https://doi.org/10.36375/prepare_u.iiche.a362

Development of Artificial Intelligent (AI)-based Model for Steel Alloys

***NUSRAT PARVEEN*[1*], *SADAF ZAIDI*[2], *MOHAMMAD DANISH*[3]**

[1]*Department of Chemical Engineering, Government Polytechnic, Bahraich, 271801, India.*
[2]*Department of Post Harvest Engineering and Technology, Aligarh Muslim University.*
[3]*Department of Chemical Engineering, Aligarh Muslim University, Aligarh, 202002, India*
**Corresponding author's mail id:* nusrataliq@gmail.com

Abstract:

Alloy steel has many forms each having different properties due to the elements added to the steel. When treated with heat, these elements impart a wide range of physical properties incomparable to any metals/alloys. In the present work, the artificial intelligent (AI) technique namely, artificial neural networks (ANN) is utilized to model the true stress (τ) of ultra-strength Cr-Mn-Si-Ni ultra-strength alloyed steel in terms of holding time, heating rate, tensile temperature (T_{ts}) and strain rate (γ). Neural networks are trained iteratively by adjusting the connections between nodes and the weight. A well-trained network can effectively predict the target. The developed ANN-based model is compared to the commonly employed multiple regression (MR) model in terms of statistical parameters. The coefficient of determination (R^2) values for the ANN and MR models are 0.9948 and 0.2924 on the other hand average absolute relative error (AARE) are observed as 3.6%, and 56.16% respectively. The results thus obtained show that the ANN-based model has higher accuracy with greater generalization.

Keywords: *Ultra-strength steel; artificial neural networks (ANN); multiple regression (MR); average absolute relative error (AARE); coefficient of determination (R^2).*

1. Introduction

Steel has several types of alloys ranging from stainless steel to high temperature steels with flat carbon products are there. Due to some enviable characteristics such as low cost, fast and easy installation, high strength, stiffness, lightness and mass production, etc., steel is utilized for the production of many products and in the construction of buildings and automotive bodies. The properties of all alloy steels depend up on the elements added to the steel. Alloy steels and plain carbon steels are mainly two kinds of steels. Metals or other elements with different proportions are combined with steel for the production of alloy steels[1,2]. Steels are basically the alloys of iron and carbon. When treated with heat, a wide range of physical properties incomparable to any metal/alloys are produced. There are numerous benefits of alloys such as the introduction of magnetic properties, corrosion and electrical resistance, hardness at red heat, low coefficient of expansion, etc. With the inclusion of one or more metals a broad span of mechanical properties are available which are likely to beat any specific qualities. Therefore, it is viable to build more elasticity, toughness or hardness. However, most important is the prospect of achieving high tensile strength but with slight decrease in ductility. Besides, for plain carbon steels with the rise in strength there is a rise in carbon resulting in a decline in ductility.

Artificial intelligence techniques such as artificial neural networks (ANN) in particular, are an efficient technique to predict optimal conditions as it does not include any assumptions or simplifications. ANN mimics the processes of human learning. Neural networks are trained iteratively by adjusting the connections between nodes and the weights. A well-trained network can effectively predict the target. For large input datasets it has many advantages over the conventional statistical methods. It does not require presumed mathematical equations to show the relationship among the model inputs and corresponding outputs and can automatically detect the hidden relationships or patterns between input and output present in the data to find out the model structure[3-5]. Any continuous function can be approximated to the desired accuracy with large number of hidden neurons by ANN. Using this intelligent method for evaluating true stress-true strain curve has not been done and to best of authors' knowledge no work in the open literature is reported till now.

In the current study, ANN has been used for the prediction of true stress for ultra-strength Cr-Mn-Si-Ni alloyed steel. ANN model has been compared with multiple regression (MR) model on the basis of statistical measures[6]. Zhang et al.[7] performed the tensile tests experiment of Cr-Mn-Si-Ni alloyed steel using the parameter were tensile temperature (°C) [25-900], Heating rate (10 °C/s), Isothermal holding time (120 s) and Strain rate (s^{-1}) [0.0005-0.01]. These

https://doi.org/10.36375/prepare_u.iiche.a362

experimental data have been utilized to develop the models and validate them.

2. Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) process the information in parallel as inspired by human nervous system. ANN was first introduced by McCulloch and Pitts.[8] ANNs have the ability to learn from the experiences like humans. ANN is composed of artificial neuron as nodes or units. Each unit is connected to other by its weight. Processing units and its weight are summed up with an adjustable bias unit. This summed response goes to some transfer (activation) function[9,10] and the output is thus produced. Figure 1 represents an artificial neuron with multiple inputs. ANN processing depends on weights or learning, network topology, and the activation function. They are also termed as the building blocks of ANN.

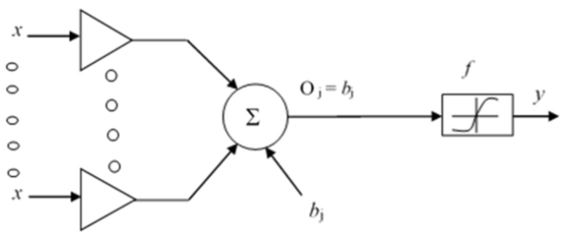

Fig. 1: An artificial neuron

2.1 ANN Modelling Procedure

The procedure needed for the development of the ANN model is described via flowchart[11-14] exhibited in Figure 2. Firstly, dataset in the form of independent and dependent variables is collected for a process and then the complete dataset is preprocessed (data normalization or scaling and sometimes outliers are also needed to be removed). The total dataset is then partitioned into two sets as training and test sets. The second step involves the optimum structure of ANN i.e., the type of ANN, number of neurons at the input layer, number of hidden layers and the output layer, type of activation function at the hidden layer and the output layer are defined. After the network structure is finalized, neural networks training is done by adjusting weights and bias, checking the generalization performance and training process is finish with post training analysis and the developed ANN model performance is analyzed using test set on the basis of some statistical parameters. ANN is now ready for prediction.

3. Development of ANN model

In the present study, the true stress of ultra-strength alloyed steel has been predicted by ANN with pure-linear, log-sigmoid and tan-sigmoid activation functions. The complete dataset consisting of 223 samples was retrieved from the available published literature[7]. Then it was partitioned into 80% (178 data points) and 20% (45 data points) as the training dataset and the test dataset, respectively. The developed ANN models have been evaluated and validated with respect to statistical measures and then their performance has been compared with the widely employed multiple regression (MR) model.

3.1. Assessment of ANN model

Artificial neural networks with forward feed were used in the present study. It is represented in Figure 3. It comprised of three layers. They were basically the input layer consisting of all the independent variables, hidden layer and the output layer consisting of the dependent variable. Neural network architecture 4-10-1 was employed for modeling the true stress of alloyed ultra-strength steel with 4 neurons for holding time (s), heating rate (˚C/s), tensile temperature (˚C), and strain rate (s^{-1}) at the input layer. The hidden layer consisted of 10 neurons and 1 neuron was for the output layer. Input and output layer neurons were fixed according to specific problem whereas hidden layer neuron was selected by trial-and-error procedure because inadequate number of neurons give rise to under fitting whereas more neurons may result in over fitting[15-17]. The ANN model training was done until the mean square error (MSE) was minimal subsequently comparing the actual experimental values with the network output[18,19]. A properly trained model achieves highest values of coefficient of correlation (R) approximately to 1 and lowest MSE. MATLAB 2015 mathematical software with the ANN Toolbox was used to carry out all the computation.

