A new approach for prediction of the wear loss of PTA surface coatings using artificial neural network and basic, kernel-based, and weighted extreme learning machine

Mustafa ULAS¹, Osman ALTAY¹, Turan GURGENC^{2,*}, Cihan ÖZEL³

¹ Department of the Software Engineering, Firat University, Elazig 23119, Turkey

² Department of the Automotive Engineering, Firat University, Elazig 23119, Turkey

³ Department of the Mechanical Engineering, Firat University, Elazig 23119, Turkey

Received: 28 June 2019 / Revised: 01 October 2019 / Accepted: 07 November 2019

© The author(s) 2019.

Abstract: Wear tests are essential in the design of parts intended to work in environments that subject a part to high wear. Wear tests involve high cost and lengthy experiments, and require special test equipment. The use of machine learning algorithms for wear loss quantity predictions is a potentially effective means to eliminate the disadvantages of experimental methods such as cost, labor, and time. In this study, wear loss data of AISI 1020 steel coated by using a plasma transfer arc welding (PTAW) method with FeCrC, FeW, and FeB powders mixed in different ratios were obtained experimentally by some of the researchers in our group. The mechanical properties of the coating layers were detected by microhardness measurements and dry sliding wear tests. The wear tests were performed at three different loads (19.62, 39.24, and 58.86 N) over a sliding distance of 900 m. In this study, models have been developed by using four different machine learning algorithms (an artificial neural network (ANN), extreme learning machine (ELM), kernel-based extreme learning machine (KELM), and weighted extreme learning machine (WELM)) on the data set obtained from the wear test experiments. The R^2 value was calculated as 0.9729 in the model designed with WELM, which obtained the best performance among the models evaluated.

Keywords: wear loss prediction; surface coating; plasma transferred arc welding; artificial neural network; extreme learning machine

1 Introduction

Industries such as transportation, power generation, and manufacturing are important to developed societies. In these industries, the surface of many moving machine parts interact with each other [1], so wear and friction are unavoidable [2]. Therefore, it is essential to make their surfaces resistant to wear [3].

Steel has been one of the most widely used building materials for many years owing to its low cost, simple manufacturing, and flexibility in design. However, the wear and surface properties of steel are not very

good. Therefore, it is necessary to improve the surface properties of steel parts intended to work in harsh environments involving high-wear conditions [4]. Surface coating is one of the most useful methods available to improve the surface properties of steel. Surface coating applications are used in a number of applications where metal surfaces will be exposed to high wear, such as crushers, metal-to-metal contact machinery, and grinders [5]. Iron-based alloys are frequently preferred for surface coatings because of their low cost and excellent performance [6]. FeCrC ferroalloys are often used to make wear-resistant

* Corresponding author: Turan GURGENC, E-mail: tgurgenc@firat.edu.tr

parts [7], as their cost is low and their mechanical properties are very useful [8]. When the surface of a metal is coated with FeCrC ferroalloy, some very hard carbide phases are formed in the coating layer, such as $(\text{Cr,Fe})_3\text{C}$, $(\text{Cr,Fe})_7\text{C}_3$, and $(\text{Cr,Fe})_{23}\text{C}_6$. These carbides increase the hardness of the coating and the wear resistance of the coated surface [9]. The hardness properties of FeCrC ferroalloys can be improved by adding FeB and FeW ferroalloys, melted and deposited together on a material's surface [10, 11].

The surface of AISI 1020 can be coated using different melting welding methods, such as gas tungsten arc welding (GTAW) [12], laser welding [13], shielded metal arc welding (SMAW) [14], and plasma transferred arc welding (PTAW) [15]. Among these methods, PTAW has some advantages such as high deposition rate and low heat input [16], high temperature (over 40,000 K) [17], excellent arc stability [18], and low thermal distortion of the parts as well as low environmental impact [19].

Veinthal et al. [20] coated the surface of 1.0037 mild steel with FeCrC using the PTAW method and investigated abrasive impact and surface fatigue wear behavior. Hornung et al. [21] coated the surface of 1.0037 mild steel with FeCrC using the PTAW method and investigated the wear behavior of coated surfaces. Gur et al. [22] coated the surface of AISI 316 steel with FeCrC and FeCrC/B₄C mixed in different ratios and investigated the resulting abrasive wear properties. An AISI 1045 steel surface was coated with FeMo using PTAW to investigate wear resistance in Ref. [23].

In the model proposed by Huang et al. [24], the surface roughness and processing parameters in a poka-yoke system were predicted online using an artificial neural network (ANN)-based model. Surface roughness is a critical quality index that determines the quality of the machined surfaces and is influenced by the cutting parameters. Zhang and Shetty [25] attempted to predict surface roughness using a support vector machine, a neural network, and variance analysis. Khanlou et al. [26] studied the surface characteristics of sandblasted and acid-etched titanium alloy using adaptive neuro fuzzy inference system to predict surface roughness. In another study on surface characteristics, the surface roughness of turning was modeled using an ANN [27]. The ANN, a discrete particle swarm

optimization (DPSO) algorithm and an algorithm-based selective ANN ensemble were compared regarding online tool wear prediction in drilling operations. As a result of the study, it was observed that the ensemble ANN obtained better results than did other algorithms [28]. In the study by Unune et al. [29], the fuzzy logic model was used to predict the amount of material loss and surface roughness due to wear.

Prediction of Mo coating wear loss quantity was performed using an ANN [30]. The erosive wear rate of coatings prepared by high velocity oxygen fuel and flame spray flexi-cord techniques has been predicted by an ANN [31]. In a model designed using 99 data samples, Gaussian process regression, linear regression, and support vector machine methods were evaluated. The best R^2 value in this study was calculated as 0.96 [32].

