Development of Regression Model to Predicting Yield Strength for Different Steel Grades

Charanjeet Singh Tumratea, Shambo Roy Chowdhuryb, Dhaneshwar Mishrac,d

aDepartment of Civil Engineering, Manipal University Jaipur, Dehmi Kalan, Jaipur, Rajasthan 303007, India
bDepartment of Mechatronics Engineering, Manipal University Jaipur, Dehmi Kalan, Jaipur, Rajasthan 303007, India
cDepartment of Mechanical Engineering, Manipal University Jaipur, Dehmi Kalan, Jaipur, Rajasthan 303007, India
dMulti Scale Simulation Research Center, Manipal University Jaipur, Dehmi Kalan, Jaipur, Rajasthan 303007, India

Abstract

Yield strength of steel is an important property in designing a steel component. The maximum load that material can sustain without undergoing permanent deformation is represented by yield strength. A robust model is purposed, which can accurately predict the yield strength of different steel grades, based on the varying chemical composition of steel from the employed database. The dataset consisting of chemical composition and experimental yield strength of steel rods from a collapsed building site. MATLAB® was used to generate a regression model such as multivariate regression and linear regression, to predict the yield strength values for different steel grades. A comparative study among the different regression models for RMS error, Max error, and highest accuracy is discussed in this paper.

Keywords: yield strength, steel grade, multivariate, linear, regression

1. Introduction

The material develops permanent deformations, once the stresses in materials reaches yield stress limit. Therefore, the material consists of residual deformation at yield point, once the applied stress is released. In absence of well-defined yield point, 0.2% offset is considered to determine the approximate yield point. In experimental/laboratory analysis to determine material properties, low sampling rate for sample testing is a real concern in determining accurate results. Statistical tool can be used to create mathematical model, in-order to ease the prediction process. The important parameters such as material’s chemical composition and its manufacturing process, are vital in predicting mechanical properties of the materials. For manufacturing high quality steel at minimum cost, it is important to understand the effect of each variable at production stage. Mathematical model has been developed to predict material properties such as yield strength, tensile strength and percentage elongation. In past, material properties of aluminum alloy, used in defense sector, were predicted by considering database including parameters of friction stir welding process. Mathematical models, with confidence level of 95%, was developed using response surface methodology with adequacy check performed by analysis of variance technique [1]. Technique such as linear regression has been applied to develop mathematical function which in turn assisted in predict mechanical properties of rebar by considering the rebar’s chemical composition and thermo-mechanical variables as input dataset [2]. The relation between tensile strength and impact strength of carbon steels, varying carbon from 0.15 to 0.3%, have also been predicted by performing correlation analysis [3]. Empirical formula to predict steel mechanical properties was established by considering chemical, mechanical and rolling parameters of industrial hot-rolled coils [4]. The low-alloyed steel plate’s tensile strength was determined by developing predicting model by considering element concentrations and rolling process variables. Special attention was also paid as large data was considered in development of statistical models. This was done so that overfitting can be avoided [5]. Multivariate adaptive regression is also used to predicted and the mechanical properties (yield, tensile, elongation) of a steel strip [6]. The steel such as Admit is roll manufactured. The effect of Adamite steel’s chemical compositions on its mechanical properties can be predicted by using the multiple regression technique. Graphical representation of correlations coefficients can be obtained by using software such as MATLAB® [7].

The present study purposes a regression technique using MATLAB®, to predict mechanical property (yield strength) of steel rods from a collapsed building site. The database [8] consist of chemical composition and mechanical properties of the scrap steel that needs to be reused as a raw material in production of reinforced steel rods. The
employed database consists of chemical composition and material properties of 12 mm and 16 mm diameter steel rods. These rods were collected from 6 different collapsed building site locations and two steel plants. Total of 11 steel rod sample’s chemical composition along with their yield strength values were mentioned in the dataset. The predicted yield strength value is obtained by using multivariate regression and linear regression using MATLAB®.

