>Feed rate</td>
<td>-42.791</td>
<td>7.725</td>
</tr>
<tr>
<td>Depth of cut</td>
<td>-1.127</td>
<td>.710</td>
</tr>
</tbody>
</table>

Table 7 Actual and RM predicted tool life results

<table>
<thead>
<tr>
<th>No.</th>
<th>Actual tool life [min.]</th>
<th>Predicted tool life (RM model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.5</td>
<td>6.6609</td>
</tr>
<tr>
<td>2</td>
<td>4.429</td>
<td>3.8974</td>
</tr>
<tr>
<td>3</td>
<td>2.672</td>
<td>3.95785</td>
</tr>
<tr>
<td>4</td>
<td>2.338</td>
<td>3.95785</td>
</tr>
<tr>
<td>5</td>
<td>7.238</td>
<td>6.72135</td>
</tr>
<tr>
<td>6</td>
<td>3.237</td>
<td>2.3818</td>
</tr>
<tr>
<td>7</td>
<td>0.492</td>
<td>1.19435</td>
</tr>
<tr>
<td>8</td>
<td>0.553</td>
<td>-0.3817</td>
</tr>
<tr>
<td>9</td>
<td>6.7226</td>
<td>5.59435</td>
</tr>
<tr>
<td>10</td>
<td>2.696</td>
<td>2.32135</td>
</tr>
<tr>
<td>11</td>
<td>3.34</td>
<td>3.95785</td>
</tr>
<tr>
<td>12</td>
<td>3.393</td>
<td>4.0183</td>
</tr>
<tr>
<td>13</td>
<td>1.449</td>
<td>1.2548</td>
</tr>
<tr>
<td>14</td>
<td>6.5</td>
<td>5.5339</td>
</tr>
</tbody>
</table>

Fig. 4 actual and predicted results of ANN models

Table (8) MSE of 3-9-1 ANN and RM models

<table>
<thead>
<tr>
<th>Models</th>
<th>3-9-1 /testing</th>
<th>RM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.22812</td>
<td>1.374958</td>
</tr>
</tbody>
</table>
VI. CONCLUSIONS

The previously presented and discussed results of the current study provided sufficient evidence of the higher capability and effectiveness of the ANN model in modeling and mapping the non linear input-output relations more than those of other traditional models. Moreover, the findings showed that the best neural network topologies chosen were (3-9-1); which reflects a single hidden layer with nine hidden neurons respectively. It was also found that the MSE of the ANN model was higher than that of the RM in testing phase since it reached 0.22812 while the MSE of the RM model was 1.374958. Despite the fact that the experiment was carried out based on low training and testing data, the effectiveness and capability of the ANN model predicting the tool life of uncoated carbide when end milling of Ti6Al4V under dry conditions in a good accuracy were evidently proved.

REFERENCES

4. M. Joseph Davidson, N. Selvakumar, “A Neural Network (NN) and Response Surface Methodology (RSM) based Prediction Model for Sintered Aluminium Forms”, presented at 2012 11th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering And Data Bases, Cambridge, UK.
5. Foad Qassemi, Ali Reza Ahmadi, Mostafa Alizadeh, Ahmad Qassemi, “Prediction of Influence of Doping of NaN03 on the Solid Phase Thermal Decomposition of Bitumen using Neural Networks”, presented at 2012 11th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering And Data Bases, Cambridge, UK.

Salah Al-Zubaidi is a BSc and MSc in Mechanical engineering from Baghdad university, Iraq. He is a PhD student at Mechanical and Material Engineering Department, Universiti Kebangsaan Malaysia. His current research interests are artificial intelligence and machining technology of aerospace materials.

Jaharah A.Ghani is a BEng and an MSc in Manufacturing Engineering from Leeds Metro University and Warwick University, UK, respectively. She received a PhD in Manufacturing Engineering from Universiti Malaya, Malaysia. Her current research interests are machining technology of tool steels, cast iron and aerospace materials.

Che Hassan Che Haron is a BS and an MS in Mechanical Engineering from University of Bridgeport, USA and University of Pittsburgh, Pennsylvania, USA, respectively. He received a PhD in Manufacturing Engineering from Coventry University, UK. His current research interests are machining technology of tool steel, titanium alloys and EDM.