The tests required to determine the wear resistance of materials are time consuming and costly. In addition, special devices are needed to determine the wear resistance of the materials. By predicting the amount of wear loss using machine learning methods, the loss of time, high experiment costs, and labor efforts can be reduced [32].

In this study, various extreme learning machine (ELM) and ANN models were applied to predict the wear loss amount. ELM, which is structurally similar to ANN, presents an advantage with respect to the elimination of training time. Different variations of ELM have been developed to improve the algorithm's prediction capabilities. In this study, the comparisons of single hidden layer feedforward neural networks (SLFNs) using different ELM algorithms are analyzed. The results obtained by applying ANN, ELM, kernel-based ELM (KELM), and weighted ELM (WELM) to predict wear amount using the experimental dataset are discussed. As a result of 189 wear experiments, wear loss data according to the applied load and sliding distance was observed and recorded. It was concluded that WELM is the most appropriate and best prediction algorithm for the dataset considered. Designing models using ELM, KELM, and WELM algorithms for wear loss prediction is a new approach to the best of our knowledge. The four methods (ANN and three ELM variations) are detailed in Section 4, where their suitability for the prediction of wear loss is discussed along with the results.

2 Data acquisition

2.1 Sample preparation

AISI 1020 (a low carbon steel) was used as substrate material because its production costs and prices are relatively inexpensive. Low carbon steels are often used as structural steels in the market, but their wear and corrosion properties are poor. Thus, low wear and corrosion properties limit their application. Improving the surface properties of low carbon steel is a very economical way to produce machine and construction parts that have high wear resistance. The AISI 1020 steel used as substrate material was acquired from the market in 20 mm×10 mm×1,000 mm sheets and was cut into 105 mm-length pieces by guillotine. To remove the oxides on the surface of the parts, 1 mm of material was removed from the surface of the parts using a computer numerical control (CNC) milling machine and then samples were machined to the dimensions shown in Fig. 1 [33]. After filing off the burrs on the surfaces of the machined samples, compressed air was applied to remove any remaining dust from the surface. After cleaning any remaining dirt and oil with acetone, samples were dried in an oven at 60 °C for 30 min to remove any moisture.

High chromium FeCrC, FeW, and FeB ferroalloys were used as surface coating powders. The chemical compositions of the ferroalloy powders are given in Table 1. Before coating, the powders were placed in separate ceramic containers to dehumidify them and were dried in an oven at 110 °C for 1 h. The powders were then weighed with precision scales and mixed at the ratios given in Table 2. Based on Ref. [34], the highest hardness value of the coating made with three different ferroalloys was obtained by using the ratios in the number 1 mixture. The highest hardness value of the FeCrC-FeB coating was obtained for the samples mixed in ratio number 3 [35]. These mixing ratios were selected as primary candidates and the other two different mixing ratios were determined by increasing the FeB ratio to obtain higher hardness (numbers 2 and 4). The experimental mixtures were separately stirred for 1 h at 150 rpm in a mechanical mixer to obtain more homogeneous mixtures. The powder mixtures were placed on the open channel (Fig. 1) and compressed. Alcohol was used so that a powder

mixture adhered to the substrate surface to prevent the coating powder from flying during welding. After this procedure, the experimental samples were placed in a furnace to remove any moisture and allowed to dry for 1 h at 100 °C. After the samples were removed from the oven, they were allowed to cool to room temperature and surface coating operations were performed with the constant parameters given in Table 3 and the heat inputs given in Table 4 using

Fig. 1 Machined substrate materials dimensions. Reproduced with permission from Ref. [33], © Firat University, 2017.

Table 1 Chemical compositions of ferroalloy powders. (wt%)

Powder	Cr	B	C	P	Si	W	Mn	Fe
FeCrC	66.77	—	7.95	0.007	0.55	—	—	Rest
FeW	—	—	0.045	0.05	0.52	79.26	—	Rest
FeB	—	18.22	0.3	0.05	0.5	—	—	Rest

Table 2 Mixing ratios of the surface coating powders. (wt%)

Powder mixture	FeCrC	FeW	FeB
1	60	20	20
2	50	20	30
3	70	—	30
4	50	—	50

Table 3 Coating constant parameters.

Coating constant parameter	Value
Voltage (V)	18–20
Plasma gas flow rate (L/min)	0.5
Shielding gas flow rate (L/min)	8
Distance between part and torch (mm)	2
Electrode type	2% thorium electrode

Table 4 Production parameters used in PTA surface coating.

Sample number	Powder mixture	Coating speed (m/min)	Current (A)	Heat input (kJ/mm)
S1	1	0.15	140	0.585
S2	1	0.15	160	0.704
S3	1	0.1	120	0.713
S4	1	0.1	140	0.878
S5	1	0.1	160	1.056
S6	2	0.15	140	0.585
S7	2	0.15	160	0.704
S8	2	0.1	120	0.713
S9	2	0.1	140	0.878
S10	2	0.1	160	1.056
S11	3	0.15	120	0.475
S12	3	0.15	140	0.585
S13	3	0.15	160	0.704
S14	3	0.1	120	0.713
S15	3	0.1	140	0.878
S16	3	0.1	160	1.056
S17	4	0.15	140	0.585
S18	4	0.15	160	0.704
S19	4	0.1	120	0.713
S20	4	0.1	140	0.878
S21	4	0.1	160	1.056

a Thermal Dynamics WC100B brand PTA welding device, shown in Fig. 2.