2. Regression Models

The material properties such as strength, hardness, ductility or malleability along with weldability are very sensitive to the carbon content in steel. The carbon content along with other elements such as chromium, manganese, molybdenum, vanadium, copper, silicon and nickel, within steel composition also affect the carbon equivalence in a sample. It becomes vital to determine the correlation among the chemical compositions, processing parameters and material properties. These relations will help in development of new and economic materials. Equivalent carbon content in steel by standard expression, given by American Welding Society (AWS), is given in equation 1.

\[ CE_{-}AWS = C + \frac{(Mn+Si)}{6} + \frac{(Cr+Mo+V)}{5} + \frac{(Cu+Ni)}{15} \]  

Multivariate regression technique is applied to predict the response variable (yield strength) value, with respect to variations in predictor variables (chemical composition). For regression models, the response variables are assumed to be a linear (or non-linear) combination of predictor variables and can be defined by equation 2. In this equation \( Y \) is the response vector and \( X \) is predictor variable matrix of dimension \( N \times P \) where \( N \) rows are equivalent to \( N \) observations and \( P \) columns are equivalent to \( P \) predictor variables. \( \beta \) is the coefficient vector and \( \varepsilon \) is the error term. The predicted output \( Y' \) is calculated as per equation 3. The regression algorithm runs iterations to minimize \( \varepsilon \). For this purpose, we have used MATLAB® inbuilt function ‘mvregress’ to generate the regression model where PLS (partial least square) as optimization parameter. Out of eleven samples, eight samples were randomly selected as training samples while for testing all 11 samples were used to calculate RMS error. Data in table 1 and table 2 were used to testing the concept discussed above.

\[ Y = X\beta + \varepsilon \]  
\[ Y' = \hat{X}\beta \]  

| SI.No. | Steel grade | Yield strength (N/mm2) |
|-------|-------------|-----------------------|
| 1     | EA          | 460.14                |
| 2     | IA          | 486.32                |
| 3     | SA          | 551.56                |
| 4     | OA          | 492.10                |
| 5     | AA          | 467.27                |
| 6     | ALA         | 466.94                |
| 7     | A12         | 405.69                |
| 8     | A16         | 389.12                |
| 9     | B10         | 410.59                |
| 10    | B12         | 404.62                |
Table 2: The X matrix of predictors

| SL.NO. | Chemical Composition (in %) | EA   | IA   | SA   | OA   | AA   | ALA  | A12  | A16  | B10  | B12  | B16  |
|--------|----------------------------|------|------|------|------|------|------|------|------|------|------|------|
| 1      | C                          | 0.339| 0.311| 0.345| 0.324| 0.351| 0.315| 0.259| 0.329| 0.330| 0.169| 0.291|
| 2      | Si                         | 0.231| 0.223| 0.206| 0.222| 0.23  | 0.216| 0.179| 0.176| 0.307| 0.228| 0.193|
| 3      | S                          | 0.080| 0.086| 0.079| 0.018| 0.049| 0.063| 0.038| 0.036| 0.040| 0.047| 0.042|
| 4      | P                          | 0.069| 0.079| 0.068| 0.075| 0.067| 0.056| 0.043| 0.042| 0.045| 0.056| 0.054|
| 5      | Mn                         | 0.983| 0.991| 0.806| 1.001| 0.963| 0.981| 0.519| 0.555| 0.727| 0.579| 0.579|
| 6      | Ni                         | 0.106| 0.107| 0.110| 0.112| 0.121| 0.123| 0.100| 0.112| 0.091| 0.085| 0.105|
| 7      | Cr                         | 0.223| 0.223| 0.225| 0.212| 0.220| 0.205| 0.154| 0.164| 0.163| 0.204| 0.271|
| 8      | Mo                         | 0.030| 0.030| 0.031| 0.023| 0.031| 0.032| 0.028| 0.023| 0.027| 0.030| 0.025|
| 9      | V                          | 0.006| 0.006| 0.006| 0.005| 0.005| 0.048| 0.005| 0.004| 0.004| 0.005| 0.004|
| 10     | Cu                         | 0.283| 0.284| 0.282| 0.281| 0.283| 0.281| 0.342| 0.261| 0.245| 0.292| 0.308|
| 11     | W                          | 0.012| 0.012| 0.011| 0.013| 0.011| 0.012| 0.010| 0.013| 0.015| 0.012| 0.011|
| 12     | Ti                         | 0.002| 0.002| 0.002| 0.002| 0.002| 0.002| 0.001| 0.002| 0.003| 0.001| 0.003|
| 13     | Sn                         | 0.016| 0.016| 0.015| 0.013| 0.016| 0.017| 0.014| 0.015| 0.017| 0.016| 0.013|
| 14     | Co                         | 0.011| 0.011| 0.011| 0.012| 0.011| 0.011| 0.012| 0.010| 0.014| 0.013| 0.011|
| 15     | Al                         | 0.006| 0.003| 0.004| 0.005| 0.005| 0.003| 0.021| 0.021| 0.021| 0.020| 0.021|
| 16     | Nb                         | 0.001| 0.001| 0.001| 0.001| 0.001| 0.002| 0.001| 0.001| 0.002| 0.002| 0.001|
| 17     | Fe                         | 97.602| 97.615| 97.618| 97.681| 97.634| 97.634| 98.273| 98.236| 97.950| 98.241| 98.068|

