Surface Topography: Metrology and Properties

PAPER

Taguchi approach and decision tree algorithm for prediction of wear rate in zinc oxide-filled AA7075 matrix composites

To cite this article: S V Alagarsamy *et al* 2021 *Surf. Topogr.: Metrol. Prop.* **9** 035005

IOP Publishing Surf. Topogr.: Metrol. Prop. 9 (2021) 035005 https://doi.org/10.1088/2051-672X/ac0f34

Surface Topography: Metrology and Properties

RECEIVED
14 December 2020

REVISED
14 June 2021

ACCEPTED FOR PUBLICATION
28 June 2021

PUBLISHED
7 July 2021

PAPER

Taguchi approach and decision tree algorithm for prediction of wear rate in zinc oxide-filled AA7075 matrix composites

S V Alagarsamy[1], R Balasundaram[2], Ravichandran M[2], V Mohanavel[3], Alagar Karthick[4] and S Sathiya Devi[5]

[1] Department of Mechanical Engineering, Mahath Amma Institute of Engineering and Technology, Pudukkottai-622 101, Tamil Nadu, India
[2] Department of Mechanical Engineering, K Ramakrishnan College of Engineering, Tiruchirappalli-621 112, Tamil Nadu, India
[3] Centre for Materials Engineering and Regenerative Medicine, Bharath Institute of Higher Education and Research, Chennai-600 073, Tamil Nadu, India
[4] Department of Electrical and Electronics Engineering, KPR Institute of Engineering and Technology, Coimbatore-641407, Tamil Nadu, India
[5] Department of Computer Science Engineering, University College of Engineering, BIT Campus, Tiruchirappalli −620 024, Tamil Nadu, India

E-mail: smravichandran@hotmail.com

Keywords: composites, aluminium alloy, zinc oxide, wear rate, optimization, taguchi method, decision tree algorithm

Abstract

This work aims to apply data-driven approach [decision tree (DT) algorithm] to analyse the wear rate (WR) of ZnO-filled AA7075 composites. The results of model-based analysis was compared with Taguchi analysis. Stir casting was used to produce the composite samples. Characterization studies were conducted to analyse the composition and morphology. The scanning electron microscopy results indicated the even dispersion of ZnO in the AA7075. The energy-dispersive x-ray spectroscopy pattern ensures the presence of matrix elements and the inclusion of reinforcement particles into the proposed composites. To minimize the number of experimentation, L27 Orthogonal array is used for finding WR. The 'DuCom' Pin-on-Disc apparatus were used to prepare WR data for the set of the proposed composites. Taguchi technique reveals the optimum level factors for obtaining the minimum 'WR' is reinforcement content of 10 wt.%, applied load (P) at 10 N, sliding velocity (V) at 1 m s^{-1} and sliding distance (D) of 1000 m. The experiments results from DT algorithm, and analysis of variance and signal-to-noise ratio analysis from Taguchi-based approach confirmed that reinforcement is the primary element for affecting wear of the composites. The reason for applying DT algorithm is that, the low-level knowledge could be converted into high-level knowledge (If-then-else rules), which can be effortlessly explicable by semiskilled personnel.

1. Introduction

The manufacturing industries always looking for adopting new technologies that could improve production process and quality. The implementation of artificial intelligence and machine learning will help the manufacturers to improve the product quality and optimize the process. Due to the continuous expansion of information technology, the interdisciplinary field is more attractive to the researchers in applying their domain knowledge. Data mining (DM), image processing and fractal dimensions from computer science domain are the indications of interdisciplinary approach to apply other domain knowledge to manufacturing systems [1]. With the information age, data sets have become increasingly rich, but the knowledge contained in the data sets have not been fully utilized or explored. DM is the amalgamation of machine learning, databases and statistics. The application of 'DM' in the production industry was begun in 1990 and progressively receiving attention to research community [2]. Piatesky-Shapiro *et al* [3] reported the DM is a very useful and emerging area for industry. Shahbaz and Harding [4] proved that many areas have been benefited from DM algorithms in manufacturing firms and there are still several fields that may benefit more. Yong-Hong Kuo and Andrew Kusiak [5] showed that data had been utilized in production research in significant ways and they reported that data-based approach in the production research field

Table 1. Elements of AA7075.

Element	Si	Fe	Cu	Mn	Mg	Cr	Zn	Ti	Al
Weight (%)	0.4	0.5	1.9	0.3	2.7	0.25	7.1	0.20	Bal

shifted the research focus from an analytical model to data-driven model. The analyses of manufacturing data have been popular in the areas of data-driven methodologies. Harding *et al* [6] displayed in their review paper on application of DM in the production process, operations, decision support systems and product quality improvements and they reported that the identified key parameters could be used to enhance the quality of the products.

Gardener and Bieker [7] reported that by using decision tree (DT) algorithms, considerable amount of time was saved for the manufacturing of semi-conductors. The task of DM are: Clustering, Classification, Association analysis and Regression analysis, etc [8]. Among these, classification models predict the categorical class label. For example, a classification model from bank data set can categorize bank loan applications as either safe or risky. The derived model can be represented in the form of If-Then-Else rules. These rules are simple and easy to understand and can be applied to new data sets if the accuracy of the classification rules is acceptable. DT algorithm has the potential to use other domains due to its simplicity and robustness.

In recent years, Taguchi method of experimentation is more attractive to researchers to limit the number of experiments and producing high-quality results. It provides a complete and coherent approach to analyse the process parameters [9]. Hence in this work, Taguchi-based L27 Orthogonal Array is used to do experimentation and to examine the wear behaviour of the proposed materials. Taguchi method of experimentation was used to verify the results obtained by DM.

Aluminium alloys are used in various sectors owing to their exceptional strength and extraordinary heat conductivity [10]. However, these alloys display underprivileged tribological properties leading to problems in confrontational conditions. To augment the wear resistance, aluminium and its alloys are reinforced with hard or self-lubricating particles such as SiC, TiC, B_4C, MgO, ZnO, TiO_2, ZrB_2, Al_2O_3 and graphite etc [11–13]. The aluminium matrix composites (AMCs) have been developed by various processing techniques. Among these routes, stir casting is the proved auspicious techniques for producing AMCs because of their easiness, suppleness, inexpensive and appropriate to large quantity fabrication [14, 15]. Poppy Puspitasari *et al* [16] studied the mechanical properties in Al-Si alloy with the presence of ZnO particles and reported that by adding ZnO, the strength and hardness of the composite have been improved. Baradeswaran *et al* [17] analysed the wear behaviour of AA6061 and AA7075 incorporated with B_4C and graphite particles. From the experiments, the authors concluded that the AA7075 composite has better wear resistance as compared to AA6061.

The exhaustive available literature revealed that, AA7075 is a potential alloy in many industries for making the high performance components such as aircraft fittings and gears. Hence, in this work, AA7075 is used. Encouraged by the advantage of machine learning for manufacturing industries, this work proposes, the combination of DT and Taguchi approach for analysing the wear rate (WR) of the composites of AA7075-ZnO composites. To study the wear behaviour, the attributes such as (i) weight percentage of ZnO, (ii) P, (ii) V and (iv) D are considered.

2. Experimental details

AA7075 alloy was chosen as matrix materials, and the ZnO was used as filler material. The elements of AA7075 alloy is provided in table 1. The AA7075 ingot was melted in an electric furnace at a temperature of 850 °C. The measured quantities of ZnO particles were added into the molten AA7075 to form AA7075-ZnO composites with altering weight fractions (0, 5 and 10 wt.%) of ZnO. The melt was stirred randomly for the duration of 10 min. Then the molten metal was poured into the metallic die and allowed to solidify, and afterwards the cast was removed. A similar procedure was repeated to produce the composites with various wt.% of ZnO. Figure 1 displays the computer-controlled stir casting setup used for the synthesizing of unreinforced and reinforced composites along with a complete work plan. The microstructure observation of produced cast samples was done by scanning electron microscopy (SEM) and energy-dispersive x-ray spectroscopy (EDS) analyses.

