Materials Today Communications 45 (2025) 112420

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

Machine learning-based prediction of bushing dimensions, surface roughness and induced temperature during friction drilling of pre-heated A356 aluminum alloy

Mahmoud Khedr [a,b,*], Ahmed Abdalkareem [a], Amr Monier [a,c], Rasha Afify [a], Tamer S. Mahmoud [a], Antti Järvenpää [d]

[a] Mechanical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo 11629, Egypt
[b] Future Manufacturing Technologies (FMT), Kerttu Saalasti Institute, University of Oulu, Pajatie 5, Nivala FI-85500, Finland
[c] Institute of Ultrasonic Technology, Shenzhen Polytechnic University, Shenzhen, 518055, P.R. China
[d] Laser processing and additive manufacturing, Mechanical Engineering Department, School of Energy Systems, Lappeenranta-Lahti University of Technology LUT, P.O. Box 20, FI-53851 Lappeenranta, Finland

ARTICLE INFO

Keywords:
Thermal drilling
Surface roughness
Bush dimensions
Machine learning
Random Forest Regressor
Gradient Boosting Regressor
Data-driven manufacturing

ABSTRACT

This study explores the application of machine learning algorithms, specifically Random Forest Regressor (RFR) and Gradient Boosting Regressor (GBR), to predict key outcomes of the friction drilling of A356 aluminum alloy. Optimizing process parameters such as rotational speed (RS), feed rate (FR), and preheat temperature (PH) is critical to achieve high-quality bushings during friction drilling. The study focused on predicting bush height (*ha*), thickness (*t*), surface roughness (*Ra*), and the induced temperature at workpiece/drilling-tool interface (*T*) through a dataset consisting of 27 experiments. The results showed that RS and PH had a significant influence on *ha* and *T*, with higher values of both parameters leading to increased bush height and induced temperature. Nevertheless, FR demonstrated a weaker effect on these responses but had a more pronounced impact on *t* and *Ra*. Feature importance analysis revealed that RS and PH were the most critical parameters for optimizing the friction drilling process, while FR had a lower effect. Additionally, the GBR model outperformed the RFR model in predicting *ha*, *t*, and *Ra*, providing more accurate results for these dimensions. Whereas the RFR exhibited a better behavior in predicting *T*, demonstrating the machine learning potential to enhance precision of the formed bushings.

1. Introduction

Friction drilling, also known as form or thermal drilling, is a non-conventional method of creating bushings and holes in thin-walled materials such as aluminum alloys [1,2]. The process employs a rotating conical or hexagonal tool to generate heat through friction, softening the material and displacing it to form the desired hole [3,4]. Unlike conventional drilling, friction drilling offers benefits such as bushing formation, which increases material strength and offers better load-bearing capacity, making it widely used in aerospace, and automotive industries [5,6].

The dimensions and quality of the formed bushings are intensively affected by the friction drilling parameters such as the drilling tool rotational speed and feed rate [7,8]. Generally, rises in the rotational speed and reductions in the feed rate led to a better surface roughness and a longer bushing height [9]. However, brittle materials in the as-cast condition are rarely investigated through friction drilling processing to avoid petal formation in the formed bushings, such as A356 aluminum alloy [10–12].

A356 aluminum alloy is commonly used in marine and automotive applications for its excellent castability, mechanical strength, and corrosion resistance [13]. Furthermore, A356 is utilized in industrial applications requiring lightweight, such as structural housings, and frames, as well as cast engine blocks, cylinder heads, and pistons due to its ability to withstand high temperatures and pressures [14]. Friction drilling is widely applied to thin housings and frames to provide extra bushing space sufficient for clamping purposes. However, achieving optimal surface quality and bushing formation during friction drilling of

* Corresponding author at: Mechanical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo 11629, Egypt.
E-mail addresses: mahmoud.khedr@feng.bu.edu.eg, mahmoud.khedr@oulu.fi (M. Khedr).

https://doi.org/10.1016/j.mtcomm.2025.112420
Received 25 December 2024; Received in revised form 22 February 2025; Accepted 1 April 2025
Available online 2 April 2025

as-cast A356 Al-alloy is difficult due to its brittleness [15]. Previous studies have demonstrated that carrying out preheating (prior to friction drilling) significantly enhances the quality and dimensions of drilled bushing [16,17]. Therefore, this study investigates the effect of preheating of as-cast A356 Al-alloy on the characteristics of the formed bushings at different rotational speeds and feed rates.

Optimizing the friction drilling parameters of A356 Al-alloy is crucial for improving the efficiency and quality of the formed bushings [18]. This can be achieved via advanced tools such as machine learning (ML) algorithms [19,20]. ML techniques could identify optimal material compositions and processing conditions corresponding to the best performance [21,22]. In industrial settings, optimizing friction drilling parameters can achieve desired outcomes through empirical methods and trial-and-error, which are resource-intensive and time-consuming [23]. With the advent of ML, predictive models have shown promise in enhancing process understanding and optimizing outcomes in complex manufacturing processes [24]. ML models, particularly Random Forests, have demonstrated robustness and adaptability in applications requiring non-linear modeling, high-dimensional data analysis, and feature interpretability [25]. However, limited research has been conducted on using ML to optimize friction drilling parameters [26], particularly for preheated A356 aluminum alloy.

This study addresses a critical research gap by investigating how variations in rotational speed, feed rate, and preheating influence key friction drilling outcomes. While prior studies have relied on empirical methods for process optimization, these approaches are often resource-intensive and time-consuming. ML approaches, such as Random Forest Regressor (RFR) and Gradient Boosting Regressor (GBR), offer a promising alternative by enabling data-driven predictions to optimize process parameters efficiently. Thus, the present study explores the application of RFR and GBR models to predict friction drilling outcomes in A356 aluminum alloy. The study aims to predict four key outcomes (bush height, bush thickness, induced temperature during friction drilling process, and surface roughness) based on tool rotational speed, feed rate, and specimen' preheat temperature. The significance of this research lies in its potential to enhance process efficiency and reduce trial-and-error in friction drilling applications. The main objective of the study is to develop ML-based predictive models, seeking to contribute to the field of data-driven manufacturing, enabling more efficient process control and quality assurance in friction drilling applications.