3. Result

The regression lines help us in predicting the average y term which is associated with the given x term. The root mean square error determines the spread of residuals i.e. the difference between the actual value and the predicted values. The average root mean square (RMS) is calculated using regression model for all eleven samples. The average percentage root mean square error for twenty different iterations was around 1.8% of the average yield strength. The minimum and maximum percentage RMS were 0.28% and 2.5% of the average yield strength. Table 3 shows the comparison of the actual yield strength values of the eleven samples from the literature with the predicted yield.
strength value. The predicted values are rounded off to four significant figures.

### Table 3: Actual and predicted yield strength value

| Sample label | Average Yield strength | Predicted yield strength |
|--------------|-------------------------|--------------------------|
| EA           | 460.14                  | 477.8                    |
| IA           | 486.32                  | 474.2                    |
| SA           | 551.56                  | 547.0                    |
| OA           | 492.1                   | 471.4                    |
| AA           | 467.27                  | 482.7                    |
| ALA          | 466.94                  | 495.6                    |
| A12          | 405.69                  | 407.1                    |
| A16          | 389.12                  | 392.4                    |
| B10          | 410.59                  | 420.0                    |
| B12          | 404.62                  | 335.0                    |
| B16          | 373.13                  | 366.9                    |

![Fig. 1: Yield strength: Predicted Vs Actual (0.28% RMS Error)](image-url)
4. Conclusion

The error term determined from eight randomly selected samples were tested for all eleven samples to calculated root mean square error. The minimum percentage root mean square were 0.28% of the average yield strength as shown in Fig. 1. This clearly shows the difference between the literature yield strength and predicted yield strength of the material is acceptable when considering 0.28% RMS error. For 0.38% RMS error the prediction in not that accurate as shown in Fig. 2. Further accuracy can be obtained, by applying multivariate regression technique, by increasing the input data in the form of chemical composition and material properties of more number of material samples.

5. Reference

[1] R. Palanivel, P. Koshy Mathews and N. Murugan, Development of mathematical model to predict the mechanical properties of friction stir welded AA6351 aluminum alloy, Journal of Engineering Science and Technology Review 4 (1), 2011, pp.25-31

[2] Raphaella H. F. Murta, Fabricio D. Braga, Pedro P. N. Maia, Otilio B. F. Diógenes and Elineudo P. de Moura, Mathematical modelling for predicting mechanical properties in rebar manufacturing, Institute of Materials, Minerals and Mining, 2020.

[3] Y. Unigovsky, Regression analysis of the mechanical properties—Composition dependencies for cast low- and medium-carbon steels, Journal of Materials Engineering and Performance volume 9, 2020.pp. 365-369.

[4] Jirathanakul Noppon and Somrerk Chandra-Ambhorn, Prediction of the Mechanical Properties of Hot-Rolled Low Carbon Steel Strips in Correlation to Chemical Compositions and Rolling Conditions, Key Engineering Materials, 2011, pp. 401-406

[5] Ilmari Juutilainen, Juha Rönöing and Lassi Myllykoski, Modelling the strength of steel plates using regression analysis and neural networks, InProceedings of International Conference on Computational Intelligence for Modelling, Control and Automation, 2003, pp.681-691

[6] A. Mukhopadhyay and A. Iqbal, Prediction of mechanical property of steel strips using multivariate adaptive regression splines, Journal of Applied Statistics, 2009, pp. 1-9

[7] Kiss I, Alexa V, Serban S, Rackov M and Čavić M, Statistical research using the multiple regression analysis in areas of the cast hipereutectoid steel rolls manufacturing, IOP Conference Series: Materials Science and Engineering, 2018, pp. 012077
[8] A.A. Adeleke, J.K. Odusote, P.P. Ikubanni, O.A. Lasode, O.O. Agboola, A. Ammasi and K.R. Ajao, Dataset on the evaluation of chemical and mechanical properties of steel rods from local steel plants and collapsed building sites, Data in Brief, 2018, pp. 1552-1557.