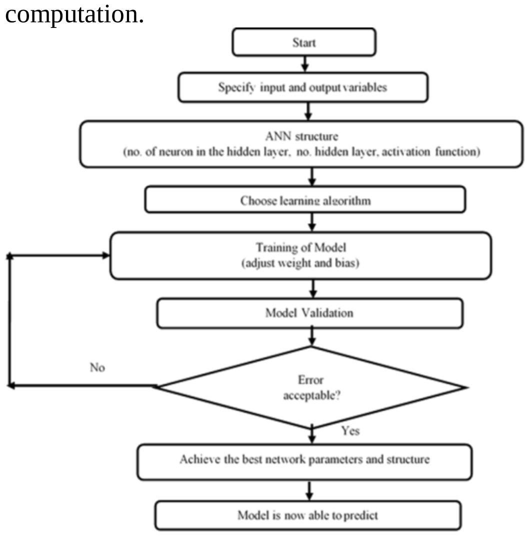

Fig. 2. ANN modelling procedure

https://doi.org/10.36375/prepare_u.iiche.a362

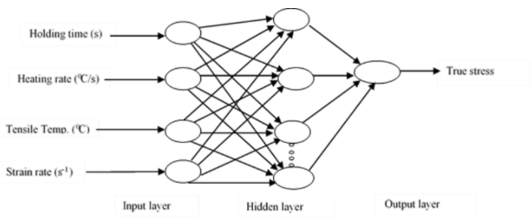

Fig. 3. Schematic topology of 4-10-1 ANN architecture.

Table 1 exhibits the MSE and R values for different network architectures. After several simulation experiments, it was found that the network structure 4-10-1 with Levenberg-Marquardt '*trainlm*' training algorithm exhibited the best predictability with minimum MSE value and the maximum value of R.

Table 1. Performance parameters of various network structures trained using Levenberg-Marquardt '*trainlm*' algorithm

Network Structure	**Training**		**Testing**	
Input-hidden-output layer	Correlation coefficient (R)	Mean square error (MSE)	Correlation coefficient (R)	Mean square error (MSE)
4-2-1	0.900264	482.1432	0.932258	274.29105
4-4-1	0.954193	212.4739	0.953369	280.40387
4-5-1	0.988089	65.9339	0.974146	95.77393
4-8-1	0.983268	72.32596	0.986103	55.98957
4-10-1	**0.990376**	**48.46865**	**0.991143**	**41.48313**
4-12-1	0.98279	69.10040	0.984184	70.16034
4-15-1	0.956008	206.61867	0.958719	263.98404
4-18-1	0.986416	72.53314	0.980613	77.64014
4-20-1	0.987465	58.33295	0.988656	79.51751

The ANN model predicted true stress values for ultra-strength alloyed steel were compared to the experimental values of the training set and the test set as revealed in the Figure 4. The predicted data points are found to lie near to the ideal fit line. The adequate selection of neuron results in much improved performance than the MR model. Further, the model output is compared with the experimental data for each run in Figures 5 and 6, respectively. The predicted data overlaps the true stress experimental data for most runs. Comparison of statistical parameters for the predicted training set and test set are depicted in Table 2. These parameters are found in close proximity to each other. On the basis of the analysis of Figures 4, 5 and 6 and the Table 2, the ANN-based model gives a superior prediction performance with high generalizability. Moreover, the test dataset results are much improved than those of the train dataset.

Table 2. Statistical measures for ANN model of training and test sets.

Statistical measures	**Train data**	**Test data**
AARE (%)	5.65	3.6
R^2	0.9897	0.9948
RMSE	7.0025	5.1485
SD	3.1417	0.0275
MRE	0.0565	0.036
Q^2_{LOO} (Train data), Q^2_{ext} (Test data)	0.9793	0.9891

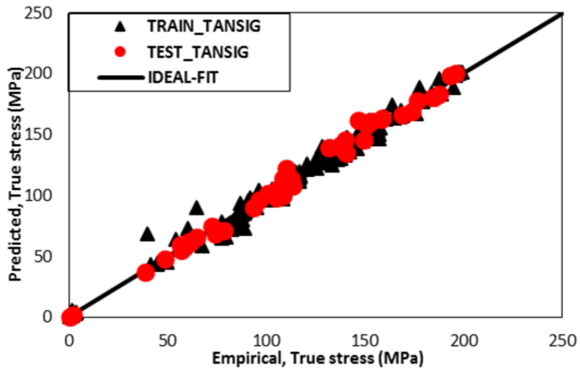

Fig. 4. ANN simulation of the true stress for ultra-strength alloyed steel using optimal parameters for training set and test set.