2.2 Microhardness measurements

The hardness of the coated surface layers was measured using an EMCO TEST brand microhardness tester from the midpoint of the top surface of the coating layer (up to 200 gf) at 0.25 mm intervals toward the substrate material and the average microhardness was obtained by calculating the average of the microhardness values measured at the different points along the coating layer. The average microhardness values of the samples are presented in Fig. 3.

2.3 Wear tests

The samples required for the wear tests were cut into 6 mm×9 mm×6 mm sections from the areas close

to the midpoint of the surface-coated samples. Before the wear tests, the coated surfaces of the samples were sanded with 400 mesh abrasive and cleaned with alcohol. Tests were performed on a block-on-disk type of wear test device, which is schematically shown in Fig. 4. Wear tests were performed at 19.62, 39.24, and 58.86 N using a load applied in the normal direction.

Fig. 2 PTA welding device. **Fig. 3** Average microhardness values of samples. **Fig. 4** Schematic of block-on-disk wear test device.

Normal loads were applied pneumatically by using a compressor and were adjusted by a pneumatic valve. AISI 52100 steel with a diameter of 15 mm was used as the abrasive. Samples were worn at a sliding distance of 900 m at each load. Weight losses were measured with a precision scale (10^{-5} g accuracy) every 300 m. All experimental procedures were performed at room temperature. The experimentally obtained wear loss data and microstructures of samples S6–S16 are described in detail in Refs. [33, 36]. For completeness, the experimentally obtained wear loss data of all samples used for the machine learning model evaluations are shown in Fig. 5.

3 Prediction models

3.1 ANN

In the 1940s, neural networks based on the brain structure emerged, and in the late 1950s, the first practical application of neural networks was introduced with the introduction of the perceptron network [37]. Artificial neural networks are successfully used to model nonlinear functions. They can estimate different non-linear functions at or close to the desired level of accuracy. The flexibility of artificial neural networks in predicting non-linear functions has transformed them into invaluable tools for data processing [38].

An ANN is composed of artificial neurons modeled similarly to human nervous system cells. Neurons are linked to each other by weights. The neurons are gathered in the layers in the network and the output from one layer serves as the input to the next layer. Thus, ANNs can learn regression problems and predict their output or outputs [39]. An ANN model in general consists of an activation function, weights, summing of calculated weights, and input and output neurons. The weights (w_{ij}) indicate the connection strength of the neurons. The value of b shows that the bias values. The net $((n)_i)$ represents as the input of the neurons. The calculation of a basic neural network model is given by the following equation:

$$(n)_j = \sum_{i=1}^n w_{ij} x_i + b \quad (1)$$

In a feed-forward network, all neurons in each

Fig. 5 Wear loss quantities at different normal loads.

layer are connected only to those in the next layer; all neurons in the same layer are independent of each other. The outputs of a layer form the inputs of the next layer. The linkage between the layers is accomplished using weights. A forward feeding ANN consists of data (non-computing) nodes serving as input neurons in the input layer, which propagate data through weights to the hidden layer(s), and an output layer [39]. The input and output layers of the ANN can have

multiple neurons depending on the desired result. There are no rigorous methods available to determine the number of hidden layers or neurons that would yield the best result. Therefore, setting the architecture of the ANN requires experience and conducting experiments to determine its optimal configuration for a given problem [40].

The capability and efficiency to obtain a useful solution to a given problem depends largely on the activation function used by the ANN intercalary to the model structure. The choice of activation functions has a considerable impact on the speed of the network. The activation function used in ANN models may vary depending on the structure of the problem [41], and many different activation functions are available. In this study, the sigmoid is used as the activation function. Where e is euler's number. Equation (2) defines the sigmoid activation function used in this study:

$$f(n)_j = 1 / (1 + e^{-\alpha(n)_j}) \quad (2)$$

The multi-layer perceptron has been shown successfully for estimating nonlinear correlations in different problems, and is one of the most popular ANNs applied to engineering problems [38]. The ANN model must be trained before it can provide a prediction. There are many different algorithms used for training neural networks and the performance of these algorithms varies according to the dataset. The backpropagation algorithm is a common and effective algorithm used to train multi-layer perceptron networks [42].

In a back-propagation ANN, all neurons in the ANN in the layers send their calculated weight values from Eq. (1) forward to the next layer. Then, the error calculated for each layer is propagated backward to the next layer. This process is known as the training or learning process. In the training process, a pair of templates is presented to the network consisting of the values of the inputs and the corresponding desired output values. The ANN calculates the actual outputs based on the weights and a model threshold. The actual output is then matched with the network prediction by transmitting the resulting error back over the network; the weight values in each layer are modified to minimize the error calculated in the output layer. The main purpose of this process is to reduce the

overall error between the predicted output and the actual output [43]. ANNs require sufficient experimental samples in the training dataset to achieve a high-accuracy performance [44].

3.2 Extreme learning machine

Feedforward artificial neural networks are widely used in many different areas owing to their capabilities. The first is using direct input samples to achieve nonlinear mapping and the second is to present a viable model based on natural and artificial classes. However, the lack of fast learning algorithms for ANN and the use of traditional methods that take hours and even days to train the network have led to the requirement for more efficient algorithms. To overcome the disadvantages of ANNs, the extreme learning machine algorithm has been developed [45].