2. Experimental setup and methodology

2.1. Materials and casting procedures

Table 1 shows the chemical composition of the investigated material which is A356 aluminum alloy. The material was supplied at the shape of ingots by the EGYPTALUM company in Egypt. The ingots were melted in an induction furnace, then degassed by Argon gas, followed by pouring in C-steel rectangular molds. Subsequently, after solidification, the rectangular mold was cut into squared sheets with dimensions of $50 \times 50 \times 3\ mm^3$.

2.2. Friction drilling processing

The friction drilling experiments were conducted on a CNC milling machine equipped with a conical friction drilling tool. Fig. 1.a displays a schematic drawing of the friction drilling tool, depicting the main regions of the tool as well as the main dimensions. Fig. 1.b shows the holding fixture of the specimen, it has a grooved slot with similar dimensions as the workpiece to allow smooth sliding of the preheated specimens, whereas two grasping pins are designed to prevent the specimen from movements during the friction drilling processing. Fig. 1.c depicts a schematic drawing of the tool penetration through the workpiece. Fig. 1.d shows the geometry of the formed bushing after finishing the friction drilling process, displaying the diameter and height of the formed bushing.

During friction drilling parameters such as rotational speed (RS), and feed rate (FR) were manipulated with operating values as listed in Table 2. Furthermore, the preheating of the specimens implied 100, 150 and 200 °C.

2.3. Bushing dimensions, surface roughness, and temperature measurements

The quality of the friction drilled bushes was evaluated according to the values of the bushing height (*ha*), which represents the height of the extruded material around the drilled hole. and bushing thickness (*t*), which represents the thickness of the bush formed during the drilling processes.

Surface roughness of the bushings (*Ra),* which is critical to the quality and functionality of the drilled hole, was measured along the internal drilled hole surface via a Mitutoyo SJ-310 surface roughness tester. The average of five readings was recorded. Whereas the peak temperature (*T*) observed at the tool-workpiece interface during friction drilling was recorded via thermal imager (FLUKE Ti32) infrared camera.

2.4. Data setting and preprocessing

The dataset used in this study consists of 27 friction drilling experiments on the A356 aluminum alloy, where each experiment records the process parameters and corresponding outcomes. The input parameters imply RS, which was set at levels of 2000, 3000, and 4000 rpm, FR, which was set at levels of 40, 60, and 80 mm/min, and the workpiece preheat temperature (PH), whereas the preheating was carried out at 100, 150, and 200 °C. The values of input parameters were chosen based on their proven effectiveness in reducing cracks and improving bushing quality in similar aluminum alloys [10,27,28]. Preheating temperatures (100–200 °C) were derived from studies showing the beneficial effects of moderate preheating on material ductility and deformation [29]. The output variables imply *ha, t, Ra, and T*.

Given the small size of the dataset, preprocessing steps were minimal but crucial to ensure model reliability. Missing values were checked, though none were found, and all variables were standardized where necessary to eliminate biases related to variable scale. The dataset was then split into training (80 %) and testing (20 %) sets to evaluate model performance on unseen data.

2.5. Model selection and training

The RFR and GBR models were selected for their ability to handle non-linear relationships and provide feature importance scores, enhancing interpretability [30–32]. These models assume that the relationships between process parameters and output responses are non-linear and complex. RFR and GBR construct an ensemble of decision trees, each trained on a random subset of the data. The final prediction is the average of predictions from all trees, which reduces overfitting and enhances generalization [33]. Hyperparameter tuning was conducted to optimize model performance. Key hyperparameters, such as the number of trees (n_estimators), maximum tree depth (max_depth), and minimum

Table 1
Chemical composition of the investigated A356 aluminum alloy.

Elements	Si	Fe	Cu	Mn	Mg	Zn	Cr	Ti	Ni	Al
Weight %	6.23	0.067	0.004	0.0029	0.337	0.0007	0.0005	0.136	0.0003	93.1

