Civil Engineering Infrastructures Journal 2024, 57(1): 189-203
DOI: 10.22059/CEIJ.2023.349334.1902

RESEARCH PAPER

Determination of DBTT of Functionally Graded Steels Using Artificial Intelligence

Anjali, K.[1*], Yesilyurt, S.N.[2], Samui, P.[3], Dalkilic, H.Y.[4] and Katipoglu, O.M.[5]

[1] Ph.D. Candidate, Research Scholar, School of Water Resources, Indian Institute of Technology, Kharagpur, India.
[2] Ph.D. Candidate, Graduate School of Natural and Applied Sciences, Department of Civil Engineering, Dokuz Eylul University, Izmir, Turkey.
[3] Associate Professor, Department of Civil Engineering, National Institute of Technology, Patna, India.
[4] Associate Professor, Department of Civil Engineering, Erzincan Binali Yildirim University, Erzincan, Turkey.
[5] Professor, Department of Civil Engineering, Erzincan Binali Yildirim University, Erzincan, Turkey.

Received: 08 Jan. 2023; Revised: 22 Jun. 2023; Accepted: 03 Jul. 2023

ABSTRACT: This study applied three Artificial Intelligence (AI) models to project the Ductile to the Brittle Transition Temperature (DBTT) of Functionally Graded Steels (FGS). These prediction models are Minimax Probability Machine Regression (MPMR) model, Genetic Programming (GP), and Emotional Neural Network (ENN) algorithms with strong prediction performance. The data of FGS type, crack tip configuration, the thickness of the graded ferritic zone, the thickness of the graded austenitic region, the distance of the notch from the Bainite or Martensite intermediate layer, and temperature were used as inputs in the establishment of the AI models. Charpy impact test (CVN) values obtained from experiments used as output. The datasets have been divided into two groups: one for training and another for testing. The performance of the established AI models was evaluated through 16 statistical indicators and graphically used regression error characteristics, an area over the curve, Taylor diagrams, and scatter plots. As a result, the GP model showed superior prediction performance to other models. The primary objective of this study was to decrease the parameter count while also facilitating model comparisons. In this way, in areas with complex studies such as civil engineering; it allows the work to be completed more practically.

Keywords: Artificial Intelligence, ENN, Genetic Programming, Minimax Probability Machine Regression, Taylor Diagram.

1. Introduction

Depending upon the composition and microstructure of steel, such as grain size, phase fractions, etc., the mechanical properties' mechanisms are highly complex. Recent Studies throw light on establishing a data-driven technique-based mechanical

* Corresponding author E-mail: kanjali19feb@kgpian.iitkgp.ac.in

property prediction model (Kumar et al., 2022; Hoang and Tran, 2023). A mechanical property prediction model is advantageous to guide the procedure design of steel at least cost in a shorter production cycle. Insufficient understanding of the physical principles governing effective mechanical properties hinders the development of a first-principle-based physical model. Consequently, models propelled by data, such as Artificial Intelligence models, are more preferable (Dastorani et al., 2018; Ziggah et al., 2022; Al Adwan et al., 2023). This paper describes research aimed at devising a reliable prediction model for predicting the DBTT of FGS. Determining the DBTT of steel is essential in manufacturing engineering and structural applications (Bae et al., 2023). Researchers generally use the Charpy impact test, alternatively referred to as the Charpy V-Notch test (CVN), to predict the DBTT of steel (Dubey et al., 2023; Switzner et al., 2023). The experimental methods always give some limitations. Hence, The CVN does not provide a reliable result.

Therefore, alternative methods are required to determine the DBTT of steel. Nazari et al. (2011) applied Artificial Neural Network (ANN) to find out the DBTT of the steel. ANN gave encouraging and comparatively better performance. However, ANN has several drawbacks, such as low generalization capability, overtraining, a black-box approach, etc. (Park and Rilett, 1999; Kecman, 2001; Kovalev, 2021). This article employs the following three Artificial Intelligence (AI) techniques for determining the DBTT of steel. It adopts widely used AI techniques such as Minimax Probability Machine Regression (MPMR), Emotional Neural Network (ENN), and Genetic Programming (GP) to determine the DBTT of steel.

Thanks to the developing AI technology, experimental studies and predictions made in many engineering fields in recent years have been more accurate, and their reliability has increased. AI techniques are closed-box models used for prediction and verification and have been used to solve many problems in the literature (Vouros, 2022; Alyousef et al., 2023). Methods such as MPMR, GA, and ENN, which are the subject of this study, have an essential place in the literature of their superior performance. MPMR model is constructed by Lanckriet et al. (2001) and it gives the bound over future prediction. There are many implementations of MPMR in literature (Deng et al., 2018; Bonakdari et al., 2019; Kar et al., 2023). Koza (1992) is credited with the development of GP, which is founded on the fundamental concept of Genetic Algorithms (GA) and functions by manipulating parse trees. GP model exhibited good performance for solving various problems (Koshiyama et al., 2015; Papa et al., 2017; Lin et al., 2018; Astarabadi and Ebadzadeh, 2019). The ENN model is built on top of the Emotional Back Propagation (EBP), which is an innovative training algorithm that incorporates emotional cues and weights to enhance individuals' learning processes and decision-making abilities (Kumar et al., 2021). Recently, there has been extensive research into the incorporation of emotions into machine learning (Babaie et al., 2008; Khashman, 2008, 2009; Biswas et al., 2019). AI models have been created based on the database compiled from the research conducted by Nazari et al. (2011) as shown in Table 1. The dataset comprises of 140 experimental results of the CVN.

This study aims to evaluate the prediction potential of experimentally obtained CVN values with GP, MPMR, and ENN algorithms. To achieve this objective, three AI algorithms are employed to analyze data encompassing FGS type, crack configuration, graded ferritic zone thickness, graded austenitic region thickness, notch-to-intermediate layer distance, and temperature parameters, ultimately yielding CVN values. Estimated and experimentally obtained CVN values are compared with 16 different statistical indicators, and Regression Error

Characteristic (REC), Area Over Curve (AOC), and Taylor diagrams, and the most successful model are decided.

2. Materials and Methods

This study estimates CVN values of FGS using various experimental parameters. Figure 1 illustrates the graphs employed to depict the interrelationships among the data sets. Key statistical measures were examined, such as mean, standard deviation, distortion and kurtosis to assess the central tendency, spread, and shape of the data to ensure a comprehensive understanding of its properties and distribution characteristics. In addition, the experimental data used in this work are shown in Table A1.

2.1. Comprehensive Overview of MPMR Model

The foundation of the MPMR model lies in the Minimax Probability Machine Classification algorithm (MPMC). It builds a regression function using the Mercer Kernel, establishing a direct boundary (±ε) on the minimum probability and maximizing it (Lanckriet et al., 2001). The MPMR model does not rely on assumptions about the distribution of data, which can result in reduced validity and generality.