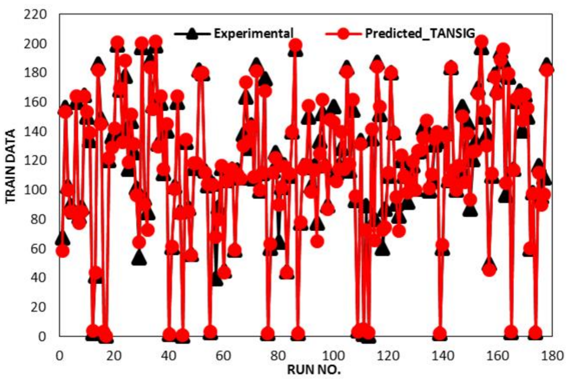

Fig. 5. Training course curve of the true stress for ultra-strength alloyed steel.

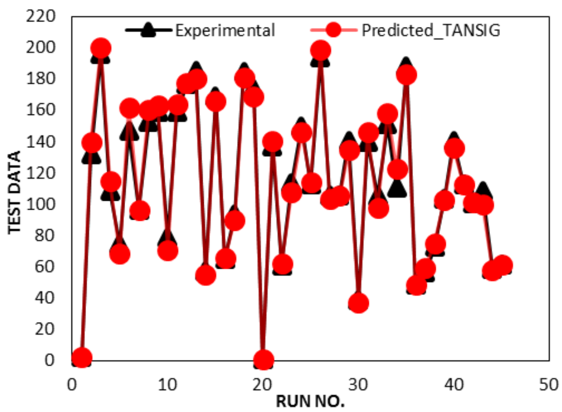

Fig. 6. Test course curve of true stress for ultra-strength alloyed steel.

https://doi.org/10.36375/prepare_u.iiche.a362

3.2. ANN Model Versus the MR Model

The MR model was developed with 178 samples of training dataset to predict the stress that depends on holding time (x_1), heating rate (x_2), tensile temperature (x_3), and strain rate (x_4).
The MR model equation thus obtained is given below:

$$y = 11.3313x_2^{-3.0866}x_3^{0.0945}x_4^{0.7613}$$

In the regression, the independent variables were observed to be significant.
The predictability of the ANN model with pure-linear, log-sigmoid, tan-sigmoid activation functions at the hidden layer is compared with the MR model in Table 3 based on the test set. The ANN model with tan-sigmoid activation function is found to have the highest R^2 value and gives the lowest MRE, SD (standard deviation), RMSE (root mean square error), and AARE (average absolute relative error) values among the four models. Further, the parity plot of the true stresses for the four models displayed in Figure 7 utilizing the test set, clearly shows that the ANN model with tan-sigmoid activation function predicts the data point more accurately than the other three models with least scatter from the ideal fit line. All this goes on to show that the ANN model with tan-sigmoid activation function is having the highest accuracy and generalizability and therefore gives superior prediction performance.

Table 3. Statistical evaluation of ANN and MR models for test set.

Statistical measures	ANN-PURELIN model	ANN-LOGSIG model	ANN-TANSIG model	MR model
AARE (%)	24.77	11.11	3.6	56.16
R^2	0.2837	0.9602	0.9948	0.2924
RMSE	42.21907	10.6853	5.1485	69.6088
SD	23.6132	4.8269	0.0275	2.3534
MRE	0.2477	0.9055	0.0358	0.5616

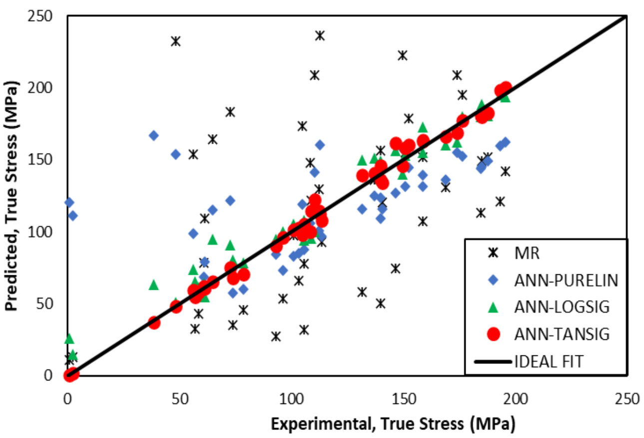

Fig. 7: Comparison of the MR with the ANN models for predicting the true stress of ultra-strength alloyed steel using test data.

