



An Approach of Data Science for the Prediction of Wear Behaviour of Hypereutectoid Steel

Poornima Hulipalled¹ · Veerabhadrapa Algur² · V. Lokesha³

Received: 19 January 2022 / Revised: 18 March 2022 / Accepted: 18 April 2022 / Published online: 8 May 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

The data science has developed into an effective approach for predicting wear behaviour of solid materials. This contemporary work has mainly focused on the prediction of wear rate for the hypereutectoid steel which are tested under different operating conditions (sliding speed, normal pressure and sliding distance). To train and forecast the wear rate of the models, supervised machine learning methods such as random forest, Gaussian process regression, and support vector machine were used. Analysis of variance yielded 63.74%, 28.52%, and 1.08% for the contributions of normal pressure, sliding speed, and sliding distance, respectively. In order of prediction accuracy, the three machine learning algorithms used are: RF, GPR, and SVM. RF outperformed all other created models in terms of R^2 (training and test), mean absolute error, and root mean squared error. At greater loads and speeds, the worn surface reveals that sub-surfaces fractured and broke, forming a plate-like structure of wear debris. The discoveries might speed up the creation of new functional hypereutectoid steel by allowing for the manufacture of hypereutectoid steel with controlled wear properties.

Keywords ANOVA · Hypereutectoid steels · Pin-on-disc · Random forest · Support vector machine · Gaussian process regression

1 Introduction

Hypereutectoid steels are mostly utilized in the fabrication of working devices such as rolling mills, forklifts, and bulldozers, as well as metal forming tools [1, 2]. These steels must meet the following criteria: adequate hardenability, wear resistance, maximum strength, fracture toughness, and high hardness. The amount of carbon in steel determines its hardness [3–7]. Wear is the most typical mode of failure for these materials. As a result, it's important to understand the wear rate under different operating conditions.

Many studies reported that wear rate of hypereutectoid steels under different operating conditions. For instance,

Etsuo Marui et al. [8] studied the wear behaviour of hypereutectoid steels at different operating temperature and prepared two types of specimens: quenched and as-received specimens for conducting the wear test. The result shows that quenched specimens have high wear resistance than the as-received specimens. Xingjian Gao et al. [9] examined the wear behaviour of full pearlite with and without thermo-mechanical treatments on bimetallic hypereutectoid steels at room temperature. The result shows that when interlamellar spacing is greater than 200 nm, delamination dominates the wear process, whereas when interlamellar spacing is less than 200 nm, substantial wear occurs.

Despite the fact that several experimental investigations have been conducted to analyse the wear properties of hypereutectoid steels under numerous operating parameters, the majority of these research are time-consuming. Thus, to reduce the experimental runs and associated costs, ML models have been adopted in prediction of wear rate of the materials. For instance, predicted wear rate for coated ferroalloy under various operating parameters by using various machine learning techniques. GPR algorithm was yielded higher accuracy in comparison with LR and SVM [10]. The wear rate of polyethylene is predicted under various

✉ Poornima Hulipalled
poornima.algur@gmail.com

¹ Department of Studies in Computer Science, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India

² Department of Mechanical Engineering, Rao Bahadur Y Mahabaleswarappa Engineering College, Ballari, Karnataka, India

³ Department of Studies in Mathematics, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India

Table 1 Chemical composition of the hyper eutectoid steel (wt%)

C	Si	Mn	P	S	Cr
0.92	0.256	0.348	0.002	0.002	1.100

Table 2 Wear test parameters

Descriptions	Value
Normal pressure (MPa)	0.1249, 0.3747, 0.6245, 0.8743
Sliding speed (m/s)	1,3,5,7
Sliding distance (m)	5,000, 7,500, 10,000
Motion	Reciprocating
Temperature (°C)	23.1
Relative humidity (RH)%	34

operating parameters. The K-nearest neighbour (KNN) technique was discovered to have the lowest prediction error when bunching into subgroups depending on radiation constraints [11]. Md Syam Hasan et al. applied ML algorithms to predict the mechanical properties of self-lubricating aluminium graphite (Al/Gr) composites [28].