3.2.1 Gradient-based solution

The gradient-based solution is traditionally used to train SLFNs such as an ELM. Specifically, it is used to find $\tilde{w}_i, \tilde{b}_i$, and $\tilde{\beta}$ ($i = 1, \dots, \tilde{N}$) values [45] using Eq. (3):

$$\begin{aligned} & \|H(\tilde{w}_1, \dots, \tilde{w}_{\tilde{N}}, \tilde{b}_1, \dots, \tilde{b}_{\tilde{N}})\tilde{\beta} - T\| \\ & = \min w_i b_i \beta \|H(w_1, \dots, w_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\beta - T\| \end{aligned} \quad (3)$$

Where H is the hidden layer output matrix of the SLFNs, $\tilde{w}_i$ is the weight vector, $\tilde{b}_i$ is the bias value of the SLFNs, $\tilde{\beta}$ is the weight vector between the i -th hidden node and the output nodes, and T represents the matrix of target values.

This equates to the following minimum cost function:

$$E = \sum_{j=1}^N \left[\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \times x_j + b_i) - t_j \right]^2 \quad (4)$$

If the value of H is not known in the gradient-based learning algorithm, the algorithm usually begins to look for a minimum $\|H\beta - T\|$ value. In the gradient-based minimization process, the weights (w_i, β_i) and the bias value are expressed as b_i , and Eq. (5) is applied to Eq. (4) for minimization purpose [45].

$$W_k = W_{k-1} - n \frac{\partial E(W)}{\partial W} \quad (5)$$

where vector W is the set of weights (w_i, β_i) and n is the training rate. The learning algorithm most used in feed forward neural networks is the computational back-propagation learning algorithm involving the distribution of gradients from output to input. However, there are a few inherent problems related to the backpropagation learning algorithm [45]:

1) When the training rate n is low, the learning algorithm runs very slowly. If n is high, the algorithm is unbalanced and produces a solution distant from the desired solution.

2) One of the factors that affect the learning algorithm is the presence of local minima. The learning algorithm can become trapped in a local minimum before finding the global minimum.

3) The regression may have been over-trained using the learning algorithm or the network may yield poor generalization performance. Therefore, appropriate stopping methods are required to achieve useful models.

4) Gradient-based learning requires significant time in most applications.

These problems have been eliminated by the extreme learning machine algorithm, which is a more effective learning algorithm for feed forward neural networks [45]. Using the least squares norm, unlike the traditional convergence theorems, only the number of neurons in the hidden layer is randomly assigned. Weight values and bias values are not randomly assigned and are normally updated only once (are learned in a single step). The purpose of almost all learning algorithms is to find the minimum error rate, but they cannot always reach the minimum error rate because of local minima and the need for essentially infinite training iterations to find a global minimum for some types of problems. ELM is intended to circumvent these issues and is applied as follows.

In a given training set $S = \{(x_i, t_i) \mid x_i \in R^n, t_i \in R^m, i = 1, \dots, N\}$, the activation function $g(x)$ and the number of hidden nodes $\tilde{N}$ are determined as follows [46].

Step 1: Make a random assignment to w_i and $b_i (i = 1, \dots, \tilde{N})$.

Step 2: Calculate the hidden layer output matrix H .

Step 3: Calculate the output $\beta = H^* \times T (T = (t_1, \dots, t_N)^T)$

and obtain the inverse of the H matrix (H^*), using the Moore-Penrose inversion.

Briefly, in the extreme learning machine algorithm, weight and bias values are produced randomly in a manner of speaking, but are learned in a single step that a non-linear system is transformed into a linear system [45].

3.2.2 KELM

The KELM algorithm developed from the ELM algorithm introduces a positive sorting coefficient to provide more stable learning [47]. The KELM algorithm is implemented in many different areas owing to its learning speed and generalizability [48]. In cases when the hidden layer property mapping $h(x)$ is not known, the kernel matrix can be defined for the ELM by the following equation [49]:

$$\delta_{\text{ELM}} = HH^T: \delta_{\text{ELM},ij} = h(x_i) \times h(x_j) = K_{(x_i, x_j)} \quad (6)$$

When the kernel is applied to the ELM algorithm, the hidden layer mappings of $h(x)$ are known to the practitioner, that is, the operator knows $h(x)$ instead of $k(u, v)$ [50]. The number of hidden nodes must also be specified in L (the dimensionality of the hidden layer) [51]. The output function of KELM is then given by the following equation:

$$f(x) = h(x)H^T \left(\frac{1}{c} + HH^T \right)^{-1} T = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{1}{c} + \delta_{\text{ELM}} \right)^{-1} T \quad (7)$$

The KELM algorithm can be implemented in a single learning step. If the value of $h(x)$ is known to the user, then according to Frénay and Verleysen [52, 53], the extreme learning machine algorithm is defined by the following equation:

$$k(u, v) = \lim_{L \rightarrow +\infty} \left[\frac{1}{L} h(u) \times h(v) \right] \quad (8)$$

The radial basis function is typically used in the KELM algorithm [54].

3.2.3 WELM

The non-weighted ELM, which has kernel-core or

kernel-hidden nodes, has been successfully applied to various datasets. However, WELM provides better results than ELM without eliminating the speed advantage of ELM [55]. The WELM algorithm is generally more successful than ELM [56]. The reasons for WELM's better results are: 1) ELM, which is based on the empirical risk reduction principle, tends to create an overly compatible (overfitted) model; 2) ELM has poor control capability because the minimum norm is used to directly calculate the least squares solutions; and 3) it can lead to less-robust estimates. Such weaknesses have been eliminated by WELM [56].

3.3 Z-score normalization

In the Z-score method, normalization is performed by calculating the mean and standard deviation of the data. Equation (9) is used to normalize a dataset [57]:

$$x'_i = (x_i - \bar{x}) / \sigma_x \quad (9)$$

where σ_x represents the standard deviation of x values, and $\bar{x}$ is the arithmetic mean of these values.