The true stress values for the ultra-strength alloyed steel as predicted via MR and ANN models in terms of AD for the training set are given in Table 4. It was observed that almost 89.33% data points predicted via ANN-based model lie below 10% AD and 96.07% of the total data points lie within 20% AD. Above 20% AD there are only 3.93% data points. However, for the MR model, 14.01% data points are predicted within 10% AD and 32.55% of the total data points lie below 20% AD. Above 20% AD there are 67.45% data points. Therefore, it can be said that the ANN-based model predicts excessively huge percentage of data with a little error margin.
The true stress values for the ultra-strength alloyed steel as predicted via MR and ANN models in terms of AD for the test set are given in Table 5. It was found that almost 97.78% data points predicted via ANN-based model lie below 10% AD and the entire data points lie below 20% AD. Whereas, 8.89% data points are predicted by MR model within 10% AD and 11.11% of the total data points lie below 20% AD. Above 20% AD there are 88.89% data points.
This depicts that the ANN has high accuracy and is very generalized for the unseen test dataset as well. This is because of the selection of optimum parameters of the developed model i.e., tan-sigmoid activation function at the hidden layer, optimal number of neurons, and the Levenberg-Marquardt training algorithm.

4. CONCLUSIONS

The comparative study, training and test course curves, and the statistical evaluation parameters values, point to the high predictability of the ANN model as compared to the MR model. It means that MR model has the poor prediction performance. It also emphasises that the ANN-based model with tan-sigmoid activation function simulates the best among the three activation functions used namely, pure-linear, log-sigmoid and tan-sigmoid.

Table 4. Percentage distribution of the true stress data via MR and ANN-based models in terms of AD for training set.

AD (%)	% of MR model predicted values	Cumu-lative score	% of ANN model predicted values	Cumu-lative score
AD< 10	14.01	14.01	89.33	89.33
10<AD< 20	18.54	32.55	6.74	96.07
AD>20	67.45	100	3.93	100
Total	100		100	

Table 5. Percentage distribution of the true stress data via MR and ANN-based models in terms of AD for test set.

AD (%)	% of MR model predicted values	Cumu -lative score	% of ANN model predicted values	Cumu -lative score
AD< 10	8.89	8.89	97.78	97.78
10<AD < 20	2.22	11.11	2.22	100
AD>20	88.89	100	0.00	
Total	100		100	

ACKNOWLEDGEMENTS: Authors truly want to thank Messrs Jing-Yu Zhang, Ping Jiang, Ze-lin Zhu, Qiang Chen, Jie Zhou and Yi Meng for their published research paper. The data for this modelling work has been collected from there. All the authors are also acknowledged whose studies have been referred to in this paper.

REFERENCES

[1] M.F. Carlson, B. V. Narasimha Rao, G. Thomas, The effect of austenitizing temperature upon the microstructure and mechanical properties of experimental Fe/Cr/C steels, Metall. Trans. A. 10(1979)1273–1284. https://doi.org/10.1007/BF02811983.

[2] W.S. Lee, T.T. Su, Mechanical properties and microstructural features of AISI 4340 high-strength alloy steel under quenched and tempered conditions, J. Mater. Process. Technol. 87(1999)198-206. https://doi.org/10.1016/S0924-0136(98)00351-3.

[3] N. Parveen, S. Zaidi, M. Danish, Artificial intelligence (AI)-based friction factor models for large piping networks, Chem. Eng. Commun. 207 (2019)213–230. https://doi.org/10.1080/00986445.2019.1578757.

[4] N. Parveen, S. Zaidi, M. Danish, Development and Analyses of Artificial Intelligence (AI)-Based Models for the Flow Boiling Heat Transfer Coefficient of R600a in a Mini-Channel, ChemEngineering.2(2018)27. https://doi.org/10.3390/chemengineering2020027.