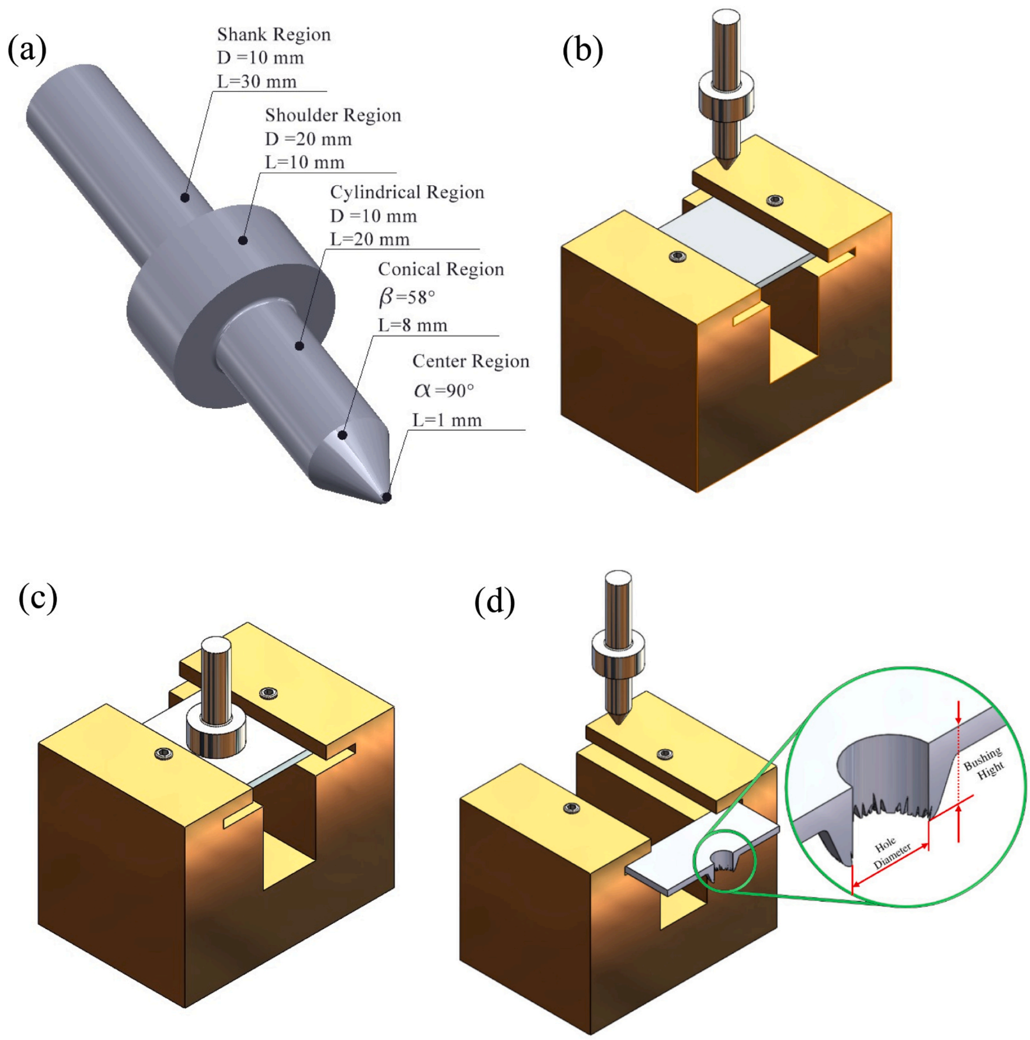

Fig. 1. Schematic drawing of (a) friction drilling tool, (b) fixing of the workpiece, (c) tool penetration through the workpiece, and (d) cross-section of the formed bushing.

Table 2
Values of friction drilling parameters.

Pre-heating temperature, °C	Rotational speed, rpm	Feed rate, mm/min
100, 150, and 200	2000, 3000, and 4000	40, 60, and 80

samples per leaf (min_samples_leaf), were explored using Grid Search with cross-validation. PFR and GBR models were processed via Jupyter notebook through Anaconda software [34].

2.6. Model evaluation metrics

To assess RFR and GBR models performance, two metrics were employed. (i) mean absolute error (MAE), which measures the average magnitude of prediction errors, providing a straightforward interpretation of accuracy. (ii) root mean square error (RMSE), which is similar to MAE but penalizes larger errors more, making it useful for understanding error distribution. By comparing MAE and RMSE for each output variable, a comprehensive understanding of the model's predictive accuracy is carried out.

3. Results and discussion

3.1. Experimental values

Table 3 shows the input parameters of *RS*, *FR*, and *PH* of the 27 experiments and the corresponding output parameters (*ha, t, Ra*, and *T*). Clearly, as RS increases from 2000 to 4000 rpm, there is a general increase in *ha* and *t*. This implies that higher rotational speeds promote more material deformation, leading to taller and thicker bushes. *T* also tends to increase significantly with higher RS, especially noticeable when RS increases from 3000 to 4000 rpm. This suggests that higher rotational speeds generate more heat due to increased friction. *Ra* generally decreases with increasing RS, especially at lower feed rates and lower preheat temperatures. This means higher speeds tend to create smoother surfaces, potentially due to the more continuous material flow at higher temperatures.

Additionally, it is clear that *ha* decreases as the feed rate increases. For example, at 2000 rpm and 100 °C, *ha* decreases from 6.80 mm at 40 mm/min to 4.90 mm at 80 mm/min. This indicates that higher feed rates may reduce the contact time for material deformation, resulting in shorter bushes. *t* also shows a slight tendency to decrease with increased FR, though this effect is less pronounced than for *ha*. *Ra* generally increases with higher feed rates. For instance, at 2000 rpm and 100 °C, *Ra* increases from 5.872 μm at 40 mm/min to 7.564 μm at 80 mm/min.

Table 3
Friction drilling parameters and corresponding output responses.

Input parameter			Output response			
RS, rpm	*FR*, mm/min	*PH*, °C	*ha*, mm	*t*, mm	*T*, °C	*Ra*, μm
2000	40	100	6.80	1.655	166.5	5.872
2000	40	150	7.50	1.760	200.7	6.09
2000	40	200	7.70	1.970	212.1	6.104
2000	60	100	5.40	1.600	154.4	6.352
2000	60	150	5.60	1.835	195.9	6.68
2000	60	200	6.10	1.930	205.8	7.122
2000	80	100	4.90	1.770	154.4	7.564
2000	80	150	5.00	2.040	185.7	7.886
2000	80	200	5.20	2.240	202	8.915
3000	40	100	8.20	1.740	189.7	3.79
3000	40	150	8.50	1.845	214	4.092
3000	40	200	8.70	1.815	267.4	4.223
3000	60	100	7.10	1.645	175.4	4.534
3000	60	150	7.30	1.690	212.2	4.85
3000	60	200	7.60	1.700	249.8	4.986
3000	80	100	6.80	1.695	171.9	5.157
3000	80	150	6.20	1.720	209.6	5.25
3000	80	200	5.90	1.725	218.8	5.686
4000	40	100	8.90	1.550	224.7	1.846
4000	40	150	9.20	1.655	276.5	1.992
4000	40	200	9.40	1.730	366.8	2.11
4000	60	100	8.20	1.625	200.9	2.362
4000	60	150	8.30	1.710	250.4	2.673
4000	60	200	7.90	1.885	334.7	2.984
4000	80	100	8.00	1.705	196.9	3.017
4000	80	150	7.80	1.715	214.3	3.278
4000	80	200	9.00	1.800	280	3.513

This leads us to conclude that higher feed rates lead to an increased cutting action, creating a rougher surface, which agrees with the results reported in the literature [10].

Regarding *PH*, *ha* and *t* generally increase with higher preheat temperatures. Heating the A356 Al-alloy softens it [17], allowing more deformation and extrusion around the drilled hole (at the tool/workpiece interface. *T* also increases with higher PH, as expected, since starting with a higher material temperature adds to the heat generated by friction. *Ra* tends to increase with PH, especially at lower rotational speeds. Preheating causes the material to become softer and more prone to roughness variations on the surface.

3.2. Correlation between input parameters and output responses

3.2.1. Heat maps

The heatmaps provide a comprehensive overview of the interactions between friction drilling parameters and output responses [20,35]. Fig. 2 depicts Pearson correlation heatmap between the input parameters (RS, FR, and PH) and the output responses (*ha*, *t*, *T*, and *Ra*). The correlation coefficients range from −1 to 1, where values closer to 1 indicate a strong positive relationship, values closer to −1 denote a strong negative relationship, and values near 0 suggest no significant linear correlation.