The prominent characteristic of the MPMR model is the regression function provides a low limit to the probability. An unspecified regression function $g: R^n \to R$ is employed to generate the training data from the data set $(u_i, v_i); i = 1,2,..n$. The interrelationship between u_i and v_i is outlined as Eq. (1).

$$v_i = g\,(u_i) + \delta_i \tag{1}$$

where δ_i: indicates the error such that $E[\delta] = 0$, $Var[\delta] = \sigma\delta^2$.

The primary objective of MPMR is to maximize the minimum probability under the error ±ε constraint.

The calculation of the bound on minimum probability (Ω) is given by Eqs. (2-3).

$$\hat{v} = \hat{g}(u) \tag{2}$$

$$\Omega = \inf P_r\,\{\,|\,\hat{v} - v| \le \varepsilon\} \tag{3}$$

Much like the kernel formulation for the MPMC, the MPMR formulation is depicted as Eq. (4).

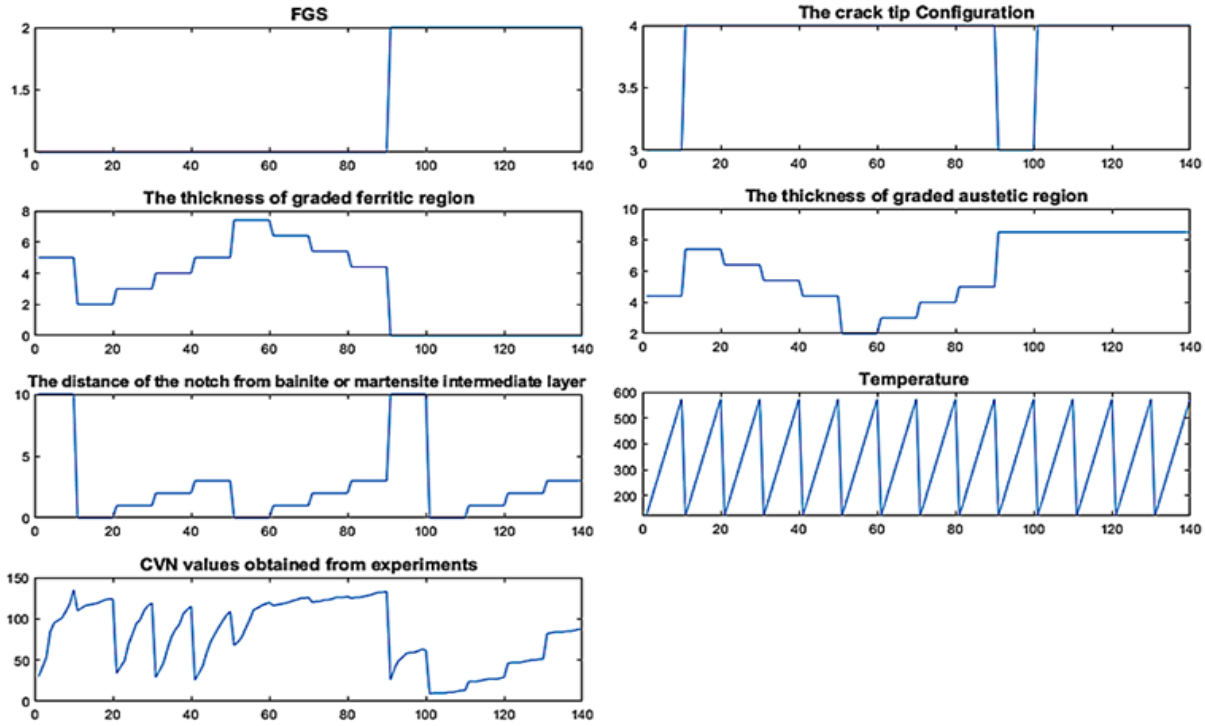

Fig. 1. Change graphs of the data used in the study

$$\hat{v} = \hat{g}(u) = \sum_{i=1}^{N} \beta_i K(u_i, u) + b \quad (4)$$

where $K(u_i, u) = \varphi(u_i).\varphi(u)$ is Mercer's Kernel, u_i: denotes the training data whereas β_i, $b \in R$ indicated the result of the MPMR algorithm. *W*: is the weight and *b*: is the bias.

For the development of the MPMR model, the dataset was split into two segments: one for training and the other for testing. Thus, this eliminates the generalization and memorization problems of the model. Also, Radial basis function was employed as a kernel function for the development of the MPMR model. The training dataset is used to learn the MPMR model. This study employs 98 datasets as the training dataset. To evaluate the conduct of the MPMR model, the test data was normalized within the range of 0 to 1. Eq. (5) was used for normalization.

$$X_{norm} = \frac{X_{act} - X_{min}}{X_{max} - X_{min}} \quad (5)$$

where X_{norm}: is the normalized value, X_{act}: is the actual value, X_{min}: is the minimum value, and X_{max}: is the maximum value.

2.2. Comprehensive Overview of ENN Model

The ENN model is the neural network that has been developed based on emotions in order to enhance decision-making and learning capabilities. Although it is not anticipated that machines will feel emotions and respond emotionally, regulatory signals and informational signals do progress within the ENN model. Emotions may be imitated in machines analogous to machine intelligence. This enhances the model's ease of use, swift learning and immediate response. This model is based on EBP and takes into account things like self-esteem, stress levels, and anxiety, also emotional weights in the cerebral emotional neurons. The neural network contains a multitude of hormone-producing glands that release virtual hormones, subsequently influencing the functioning of specific nodes within the neural system (Babaie et al., 2008).

An ENN is composed of three layers according to the course of information: the input layer, the hidden layer, and the output layer. The input, hidden, and output layers each consists i, j, and h neurons. Here, X_i and Y_i: denote input and output values of i neurons, respectively, in a way that $X_i = Y_i$. Input and output estimates of the hidden neurons are designated by X_h and Y_h, respectively, and are interconnected by Eq. (6). Input and output which estimates of output neurons are designated by X_j and Y_j, respectively, and are interconnected by Eq. (7).

$$Y_h = \frac{1}{1 - \exp(-X_h)} \quad (6)$$

$$Y_j = \frac{1}{1 + \exp(-X_j)} \quad (7)$$

The ENN model employs the same training data set, testing data set, input, output, and standardization approach as the MPMR model does.

2.3. Comprehensive Overview of GP Model

GP is an AI-based method used to formulate experimental studies with multivariate parameters for which analytical models are unavailable (Cevik and Sonebi, 2008). GP imitates the biological advancement of living creatures and creates a computer model of functions and terminals known as chromosomes (Kohestani et al., 2017)

There are 5 stages involved in GP model to lead to the solution as follows:

- Stage 1: A population of randomly generated programs is created.
- Stage 2: The programs with higher fitness/better accuracy concerning the output are chosen using any methods such as Roulette Wheel Selection, Tournament, Ranking, etc.
- Stage 3: Two selected winner programs are copied onto the next generation by

exchanging parts to generate cross overs and then randomly changing the winner programs. Only similar parts of the programs can be replaced.