To appraise the sliding wear behaviour of produced composites, the Pin-on-Disc (POD) test was used. The sample size of 10 mm dia × 30 mm length of the pin was cut from the cast composites by using wire EDM. The tests were performed as per ASTM G99 standard using POD wear measuring instruments (DUCOM, Bangalore) in dry sliding circumstances. Figure 2 displays the experimental setup of POD apparatus along with sample size. EN31 hardened steel was used to make the disc material with a hardness of 60 HRC. Initially, the composite samples were prepared with acetone and the disc was polished to get a neat surface. The pins were tested as per the experimental layout, and the mass of each pin was

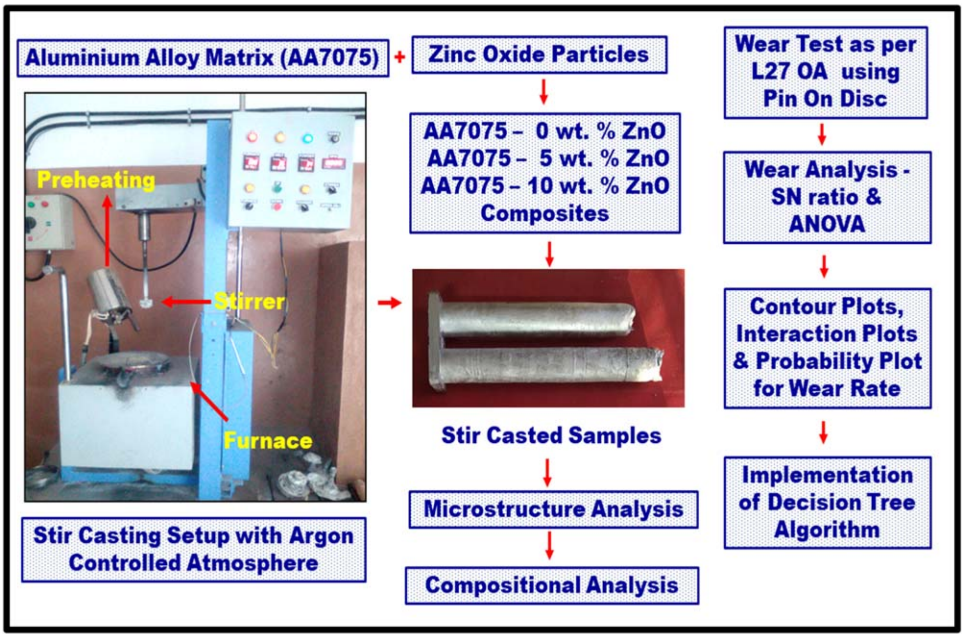

Figure 1. Stir casting furnace used for casting and work plan.

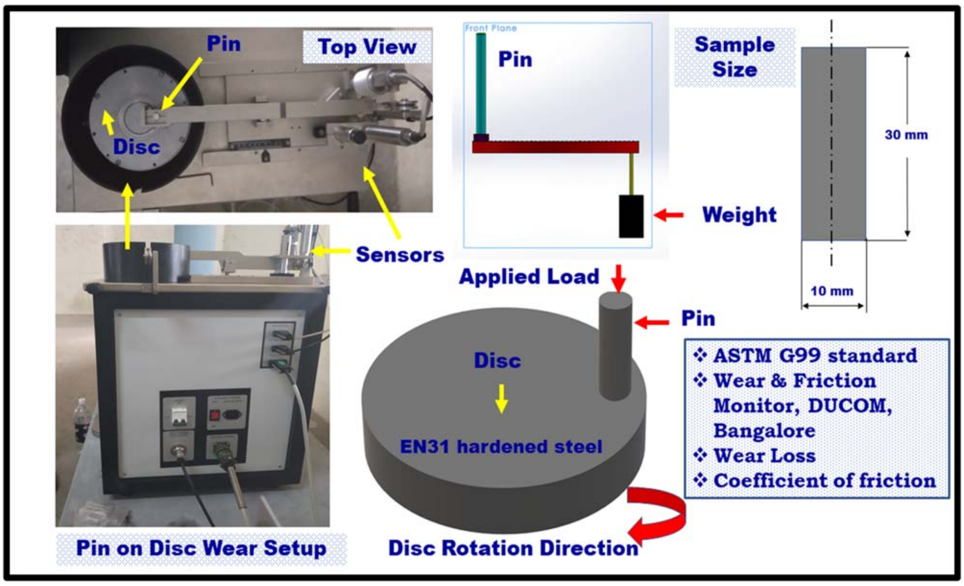

Figure 2. POD experimental set up.

acquired with an accuracy of 0.0001 g. The 'WR' was computed by using equation (1) [18].

$$\text{Wear rate}(\text{mm}^3/\text{m}) = \frac{m}{\rho D} \quad (1)$$

where m is the loss of materials in pins (g), ρ is the density of composite specimen (g mm^{-3}) and D is the 'D' (m).

The experiments were steered to assess the effects of control factors on 'WR' of AA7075-ZnO composites. From the extensively available literature studies, it is observed that there are many factors that individually influence the WR of AMCs [19, 20]. In the present work, four control factors with three levels were chosen and is provided in table 2. For this experimental work, L27 orthogonal array was applied [21]. Taguchi approach has been used to predict the optimum level of factors for the 'WR'. The experimental orthogonal array design, calculated output responses and their signal-to-noise (SN) ratios are presented in table 3.

In this work, an effort is made to propose a set of rules in the form of *If-then else* from L_{27} orthogonal array data set. These rules are easy to understand by the semi-skilled labour and they can classify or predict the future data without depending on technical expertise. The different algorithms used for classifications are (i) Naive Bayes classifier, (ii) support vector

Table 2. Wear control parameters and levels.

Control parameters	Level 1	2	3
Reinforcement (wt.%)	0	5	10
P (N)	10	20	30
V (m/s)	1	2	3
D (m)	1000	1500	2000

Table 3. L27 orthogonal array design, output response and SN ratios.

Sl. no	Reinforcement (wt.%)	P (N)	V (m/s)	D (m)	WR (mm^3/m)	SN ratio (dB)
1	0	10	1	1000	0.00294	50.6331
2	0	10	2	1500	0.00343	49.2941
3	0	10	3	2000	0.00386	48.2683
4	0	20	1	1500	0.00318	49.9515
5	0	20	2	2000	0.00367	48.7067
6	0	20	3	1000	0.00441	47.1112
7	0	30	1	2000	0.00514	45.7807
8	0	30	2	1000	0.00477	46.4296
9	0	30	3	1500	0.00465	46.6509
10	5	10	1	1500	0.00266	51.5024
11	5	10	2	2000	0.00309	50.2008
12	5	10	3	1000	0.00254	51.9033
13	5	20	1	2000	0.00327	49.7090
14	5	20	2	1000	0.00315	50.0338
15	5	20	3	1500	0.00290	50.7520
16	5	30	1	1000	0.00347	49.1934
17	5	30	2	1500	0.00363	48.8019
18	5	30	3	2000	0.00381	48.3815
19	10	10	1	2000	0.00179	54.9429
20	10	10	2	1000	0.00143	56.8933
21	10	10	3	1500	0.00167	55.5457
22	10	20	1	1000	0.00215	53.3512
23	10	20	2	1500	0.00239	52.4320
24	10	20	3	2000	0.00269	51.4050
25	10	30	1	1500	0.00335	49.4991
26	10	30	2	2000	0.00359	48.8981
27	10	30	3	1000	0.00287	50.8424

machines, (iii) decision trees and (iv) artificial neural network. The Naive Bayes, support vector machines and neural network algorithms generate black box patterns and, therefore, their interpretability becomes low. Whereas the DT algorithm produces high interpretability. Here the dataset has been converted into tree structure as per information gain. It is nothing but the amount of disorder existing in the data set. Quinlan [22] created an algorithm for DT named iterative dichotomizer 3(ID3). The later version of the same is enhanced version C4.5. It hunts through the attributes of occurrences in the dataset and selects the utmost outstanding splitting attributes based on information gain. DT algorithm provides output as in the tree arrangement. The high-level knowledge is obtained from the low-level knowledge. L27 OA was used to apply the C4.5 algorithm for the newly developed composite materials.