3.4 Evaluation metrics

Three different evaluation criteria were applied to test the designed models. The applied criteria are R^2 , the root mean square error (RMSE), and mean absolute error (MAE). The equations for the evaluation criteria are presented below.

MAE: This metric records the overall level of agreement between the observed and modeled datasets in actual units. It is a non-negative measurement with no upper limit and the result is zero for a perfect model. The equation is given below:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \bar{f}_i| \quad (10)$$

RMSE: It is frequently used to measure the differences between the predicted values and the observed values in models using machine learning systems [58]. RMSE represents the square root of the second sampling moment or the second order square of the differences between the predicted values and the observed values. The equation is as follows.

$$\text{RMSE} = \sqrt{\left(\frac{1}{N} \right) \times \sum_j |t_j - \theta_j|^2} \quad (11)$$

R^2 : Based on the total rate of variation of the results described by the model, R^2 provides a measure of how well the model reproduces the observed results. The equation is given below:

$$R^2 = 1 - \left[\frac{\sum_j (t_j - o_j)^2}{\sum_j (t_j - \hat{t})^2} \right] \quad (12)$$

4 Results and discussion

In this study, we predict the amount of wear loss (mg) using machine learning methods based on powder mixture (wt%), average microhardness (HV), normal load (N), and sliding distance (m) as input parameters. Z-score normalization is applied to the input parameters to improve the performance of the models. The dataset consists of 189 results consisting of the wear losses obtained by abrading 21 samples, which were coated with different mixtures and measured at three different sliding distances and three different normal loads separately. Models have been designed using ANN, ELM, KELM, and WELM algorithms to predict the amount of wear loss. The number of hidden neurons, the number of input neurons, the number of output neurons, and the activation functions used in ANN, ELM, and WELM are listed in Table 5. The number of kernel parameters, the number of input neurons, the number of output neurons, and the activation function used in KELM are given in Table 6.

Table 5 ANN, ELM, and WELM model architectures.

Model	Hidden neuron	Activation function	Output neuron	Input neuron
ANN	10	Sigmoid	1	4
ELM	70	Sigmoid	1	4
WELM	500	Sigmoid	1	4

Table 6 KELM model architectures.

Model	Kernel parameter	Activation function	Output neuron	Input neuron
KELM	6	RBF kernel	1	4

To test the models designed to predict the wear loss as an amount, the dataset was divided into two parts: training and test datasets. The training data consisted of 151 samples (80% of all data) and the test data consisted of 38 samples (20% of all data). The training and test data were distributed randomly. The same training and test datasets have been used to provide fair comparisons among the four different methods. Based on the results of the study, WELM and ANN achieved results very close to each other, followed by KELM and ELM, respectively.

The values obtained from the experimental and

predicted values of the designed models are shown in the graphs in Fig. 6 according to the test data index numbers. Figure 6 shows the comparison of the experimental results to the predicted results for the test data; it indicates the error band between the target and actual prediction, and indicates the predicted amount of wear loss. From the graphs, all designed models can predict the experimental results with high accuracy. For indices 5, 19, 26, and 32 in the test data, the KELM algorithm exhibits the worst overall performance, while the WELM algorithm exhibits the best overall performance and predicts wear amount

Fig. 6 Observed and predicted data with errors.

values closer to the target values. The target values in these indices are 4.26, 13.56, 4.24, and 6.55, respectively. The predicted values by the KELM algorithm are 6.12, 11.81, 2.72, and 5.33, respectively. Predicted values by the WELM algorithm are 5.69, 12.78, 3.43, and 6.51, respectively. The error rates are shown in Figs. 6(c) and 6(d). These four indices represent the worst predicted values, with more than twice the MAE error than other indices in the designed models.

To observe the errors more clearly, an error bar graph is presented in Fig. 7. In Fig. 7, the error values

for the prediction values from each model can be evaluated in detail. The prediction by the WELM model in Fig. 7(d) produces the lowest error values. When the MAE values obtained by dividing the absolute errors totals of the models by the total number of samples were compared, the MAE was calculated as 0.4428 for the ANN model (Fig. 7(a)), 0.4797 for the ELM model (Fig. 7(b)), and 0.5298 for the KELM model (Fig. 7(c)). The best MAE value of 0.4369 was calculated for the WELM model (Fig. 7(d)). Figure 8 demonstrates the accuracy of the models during the

Fig. 7 Error bar graph of models designed with machine learning algorithms.

Fig. 8 Scatterplots of the ANN, ELM, KELM, and WELM models.

testing periods using scatterplots.

RMSE, MAE, and R^2 values were calculated separately for each model to evaluate their performance. These values are presented in Table 7 for ANN, ELM, KELM, and WELM. The R^2 values are 0.9728 for ANN, 0.9690 for ELM, and 0.9565 for the KELM model, respectively. In the model designed with WELM, the R^2 value was calculated as 0.9729, indicating better performance compared to that of other models. The lowest RMSE value was calculated as 0.5515 for WELM. The RMSE value of WELM is very similar to that of ANN, and is 6.5% lower than that of ELM, and 21.2% lower than that of KELM. For the WELM algorithm, the MAE value was the lowest at 0.4369. The WELM algorithm achieved a 17.5% lower result than did the KELM algorithm, which exhibited the worst performance regarding MAE. For the WELM and ANN algorithms, although the R^2 values are very similar, the WELM algorithm performed the best for the data set in terms of RMSE and MAE values. From the results of two different methods (ANN and ELM) used to construct models designed to predict wear loss, the model created using the WELM algorithm is the most suitable because of WELM's training speed and stability advantages.