[5] N. Parveen, S. Zaidi, M. Danish, Comparative analysis for the prediction of boiling heat transfer coefficient of R134a in micro/mini channels using artificial intelligence (AI)-based techniques, Int. J. Model.Simul.40(2020)114–129. https://doi.org/10.1080/02286203.2018.156480

[6] S. Zaidi, Support Vector Regression (SVR) Model to Predict the Boiling/Non-boiling Length of the Heated Tube in a Vertical Tube Thermosiphon Reboiler, in: B.B. Singh, G.C. Mishra, S.K. Yadav (Eds.), 3rd Int. Conf. "Innovative Approach Appl. Phys. , Math. / Stat. , Chem. Sci. Emerg. Energy Technol. Sustain. Dev., "Social Welfare Foundation" In Association with "Krishi Sanskriti," Jawaharlal Nehru University, New Delhi, India, 2014: pp. 1–139.

[7] J.Y. Zhang, P. Jiang, Z. lin Zhu, Q. Chen, J. Zhou, Y. Meng, Tensile properties and strain hardening mechanism of Cr-Mn-Si-Ni alloyed ultra-strength steel at different temperatures and strain rates, J. Alloys Compd. 842 (2020) 155856. https://doi.org/10.1016/j.jallcom.2020.155856.

[8] W.S. Mcculloch, W. Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, Bull. Math. Biophys. 5 (1943) 115–133.

[9] Ö.F. Ertuğrul, A novel type of activation function in artificial neural networks: Trained activation function, Neural Networks. 99 (2018) 148–157. https://doi.org/10.1016/j.neunet.2018.01.007.

[10] J. Feng, S. Lu, Performance Analysis of Various Activation Functions in Artificial Neural Networks, J. Phys. Conf. Ser. 1237 (2019) 1237. https://doi.org/10.1088/1742-6596/1237/2/022030.

[11] V. Devabhaktuni, M. Yagoub, Y. Fang, J. Xu, Q.-J. Zhang, Neural networks for microwave modeling: Model development issues and nonlinear modeling techniques, Int. J. RF Microw. Comput. Eng. 11 (2001) 4–21.

[12] N. Parveen, S. Zaidi, M. Danish, Development and analyses of data-driven models for predicting the bed depth profile of solids flowing in a rotary kiln, Adv. Powder Technol. 31 (2020) 678–694.

[13] N. Parveen, S. Zaidi, M. Danish, Support Vector Regression Prediction and Analysis of the Copper (II) Biosorption Efficiency, Indian Chem. Eng. 59 (2017) 295–311. https://doi.org/10.1080/00194506.2016.1270778.

[14] N. Parveen, S. Zaidi, M. Danish, Development of SVR-based model and comparative analysis with MLR and ANN models for predicting the sorption capacity of Cr(VI), Process Saf. Environ. Prot. 107 (2017) 428–437.

https://doi.org/10.1016/j.psep.2017.03.007.
[15] M. Massinaei, Estimation of metallurgical parameters of flotation process from froth visual features, Int. J. Min. Geo-Eng. 49 (2015) 75–81.
[16] S. Karsoliya, Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture, Int. J. Eng. Trends Technol. 3 (2012) 714–717.
[17] H. Merdun, Ö. Çinar, Artificial neural network and regression techniques in modelling surface water quality, Environ. Prot. Eng. 36 (2010) 95–109.
[18] S. Aber, a. R. Amani-Ghadim, V. Mirzajani, Removal of Cr(VI) from polluted solutions by electrocoagulation: Modeling of experimental results using artificial neural network, J. Hazard. Mater. 171 (2009) 484–490. https://doi.org/10.1016/j.jhazmat.2009.06.025.
[19] A. Verikas, M. Bacauskiene, Using artificial neural networks for process and system modelling, Chemom. Intell. Lab. Syst. 67 (2003), 187–191. https://doi.org/10.1016/S0169-7439(03)00093-5.