The analysis reveals several key patterns. A strong positive correlation exists between *ha* and PH, as well as between *T* and PH. This suggests that higher preheating temperatures lead to increased material deformation and induced temperature at the tool/workpiece interface, directly influencing the height and heat generation during the friction drilling process. Additionally, there is a moderate positive correlation between RS and both *ha* and *T*, indicating the significant role of rotational speed in influencing the thermal and dimensional characteristics of the formed bushing. In contrast, FR demonstrates relatively weaker correlations with the output responses, highlighting that variations in FR have less influence on the outputs.

In contrast, negative correlations are observed between RS and *Ra*, signifying that increased rotational speeds tend to improve surface finish, and thus reducing roughness. Similarly, *Ra* exhibits weaker positive

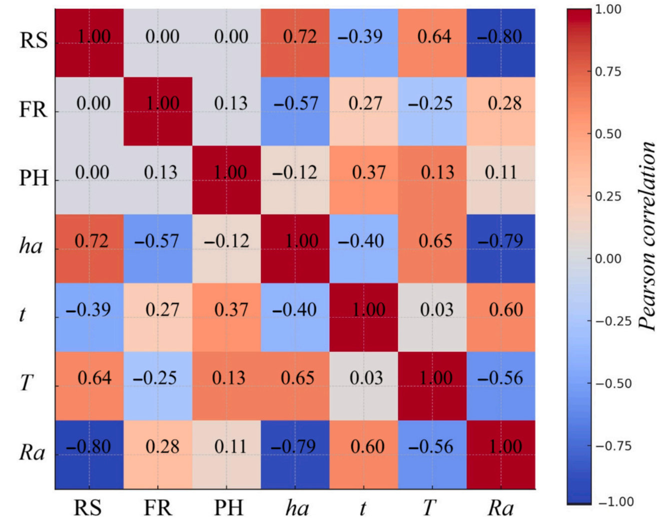

Fig. 2. Pearson correlation heatmap showing the linear relationships between input parameters and output responses.

correlations with *T*, suggesting that excessive heat slightly degrades the surface quality. There exists a moderate correlation between *t* and PH, RS, and *T*, reflecting the complex thermal and mechanical interactions during material deformation.

3.2.2. Main effects plot for ha, t, Ra, and T

Fig. 3 depicts the main effects plots for *ha, t, Ra*, and *T* (outputs) corresponding to variations in RS, FR, and PH (inputs). A main effects plot shows the relationship between each input variable and the output while averaging the effect of other variables. This helps to understand the individual impact of each input on the output without interactions with other variables. It is apparent that both *ha* and *T* show similar behaviors (Fig. 3.a and d). Clearly, both *ha* and *T* are increased with increases in RS and PH, but decrease with increasing FR, which could be attributed to enhanced plastic deformation resulting from increased dislocation mobility [36] under severe drilling conditions induced by high rotaional speed and preheating [17]. On the other hand, *t* and *Ra* decrease with increases in FR and PH, whereas they increase with RS increase (Fig. 3.b and c). The present results agree with the literature [37,38], displaying that rising the tool rotational speed, improves bushing heights but decreases bushing thicknesses, and vice versa. On the other hand, increased FR results in decreasing bush heights but increases the thickness. Additionally, the increase in PH enhances the induced temperature during friction drilling, which improves height and thickness of the formed bushes, but it deteriorates the quality of the surface (surface roughness values are increased).

3.3. Comparing experimental and predicted values of ha, t, Ra, and T

Fig. 4 depicts a comparison of the actual and predicted values for *ha*, *t*, *Ra*, and *T* corresponding to variations in RS, FR, and PH of the 27 experiments through RFR and GBR models. Clearly, the RFR model demonstrates a generally good fit with the experimental data for *ha*. However, there are noticeable deviations, particularly at higher bush height values. These discrepancies suggest that while RFR captures the overall trend well, it may experience some overfitting or underfitting in certain regions. In contrast, the GBR model shows a closer alignment with the experimental data, particularly at the higher bush height values, indicating that GBR may be better at capturing the variations in the larger bush heights. However, GBR does show some minor deviations at lower bush heights, suggesting that it may be more sensitive to smaller variations in the data.