In present work, dry sliding rate of wear of hypereutectoid steel under varying operating parameters was investigated by using a testing machine called pin-on-disc. To better understand the wear mechanism, a scanning electron microscope was utilized to investigate the worn surface and debris. ANOVA method was adopted to calculate the contribution performance of wear rate for each parameter. Machine learning algorithms, namely SVM, GPR and RF were used to predict the wear rate of hypereutectoid steel.

2 Methodology

2.1 Data Acquisition

For the current work, 0.92 wt% carbon (C) of hypereutectoid steel was used for wear test and chemical composition is shown in Table 1.

Dry sliding wear tests had been accomplished using friction and wear testing machine by varying operating parameters [12]. Table 2 lists the wear test parameters that were employed in this study.

2.2 Machine Learning Algorithm

To train the ML algorithms, input variables are namely sliding distance, normal pressure and sliding speed where wear rate is the output parameter considered from the experimental results [27, 28]. ML models predicts the wear rate based on the input parameters. Therefore, overall, 48 experiments

were conducted to collect the datasets of hypereutectoid steel. Generally, ML Algorithms divides the dataset arbitrarily into training set and testing set along with input and output parameters [13, 14]. To train the model's, training set was used to optimize the characteristics associated with inputs to outputs, while the testing set was used to evaluate the model's performance by comparing the model's output to the results of numerous machine learning approaches discussed in the sections below. The dataset was split into two parts: a training set (70%) and a testing set (30%).

2.2.1 Support Vector Machine (SVM)

SVM algorithm derive the relationship between output and input values with the regression method [22]. The best fit line is derived against maximum number of points over the hyperplane by the linear kernel function. [15].

2.2.2 Gaussian Process Regression (GPR)

The GPR technique is a flexible tool that can generate anticipated outcomes as well as a confidence interval, which can be used to assess prediction reliability [16, 17].

2.2.3 Random Forest (RF)

Multiple decision trees are used in RF, which is a supervised learning technique. The main premise of RF is to combine several decision trees to decide the final result, rather than depending on individual decision trees. It uses ensemble learning, which would be the technique of combining several classifiers to tackle a complicated problem and enhance the performance of the model [18, 19].

Finally, to create effective ML models, a k-fold cross-validation approach was utilized with a fold value of 10 with three assessment measures: mean absolute error (MAE), R -squared (R^2), and root mean square error (RMSE) [20, 27].

3 Results and Discussion

3.1 Wear Rate

Figure 1 depicts the wear performance of hypereutectoid steel as a combination of normal pressure, sliding speed and distance. It was perceived that wear performance of hypereutectoid steel changed significantly reliant on normal pressure, sliding speed and distance. As normal pressure increases, there is increase in wear rate for utmost all the sliding speeds. Wear rate is decreased from 1 to 3 m/s sliding speed, latter it is increased with raise in sliding speed for every sliding distance for all the normal pressures. The