3. Results and discussion

3.1. Microstructure analysis of produced composites

The microstructure of the developed composites was investigated by SEM coupled with EDS. Figures 3(a)–(c) illustrates the SEM image of AA7075 and ZnO-reinforced composites. Figure 3(a) display the SEM image of AA7075 alloy and it is witnessed the nonappearance of ZnO particles. From figures 3(b) and (c), it can be witnessed that ZnO particles are equivalently distributed over the aluminium matrix. It also reveals that there is no agglomeration and voids are exited in the proposed composites because of properly chosen casting parameter. The results of EDS analysis of unreinforced AA7075, and ZnO-reinforced composites is presented in figures 4(a)–(c), and it is evident that the appearance of peaks confirms the existence of matrix elements and reinforcements. From the figures, the high peaks specify

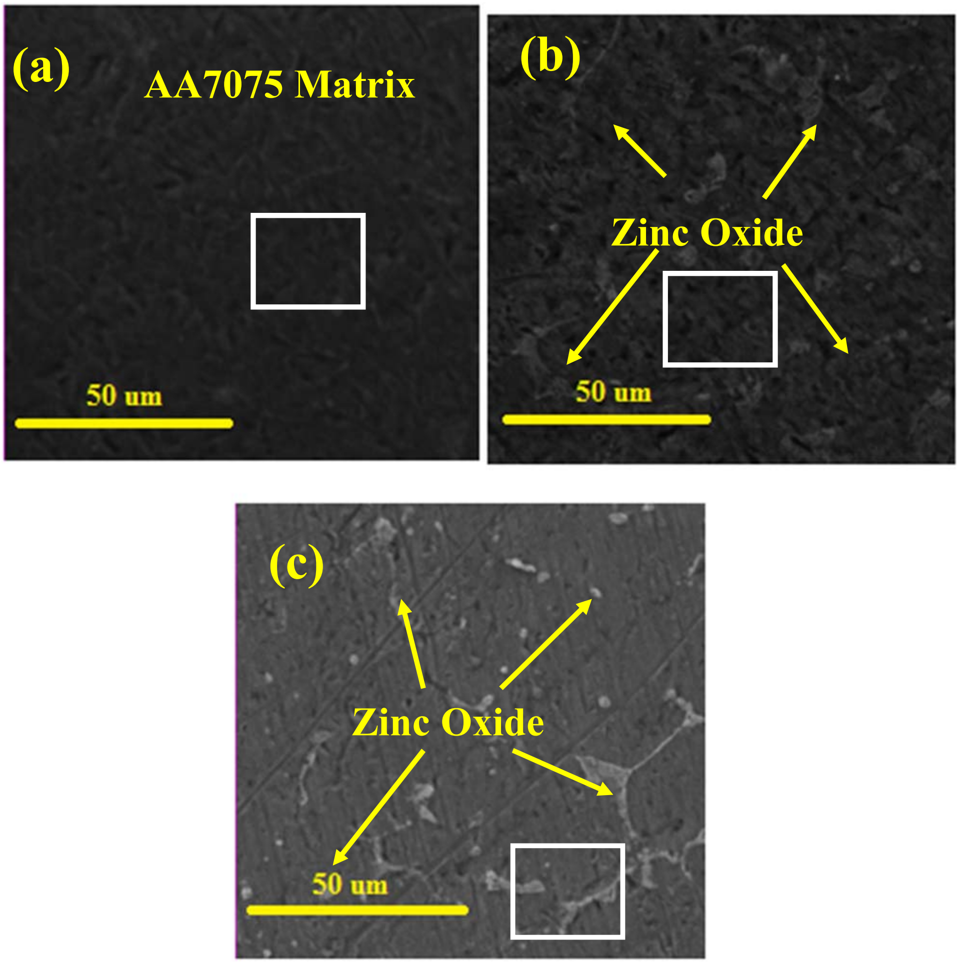

Figure 3. SEM image of (a) AA7075-0 wt.% ZnO, (b) AA7075-5wt.% ZnO and (c) AA7075-10 wt.% ZnO composites.

Al and the minor peaks show the existence of ZnO particles and other elements of the matrix such as Cu, Zn, Mg and O. It is understood that rise in Zn peaks with a rise in wt.% of ZnO in the matrix.

3.2. SN ratio and mean analysis

To find the optimum level of wear factors for the response, SN ratio was performed during the Taguchi analysis. SN ratio displays the sensitivity of the output factors to be analysed [23]. A smaller the better SN ratio was considered for this investigation since the minimum 'WR' for produced composites to be achieved. The SN ratio and means values for the 'WR' are shown in tables 4 and 5. The effect of the control factors on 'WR' is found in the tables by the specified rank. This rank is assigned by means of delta value which is computed between the higher and lower value of the concerned column of parameter. Based on the rank obtained (tables 4 and 5), the reinforcement wt.% was indicated as the most predominant factor on 'WR' subsequently by 'P', whereas 'V' was revealed as the insignificant factor.

Figures 5 and 6 display the main effect plot of SN ratios and mean values of 'WR' with reverence to various level of control factors. These plots observed the optimum level of parameters and also describes the impact of each level of factors on the 'WR' of produced composites. According to figure 5, exposed the optimum level of parameters to obtain the least 'WR' of composites at 10 wt.% of ZnO, 'P' at 10 N, 'V' of 1 m s^{-1} and 'D' at 1000 m.

In figure 5, it has been revealed that, the maximum SN ratio of reinforcement content is the primary dominant factor for affecting the 'WR' subsequently by 'P' and 'D'. In figure 6, it is clearly noticed that the 'WR' of unreinforced composite is higher than the ZnO-reinforced composites. Because the deficient hardness property of matrix alloy produces more 'WR'. But the inclusion of ZnO particles AA7075 matrix composites reduce the 'WR'. It also has been understood that the 'WR' of produced composites declines with an increase in wt.% of ZnO particles into the AA7075. The reason is that, the incorporation of ZnO particles enhance the hardness property and thus improved the wear resistance of fabricated composite samples [24]. The variation in 'WR' of the tested composites with 'P' is shown in figure 6. It clearly stated that, at low P condition, the less 'WR' is obtained due to the formation of thin oxide layer prevents direct contact of composite pin surface with the counter disc

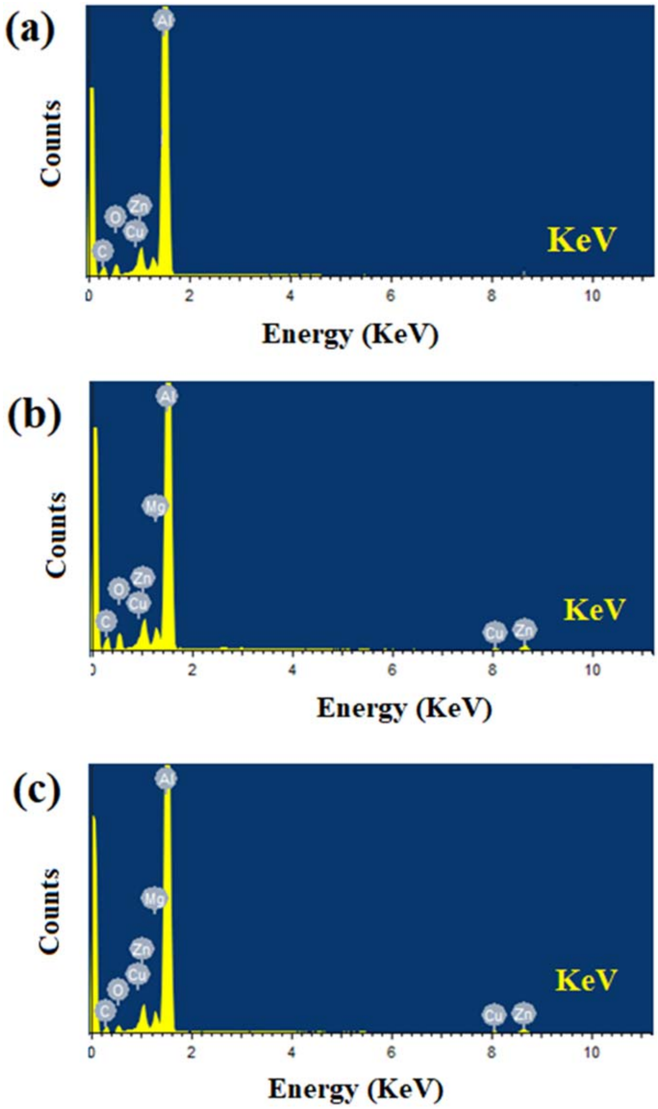

Figure 4. EDS pattern of (a) AA7075-0 wt. % ZnO, (b) AA7075-5wt.% ZnO and (c) AA7075-10 wt.% ZnO composites.

Table 4. SN ratio of 'WR'.

Level	Reinforcement (wt. %)	P (N)	V (m/s)	D (m)
1	47.77	52.32	50.49	50.72
2	50.05	50.38	50.19	50.49
3	52.65	48.28	50.10	49.59
Delta	4.87	4.04	0.40	1.13
Rank	1	2	4	3

Table 5. Means of 'WR'.