In the case of *t*, the RFR model tends to over-predict at lower

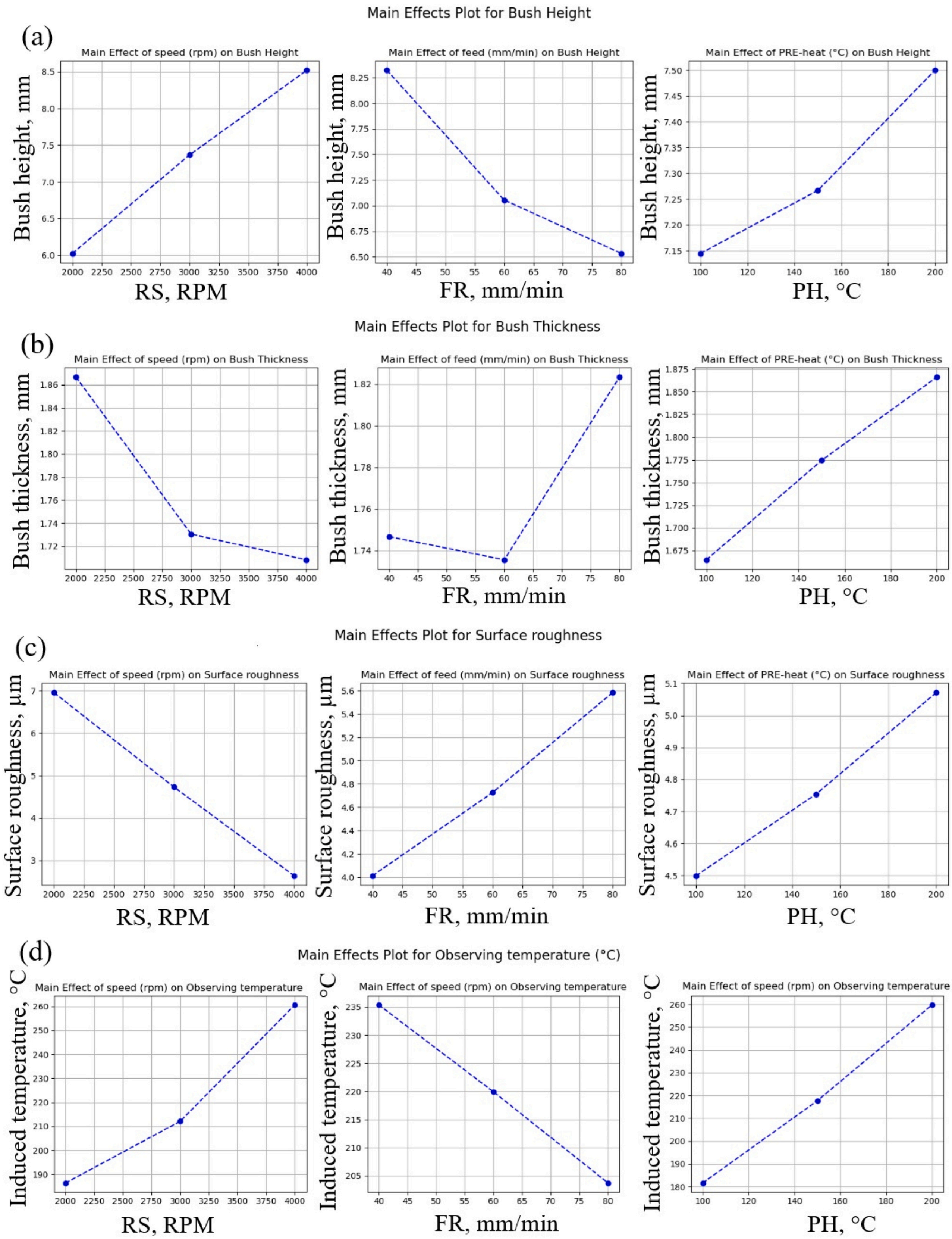

Fig. 3. Main effect plots for: (a) bush height, (b) bush thickness, (c) surface roughness, and (d) induced temperature during friction drilling processes at different values of RS, FR, and PH.

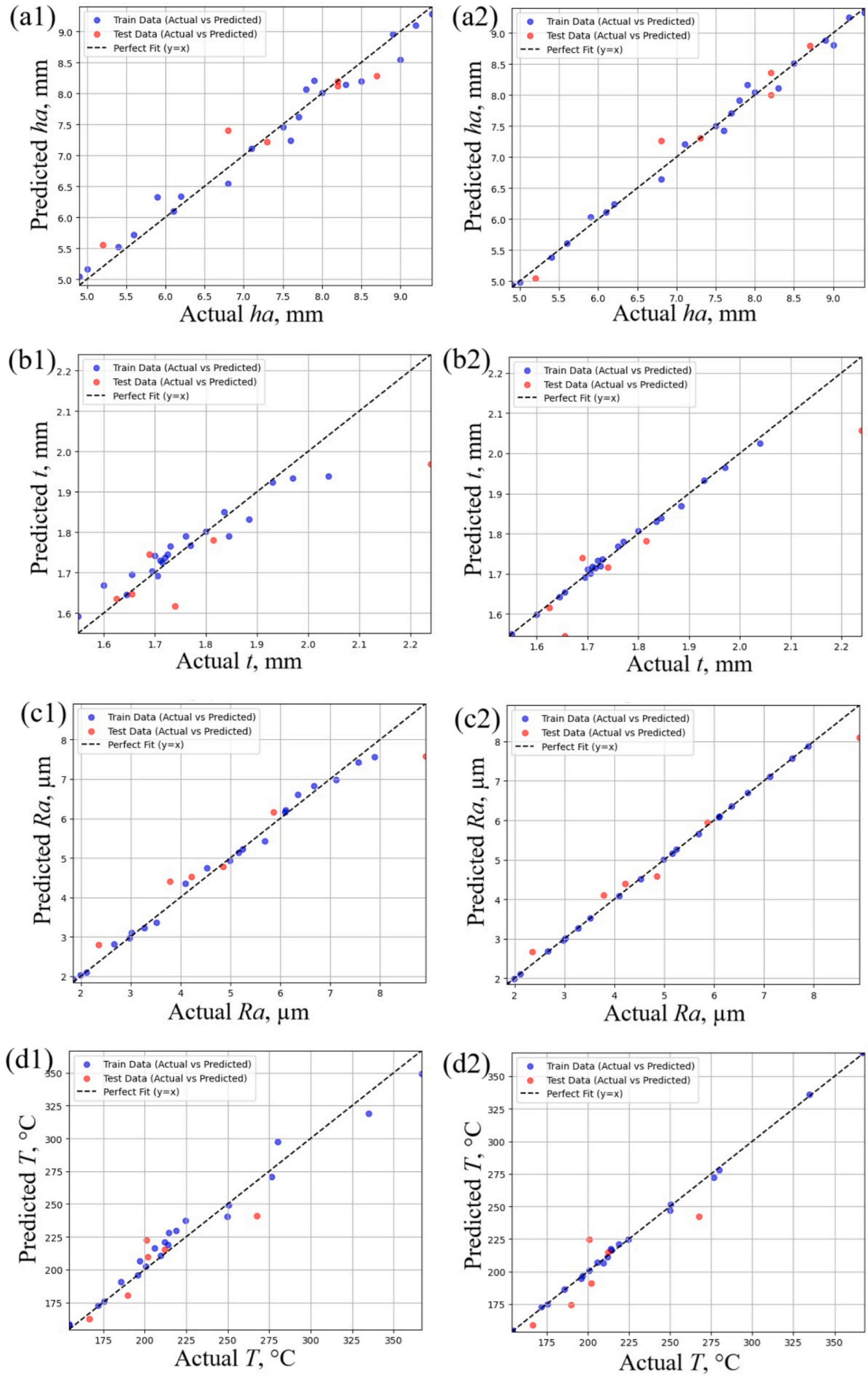

Fig. 4. Comparison of (1) RFR and (2) GBR results for the predicted and experimental values of: (a) bush height, (b) bush thickness, (c) surface roughness, and (d) induced temperature.

thickness values but provides relatively consistent predictions as the thickness increases. The RFR model's predictions show fewer outliers when compared to GBR. Conversely, the GBR model exhibits some stronger outliers, particularly at the higher bush thickness values, suggesting that GBR may under-predict the thickness in extreme cases. This indicates that GBR might struggle to capture the full range of bush thickness variations, particularly at the extremes conditions.