Table 7 Performance evaluation of the designed machine learning models.

Model	ANN	Extreme learning machine	Kernel-based extreme learning machine	Weighted extreme learning machine
RMSE	0.5525	0.5901	0.6996	0.5515
MAE	0.4428	0.4797	0.5298	0.4369
R^2	0.9728	0.9690	0.9565	0.9729

FeCrC, FeW, and FeB ferroalloys were mixed in different proportions and were used to coat low carbon steel surfaces by the PTA method. For the prediction of wear amounts, models constructed using ANN, ELM, KELM, and WELM algorithms, representing various machine learning methods were applied to 189 wear loss data obtained experimentally in the laboratory. In the study, it was observed that all of the designed models achieved high performance, and can be successfully used in various industries (such as rolling, mining, and agricultural machinery) where metal parts are subject to high wear.

5 Conclusions

Four different methods have been used to create models to evaluate their predictions regarding wear amounts of surface coatings applied by the welding melting method, which is frequently used in the industry to produce wear-resistant metal surfaces. This study aims to eliminate the time and labor associated with lengthy and costly wear tests. Prediction of wear loss was achieved using models created by the ANN, ELM, KELM, and WELM algorithms, where the wear loss dataset used in this study was obtained by some researchers in our group.

The study provides novelty in terms of the use of ELM, KELM, and WELM algorithms for predicting the wear loss amounts of metal-coated surfaces under a set of given conditions. This is clearly supported by the results and in the discussion, which demonstrate that the machine learning algorithms successfully predict the wear loss amounts of different coated surfaces to varying degrees. In this study, the models designed by using ANN, ELM, KELM, and WELM algorithms were compared to each other, and WELM achieved the best R^2 value of 0.9729 for wear loss.

When the same data model is constructed using different variations of the ELM algorithm, different results are obtained. WELM achieved an MAE value 17.5% lower than that of KELM.

The wear loss of coatings made with ferroalloy powders mixed at different ratios and with different parameters have been successfully predicted. The success rates achieved indicate that the models evaluated can eliminate the loss of time and the labor associated with lengthy and costly wear tests by using machine learning algorithms in the prediction of wear loss quantities of surface coatings applied to metal surfaces using welding melting methods with high success.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- [1] Holmberg K, Erdemir A. Influence of tribology on global energy consumption, costs and emissions. *Friction* **5**(3): 263–284 (2017)
- [2] Khadem M, Penkov O V, Yang H K, Kim D E. Tribology of multilayer coatings for wear reduction: A review. *Friction* **5**(3): 248–262 (2017)
- [3] Luo W, Selvadurai U, Tillmann W. Effect of residual stress on the wear resistance of thermal spray coatings. *J Therm Spray Technol* **25**(1–2): 321–330 (2016)
- [4] Paul T, Alavi S H, Biswas S, Harimkar S P. Microstructure and wear behavior of laser clad multi-layered Fe-based amorphous coatings on steel substrates. *Lasers Manuf Mater Process* **2**(4): 231–241 (2015)
- [5] Azimi G, Shamanian M. Effect of silicon content on the microstructure and properties of Fe–Cr–C hardfacing alloys. *J Mater Sci* **45**(3): 842–849 (2010)
- [6] Zahiri R, Sundaramoorthy R, Lysz P, Subramanian C. Hardfacing using ferro-alloy powder mixtures by submerged arc welding. *Surf Coat Technol* **260**: 220–229 (2014)
- [7] Fan C, Chen M C, Chang C M, Wu W T. Microstructure change caused by (Cr,Fe)₂₃C₆ carbides in high chromium Fe–Cr–C hardfacing alloys. *Surf Coat Technol* **201**(3–4): 908–912 (2006)
- [8] Yang J, Hou X R, Zhang P, Zhou Y F, Yang Y L, Ren X J, Yang Q X. Mechanical properties of the hypereutectoid Fe–Cr–C hardfacing coatings with different nano-Y₂O₃ additives and the mechanism analysis. *Mater Sci Eng: A* **655**: 346–354 (2016)
- [9] Zhou Y F, Qin G K, Jiang P J, Wang S F, Qi X W, Xing X L, Yang Q X. Dry sliding wear behavior of (Cr,Fe)₇C₃₋₇(Cr,Fe) metal matrix composite (MMC) coatings: The influence of high volume fraction (Cr,Fe)₇C₃ carbide. *Tribol Lett* **66**(3): 108 (2018)
- [10] Durmuş H, Çömez N, Gül C, Yurddaşkal M, Yurddaşkal M. Wear performance of Fe–Cr–CB hardfacing coatings: Dry sand/rubber wheel test and ball-on-disc test. *Int J Refract Met Hard Mater* **77**: 37–43 (2018)
- [11] Yilmaz S O, Özenbaş M, Yaz M. FeCrC, FeW, and NiAl modified iron-based alloy coating deposited by plasma transferred arc process. *Mater Manuf Processes* **26**(5): 722–731 (2011)
- [12] Teker T, Karataş S, Yilmaz S O. Microstructure and wear properties of AISI 1020 steel surface modified by HARDOX 450 and FeB powder mixture. *Prot Met Phys Chem Surf* **50**(1): 94–103 (2014)
- [13] Masanta M, Shariff S M, Choudhury A R. Evaluation of modulus of elasticity, nano-hardness and fracture toughness of TiB₂–TiC–Al₂O₃ composite coating developed by SHS and laser cladding. *Mater Sci Eng: A* **528**(16–17): 5327–5335 (2011)
- [14] Eroglu M. Boride coatings on steel using shielded metal arc welding electrode: Microstructure and hardness. *Surf Coat Technol* **203**(16): 2229–2235 (2009)
- [15] Reinaldo P R, D'Oliveira A S C M. NiCrSiB coatings deposited by plasma transferred arc on different steel substrates. *J Mater Eng Perform* **22**(2): 590–597 (2013)
- [16] Hou Q Y, Gao J S, Zhou F. Microstructure and wear characteristics of cobalt-based alloy deposited by plasma transferred arc weld surfacing. *Surf Coat Technol* **194**(2–3): 238–243 (2005)