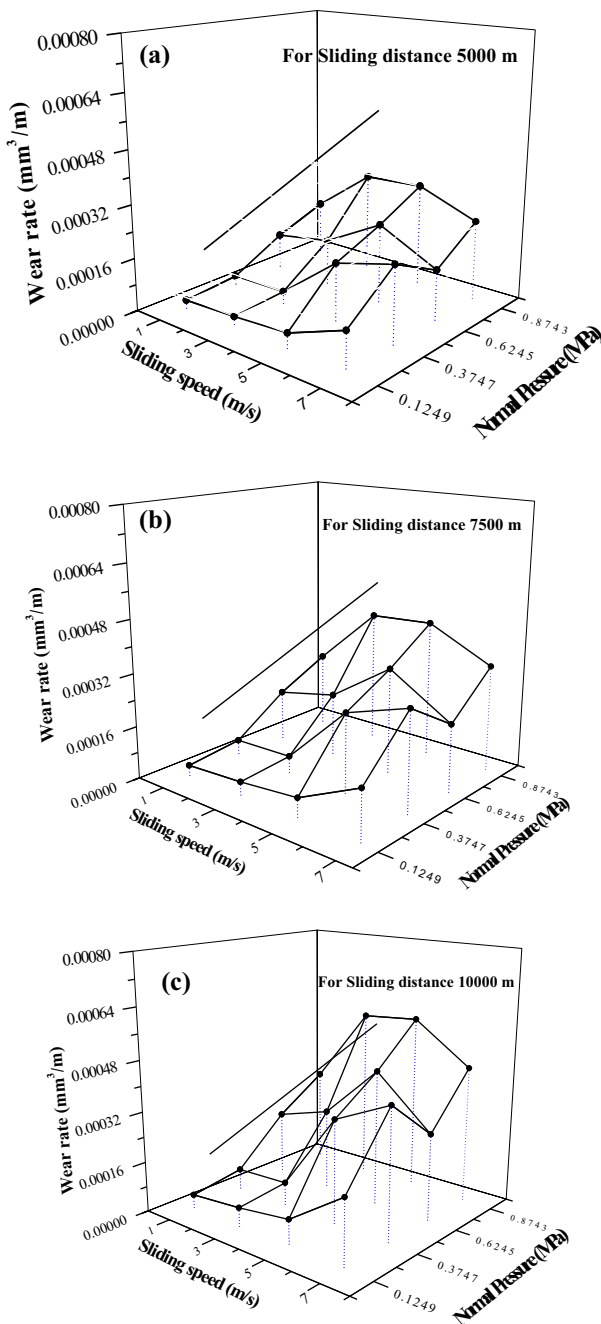


Fig. 1 Wear rate values of hypereutectoid steel at different sliding distances: **a** 5000 m, **b** 7500 m, **c** 10,000 m

normal pressure is increased at 3 m/s sliding speed, the results of wear rate is nearly constant for all the sliding distances. In the peak operational circumstances at 7 m/s with 0.8743 MPa, wear rate will be high when compared among the other operational conditions. Wear rate reduced from 1 to 3 m/s then the same increases from 3 to 7 m/s. At 3 m/s, the result of wear rate is utmost least and roughly steady with every normal pressure and for all the sliding speeds. Wear mechanism varies from mild wear to severe wear, when

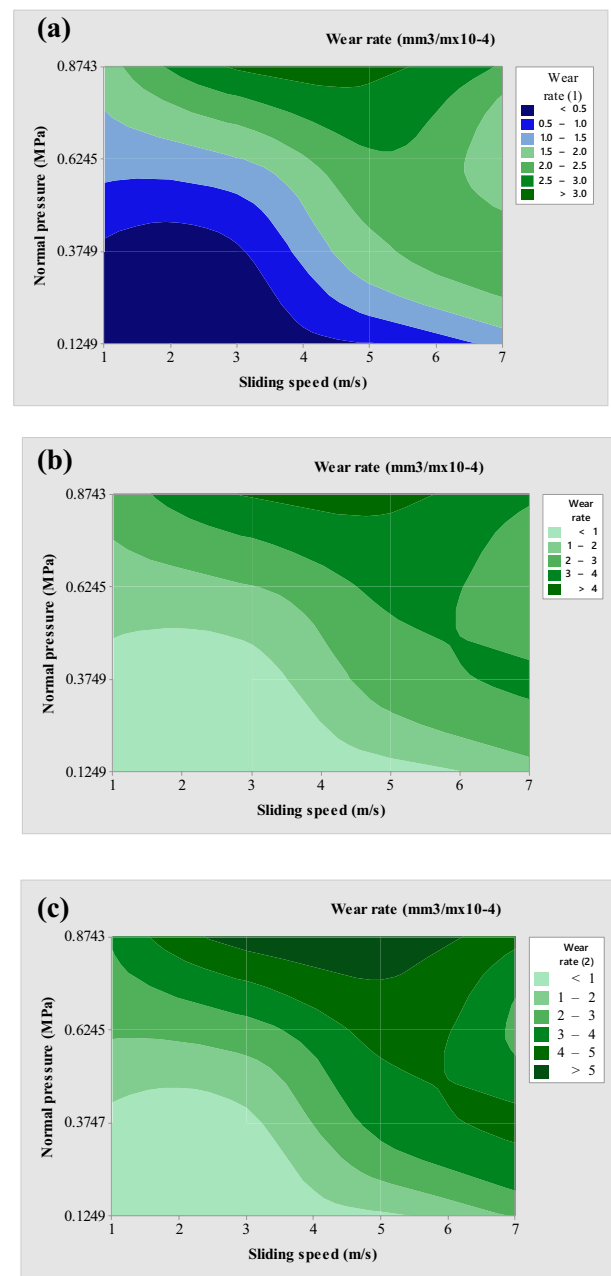


Fig. 2 Contour maps of wear rate of hypereutectoid steel with different sliding distances: **a** 5000 m, **b** 7500 m, **c** 10,000 m

sliding speed increased from 3 m/s. Hence, 3 m/s can be considered as the critical sliding speed for hypereutectoid steel.