Level	Reinforcement (wt.%)	P (N)	V (m/s)	D (m)
1	0.004139	0.002559	0.003126	0.003099
2	0.003169	0.003090	0.003239	0.003096
3	0.002437	0.003920	0.003267	0.003434
Delta	0.001702	0.001361	0.000140	0.000339
Rank	1	2	4	3

surface. With an increase in 'P' and 'V', the 'WR' of produced composites steadily increases and it is evident from Archard's law. While considering the 'D', the higher wear loss is obtained at the maximum 'D' conditions.

3.3. Contour plots

Figures 7(a)–(f) illustrate the combined effect of wear control parameters on the 'WR' of tested composite samples. Figures 7(a)–(c) depict the influence of reinforcement content on 'WR' over the other parameters such as (a) P, (b) V and (c) D. It has been noticed that the higher amount of reinforcements (10 wt.% ZnO particles) produces low 'WR' due to high hardness property. The high level of 'P' (30 N) and low level of 'V' (1 m s^{-1}) and D (1000 m) give more 'WR' because an increase in 'P' creates more contact between the pin surface and disc. Thus, the high temperature is developed, which results in increased

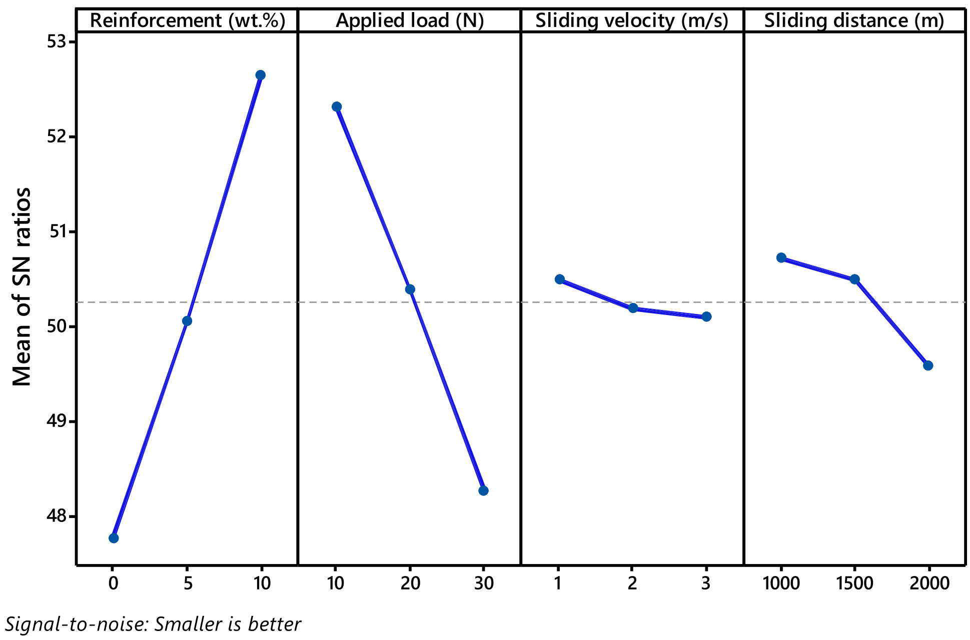

Figure 5. Main effect graph of SN ratio for 'WR'.

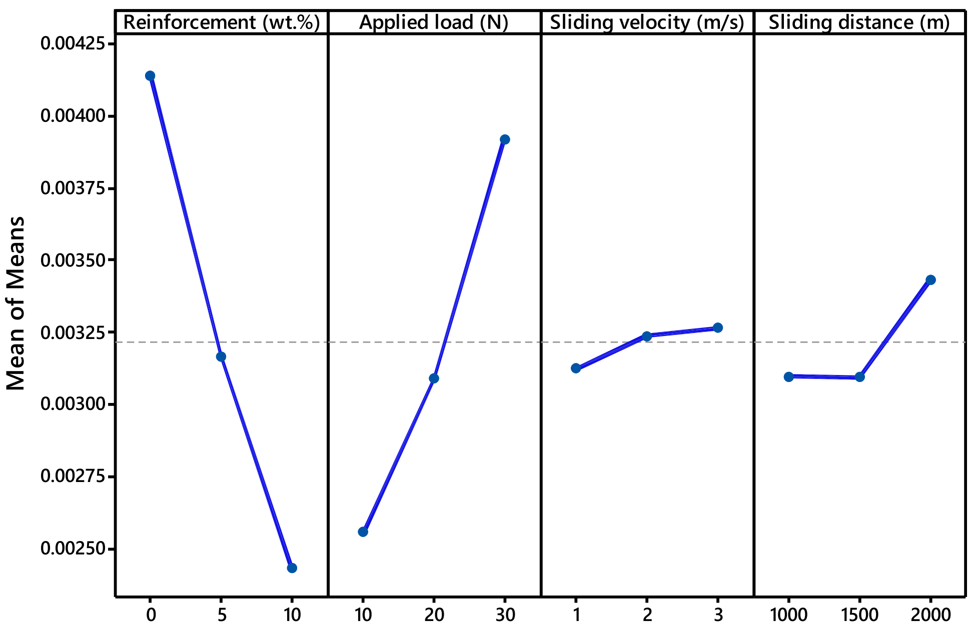

Figure 6. Main effect graph of mean for 'WR'.

'WR' of composite samples. Figures 7(d) and (e) demonstrates the influence of 'P' on 'WR' with esteem to 'V' and 'D'. The higher P (30 N) with V from 1 m s^{-1} to 3 m s^{-1} produces the maximum 'WR' 0.0042 mm^3/m. The 'WR' of composites slightly increases when 'P' and 'D' are at a moderate level. While considering the 'D' with 'V' shown in figure 7(f), it was revealed that the moderate 'WR' of 0.0030 mm^3/m is obtained at an initial level of 'V' (1 m s^{-1}) and 'D' (1000 m).

3.4. ANOVA analysis

Analysis of variance (ANOVA) is a mathematical tool for assessing the output of the examined experimental results on a group of independent variables. The impact of control parameters such as reinforcement, 'P', 'V' and 'D' were identified for the 'WR' of produced composites through ANOVA analysis. During this analysis, the *P*-value of the factor is less than 0.05, which is said to be the most dominant parameter on response [25]. ANOVA result and contribution of