Regarding *Ra*, the RFR model performs reasonably well, but it tends to predict slightly higher roughness values, especially at lower *Ra* values. The predictions from RFR are generally scattered around the line of perfect fit, without exhibiting any clear bias. On the other hand, the GBR model appears to be more accurate in predicting surface roughness, showing fewer extreme outliers compared to RFR. However, GBR does struggle to predict values in the mid-range of *Ra*, where it experiences some small deviations.

Finally, in terms of *T*, the RFR model generally provides a good accuracy, although it slightly under-predicts the induced temperature at higher values. This under-prediction at extreme temperatures suggests that RFR may not be fully capturing the relationship in the temperature extremes. In contrast, the GBR model offers better predictions for high-temperature values, with less under-prediction than RFR. However, GBR tends to over-predict temperatures in the lower range, particularly in the intermediate temperature values. This highlights GBR's sensitivity to temperature extremes but also its tendency to overestimate temperatures at lower values. In the following section, we will evaluate the model accuracy through assessing its performance and metrics.

3.4. RFR and GBR models performance and metrics

Fig. 5 displays the evolution of the actual and predicted values of the test data for *ha*, *t*, *Ra*, and *T* (outputs) corresponding to variations in RS, FR, and PH (inputs) via the RFR and GBR models. The actual values display similar trends to the predicted outputs. However, there exists small variations between the actual values and the corresponding predicted ones, which arise for several reasons related to the model's structure, the nature of the data, and the complexity of the relationships being modelled [39]. Furthermore, small datasets or limited variability in the data can limit the model's ability to generalize. In our case, working with only 27 experiments, this is a relatively small dataset for RFR and GBR models, which thrives on large datasets to build diverse trees.

The GBR model depicts more accurate behavior than the RFR algorithm regarding *ha*, *t*, and *Ra*, whereas it has a similar accuracy for *T*. Although Random Forests generally resist overfitting, they can still overfit if too many trees or overly deep trees are used, especially with a small dataset. Overfitting can make the model too specialized to the training data, leading to less accurate predictions on new data. Conversely, if the Random Forest has too few trees or shallow trees, it might underfit, meaning it captures only basic patterns and misses finer details [35]. Therefore, the RFR exhibited lower discrepancies between actual and predicted values, which could be assessed through the values of MAE and RMSE.

The FRF and GBR models achieved varying levels of accuracy across the four output variables. Table 4 summarizes the MAE and RMSE values for each output, demonstrating that the RFR model is particularly effective at predicting observing temperature, while predictions for bush dimensions and surface roughness showed slightly higher errors than using of the GBR model. This discrepancy may reflect underlying complexities in the process or data limitations for those specific outcomes. It could be concluded that GBR model is more effective in predicting bushing dimensions and surface roughness than RFR, as it exhibited reduced error difference for *ha, t*, and *Ra*, since R^2 was greater for the GBR model in predicting their values.

3.5. Feature importance and interpretation of results

The feature importance analysis from the RFR and GBR models reveals that FR demonstrates weaker correlations with the output variables, especially when compared to RS and PH. Its effect on *ha* and *T* is less pronounced, but it has a noticeable negative impact on both *ha* and *T* as the feed rate increases. The FR has a more significant influence on *t*), where an increase in FR reduces the thickness, likely due to the more rapid material movement and less time for deformation at higher feed rates, in agreement with the literature [10].

RS plays a critical role in determining several output variables, particularly *ha* and *T*. This is consistent with the observations that higher rotational speeds lead to greater material deformation, increased heat generation, and consequently higher bush height and temperature, in agreement with the literature [40,41]. Additionally, RS has a moderate influence on *Ra*, with higher speeds resulting in smoother surfaces.

PH shows a strong positive influence on both *ha* and *T*, suggesting that preheating the workpiece leads to better material flow and higher thermal generation. This is in line with the findings that an increased PH results in higher bush height and temperature but deteriorates surface quality, causing an increase in surface roughness *Ra* [17]. PH has a relatively moderate negative effect on *t*, likely due to the complex thermal and mechanical interactions during material deformation.

For achieving smoother surface finishes, adjustments to FR are essential, with lower values generally producing smoother surfaces. RS can also be adjusted to control heat generation, preventing thermal damage and maintaining surface integrity. Temperature management can be optimized through careful control of RS, balancing heat production to facilitate material flow without compromising material quality. Bush height and thickness can be fine-tuned by adjusting FR and PH, allowing for precise control over bushing dimensions based on application needs.

By focusing on the most influential parameters, this approach enables a data-driven method for friction drilling optimization, reducing the need for trial-and-error and ensuring high process efficiency. This level of control is especially beneficial for industries such as aerospace and automotive manufacturing, where specific performance standards are essential.

3.6. Limitations and future work

Although the RFR and GBR models demonstrate promising predictive accuracy, certain limitations should be acknowledged. First, the dataset size is small, which may limit generalizability. Increasing the dataset and incorporating real-time sensor data could improve model robustness. Additionally, exploring other machine learning models, such as Gradient Boosting or Neural Networks, may yield improved accuracy and insights. Future studies could also investigate additional friction drilling parameters, such as tool geometry or lubrication conditions, to refine predictions further.

4. Conclusions

This study investigated the usage of RFR and GBR models for predicting key friction drilling outcomes in A356 aluminum alloy, specifically bush height (*ha*), bush thickness (*t*), surface roughness (*Ra*), and induced temperature at the workpiece/tool interface (*T*). Through a series of 27 friction drilling experiments, the effects of rotational speed (RS), feed rate (FR), and preheat temperature (PH) on these outputs were analyzed and modeled. By modelling the relationships between process parameters and drilling outcomes, this approach facilitates data-driven optimization in manufacturing. This research contributes to the growing body of literature on machine learning in manufacturing and demonstrates a practical application of predictive modelling for friction drilling. The following conclusions were drawn out:

1. Experimentally, it was found that increases in bush height with

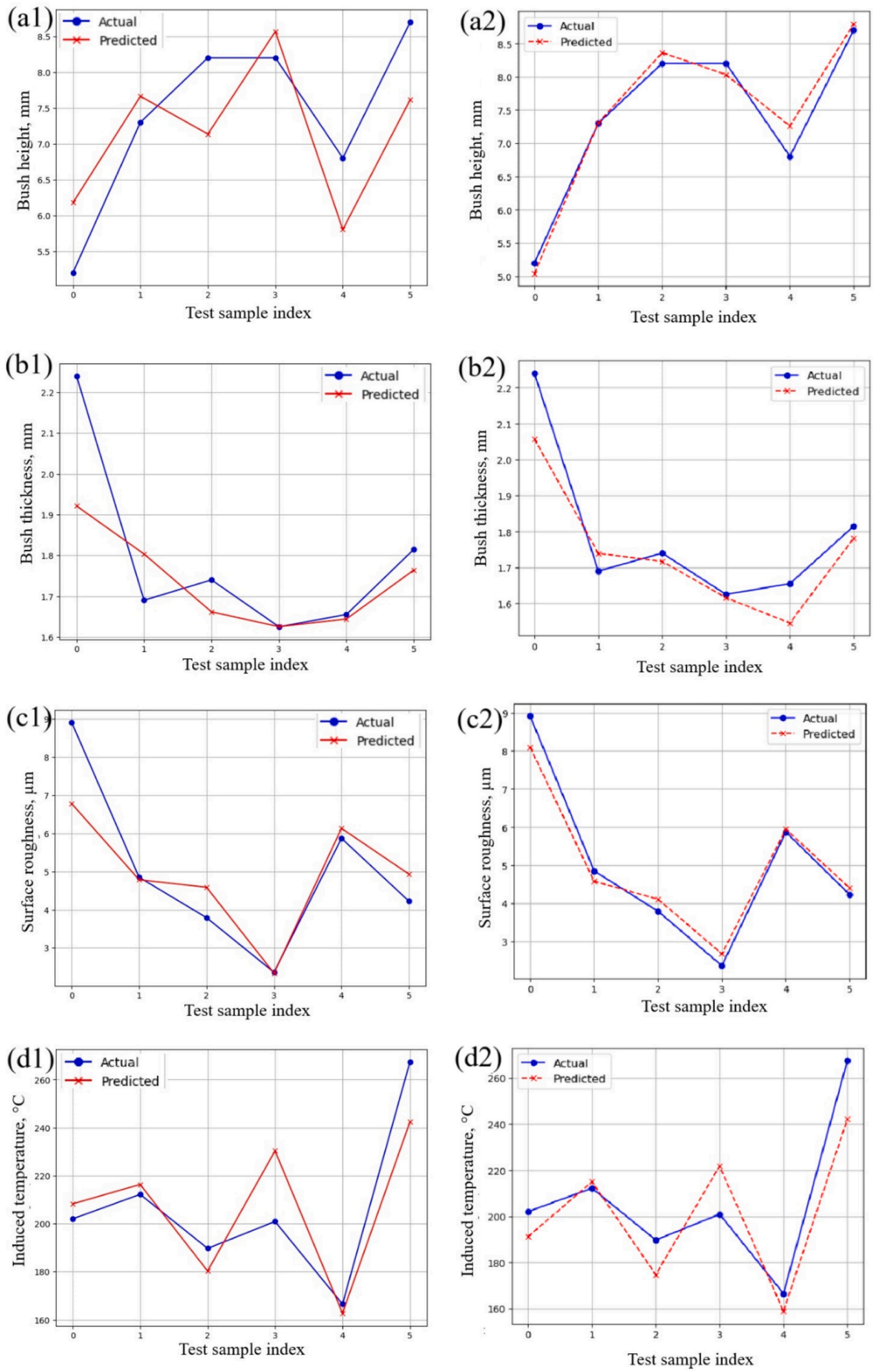

Fig. 5. Actual and predicted trends of the tested data for: (a) bush height, (b) bush thickness, (c) surface roughness, and (d) induced temperature using: (1) Random Forest Regression, and (2) Gradient Boosting Regressor model.

Table 4
Model accuracy for predicting friction drilling outputs with RFR and GBR.

Output variable	MAE		RMSE		R^2	
	RFR	**GBR**	**RFR**	**GBR**	**RFR**	**GBR**
Bush height, mm	0.257	0.175	0.336	0.222	0.917	0.964
Bush thickness, mm	0.0838	0.068	0.125	0.091	0.642	0.810
Surface roughness, µm	0.507	0.326	0.649	0.403	0.900	0.961
Induced temperature, °C	12.1	13.736	14.968	15.723	0.763	0.739

rising the tool rotational speed and preheating temperature. The bush thickness increased with increasing the feed rate and preheating temperature. However, surface roughness was deteriorated with increasing the feed rate and preheating temperature.

2. The main effects plots displayed that the feed rate is the most critical parameter influencing surface roughness and bush height, while rotational speed primarily impacts the temperature generated during drilling. Preheat temperature was also significant, particularly in enhancing bush thickness and facilitating material flow.

3. The GBR model outperforms the RFR model in predicting the bushing dimensions (*ha* and *t*) and surface roughness (*Ra*).

4. The RFR model, while effective at predicting the induced temperature (*T*), showed higher errors for bush height and surface roughness. This discrepancy may be attributed to overfitting or underfitting in regions with extreme values.

Despite the promising results of the present study, limitations remain due to the small dataset size, which may restrict the model's generalizability. Future studies should focus on expanding the dataset and incorporating real-time sensor data for dynamic modeling. Additionally, exploring alternative models, such as Bayesian Ridge regression and Neural Networks, may further enhance predictive accuracy and reveal deeper insights into the interactions between drilling parameters. Overall, this study highlights the value of machine learning in friction drilling optimization, underscoring its potential to improve friction drilling parameters and the quality of the produced bushings for specific manufacturing applications.

CRediT authorship contribution statement

Mahmoud Khedr: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ahmed Abdalkareem:** Methodology. **Amr Monier:** Writing – review & editing. **Rasha Afify:** Supervision. **Tamer S. Mahmoud:** Supervision. **Antti Järvenpää:** Resources, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors are grateful for the financial support of the TeräsPuu Project No. 2430496212. Open access to this article was provided by the FinELib agreement, University of Oulu.

Data availability

Data will be made available on request.