- Stage 4: The programs that relatively lose/have higher error are replaced by the transformed winner programs obtained in Stage 3.
- Stage 5: Stages 2 to 4 are repeated until the consecutive error values do not have an appreciable reduction in them, i.e., many generations rise till the convergence of predicted value and the actual output leads to a satisfactory limit. The GP model utilizes identical training and testing datasets, input features, output variables, and normalization techniques as those employed by the MPMR and ENN models.

2.4. Performance Indicators

This study utilized a vast array of statistical indicators to assess the efficacy of Artificial Intelligence models employed. The success of the model was made according to the most appropriate values of the estimation criteria obtained during the training and testing stages. Table 1 displays the equations of the parameters used in model selection.

Table 1. Statistical parameters

Statistical parameter	Equation	
Nash-Sutcliffe Efficiency (NS)	$NS = 1 - \frac{\sum_{i=1}^{n}(d_i - y_i)^2}{\sum_{i=1}^{n}(d_i - d_{mean})^2}; -\infty < NS \leq 1$	(8)
Variance Account Factor (VAF)	$\mathrm{VAF} = \left(1 - \frac{\mathrm{var}(d_i - y_i)}{\mathrm{var}(d_i)}\right) \times 100$	(9)
Coefficient of Determination (R^2)	$R^2 = \frac{\sum_{i=1}^{n}(d_i - d_{mean})^2 - \sum_{i=1}^{n}(d_i - y_i)^2}{\sum_{i=1}^{n}(d_i - d_{mean})^2}$	(10)
Adjusted Determination Coefficient (Adj. R^2)	$AdjR^2 = 1 - \frac{(n-1)}{(n-p-1)}(1 - R^2)$	(11)
Performance Index (PI)	$PI = adj \cdot R^2 + 0.01VAF - RMSE$	(12)
Bias Factor	$\text{Bias Factor} = \frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{d_i}$	(13)
Normalized Mean Bias Error (NMBE)	$NMBE(\%) = \frac{\frac{1}{N}\sum_{i=1}^{n}(y_i - d_i)}{\frac{1}{N}\sum_{i=1}^{n} d_i} \times 100$	(14)
Mean Absolute Percentage Error (MAPE)	$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left\lvert\frac{d_i - y_i}{d_i}\right\rvert$	(15)
Relative Percent Difference (RPD)	$RPD = \frac{SD}{RMSE}$	(16)
Willmott's Index of agreement (WI)	$WI = 1 - \left[\frac{\sum_{i=1}^{N}(d_i - y_i)^2}{\sum_{i=1}^{N}(\lvert y_i - d_{\text{mean}}\rvert + \lvert d_i - d_{\text{mean}}\rvert)^2}\right], \quad 0 < WI \leq 1$	(17)
Mean Absolute Error (MAE)	$MAE = \frac{1}{N}\sum_{i=1}^{n}\lvert(y_i - d_i)\rvert$	(18)
Mean Bias Error (MBE)	$MBE = \frac{1}{N}\sum_{i=1}^{n}(y_i - d_i)$	(19)
Legate and McCabe's Index (LMI)	$LMI = 1 - \left[\frac{\sum_{i=1}^{N}\lvert d_i - y_i\rvert}{\sum_{i=1}^{N}\lvert d_i - d_{\text{mean}}\rvert}\right], 0 < LMI \leq 1$	(20)
Expanded Uncertainty (U_{95})	$U_{95} = 1.96(SD^2 + RMSE^2)^{1/2}$	(21)
t-statistic (t-stat)	$t\text{-stat} = \sqrt{\frac{(N-1)MBE^2}{RMSE^2 - MBE^2}}$	(22)
Global Performance Indicator (GPI)	$\mathrm{GPI} = \mathrm{MBE} \times \mathrm{RMSE} \times U_{95} \times t_{\text{stat}} \times (1 - R^2)$	(23)

In Table 1, d: is the observed value and y: is the predicted value, d_i and y_i are the the observed and predicted i^{th} value, d_{mean}: is the mean of observed, R^2: is the coefficient of determination, n: is the number data samples and p: is the model input quantity which is equal to 3, *RMSE*: is the Root mean Square Error, *VAF*: is the Variance Account Factor. In determining the best model, especially the models with the lowest error and high R^2 and Willmott's Index of agreement (*WI*) values were taken into account. Comparing sixteen parameters and streamlining the parameter selection process is one of the primary objectives of this study. Thus, the absence of conflict between these parameters is essential and indicates that the two conflicting parameters cannot be substituted for one another. In essence, the study has concluded that parameters yielding comparable results to all sixteen parameters can be employed. This approach enables achieving similar outcomes in a more straightforward manner, avoiding the complexity associated with the original set of parameters.

2.5. Taylor Diagrams

Taylor Diagrams (Taylor, 2001) is a diagram of the interrelationship between predicted and observed values. It facilitates statistical comparison of different models by plotting the normal deviation of simulated values versus observed values; the correlation coefficient between observed values and simulated values; and the averaged mean square difference. Taylor diagram provides a systematic and mathematical way of demonstrating goodness of fit measures. Following Ee. (24) is the underlying mathematical relationship of Taylor diagram.

$$E = \sigma_o^2 + \sigma_s^2 - 2\sigma_o\sigma_s\rho \quad (24)$$

where E: is the averaged mean square difference, σ_o: is the normal deviation of observed values, σ_s: is the normal deviation of simulated values, and ρ: is the correlation coefficient.

2.6. Rank Analysis

When performing rank analysis, each performance parameter is assigned a rank. In models where multiple statistical indicators are combined, it is difficult to determine the best model. Therefore, rank values of statistical indicators were calculated separately in this work, and the most efficient model was marked based on the total rank value. In this study, rankings ranged from a maximum of three models to a minimum of one. The model with the greatest performance is given the third position, and the model with the worst functionality takes up the first position. The model ranked the highest in total score represents the best performance, whereas the model ranked lowest indicates the poorest performance (Zhang et al., 2020).

3. Results and Discussion

In present study, it is aimed to estimate experimental CVN values with GP, MPMR, ENN models and to compare model performances based on statistical parameters. For this purpose, many experimental results are presented to AI algorithms and an equation is proposed to determine CVN values.

The implementation of the MPMR model, GP and ENN algorithms for predicting the DBTT of FGS was accomplished using MATLAB, a powerful numerical computing software which offers a comprehensive environment for implementing and experimenting with various AI models. Depending on factors such as the specific requirements of the project, familiarity of the researchers, and availability of suitable libraries and tools, the choice of software was made. It provides extensive libraries and toolboxes for machine learning and optimization, enabling the implementation of MPMR, GP and ENN algorithms for DBTT prediction.

Figure 2 shows the correlation coefficients connecting inputs and outputs used in the setup of AI models. Accordingly, while there was a very high

positive correlation between CVN data and the thickness of graded ferritic region values, it was determined that there was a negative and very high relationship with FGS and the thickness of graded austenitic region values.