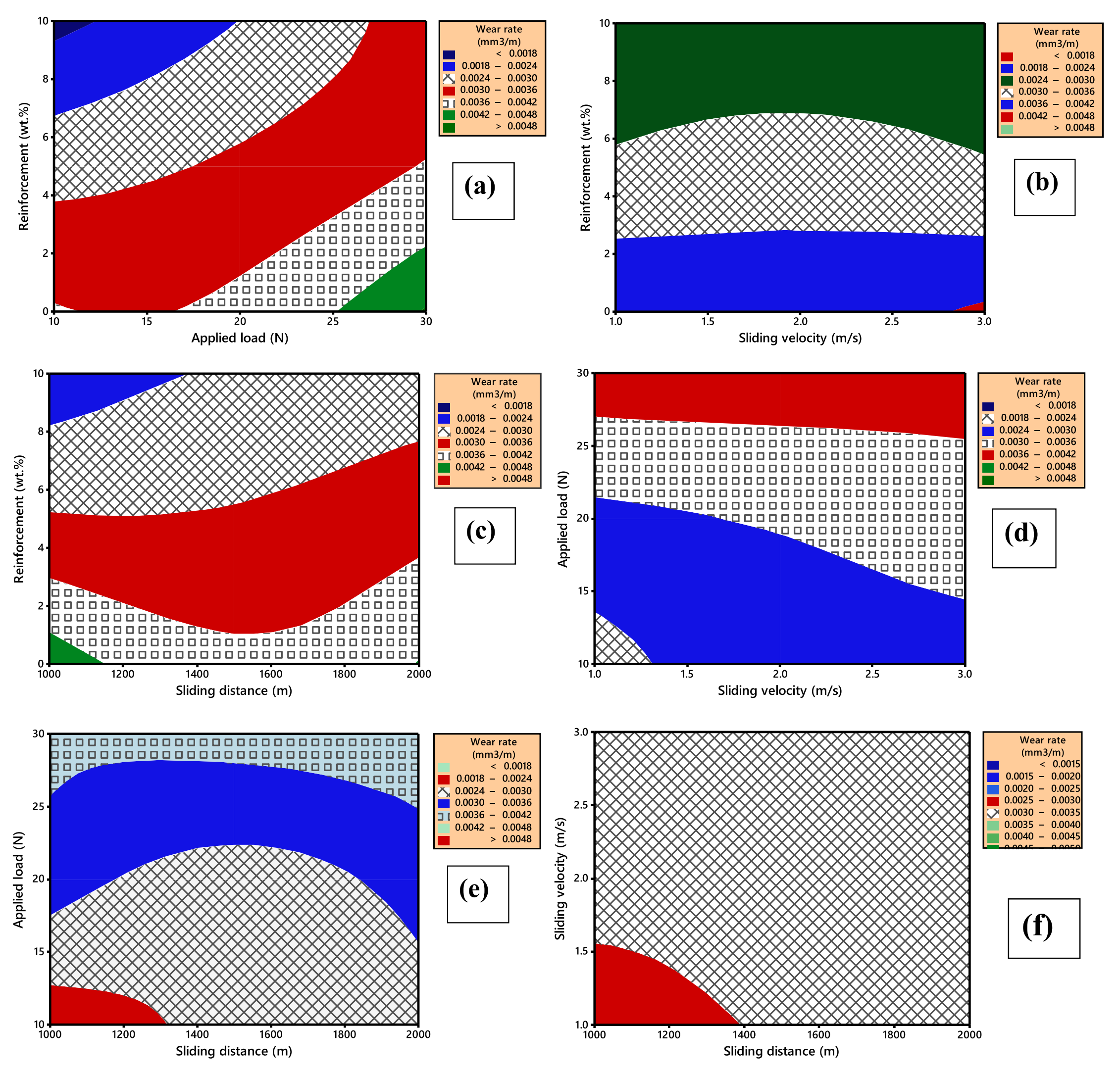

Figure 7. Contour plot for 'WR' (a) reinforcement vs. P, (b) reinforcement versus V, (c) reinforcement versus D, (d) P versus V, (e) P versus D and (f) V versus D.

Table 6. ANOVA for 'WR'.

Source	DF	Adj SS	Adj MS	*F*-value	*P*-value
Reinforcement (wt.%)	2	0.000011	0.000005	51.76	0.000
P (N)	2	0.000007	0.000004	34.29	0.000
V (m/s)	2	0.000000	0.000000	0.31	0.735
D (m)	2	0.000001	0.000000	3.02	0.075
Error	17	0.000002	0.000000		
Total	25	0.000022			

$S = 0.0003224$; $R^2 = 91.87\%$; R^2 (adj) = 88.04%; R^2 (pred) = 80.99%

parameters for 'WR' is provided in table 6 and figure 8. From the table, it was evident that the reinforcement ($P = 0.000$) and 'P' ($P = 0.000$) are the most impact parameters on the 'WR' of fabricated AA7075-ZnO composites with the contribution of 50% and 31.81%. The 'V' ($P = 0.735$) was the least impact parameter. Similar findings were previously stated by Alagarsamy *et al* [26] during the wear process of AA7075-TiO_2 composites. The obtained response values are scattered normally within the limits as shown in figure 9.

Figure 10 displays the interaction of the control parameters on 'WR'. This plot provides the information about the intersection impact of each parameter over the others. The parallel lines indicate no interaction between the parameters, whereas the non-parallel lines indicate the occurrence of interaction between the selected parameters. Figure 10 shows the interaction of reinforcement with other parameters such as 'P', 'V' and 'D' on the 'WR' of fabricated composites. It has been understood that the creation of parallel lines

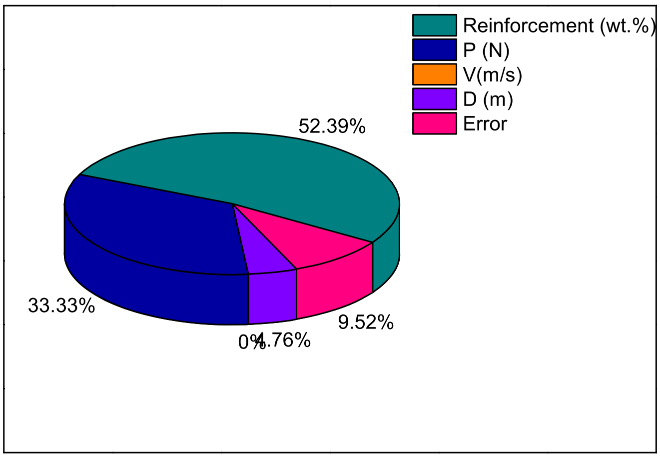

Figure 8. Contribution of parameters on 'WR'.

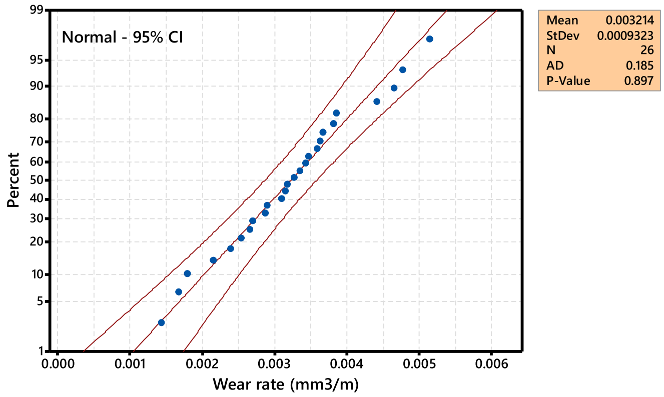

Figure 9. Probability plot for 'WR'.

ensures the absence of significant interaction of all the parameters with reinforcement content.

3.5. Prediction of 'WR' using DT algorithm

DT algorithms are termed as trained learning algorithms. It can be implemented to solve the classification problems. In the DT approach, the decision rules are framed from the training data, which is used to predict the target parameters. In the present work, reinforcement wt.%, 'P', 'V' and 'D' are used as input parameters. The response is 'WR' which is called as output parameters. The input parameters are called as predictor parameters.

The values of 'WR' as per L27 OA is provided in table 7 along with the input parameters. The response is categorized as low and high in L27 OA. The values from 0.00143 to 0.00315 are categorized as 'Low' and values above 0.00315 to 0.00514 are categorized as 'High'. Entropy is a measure of the quantity of impurity available in the data. The entropy equation is provided in equation (2) as per Claude E. Shannon.

$$\text{Entropy } (P_1, P_2 \ldots P_n) = -P_1 \log_2 P_1 - P_1 \log_2 P_{21} - \cdots\cdots - P_n \log_2 P_n \qquad (2)$$

The total number of the low attribute is 13 and the high attribute is 14. Therefore, the information gain for the class attribute is, as per equation (2):

$$\text{Info}(13, 14) = -\frac{13}{24} X \log_2 \frac{13}{27} X \log_2 \frac{14}{17} = 0.998$$

The info (13, 14) shows the overall entropy function.

3.5.1. Information gain for input variables (reinforcement wt.%, P, V and D)

Table 8 shows the details of reinforcement wt.% and information gain values for the reinforcement wt.%. The calculated number attributes (Low and High) for reinforcement wt.%, P, V and D are provided in table 8. Figure 11(a) shows the tree structure for reinforcement wt.%. The tree structure is constructed from the split value which is converted from the values in table 8. From table 8, it is understood that, the attribute '0' provides maximum gain when compared with the other reinforcement wt.%.