Figure 2 illustrates the contour maps of wear rate of hypereutectoid steel based on various sliding distances. The graph is visualized to study the impacts of sliding distance, speed and normal pressure on wear rate. From Fig. 2a and b, it was noticed that higher value of wear rate at sliding speeds of 3–5 m/s and normal pressure of 0.8743 MPa for both 5000 m and 7500 m sliding distances. From Fig. 2c, it was observed higher value of wear rate at sliding speeds

of 2–6 m/s. Though, the least value of wear rate was observed at sliding speeds of 1–4 m/s and normal pressure of 0.1249 MPa.

3.2 Worn Surface and Debris

During the wearing under high normal pressure, the asperity gets weld with the sliding disc and the disc behaves like abrasive particle over the wearing surface and results in two body abrasive wear. The welded asperity will cut the specimen like single point cutting tool and the groove appears on the wearing surface shown in Fig. 3, and broken type wear debris appeared in Fig. 4. The corresponding XRD is shown in Fig. 5.

Thus, under high speed, wearing surface becomes soft enough and deforms easily which leads to formation of cracks. Hence, plate-like wear debris produce due to the nucleation, growth and finally joining each other [21].

3.3 Regression Technique and ANOVA

3.3.1 Regression Technique

In the current work, regression technique was adopted to determine the wear rate with input parameters like normal pressures, sliding speeds and sliding distances using Minitab software. The equation for regression technique is as follows:

$$W = -0.839 + 3.494P + 0.251S + 0.000012D \quad (1)$$

where W , P , S and D represent the wear rate, normal pressure, sliding speed and sliding distance, respectively.

3.3.2 ANOVA Results

The impacts of wear characteristics (input) on hypereutectoid steel wear rate (output) were investigated using analysis



Fig. 3 SEM micrograph of hypereutectoid steel, Speed-7 m/s, Normal Pr. 0.8749 MPa

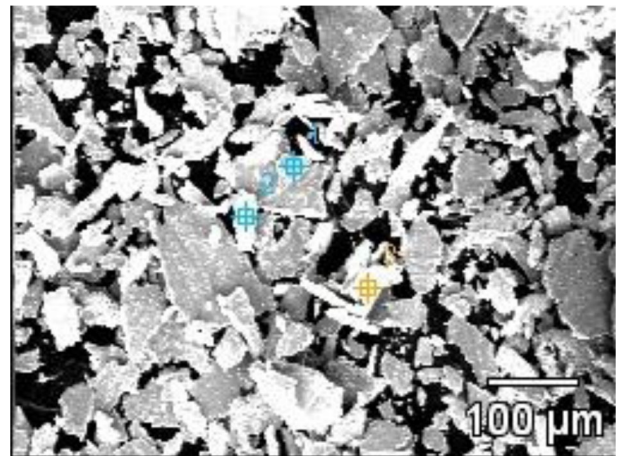


Fig. 4 Wear debris of hypereutectoid steel, Speed-7 m/s, Normal Pr. 0.8749 MPa

of variance (ANOVA). Obtained R^2 value is 93.35%. The percentage contributions of input features like sliding distance, sliding speed and normal pressure, are 1.08%, 28.52% and 63.74%, respectively. In comparison with all input features, the most significant parameter is normal pressure (Table 3).

3.4 Machine Learning Algorithms

The training data and test data are signified as red and blue scatter points, the predicted train data and predicted test data are signified as red and blue dashed lines which are plotted for all adopted ML algorithms.