Figure 3 shows the comprehensive evaluation of the algorithms with respect to various influence factors.

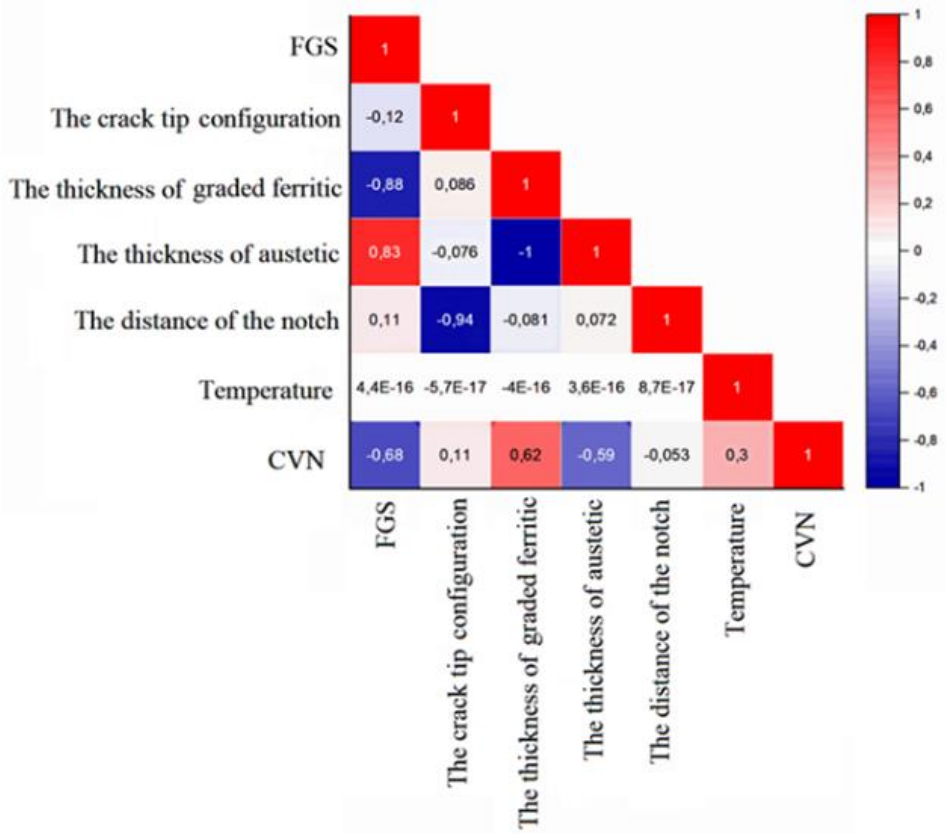

Fig. 2. Correlation matrix of the created models

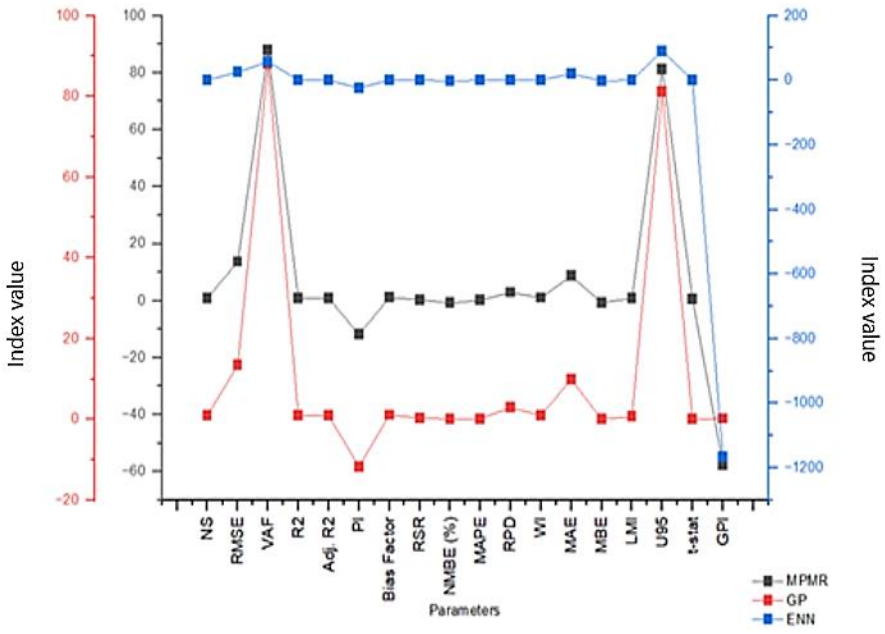

Fig. 3. Performance of different algorithm considering multiple influence factors

Figure 4 illustrates the effectuation of the training dataset. It has been clear from Figure 4 that the value of the correlation coefficient (*R*) is approximately one, and the predicted and actual values are scattered around the 45-degree linear line. Hence, the developed MPMR model gives good performance for the training dataset. The performance of the testing dataset has been depicted in Figure 5. The value of the R is approximately one and the predicted and actual values are scattered around the 45-degree linear line for the testing dataset also. Hence, the developed GP model has the ability to predict CVN value. In addition, Figures 4 and 5 illustrate the performance of ENN model as well. The *R* value is not near to one for training and testing dataset. Hence, the ENN model does not give a reliable result.

The developed GP model gives the best performance for the size of population = 800 and the size of generation = 400. The obtained results hold validity not only for the specific data set used in this study but also for data sets exhibiting similar characteristics. This suggests that the findings can be generalized to other datasets sharing comparable traits, enhancing the applicability and relevance of the study's conclusions. The performance of GP model has been shown in Figures 3 and 4. The value of *R* is close to one for training as well as testing dataset. Hence, the developed GP has successfully captured the relationship between inputs and output. The developed GP gives Eq. (25) for the prediction of CVN value.

$$\begin{aligned} CVN = {} & 376.5\, C^2G^4D^2 - 2.203(F \\ & + 8.651)^2 \\ & - 368.3F^4C^2D^2 \\ & - 639A^4G^4 \\ & + \ 0.5171\cos(CG) \\ & + 0.5171\ \sin(F+T) \\ & + 39.7A^2G^2 \\ & - 2.303CA(A+D) \\ & + 164.8 \end{aligned} \tag{25}$$

where *A*: is the thickness of graded ferritic region, *G*: is the thickness of graded austetic region, *C*: is the crack tip configuration, D is the distance of the notch from bainite or martensite intermediate layer and *T*: is temperature.

An attempt has been made for constructing REC of the developed models. Figure 6 displays the REC curves of the developed MPMR, GP and ENN models. Figure 7 shows the bar chart of AOC values of the developed MPMR, ENN and GP models. For a better model, the AOC value should be small. The value of AOC is almost same for MPMR and GP models. The performance of MPMR and GP is comparatively better than the ENN models.