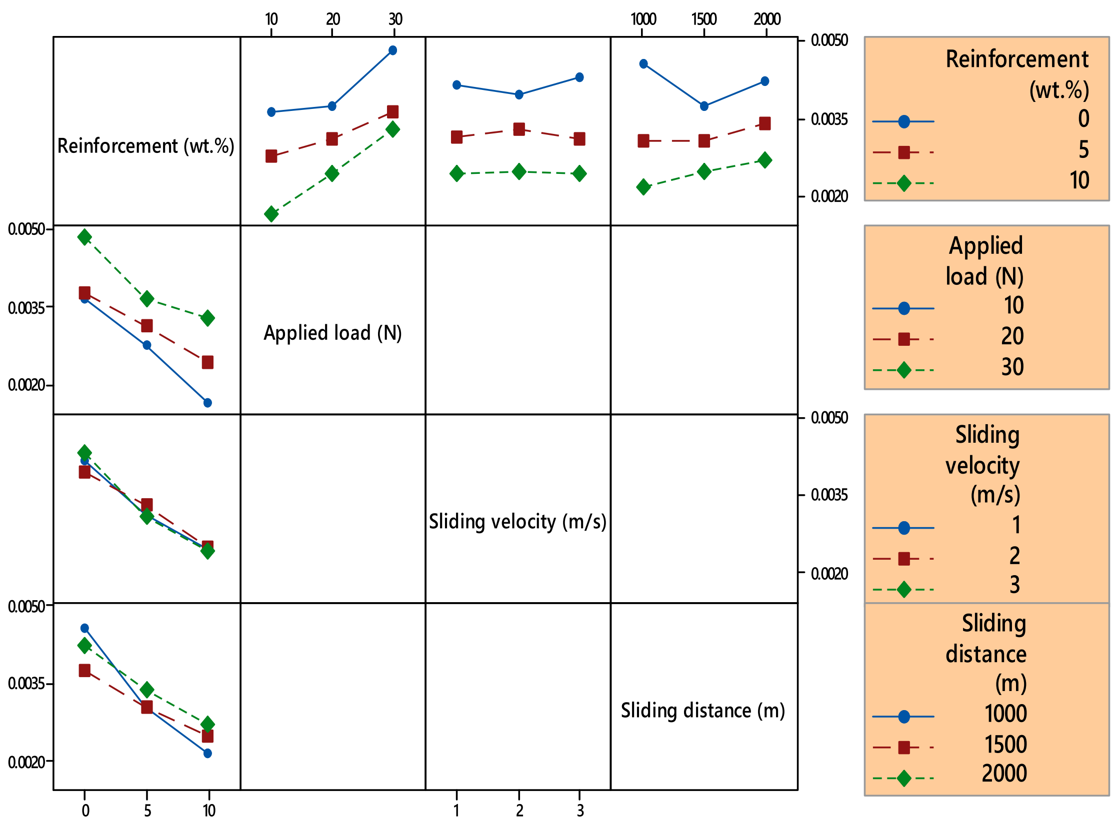

Figure 10. Interaction plot for 'WR'.

Table 7. Input and output variable for DT analysis.

	Input variables (predictor attributes)				Output variable (class attribute)	
Instance no.	Reinforcement (wt.%)	P (N)	V (m/s)	D (m)	'WR' (mm^3/m)	Categorical attribute
1	0	10	1	1000	0.00294	Low
2	0	10	2	1500	0.00343	High
3	0	10	3	2000	0.00386	High
4	0	20	1	1500	0.00318	High
5	0	20	2	2000	0.00367	High
6	0	20	3	1000	0.00441	High
7	0	30	1	2000	0.00514	High
8	0	30	2	1000	0.00477	High
9	0	30	3	1500	0.00465	High
10	5	10	1	1500	0.00266	Low
11	5	10	2	2000	0.00309	Low
12	5	10	3	1000	0.00254	Low
13	5	20	1	2000	0.00327	High
14	5	20	2	1000	0.00315	Low
15	5	20	3	1500	0.00290	Low
16	5	30	1	1000	0.00347	High
17	5	30	2	1500	0.00363	High
18	5	30	3	2000	0.00381	High
19	10	10	1	2000	0.00179	Low
20	10	10	2	1000	0.00143	Low
21	10	10	3	1500	0.00167	Low
22	10	20	1	1000	0.00215	Low
23	10	20	2	1500	0.00239	Low
24	10	20	3	2000	0.00269	Low
25	10	30	1	1500	0.00335	High
26	10	30	2	2000	0.00359	High
27	10	30	3	1000	0.00287	Low

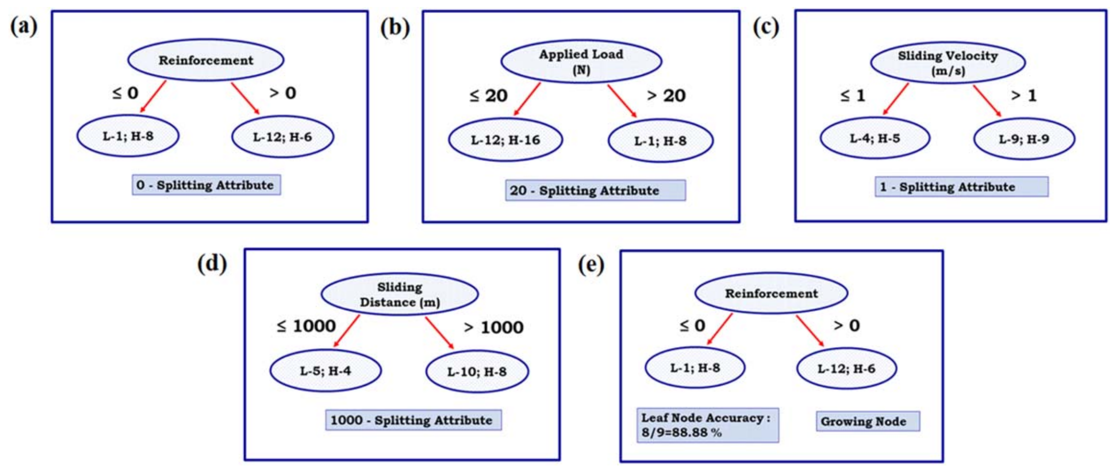

Figure 11. (a)–(e) Tree structure for reinforcement, P, V, D and reinforcement.

Table 8. Class attributes for reinforcement wt.%, P, V and D.

Factors	Levels	No. of low (L)	No. of high (H)
Reinforcement wt.%	0	1	8
	5	5	4
	10	7	2
P	10	7	2
	20	5	4
	30	1	8
V	1	4	5
	2	4	5
	3	5	4
D	1000	5	4
	1500	4	5
	2000	6	3

$$\text{Info}(1,\,8) = -\frac{1}{9}\log_2\frac{1}{9} - \frac{8}{9}\log_2\frac{8}{9} = 0.5026$$

$$\text{Info}(12,\,6) = -\frac{12}{18}\log_2\frac{12}{18} - \frac{6}{18}\log_2\frac{6}{18} = 0.9182$$

$$\text{The combination info }((1,\,8),\,(12,\,6)) = \frac{9}{27}X0.5026 + \frac{18}{27}X0.9182 = 0.7796$$

The gain of attribute reinforcement)
= Overall gain class attribute (0.998)
− Combined info of reinforcement (0.7796) = 0.218

The maximum gain value is observed for the attribute 20 when compared with other 'P'. The same info (12,6) and info (1,8) are observed as that of reinforcement calculation. Therefore, the overall gain of 'P' is **0.218.** The tree structure for 'P' is shown in figure 11(b). From table 8, it is observed that attribute '1' provides maximum gain while compared with the remaining 'V'. Tree structure for 'V' is shown in figure 11(c).

$$\text{Info}(4,\,5) = -\frac{4}{9}\log_2\frac{4}{9} - \frac{5}{9}\log_2\frac{5}{9} = 0.991$$

Table 9. The details of information of gain.

Sl. no.	Predictor attributes	Information gain
1	Reinforcement	0.218
2	Applied load	0.218
3	Sliding velocity	0.001
4	Sliding distance	0.007

$$\text{Info}(9,\,9) = -\frac{9}{18}\log_2\frac{9}{18}X2 = 1$$

$$\text{The combined info }((4,\,5),\,(9,\,9)) = \frac{9}{27}X0.991 - \frac{18}{27}X1 = 0.997$$

The gain of attribute (sliding velocity)
= 0.998 − 0.997 = 0.001

From table 8, attribute '1000' provides a higher gain, while comparing to the other 'D'. The tree structure for 'D' is shown in figure 11(d).

$$\text{Info}(5,\,4) = 0.991$$

$$\text{Info}(10,\,8) = -\frac{10}{18}\log_2\frac{10}{18} - \frac{8}{18}\log_2\frac{8}{18} = 0.4711 + 0.5199 = 0.991$$

$$\text{The combined info }((5,\,4),\,(10,\,8)) = \frac{9}{27}X0.991 - \frac{18}{27}X0.991 = 0.991$$

The gain of attribute (sliding distance)
= 0.998 − 0.991 = 0.007

Table 9 shows the values of information gain of all predictor attributes (input variables) along with information gain. From the table, it is inferred that the attributes reinforcement and applied load are significant parameters to decide on 'WR'. Whereas the sliding velocity and sliding distance least factor or even it cannot be considered on 'WR'. Since the information gain