Figure 6 displays the predicted wear rate of hypereutectoid steel, along with the measured values using SVM kernel function. The predicted R -squared values of training set and test set are ~ 0.895 and ~ 0.832 over the kernel SVM fit model. The MAE and RMSE values of training set of hypereutectoid steel are 0.307 and 0.454, respectively (Table 4). The specialty of SVM algorithm is that, it can apply on both linear and non-linear data to train the model

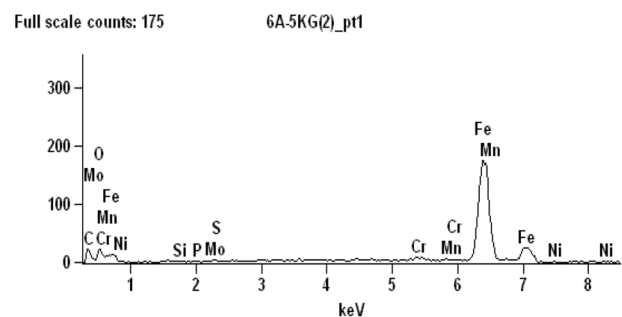
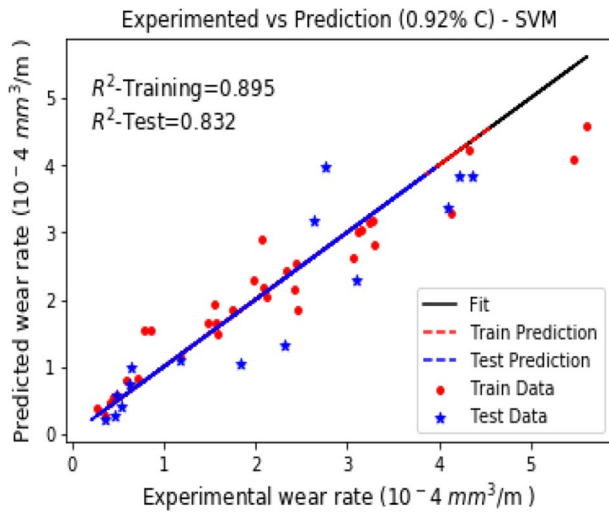
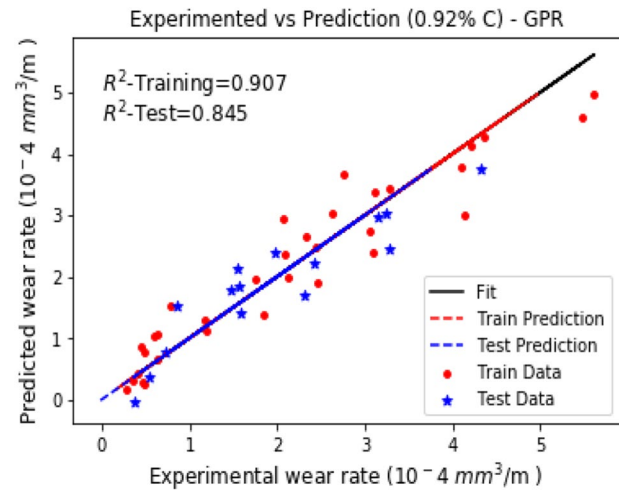


Fig. 5 XRD result of hypereutectoid steel, Speed-7 m/s, Normal Pr. 0.8749 MPa

Table 3 ANOVA results of hypereutectoid steel

Source	DF	Adj SS	Adj MS	F value	P value	Contribution (%)
Normal pressure	2	4.81446	2.40723	9.58	0.094	63.74
Sliding speed	2	2.15414	1.07707	4.29	0.189	28.52
Sliding distance	2	0.08175	0.04087	0.16	0.860	1.08
Error	2	0.50244	0.25122			
Total	8	7.55279				

$$R^2 = 93.35\%; R^2(\text{adj}) = 73.39\%$$

**Fig. 6** Experimental value against SVM predictions**Fig. 7** Experimental value against GPR predictions**Table 4** Performance of ML algorithms

	R^2 -training	R^2 -test	MAE	RMSE
SVM	0.895	0.832	0.307	0.454
GPR	0.907	0.845	0.362	0.466
RF	0.954	0.940	0.237	0.294

and the prediction of non-linear data is better since it can handle more than two predictor variables [22]. The erroneous usage of a simple kernel function accounts for the SVM algorithm's low prediction, and a portion of the dataset may be eliminated during the testing procedure [23].