Figures 8 and 9 shows the cumulative probability plot and probability density function plots of predicted/actual of the MPMR, GP and ENN models. It is clear from Figures 8 and 9 that the developed MPMR and GP give reasonable performance.

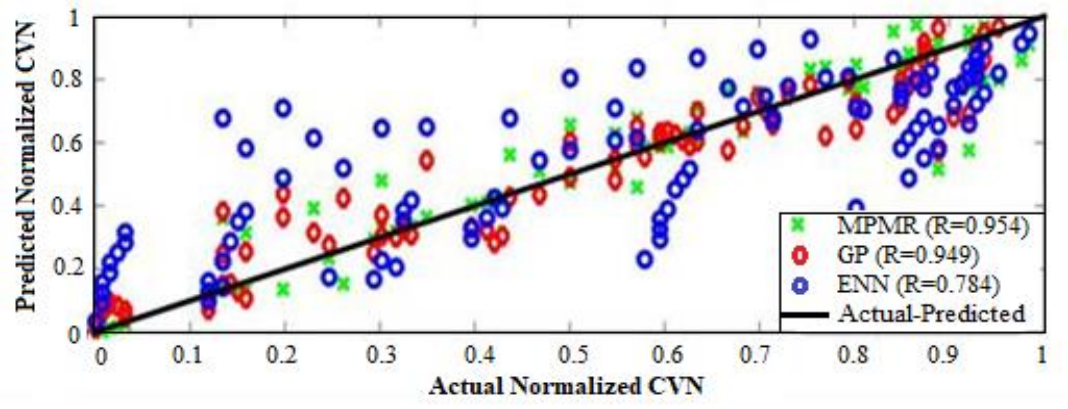

Fig. 4. Scatter plot of training dataset

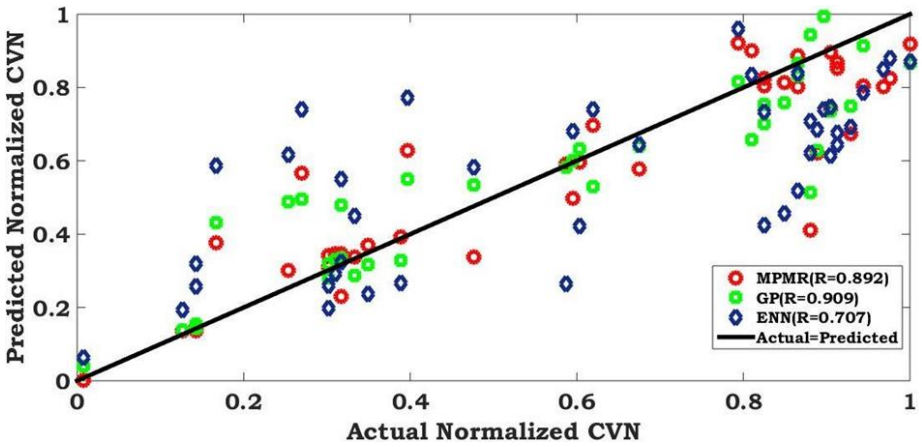

Fig. 5. Scatter plot of testing dataset

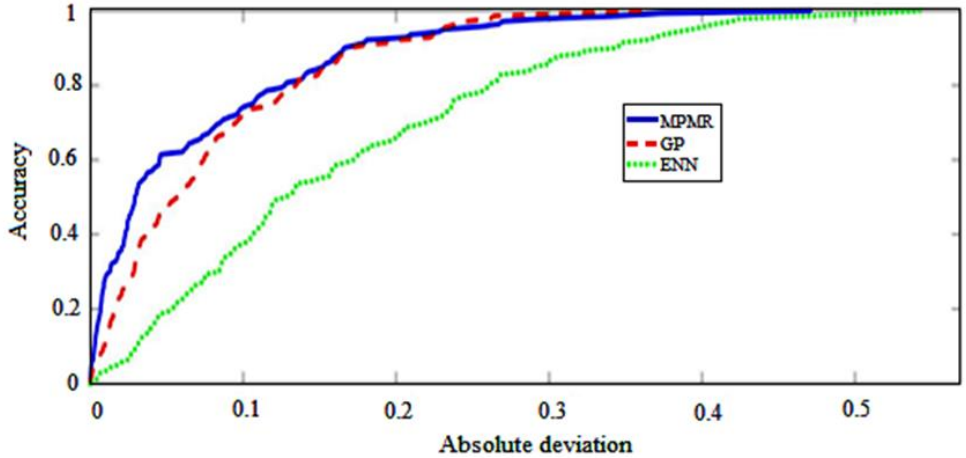

Fig. 6. REC curves of the developed MPMR, GP and ENN models

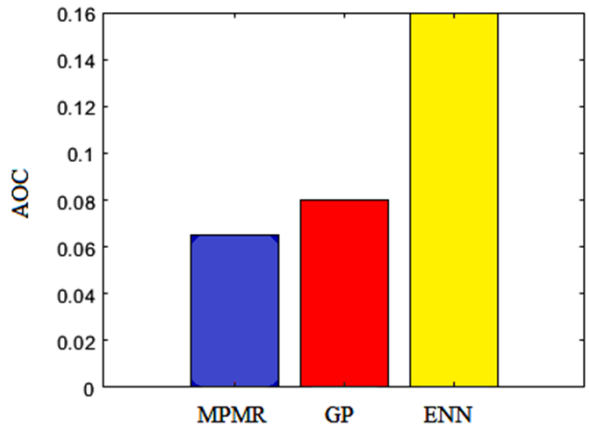

Fig. 7. Bar chart of AOC values of the different models

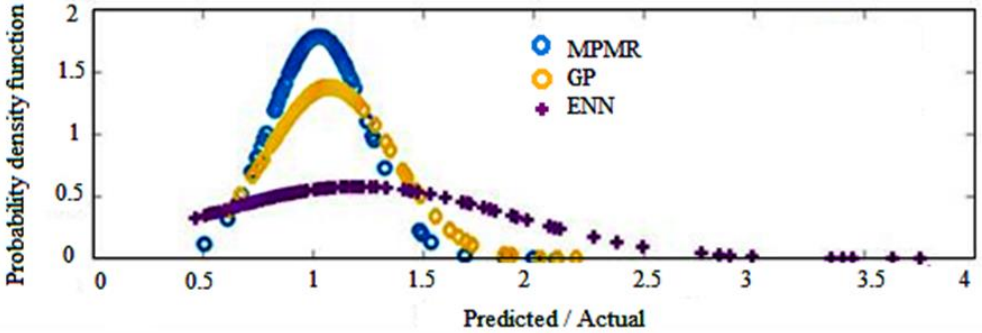

Fig. 8. Cumulative probability plots of predicted/actual of the MPMR, GP and ENN models