- [17] Liu Y F, Han J M, Li R H, Li W J, Xu X Y, Wang J H, Yang S Z. Microstructure and dry-sliding wear resistance of PTA clad (Cr, Fe)-C₃/γ-Fe ceramal composite coating. *Appl Surf Sci* **252**(20): 7539–7544 (2006)
- [18] Ozel S, Kurt B, Somunkiran I, Orhan N. Microstructural characteristic of NiTi coating on stainless steel by plasma transferred arc process. *Surf Coat Technol* **202**(15): 3633–3637 (2008)
- [19] Fernandes F, Lopes B, Cavaleiro A, Ramalho A, Loureiro A. Effect of arc current on microstructure and wear characteristics of a Ni-based coating deposited by PTA on gray cast iron. *Surf Coat Technol* **205**(16): 4094–4106 (2011)
- [20] Veinthal R, Sergejev F, Zikin A, Tarbe R, Hornung J. Abrasive impact wear and surface fatigue wear behaviour of Fe–Cr–C PTA overlays. *Wear* **301**(1–2): 102–108 (2013)
- [21] Hornung J, Zikin A, Pichelbauer K, Kalin M, Kirchgaßner M. Influence of cooling speed on the microstructure and wear behaviour of hypereutectic Fe–Cr–C hardfacings. *Mater Sci Eng: A* **576**: 243–251 (2013)
- [22] Gur A K, Ozay C, Orhan A, Buytoz S, Caligulu U, Yigitturk N. Wear properties of Fe–Cr–C and B₄C powder coating on AISI 316 stainless steel analyzed by the Taguchi method. *Mater Test* **56**(5): 393–398 (2014)
- [23] Deng X K, Zhang G J, Wang T, Ren S, Bai Z L, Cao Q. Investigations on microstructure and wear resistance of Fe–Mo alloy coating fabricated by plasma transferred arc cladding. *Surf Coat Technol* **350**: 480–487 (2018)
- [24] Huang B P, Chen J C, Li Y. Artificial-neural-networks-based surface roughness Pokayoke system for end-milling operations. *Neurocomputing* **71**(4–6): 544–549 (2008)
- [25] Zhang N, Shetty D. An effective LS-SVM-based approach for surface roughness prediction in machined surfaces. *Neurocomputing* **198**: 35–39 (2016)
- [26] Khanlou H M, Ang B C, Barzani M M, Silakhori M, Talebian S. Prediction and characterization of surface roughness using sandblasting and acid etching process on new non-toxic titanium biomaterial: Adaptive-network-based fuzzy inference System. *Neural Comput Appl* **26**(7): 1751–1761 (2015)
- [27] Pal S K, Chakraborty D. Surface roughness prediction in turning using artificial neural network. *Neural Comput Appl* **14**(4): 319–324 (2005)
- [28] Yu J B. Online tool wear prediction in drilling operations using selective artificial neural network ensemble model. *Neural Comput Appl* **20**(4): 473–485 (2011)
- [29] Unune D R, Barzani M M, Mohite S S, Mali H S. Fuzzy logic-based model for predicting material removal rate and average surface roughness of machined Nimonic 80A using abrasive-mixed electro-discharge diamond surface grinding. *Neural Comput Appl* **29**(9): 647–662 (2018)
- [30] Çetinel H, Öztürk H, Çelik E, Karlık B. Artificial neural network-based prediction technique for wear loss quantities in Mo coatings. *Wear* **261**(10): 1064–1068 (2006)
- [31] Mojena M A R, Roca A S, Zamora R S, Orozco M S, Fals H C, Lima C R C. Neural network analysis for erosive wear of hard coatings deposited by thermal spray: Influence of microstructure and mechanical properties. *Wear* **376–377**: 557–565 (2017)
- [32] Altay O, Gurgenc T, Ulas M, Özel C. Prediction of wear loss quantities of ferro-alloy coating using different machine learning algorithms. *Friction*, in press, DOI 10.1007/s40544-018-0249-z.
- [33] Gürgeç T, Özel C. Effect of heat input on microstructure, friction and wear properties of Fe–Cr–BC coating on AISI 1020 surface coated by PTA method. *Firat Univ Turkish J Sci Technol* **12**(2): 43–52 (2017)
- [34] Teker T, Karataş S, Yilmaz S O. Microstructure and wear properties of FeCrC, FeW and feti modified Iron based alloy coating deposited by PTA process on AISI 430 steel. *Arch Metall Mater* **59**(3): 925–933 (2014)
- [35] Yüksel N, Şahin S. Wear behavior–hardness–microstructure relation of Fe–Cr–C and Fe–Cr–C–B based hardfacing alloys. *Mater Des* **58**: 491–498 (2014)
- [36] Özel C, Gürgeç T. Effect of heat input on microstructure, wear and friction behavior of (wt.-%) 50FeCrC-20FeW-30FeB coating on AISI 1020 produced by using PTA welding. *PLoS One* **13**(1): e0190243 (2018)
- [37] Esfe M H, Ahangar M R H, Rejvani M, Toghraie D, Hajmohammad M H. Designing an artificial neural network to predict dynamic viscosity of aqueous nanofluid of TiO₂ using experimental data. *Int Commun Heat Mass Transfer* **75**: 192–196 (2016)
- [38] Esfe M H, Wongwises S, Naderi A, Asadi A, Safaei M R, Rostamian H, Dahari M, Karimipour A. Thermal conductivity of Cu/TiO₂–water/EG hybrid nanofluid: Experimental data and modeling using artificial neural network and correlation. *Int Commun Heat Mass Transfer* **66**: 100–104 (2015)
- [39] Açikgeç M, Ulaş M, Alyamaç K E. Using an artificial neural network to predict mix compositions of steel fiber-reinforced concrete. *Arab J Sci Eng* **40**(2): 407–419 (2015)
- [40] Mukherjee A, Biswas S N. Artificial neural networks in prediction of mechanical behavior of concrete at high temperature. *Nuclear Eng Des* **178**(1): 1–11 (1997)
- [41] Yu X H, Ye C, Xiang L B. Application of artificial neural network in the diagnostic system of osteoporosis. *Neurocomputing* **214**: 376–381 (2016)
- [42] Simpson P K. *Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations*. New York (USA): Pergamon, 1990.