Figure 7 illustrates the predicted wear rate of hypereutectoid steel against the experimental values using GPR algorithm. The predicted R -squared values of training set and testing set are ~ 0.907 , ~ 0.845 , respectively. These data evidently reveal that GPR algorithm outperforms the SVM algorithms in terms of experiment prediction accuracy. This may be due to GPR algorithm's nonparametric character [24]. This is in line with ML of ferro-alloys coating wear volume on steel surfaces, where the GPR approach outperforms the LR and SVM algorithms in terms of prediction accuracy [10]. Table 4 shows the quadratic rational GPR

algorithm outperforms the SVM algorithm. The MAE and RMSE of the training set of hypereutectoid steel are 0.362 and 0.466, respectively. During the prediction process of the GPR method, the dataset is regarded as a Gaussian distribution with mean and standard deviation. The standard deviation represents the prediction's confidence intervals, whereas the mean value is the prediction itself [25, 26]. Furthermore, the GPR approach is a nonparametric process that is unaffected by dataset structure constraints. In comparison to the SVM algorithm, the GPR approach may provide more trustworthy predictions of wear rate of hypereutectoid steel (Figs. 6, 7 and Table 4).

The predicted wear rate of hypereutectoid steel using RF algorithm are illustrated in Fig. 8. The predicted R -squared values of training set and test set are ~ 0.954 and ~ 0.940 , respectively. Hence this comparative study among the adopted ML algorithms states that, RF is the best suitable algorithm for this dataset. This is achieved by the use of random space method and bagging to prevent overfitting and the built-in automated feature selection capacity. The corresponding MAE and RMSE values of training and test dataset for hypereutectoid steel are 0.237 and 0.294, respectively. These lesser values of errors indicate the model is more efficient.

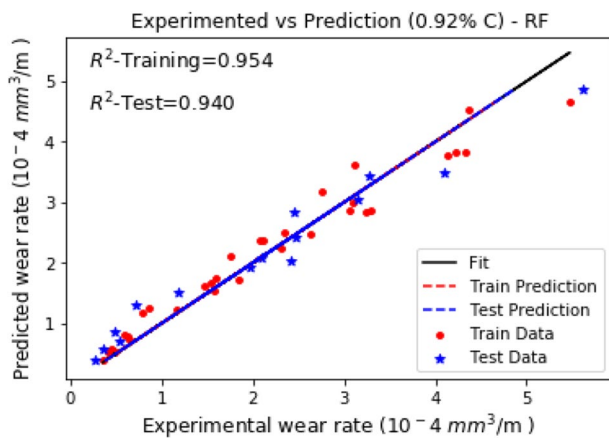


Fig. 8 Experimental value against RF predictions

4 Conclusions

1. With various sliding speeds and distances, the wear rate of hypereutectoid steel has enhanced in increased normal pressure.
2. Worn out surface shows that sub-surface crack breaks and makes the plate-like wear wreckage at higher speed and load.
3. Normal pressure is the crucial parameter in finding the wear rate of hypereutectoid steel, which is proved by the ANOVA technique.
4. RF is the most efficient ML algorithm to predict the wear performance of hypereutectoid steel with the highest accuracy of training/test (0.954/0.940).
5. The findings might enhance the innovative production of functional hypereutectoid steels with controlled wear properties.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40735-022-00668-y>.

Funding There were no research Grants for this work from any funding agencies.

Data Availability All data generated or analysed during this study are included in this published article (and its supplementary information files).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Dąbrowski R, Pacyna J, Krawczyk J (2007) New high hardness Mn-Cr-Mo-V tool steel. *Arch Metall Mater* 52(1):87–92