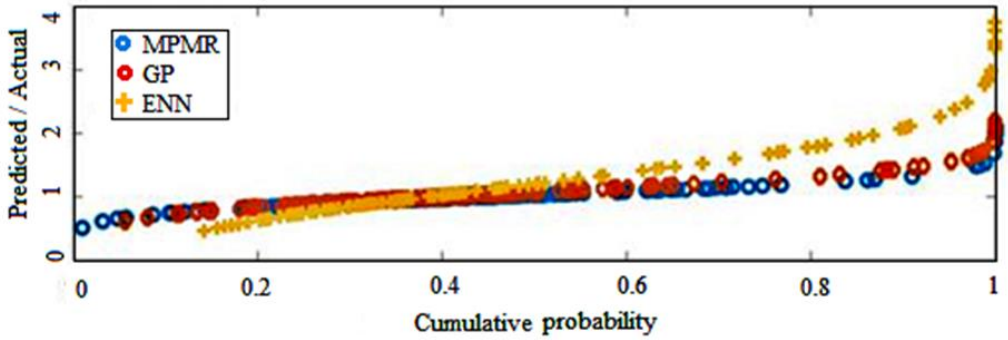

Fig. 9. Probability density function plots of predicted/actual of the MPMR, GP and ENN model

Figure 10 shows the Taylor diagrams of the MPMR, GP and ENN models. It can be seen from Figure 3 that the developed GP and MPMR models have lower RMSE, higher R^2 and closer to reference point than the ENN model. Therefore, GP and MPMR models outperform the ENN model. In addition, the almost overlapping performance points of the GP and MPMR models indicate that the success of the two models is very close. However, the fact that the GP model is slightly closer to the reference line indicates that the standard deviation of the model is closer to the actual data. This concludes that the GP model is slightly better than the MPMR.

In Table 2, the developed models were assessed according to different statistical indicators and the success order of these indicators. In addition, the ideal values of the statistical parameters used are expressed. It is clear from Table 2 that the developed GP and MPMR models predict CVN values quite well from the ENN model. When the prediction performances of all my models were sorted according to rank analysis, it was found as GP > MPMR > ENN.

4. Conclusions

This article inspected the applicability of MPMR, GP and ENN models for the prediction of CVN values of functionally graded steels. The performance and accuracy of the developed models were tested according to statistical and graphical approaches such as different statistical criteria, rank analysis and scatter diagrams, REC analysis. The major findings of the study are listed as follows:

- The performance of MPMR and GP is almost same.
- Users can use the developed GP equation for practical purposes.
- The performance of ENN is not comparatively good.

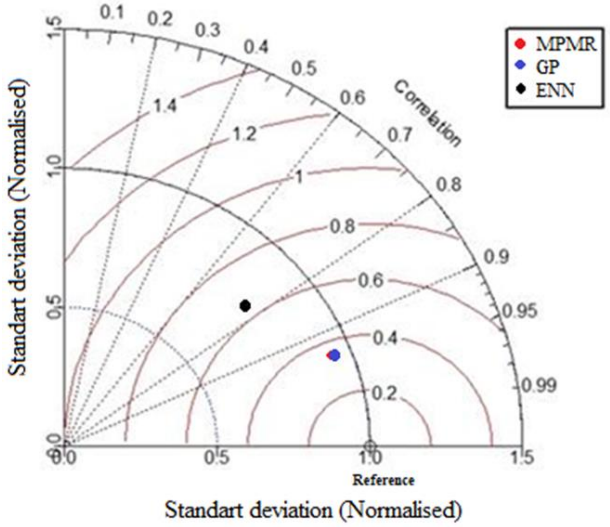

Fig. 10. Taylor diagrams of the developed GP, MPMR and ENN models

Table 2. Various parameters of the developed models

Parameters	MPMR	GP	ENN	Ideal Values
NS	0.8786	0.88147	0.57904	**1**
(Rank)	**-2**	**-3**	**-1**	
RMSE	13.6594	13.4967	25.4352	**0**
(Rank)	**-3**	**-2**	**-1**	
VAF	87.8917	88.1472	58.0706	**100%**
(Rank)	**-2**	**-3**	**-1**	
R2	0.8786	0.88147	0.57904	**1**
(Rank)	**-2**	**-3**	**-1**	
Adj. R^2	0.87682	0.87974	0.57289	**1**
(Rank)	**-2**	**-3**	**-1**	
PI	-11.904	-11.736	-24.282	**> 1.0**
Bias Factor	1.02611	1.07368	1.20422	**1**
(Rank)	**-3**	**-2**	**-3**	
RSR	0.34843	0.34428	0.64881	**0**
(Rank)	**-2**	**-3**	**-1**	
NMBE (%)	-0.8707	0.05263	-1.9827	**0**
(Rank)	**-2**	**-3**	**-1**	
MAPE	0.12394	0.17859	0.42507	**0**
(Rank)	**-3**	**-2**	**-1**	
RPD	2.87001	2.90461	1.54127	**> 2.5**
WI	0.9997	0.9997	0.99875	**1**
(Rank)	**-2**	**-3**	**-1**	
MAE	8.63595	9.86508	20.49	**0**
(Rank)	**-3**	**-2**	**-1**	
MBE	-0.7027	0.04248	-1.6002	**0**
(Rank)	**-2**	**-3**	**-1**	
LMI	0.75127	0.71587	0.40986	**1**
(Rank)	**-3**	**-2**	**-1**	
U95	81.3679	81.2635	91.5931	**0**
(Rank)	**-2**	**-3**	**-1**	
t-stat	0.60734	0.03711	0.7432	**Smaller value**
(Rank)	**-2**	**-3**	**-1**	
GPI	-57.589	0.2049	-1166.3	**Higher value**
(Rank)	**-2**	**-3**	**-1**	
Total Rank	37	**43***	18	

Note: * and red indicates the best model

- It has been concluded that the developed GP and MPMR models can provide high accuracy in the determination of CVN values of functionally graded steels. Thus, CVN values can be easily calculated when different the data of FGS type, crack type configuration, the thickness of the graded ferritic zone, the thickness of the graded austenitic region, the distance of the notch from the Bainite or Martensite intermediate layer and temperature parameters are used.
- The equations obtained from the study will contribute to the determination of CVN values with theoretical approaches by reducing the number of experiments and analyzes.
- The study is a reference for studies on data analysis and modeling in civil engineering. This study proposes to make the operations more practical by using a small number of parameters instead of using all parameters.

5. Acknowledgements

The data used in the paper were obtained from the study of Nazari et al., (2012). The help of P. Samui and A. Kumari; is appreciated for providing data in this study. The statistical analysis part was performed by H.Y. Dalkılıç, S.N. Yesilyurt and O.M. Katipoglu.

6. References

Al Adwan, J., Alzubi, Y., Alkhdour, A. and Alqawasmeh, H. (2023). "Predicting compressive strength of concrete using

histogram-based gradient boosting approach for rapid design of mixtures", *Civil Engineering Infrastructures Journal*, 56(1), 159-172, https://doi.org/10.22059/CEIJ.2022.337777.1811.