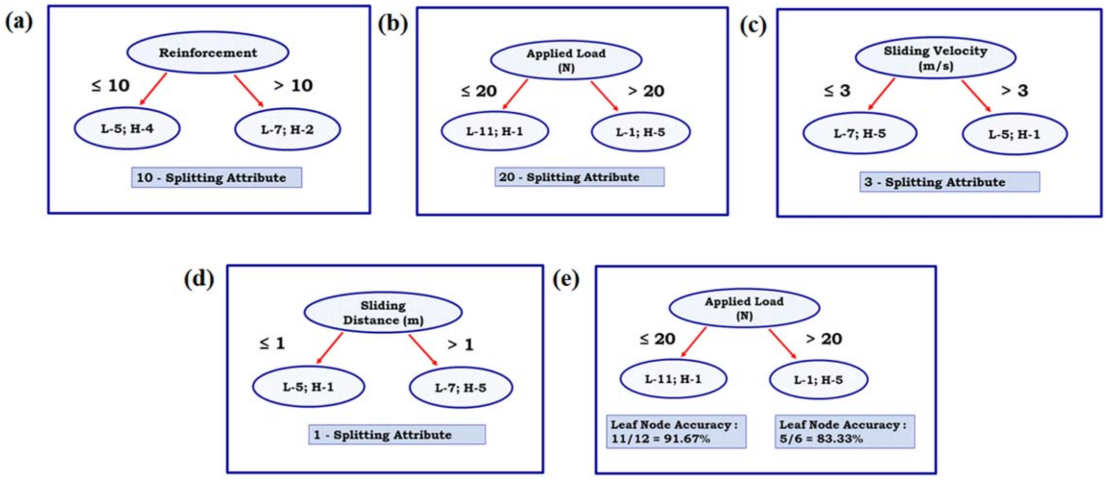

Figure 12. Tree structure for (a) Reinforcement wt.%, (b) P, (c) V, (d) D and (e) information gain for P.

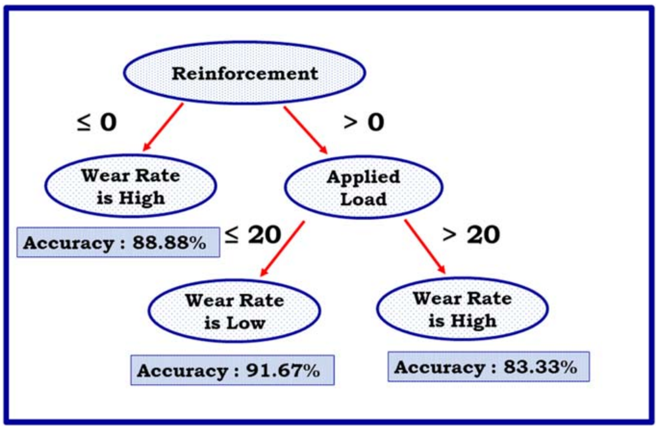

Figure 13. Decision tree for L27 Orthogonal array.

Table 10. Remaining instances after the root node.

	Input variables (predictor attributes)				Output variable (class attribute)	
Instance no.	Reinforcement (wt.%)	P (N)	V (m/s)	D (m)	'WR' (mm^3/m)	Categoricalattribute
10	5	10	1	1500	0.00266	Low
11	5	10	2	2000	0.00309	Low
12	5	10	3	1000	0.00254	Low
13	5	20	1	2000	0.00327	High
14	5	20	2	1000	0.00315	Low
15	5	20	3	1500	0.00290	Low
16	5	30	1	1000	0.00347	High
17	5	30	2	1500	0.00363	High
18	5	30	3	2000	0.00381	High
19	10	10	1	2000	0.00179	Low
20	10	10	2	1000	0.00143	Low
21	10	10	3	1500	0.00167	Low
22	10	20	1	1000	0.00215	Low
23	10	20	2	1500	0.00239	Low
24	10	20	3	2000	0.00269	Low
25	10	30	1	1500	0.00335	High
26	10	30	2	2000	0.00359	High
27	10	30	3	1000	0.00287	Low

of reinforcement and applied are equal any one can be taken as root node. Let take reinforcement as a root node. The tree structure is shown in figure 11(e).

The rule derived from the root node is:

If reinforcement is $\leqslant 0$, the wear rate is 'High' – Rule (1)

The left side of the tree cannot be split further, hence it is called 'leaf node'. Whereas in the right side of tree, it can be possible to split further, hence it is called as 'growing node'. The rule derived from the leaf node is 'if reinforcement is $\leqslant$ 0, the 'WR' is High'. The above rule correctly classifies the instances no.1, 2, 3, 4, 5, 6, 7, 8 and 9 and incorrectly classifies the instance no.1 in table 10. Therefore, the accuracy of the tree is 88.88%.

3.5.2. Construction of sub node

The instances 1 to 9 are removed and the remaining instances are shown in table 10.

For remaining instances, once again information gain of both input and output variables are calculated. The tree will be processed recursively. The information gain of the class attribute is calculated as follows: Total number of Low –12; Total number of High –6; Therefore,

$$\text{Info}(12,\ 6) = -\frac{12}{18}\log_2\frac{12}{18} - \frac{6}{18}\log_2\frac{6}{18} = 0.917$$

Information gain of input variables (reinforcement wt. %, 'P', 'V' and 'D')

Table 10 shows the number of class attributes (low and high) for corresponding reinforcement wt.%, P, V and D. The attribute '10' provides maximum gain. Figure 12(a) displays the tree structure for reinforcement wt.%.

$$\text{Info}(5,\ 4) = -\frac{5}{9}\log_2\frac{5}{9} - \frac{4}{9}\log_2\frac{4}{9} = 0.991$$

$$\text{Info}(7,\ 2) = -\frac{7}{9}\log_2\frac{7}{9} - \frac{2}{9}\log_2\frac{2}{9} = 0.763$$

$$\text{The combined info } ((5,\ 4),\ (7,\ 2)) = 0.877$$

$$\text{The gain of attribute (reinforcement)} = 0.917 - 0.877 = 0.04$$

In table 10, attribute '20' provides maximum gain while comparing other 'P'. The tree structure for 'P' is shown in figure 12(b).

$$\text{Info}(11,\ 1) = 0.413$$

$$\text{Info}(1,\ 5) = 0.6498$$

$$\text{The combined info } ((11,\ 1),\ (1,\ 5)) = 0.4918$$

$$\text{The gain of attribute (applied load)} = 0.917 - 0.491 = 0.4251$$

From table 10, attribute '1' provides maximum gain while comparing to the other 'V'. The tree structure for 'V' is shown in figure 12(c).

$$\text{Info } (7,\ 5) = 0.9798$$

$$\text{Info } (5,\ 1) = 0.6498$$

$$\text{The combined info } ((7,\ 5),\ (5,\ 1)) = 0.8698$$

$$\text{The gain of attribute (sliding velocity)} = 0.917 - 0.8698 = 0.0472$$

From table 10, attribute '1' provides maximum gain while comparing to the other 'D'. The tree structure for 'D' is shown in figure 12(d).

The info (5, 1) and info (7, 5) are the same as that of the 'V' calculation. Therefore, the overall gain of 'D' is also the same as that of 'V'.

The gain of the attribute ('D') = 0.0472.

From the information gain, it is understood that 'P' possess maximum gain. The tree structure is shown in figure 12(e).

The rule derived from the above tree is:

If applied load is $\leqslant$ 20 wear rate is 'Low'

else wear rate is 'High' – Rule (2)

The above rule correctly classifies all the instances nos except instances no. 13 and 27. Since both sides of the tree are not possible to grow further, tree formation is stopped. The final decision tree is shown in figure 13.

The decision rules are:

If reinforcement is $\leqslant$ 0, the wear rate is 'High' – Rule (1)

If reinforcement is $>$ 0 and Applied load is $\leqslant$ 20 wear rate is 'Low'

else wear rate is 'High' – Rule (2)

The above rules correctly classify all the instances nos except instances nos 1, 13 and 27.

$$\text{Theaccuracyoftreeis}\frac{24}{27} = 88.88\%$$

From the DM approach, it is understood that both the reinforcement wt.% and 'P' are the significant factors (both having equal information gain as shown in table 9) for the prediction of 'WR' for this composition. The application of the DM algorithm showed that the low-level knowledge of data is converted in the form of high-level knowledge (if-then-else rules). From this high-level knowledge (if-then rules), even the semi-skilled workers can also classify the data and predict the responses and in this case, it is 'WR' of the newly developed composite material. When we look at the results of previous researchers' who reported for the optimization of wear behaviour of the composite, the present results are well agreed with previous findings. Baradeswaran *et al* found the optimum conditions for minimum 'WR' for hybrid composites. They noticed that 'P' is the predominant factor for influencing the 'WR' of the composites [27]. Similarly, Baskaran *et al* used the Taguchi technique to find the significant factor for the 'WR' and reported that, load and sliding velocity is the highly noteworthy factors on the 'WR' of the TiC *in situ* AMCs [28]. Furthermore, numerous statistical methods such as the Taguchi technique, Grey relational analysis, response surface methodology, Genetic algorithm, particle swarm optimization and teaching-learning-based optimization techniques are being used to optimize the wear parameters of aluminium-based composites. However, they reported that, 'P' is the significant parameter to affect the 'WR'. These results are well agreed with the present result arrived from 'DT' approach. Hence we conclude that 'DT' can be effectively used for the analysis of wear parameters of aluminium-based composites.