2. Hosmani SD, Kurhatti RV, Kabadi VK (2017) Wear behavior of spheroidized cementite in hyper eutectoid plain carbon steel. *Int Adv Res J Sci Eng Technol* 4(7):257–262
3. Wang Y, Lei T, Liu J (1999) Tribo-metallographic behavior of high carbon steels in dry sliding: II. Microstructure and wear. *Wear* 231(1):12–19
4. Liu KP, Dun XL, Lai JP, Liu HS (2011) Effect of modification on microstructure and properties of ultra-high carbon (1.9 wt% C) steel. *Mater Sci Eng A* 528(28):8263–8268
5. Sharanabasappa M, Kabadi VR, Algur V (2014) Some investigation on dry sliding wear behaviour of ultra high carbon steel. *Int J Mech Eng Res* 4(1):75–82
6. Sharanabasappa M, Kabadi VR, Algur V (2015) The effect of pearlite, cementite and martensite phases on volumetric wear rate of hypereutectoid steel under dry sliding conditions. *Int J Metall Mater Sci Eng* 5(1):31–38
7. Sharanabasappa M, Kabadi VR, Shetty PB, Algur V (2015) Dry sliding wear behaviour of hypereutectoid steel under the influence of microstructures, sliding speeds and normal pressures. *Int J Mech Eng Robot Res* 4(2):1–12
8. Marui E, Hasegawa N, Endo H, Tanaka K, Hattori T (1997) Research on the wear characteristics of hypereutectoid steel. *Wear* 205:186–199
9. Gao X, Jiang Z, Wei D, Kosasih B, Jiao S, Chen D (2014) Dry sliding wear behaviour of full pearlite obtained by cladding low carbon steel to hypoeutectoid steel. *Adv Mater Res* 1017:147–153
10. Altay O, Gurgenc T, Ulas M, Ozel C (2020) Prediction of wear loss quantities of ferro-alloy coating using different machine learning algorithms. *Friction* 8(1):107–114
11. Borjali A, Monson K, Raeymaekers B (2019) Predicting the polyethylene wear rate in pin-on-disc experiments in the context of prosthetic hip implants: deriving a data-driven model using machine learning. *Tribol Int* 133:101–110
12. VeerabhadrapprAlgur VR, Kabadi GSM, Chavan VR (2017) Effect of Mn content on tribological wear behaviour of ZA-27 alloy. *Mater Today Proc* 4:10927–10934
13. Specht DF (1991) A general regression neural network. *IEEE Trans Neural Netw* 2(6):568–576
14. Scherbela M, Hormann L, Jeindl A, Obersteiner V (2018) Charting the energy landscape of metal/organic interfaces via machine learning. *Phys Rev Mater* 2(4):043803
15. Gunn S (1997) Support Vector Machines for Classification and Regression. ISIS Technical Report ISIS-1-98. Image Speech & Intelligent Systems Research Group. University of Southampton, Southampton, UK
16. Wang B, Chen T (2015) Gaussian process regression with multiple response variables. *Chemom Intell Lab Syst* 142:159–165
17. Aye S, Heyns P (2017) An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. *Mech Syst Signal Process* 84:485–498
18. Breiman L (2001) Random forests. *Mach Learn* 45:5–32
19. James G, Witten D, Hastie T, Tibshirani R (2019) An introduction to statistical learning. Springer Science and Business Media LLC, Cham
20. Kong D, Chen Y, Li N (2018) Gaussian process regression for tool wear prediction. *Mech Syst Signal Process* 104:556–574
21. Rai D, Pathak JP (2004) Influence of sliding velocity on wear behaviour of different microstructures of Ni-Cr-Mo-V steel. *Indian J Eng Mater Sci* 11:113–120
22. Singh Y, Kaur A, Malhotra R (2009) Comparative analysis of regression and machine learning methods for predicting fault proneness models. *Int J Comp Appl Technol* 35(2/3/4):183–193
23. Smola AJ, Scholkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14:199–222

24. Yun Y, Kim HC, Shin SY, Lee J, Deshpande AD, Kim C (2014) Statistical method for prediction of gait kinematics with Gaussian process regression. *J Biomech* 47(1):86–192
25. Kodamana H, Huang B, Ranjan R, Zhao Y, Tan R, Sammaknejad N (2018) Approaches to robust process identification: a review and tutorial of probabilistic methods. *J Process Control* 66:68–83
26. Daemi A, Kodamana H, Huang B (2019) Gaussian process modelling with Gaussian mixture likelihood. *J Process Control* 81:209–220
27. Algur V, Hulipalled P, Lokesha V, Nagaral M, Auradi V (2022) Machine learning algorithms to predict wear behavior of modified ZA-27 alloy under varying operating parameters. *J Bio-and Tribo-Corros* 8(1):1–10
28. Hasan MS, Kordijazi A, Rohatgi PK, Nosonovsky M (2022) Triboinformatics approach for friction and wear prediction of Al-graphite composites using machine learning methods. *J Tribol* 144(1):011701

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.