Alyousef, R., Rehman, M.F., Khan, M., Fawad, M., Khan, A.U., Hassan, A.M. and Ghamry, N.A. (2023). "Machine learning-driven predictive models for compressive strength of steel fiber reinforced concrete subjected to high temperatures", *Case Studies in Construction Materials*, 19, e02418, https://doi.org/10.1016/j.cscm.2023.e02418.

Astarabadi, S.S.M. and Ebadzadeh, M.M. (2019). "Genetic programming performance prediction and its application for symbolic regression problems", *Information Sciences*, 502, 418-433, https://doi.org/10.1016/j.ins.2019.06.040.

Babaie, T., Karimizandi, R. and Lucas, C. (2008). "Learning based brain emotional intelligence as a new aspect for development of an alarm system", *Soft Computing*, 12, 857-873, https://doi.org/10.1007/s00500-007-0258-8.

Bae, K.O., Nguyen, T.T., Park, J., Park, J.S. and Baek, U.B. (2023). "Temperature dependency of hydrogen-related impact energy degradation of type 304 austenitic stainless steel", *Journal of Mechanical Science and Technology*, 37(6), 2891-2901, https://doi.org/10.1007/s12206-023-0515-5.

Biswas, R., Samui, P. and Rai, B. (2019). "Determination of compressive strength using relevance vector machine and ENN", *Asian Journal of Civil Engineering*, 20(8), 1109-1118, https://doi.org/10.1007/s42107-019-00171-9.

Bonakdari, H., Ebtehaj, I., Samui, P. and Gharabaghi, B. (2019). "Lake water-level fluctuations forecasting using minimax probability machine regression, relevance vector machine, Gaussian process regression, and extreme learning machine", *Water Resources Management*, 33(11), 3965-3984, https://doi.org/10.1007/s11269-019-02346-0.

Cevik, A. and Sonebi, M. (2008). "Modelling the performance of self-compacting SIFCON of cement slurries using genetic programming technique", *Computers Concrete*, 5(5), 475-490, https://doi.org/10.12989/cac.2008.5.5.475.

Dastorani, M.T., Mahjoobi, J., Talebi, A. and Fakhar, F. (2018). "Application of machine learning approaches in rainfall-runoff modeling (Case Study: Zayandeh_Rood basin in Iran)", *Civil Engineering Infrastructures Journal*, 51(2), 293-310, https://doi.org/10.7508/CEIJ.2018.02.004.

Deng, Z., Chen, J., Zhang, T., Cao, L. and Wang, S. (2018). "Generalized hidden-mapping minimax probability machine for the training and reliability learning of several classical intelligent models", *Information Sciences*, 436, 302-319, https://doi.org/10.1016/j.ins.2018.01.034.

Dubey, Y., Sharma, P. and Singh, M.P. (2023). "Optimization using genetic algorithm of GMAW parameters for Charpy impact test of 080M40 steel", *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 1-11, https://doi.org/10.1007/s12008-023-01371-z.

Hoang, N.D. and Tran, D.V. (2023). "Machine learning-based estimation of concrete compressive strength: a multi-model and multi-dataset study", *Civil Engineering Infrastructures Journal*, https://doi.org/10.1007/s12008-023-01371-z (Online).

Kar, S.S., Athawale, A.A., Bhushan, M. and Roy, L.B. (2023). "Dynamic soil-structure interaction of multi-story buildings using the finite element method and minimax probability machine regression", *Engineering, Technology & Applied Science Research*, 13(4), 11170-11176, https://doi.org/10.48084/etasr.5870.

Kecman, V. (2001). *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*, MIT press.

Khashman, A. (2008). "A modified backpropagation learning algorithm with added emotional coefficients", *IEEE Transactions on Neural Networks*, 19(11), 1896-1909, https://doi.org/10.1109/TNN.2008.2002913.

Khashman, A. (2009). "Blood cell identification using emotional neural networks", *Journal of Information Science & Engineering*, 25(6), 1737-1751.

Kohestani, V.R., Vosoghi, M., Hassanlourad, M. and Fallahnia, M. (2017). "Bearing capacity of shallow foundations on cohesionless soils: A random forest based approach", *Civil Engineering Infrastructures Journal*, 50(1), 35-49, https://doi.org/10.7508/ceij.2017.01.003.

Koshiyama, A.S., Vellasco, M.M. and Tanscheit, R. (2015). "GPFIS-CLASS: A genetic fuzzy system based on Genetic Programming for classification problems", *Applied Soft Computing*, 37, 561-571, https://doi.org/10.1016/j.asoc.2015.08.055.

Kovalev, M. (2021). "Ontology-based representation of an Artificial Neural Networks", *International Conference on Open Semantic Technologies for Intelligent Systems*, (pp. 132-151), Springer International Publishing, Cham, https://doi.org/10.1007/978-3-031-15882-7_8.

Koza, J.R. (1994). "Genetic programming as a means for programming computers by natural selection", *Statistics and C mputing*, 4, 87-112, https://doi.org/10.1007/BF00175355.

Kumar, C., Vardhan, H. and Murthy, C.S. (2022). "Artificial neural network for prediction of rock properties using acoustic frequencies recorded during rock drilling operations", *Modeling Earth Systems and Environment*, 8(1), 141-161, https://doi.org/10.1007/s40808-021-01103-w.

Kumar, M., Bardhan, A., Samui, P., Hu, J.W. and Kaloop, M.R. (2021). "Reliability analysis of pile foundation using soft computing techniques: A comparative study", *Processes*, 9(3), 486, https://doi.org/10.3390/pr9030486.

Lanckriet, G., Ghaoui, L., Bhattacharyya, C. and Jordan, M. (2001). "Minimax probability machine", *Advances in Neural Information Processing Systems*, 14.

Lin, C.C., He, R.X. and Liu, W.Y. (2018). "Considering multiple factors to forecast CO2 emissions: A hybrid multivariable grey forecasting and genetic programming approach", *Energies*, 11(12), 3432, https://doi.org/10.3390/en11123432.

Nazari, A., Milani, A.A. and Zakeri, M. (2011). "Modeling ductile to brittle transition temperature of functionally graded steels by artificial neural networks", *Computational Materials Science*, 50(7), 2028-2037.

Papa, J.P., Rosa, G.H. and Papa, L.P. (2017). "A binary-constrained geometric semantic GP for feature selection purposes", *Pattern Recognition Letters*, 100, 59-66, https://doi.org/10.1016/j.patrec.2017.10.002.

Park, D. and Rilett, L.R. (1999). "Forecasting freeway link ravel times with a multi-layer feed forward neural network", *Computer Aided Civil and Infrastructure Engineering*, 14, 358-367, https://doi.org/10.1111/0885-9507.00154.

Switzner, N.T., Anderson, J., Ahmed, L.A., Rosenfeld, M. and Veloo, P. (2023). "Algorithms to estimate the ductile to brittle transition temperature, upper shelf energy, and their uncertainties for steel using Charpy V-Notch shear area and absorbed energy data", *Metals*, 13(5), 877, https://doi.org/10.3390/met13050877.