4. Conclusions

- AA7075 matrix filled with varying weight percentages (0, 5 and 10 wt.%) of ZnO particles were effectively fabricated by using the stir casting route.
- The SEM micrograph reveals the microstructure of the produced composites and it ensured the even spreading of ZnO reinforcement particles over the AA7075 matrix. It is also observed that there is no agglomeration and voids are exited in the proposed composites.
- The EDS pattern approves the occurrence of matrix elements such as Cu, Zn, Mg and O and ZnO particles into the developed composites. It is also understood that rise in Zn peaks with the rise in weight percentage of ZnO in the matrix.
- The DT algorithm from machine learning and Taguchi approach was applied to analyse the effect of wear control parameters on the wear rate of produced composites.
- The main effect plot shows that the reinforcement content of 10 wt.%, 'P' of 10 N, 'V' of 1 m s^{-1} and 'D' at 1000 m produce the less wear rate of fabricated composites.
- From S/N ratio, it is observed that reinforcement is ranked first for causing wear of the proposed composite followed with 'P', 'V' and 'D'.
- The contribution plot of ANOVA showed that reinforcement contributed 50%, 'P'−31.81%, 'D'−4.5% for causing wear and there is no effect by 'V'.
- DT algorithm on the wear data set of the proposed composite showed that the information of gain of reinforcement and 'P' are the same and it is very minimum in the case of 'V'. From the decision rules derived, it is understood that the primary element for causing wear is reinforcement and followed with 'P'.
- The advantage of machine learning for the prediction of wear is that, the 'WR' of the composite can be identified from 'If-then-else' rules. From these rules, even semi-skilled labour can identify 'WR' fora particular parameter. The results from S/N ratio and ANOVA confirmed the results of the machine learning algorithm. These rules are sufficient to classify the future composite without any effort where it is not possible for other methods.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

S V Alagarsamy https://orcid.org/0000-0002-6564-0346
Ravichandran M https://orcid.org/0000-0003-3886-283X
V Mohanavel https://orcid.org/0000-0001-9277-4190
Alagar Karthick https://orcid.org/0000-0002-0670-5138

References

[1] Djellouli A, Benyelloul K, Aourag H, Bekhechi S, Adjadj A, Bouhadda Y and ElKedim O 2018 A datamining approach to classify, select and predict the formation enthalpy for intermetallic compound hydrides *Int. J. Hydrogen Energy* **43** 19111–20
[2] Piatetsky Shapiro G 1999 The data mining industry coming of age *IEEE Intell. Syst.* **146** 32–4
[3] Shahbaz M and Harding J A 2003 An integrated data mining model for manufacturing enterprises *Int. Conf. on Manufacturing Research (Glasgow)* 539–45
[4] Kuo Y-H and Kusiak A 2019 From data to big data in production research: the past and future trends *Int. J. Prod. Res.* **57** 4828–53
[5] Harding J A, Shahbaz M, Srinivas and Kusiak A 2018 A data mining in manufacturing: a review *J. Manuf. Sci. Eng.* **128** 969–76
[6] Gardner M and Bieker J 2000 Data mining solves tough semi conductor problems *KDD '00 Proc. of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 376–83
[7] Han J, Kamber M and Pei J 2011 *Data Mining: Concepts and Techniques* 3rd ed. (United States: The Morgan Kaufmann Series in Data Management Systems Morgan Kaufmann Publishers)
[8] Alagarsamy S V, Ravichandran M, Meignanamoorthy M, Sakthivelu S and Dineshkumar S 2020 Prediction of surface roughness and tool wear in milling process on brass (C26130) alloy by Taguchi technique *Mater. Today Proc.* **21** 189–93
[9] Venkat Prasat R, Subramanian N, Radhika N and Anandavel B 2011 Dry sliding wear and friction studies on AlSi10Mg–fly ash–graphite hybrid metal matrix composites using Taguchi method *Tribology* **5** 72–81
[10] Baradeswaran A and Perumal E 2014 A study on mechanical and wear properties of Al 7075/Al_2O_3/graphite hybrid composites *Composite Part B.* **56** 464–71

[11] Chelladurai S J S and Arthanari R 2017 Dry sliding wear characterization of squeeze cast LM13/FeCu composite using response surface methodology *China Foundry* **14** 525–33
[12] Veeresh Kumar G B, Rao C S P and Selvaraj N 2011 Mechanical and dry sliding wear behavior of Al7075 alloy-reinforced with SiC particles *J. Compos. Mater.* **46** 1–9
[13] Ravichandran M and kumar S D 2016 Experimental investigations of Al–TiO_2–Gr hybrid composites fabricated by stir casting *Mater. Test.* **58** 211–7
[14] Alagarsamy S V and Ravichandran M 2019 Synthesis, microstructure and properties of TiO_2 reinforced AA7075 matrix composites via stir casting route *Mater. Res. Express* **6** 1–14
[15] Puspitasari P, Soepriyanto O R, Sasongko M I N, Dika J W and Andoko 2018 Mechanical and physical properties of aluminium-silicon (Al–Si) casting alloys reinforced by Zinc Oxide (ZnO) *MATEC Web of Conferences* **204** 05003
[16] Baradeswaran A, Vettivel S C, Elaya P and Selvakumar A N 2014 Experimental investigation on mechanical behaviour, modelling and optimization of wear parameters of B_4C and graphite reinforced aluminium hybrid composites *Mater. Des.* **63** 620–32
[17] Radhika N and Raghu R 2015 Dry sliding wear behaviour of Aluminium Al–Si_2Cu/TiB_2 metal matrix composite using response surface methodology *Tribology. Letter* **59** 1–9
[18] Baskaran S, Anandakrishnan V and Duraiselvam M 2014 Investigations on dry sliding wear behavior of *in situ* casted AA7075–TiC metal matrix composites by using taguchi technique *Mater. Des.* **60** 184–92
[19] Thirumalai Kumaran S, Uthayakumar M and Aravindan S 2014 Analysis of dry sliding friction and wear behaviour of AA6351–SiC–B_4C composites using grey relational analysis *Tribology* **8** 187–93
[20] Dharmalingam S, Subramanian R, Somasundara Vinoth K and Anandavel B 2011 Optimization of tribological properties in aluminium hybrid metal matrix composites using gray-taguchi method *J. Mate. Engg. Perform.* **20** 1457–66
[21] Quinlan J R 1986 Induction of decision trees *Mach. Learn.* **1** 181–106
[22] Alagarsamy S V and Ravichandran M 2020 Parametric studies on dry sliding wear behaviour of A1-7075 alloy matrix composite using S/N ratio and ANOVA analysis *Mater. Res. Express* **7** 1–17
[23] Fatile B O, Adewuyi B O and Owoyemi H T 2017 Synthesis and characterization of ZA-27 alloy matrix compositesreinforced with zinc oxide nanoparticles *Engineering Science and Technology, an International Journal* **20** 1147–54
[24] Marichamy S, Saravanan M, Ravichandran M and Veerappan G 2016 Parametric optimization of EDM process on α–β brass using Taguchi approach. Russian *Journal of Non-Ferrous Metals.* **57** 586–98
[25] Meignanamoorthy M and Ravichandran M 2020 Microstructure analysis and optimization of parameters for improved properties in high strength AA8079 (Al–Fe–Cu–Si–Zn) *The Phys. Met. Metall.* **121** 374–81
[26] Alagarsamy S V and Ravichandran M 2019 Investigations on tribological behaviour of AA7075-TiO_2 composites under dry sliding conditions *Industrial Lubrication and Tribology* **71** 1064–71
[27] Baradeswaran A, Vettivel S C, Elaya Perumal A, Selvakumar A N and Franklin Issac R 2014 Experimental investigation on mechanical behaviour, modelling and optimization of wear parameters of B_4C and graphite reinforced aluminium hybrid composites *Mater. Design* **63** 620–30
[28] Baskaran S, Anandakrishnan V and Duraiselvam M 2014 Investigations on dry sliding wear behavior of *in situ* casted AA7075–TiC metal matrix composites by using Taguchi technique *Mater Design* **6** 0184–92