- [43] Momeni E, Armaghani D J, Hajihassani M, Amin M F M. Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement* **60**: 50–63 (2015)
- [44] Dreyfus G. *Neural Networks: Methodology and Applications*. Berlin (Germany): Springer Science & Business Media, 2005.
- [45] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications. *Neurocomputing* **70**(1–3): 489–501 (2006)
- [46] Bilhan O, Emiroglu M E, Miller C J, Ulas M. The evaluation of the effect of nappe breakers on the discharge capacity of trapezoidal labyrinth weirs by ELM and SVR approaches. *Flow Meas Instrum* **64**: 71–82 (2018)
- [47] Huang W M, Li N, Lin Z P, Huang G B, Zong W W, Zhou J Y, Duan Y P. Liver tumor detection and segmentation using kernel-based extreme learning machine. In *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013: 3662–3665.
- [48] Yang Z, Ce L, Lian L. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Appl Energy* **190**: 291–305 (2017)
- [49] Altay O, Ulas M. The use of kernel-based extreme learning machine and well-known classification algorithms for fall detection. In *Advances in Computer Communication and Computational Sciences*. Bhatia S K, Tiwari S, Mishra K K, Trivedi M C, Eds. Singapore: Springer, 2019: 147–155.
- [50] Wang X Z. International journal of machine learning and cybernetics. *Int J Mach Learn Cybern* **1**(1–4):1–2 (2010).
- [51] Huang G B, Zhou H M, Ding X J, Zhang R. Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst, Man, Cybern, Part B (Cybern)* **42**(2): 513–529 (2012)
- [52] Frénay B, Verleysen M. Using SVMs with randomised feature spaces: An extreme learning approach. In *Proceedings of the 18th ESANN*, Bruges, Belgium: 2010.
- [53] Frénay B, Verleysen M. Parameter-insensitive kernel in extreme learning for non-linear support vector regression. *Neurocomputing* **74**(16): 2526–2531 (2011)
- [54] Huang G B, Wang D H, Lan Y. Extreme learning machines: A survey. *Int J Mach Learn Cybern* **2**(2): 107–122 (2011)
- [55] Zong W W, Huang G B, Chen Y Q. Weighted extreme learning machine for imbalance learning. *Neurocomputing* **101**: 229–242 (2013)
- [56] Deng W Y, Zheng Q H, Chen L. Regularized extreme learning machine. In *Proceedings of 2009 IEEE Symposium on Computational Intelligence and Data Mining*, Nashville, USA, 2009: 389–395.
- [57] Altay O, Ulas M. Location determination by processing signal strength of Wi-Fi routers in the indoor environment with linear discriminant classifier. In *Proceedings of the 6th International Symposium on Digital Forensic and Security (ISDFS)*, Antalya, Turkey, 2018: 1–4.
- [58] Hyndman R J, Koehler A B. Another look at measures of forecast accuracy. *Int J Forecast* **22**(4): 679–688 (2006)

Turkey, 2011. From 2004 to 2012, he worked as a

Mustafa ULAS. He is an assistant professor in Department of the Software Engineering, Firat University, Turkey. He received his B.S. degree in 2003 and the Ph.D. degree in Electric and Electronics Engineering Department in Firat University,

lecturer and a software developer in the Department of Informatics. He was vice president of Computer Center and Department of Informatics from 2008 to 2012, Turkey. He worked in the University of Michigan, USA, as a visitor researcher in 2013. He has studied on data mining, machine learning algorithms, big data, and augmented reality.

Osman ALTAY. He received his B.S. degree in Department of Electronic Computer Education of Selcuk University, Turkey, in 2011. Currently, he is continuing his Ph.D. in software

engineering at Firat University, Turkey. He is a research assistant at the Department of Software Engineering, Manisa Celal Bayar University, Turkey. His research interests include data mining, bioinformatics, machine learning, and data science.

Turan GURGENC. He received his Ph.D. degree in mechanical engineering from Firat University, Turkey, in 2017. He is a research

assistant in Automotive Engineering Department, Firat University, Turkey. His research interests include surface coating, wear analysis, sintering, manufacturing, and computational intelligence.

Cihan ÖZEL. He received his Ph.D. degree in mechanical engineering from Firat University, Turkey, in 2000. He is currently an associate

professor in Mechanical Engineering Department, Firat University, Turkey. His research interests cover CNC machining, surface coating, wear, and manufacturing.