Taylor, K.E. (2001). "Summarizing multiple aspects of model performance in a single diagram", *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183-7192, https://doi.org/10.1029/2000JD900719.

Vouros, G.A. (2022). "Explainable deep reinforcement learning: state of the art and challenges", *ACM Computing Surveys*, 55(5), 1-39, https://doi.org/10.1145/3527448.

Zhang, H., Zhou, J., Jahed Armaghani, D., Tahir, M.M., Pham, B.T. and Huynh, V.V. (2020). "A combination of feature selection and random forest techniques to solve a problem related to blast-induced ground vibration", *Applied Sciences*, 10(3), 869, https://doi.org/10.3390/app10030869.

Ziggah, Y.Y., Issaka, Y. and Laari, P.B. (2022). "Evaluation of different artificial intelligent methods for predicting dam piezometric water level", *Modeling Earth Systems and Environment*, 8(2), 2715-2731, https://doi.org/10.1007/s40808-021-01263-9.

Appendix

Table A1. Dataset used in this study

FGS Type (F)	The crack tip configuration (C)	The thickness of graded ferritic region (A)	The thickness of graded austetic region (G)	The distance of the notch from bainite or martensite intermediate layer (D)	Temperature (T)	CVN values obtained from experiments.
1	3	5	4.4	10	123	30
1	3	5	4.4	10	173	41
1	3	5	4.4	10	223	53
1	3	5	4.4	10	273	84
1	3	5	4.4	10	323	95
1	3	5	4.4	10	373	98
1	3	5	4.4	10	423	101
1	3	5	4.4	10	473	109
1	3	5	4.4	10	523	118
1	3	5	4.4	10	573	135
1	4	2	7.4	0	123	110
1	4	2	7.4	0	173	113
1	4	2	7.4	0	223	116
1	4	2	7.4	0	273	117
1	4	2	7.4	0	323	118
1	4	2	7.4	0	373	119
1	4	2	7.4	0	423	121
1	4	2	7.4	0	473	123
1	4	2	7.4	0	523	124
1	4	2	7.4	0	573	124
1	4	3	6.4	1	123	34

1	4	3	6.4	1	173	42
1	4	3	6.4	1	223	49
1	4	3	6.4	1	273	69
1	4	3	6.4	1	323	81
1	4	3	6.4	1	373	94
1	4	3	6.4	1	423	99
1	4	3	6.4	1	473	110
1	4	3	6.4	1	523	116
1	4	3	6.4	1	573	119
1	4	4	5.4	2	123	29
1	4	4	5.4	2	173	38
1	4	4	5.4	2	223	47
1	4	4	5.4	2	273	64
1	4	4	5.4	2	323	78
1	4	4	5.4	2	373	87
1	4	4	5.4	2	423	93
1	4	4	5.4	2	473	106
1	4	4	5.4	2	523	111
1	4	4	5.4	2	573	115
1	4	5	4.4	3	123	26
1	4	5	4.4	3	173	34
1	4	5	4.4	3	223	43
1	4	5	4.4	3	273	59
1	4	5	4.4	3	323	72
1	4	5	4.4	3	373	81
1	4	5	4.4	3	423	89
1	4	5	4.4	3	473	97
1	4	5	4.4	3	523	104
1	4	5	4.4	3	573	109
1	4	7.4	2	0	123	68
1	4	7.4	2	0	173	72
1	4	7.4	2	0	223	78
1	4	7.4	2	0	273	89
1	4	7.4	2	0	323	99
1	4	7.4	2	0	373	111
1	4	7.4	2	0	423	113
1	4	7.4	2	0	473	116
1	4	7.4	2	0	523	118
1	4	7.4	2	0	573	120
1	4	6.4	3	1	123	116
1	4	6.4	3	1	173	117
1	4	6.4	3	1	223	118
1	4	6.4	3	1	273	119
1	4	6.4	3	1	323	120
1	4	6.4	3	1	373	122
1	4	6.4	3	1	423	123
1	4	6.4	3	1	473	125
1	4	6.4	3	1	523	125
1	4	6.4	3	1	573	126
1	4	5.4	4	2	123	120
1	4	5.4	4	2	173	121
1	4	5.4	4	2	223	121
1	4	5.4	4	2	273	123
1	4	5.4	4	2	323	123
1	4	5.4	4	2	373	124
1	4	5.4	4	2	423	126
1	4	5.4	4	2	473	126
1	4	5.4	4	2	523	126
1	4	5.4	4	2	573	127
1	4	4.4	5	3	123	125
1	4	4.4	5	3	173	126
1	4	4.4	5	3	223	126
1	4	4.4	5	3	273	127
1	4	4.4	5	3	323	128

1	4	4.4	5	3	373	129
1	4	4.4	5	3	423	131
1	4	4.4	5	3	473	132
1	4	4.4	5	3	523	132
1	4	4.4	5	3	573	133
2	3	0	8.5	10	123	26
2	3	0	8.5	10	173	40
2	3	0	8.5	10	223	49
2	3	0	8.5	10	273	53
2	3	0	8.5	10	323	58
2	3	0	8.5	10	373	59
2	3	0	8.5	10	423	59
2	3	0	8.5	10	473	61
2	3	0	8.5	10	523	63
2	3	0	8.5	10	573	62
2	4	0	8.5	0	123	9
2	4	0	8.5	0	173	10
2	4	0	8.5	0	223	10
2	4	0	8.5	0	273	10
2	4	0	8.5	0	323	10
2	4	0	8.5	0	373	11
2	4	0	8.5	0	423	11
2	4	0	8.5	0	473	12
2	4	0	8.5	0	523	13
2	4	0	8.5	0	573	13
2	4	0	8.5	1	123	24
2	4	0	8.5	1	173	24
2	4	0	8.5	1	223	24
2	4	0	8.5	1	273	25
2	4	0	8.5	1	323	26
2	4	0	8.5	1	373	27
2	4	0	8.5	1	423	27
2	4	0	8.5	1	473	27
2	4	0	8.5	1	523	28
2	4	0	8.5	1	573	29
2	4	0	8.5	2	123	46
2	4	0	8.5	2	173	47
2	4	0	8.5	2	223	47
2	4	0	8.5	2	273	47
2	4	0	8.5	2	323	48
2	4	0	8.5	2	373	49
2	4	0	8.5	2	423	50
2	4	0	8.5	2	473	50
2	4	0	8.5	2	523	51
2	4	0	8.5	2	573	51
2	4	0	8.5	3	123	82
2	4	0	8.5	3	173	83
2	4	0	8.5	3	223	84
2	4	0	8.5	3	273	84
2	4	0	8.5	3	323	84
2	4	0	8.5	3	373	85
2	4	0	8.5	3	423	85
2	4	0	8.5	3	473	86
2	4	0	8.5	3	523	87
2	4	0	8.5	3	573	88