References

[1] A. Eliseev, E. Kolubaev, Friction drilling: a review, Int. J. Adv. Manuf. Technol. 116 (2021) 1391–1409, https://doi.org/10.1007/s00170-021-07544-y.
[2] M. Khedr, N. Hamzawy, A. Hamada, M. Jaskari, T.S. Mahmoud, I. El-Mahallawi, et al., Improved mechanical behavior of friction stir drilled 6082 aluminum alloy via T6 treatment, J. Mater. Res. Technol. 31 (2024) 2774–2785, https://doi.org/10.1016/j.jmrt.2024.06.189.
[3] R. Kumar, N.R.J. Hynes, A. Khan, Simulation of thermo-mechanical behaviour of friction drilling process, Int. J. Comput. Mater. Sci. Surf. Eng. 9 (2020) 70, https://doi.org/10.1504/IJCMSSE.2020.107401.
[4] M. Alphonse, V.K. Bupesh Raja, K. Palanikumar, D.R. Sai Krishna Sanjay, B. Venkata Subbaiah, L. Vinay Rama Chandra, Highlights of non-traditional friction drilling process a review, Mater. Today Proc. 46 (2021) 3582–3587, https://doi.org/10.1016/j.matpr.2021.01.336.
[5] Z. Demır, A geometrical approach to deriving a bushing height equation for thermal friction drilling, Mater. Tehnol. 53 (2019) 215–224, https://doi.org/10.17222/mit.2018.134.
[6] A. Guzanová, E. Janoško, D. Draganovská, M. Vrabeľ, M. Tomáš, P. Horňak, et al., Investigation of applicability flowdrill technology for joining thin-walled metal sheets, Metals (Basel) 12 (2022) 540, https://doi.org/10.3390/met12040540.
[7] H. Wu, R. Clarke, M. Porter, R. Ward, J. Quinn, C. McGarrigle, et al., Thread-stripping test procedures leading to factors of safety data for friction-drilled holes in thin-section aluminium alloy, Thin-Walled Struct. 163 (2021) 107653, https://doi.org/10.1016/j.tws.2021.107653.
[8] S. Dehghan, R. Abbasi, B.T. Hang. T. b. Baharudin, M.L. Mousavi, E. Soury, A novel approach to friction drilling process: experimental and numerical study on friction drill joining of dissimilar materials AISI304/AL6061, Metals 12 (2022) 920, https://doi.org/10.3390/met12060920.
[9] K.-Y. Su, T. Welo, J. Wang, Improving friction drilling and joining through controlled material flow, Procedia Manuf. 26 (2018) 663–670, https://doi.org/10.1016/j.promfg.2018.07.077.
[10] Y.E. Albarbary, R. Afify, E.H. Mansour, T.S. Mahmoud, M. Khedr, The effect of pre-drilling on the characteristics of friction drilled A356 cast aluminum alloy, J. Manuf. Process 82 (2022) 646–656, https://doi.org/10.1016/j.jmapro.2022.08.040.
[11] S.A. El-Bahloul, Pre-drilling effect on thermal friction drilling of cast aluminum alloy using thermo-mechanical finite element analysis, J. Appl. Comput. Mech. 6 (2020) 1371–1379, https://doi.org/10.22055/jacm.2020.34682.2455.
[12] S.A. El-Bahloul, Friction drilling of cast aluminum alloy A380 without significant petal formation and radial fracture, Int. J. Precis. Eng. Manuf. 20 (2019) 45–52, https://doi.org/10.1007/s12541-019-00039-7.
[13] E. Erzi, M. Tiryakioğlu, A simple procedure to determine incoming quality of aluminum alloy ingots and its application to A356 alloy ingots, Int. J. Met. 14 (2020) 999–1004, https://doi.org/10.1007/s40962-020-00414-5.
[14] T. Kimura, T. Nakamoto, Microstructures and mechanical properties of A356 (AlSi7Mg0.3) aluminum alloy fabricated by selective laser melting, Mater. Des. 89 (2016) 1294–1301, https://doi.org/10.1016/j.matdes.2015.10.065.
[15] M. Azadi, M. Zomorodipour, A. Fereidoon, Sensitivity analysis of mechanical properties and ductile/brittle behaviors in aluminum-silicon alloy to loading rate and nano-particles, considering interaction effects, Eng. Rep. 3 (2021), https://doi.org/10.1002/eng2.12341.
[16] M. Bilgin, Minimum quantity lubrication and heat-assisted friction drilling of AA7075-T6 aluminum alloy, CIRP J. Manuf. Sci. Technol. 35 (2021) 819–829, https://doi.org/10.1016/j.cirpj.2021.09.011.
[17] A. Abdalkareem, R. Afify, N. Hamzawy, T.S. Mahmoud, M. Khedr, Thermal management of Friction-Drilled A356 aluminum alloy: a study of preheating and drilling parameters, J. Manuf. Mater. Process. 8 (2024) 251, https://doi.org/10.3390/jmmp8060251.
[18] D.O.R. Riswanto, A. Rohman, S. Pramono, S. Martono, Application of response surface methodology as mathematical and statistical tools in natural product research, J. Appl. Pharm. Sci. 9 (2019) 125–133, https://doi.org/10.7324/JAPS.2019.91018.
[19] P. Kathmore, B. Bachchhav, S. Nandi, S. Salunkhe, P. Chandrakumar, E.A. Nasr, et al., Prediction of thrust force and torque for high-speed drilling of AL6061 with TMPTO-based bio-lubricants using machine learning, Lubricants 11 (2023) 356, https://doi.org/10.3390/lubricants11090356.
[20] S. Jain, R. Jain, K.R. Rao, A. Bhowmik, Leveraging machine learning to minimize experimental trials and predict hot deformation behaviour in dual phase high entropy alloys, Mater. Today Commun. 41 (2024) 110813, https://doi.org/10.1016/j.mtcomm.2024.110813.
[21] C. Liu, X. Wang, W. Cai, Y. He, H. Su, Machine learning aided prediction of glass-forming ability of metallic glass, Processes 11 (2023) 2806, https://doi.org/10.3390/pr11092806.
[22] C. Liu, S. Min, Y. Pang, Y. Chai, The meshfree radial point interpolation method (RPIM) for wave propagation dynamics in non-homogeneous media, Mathematics 11 (2023) 523, https://doi.org/10.3390/math11030523.
[23] R. Agarwal, V. Gupta, J. Singh, V. Jain, Prediction of surface roughness and cutting force induced during rotary ultrasonic bone drilling via statistical and machine learning algorithms, Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci. 236 (2022) 11123–11135, https://doi.org/10.1177/09544062221110407.
[24] C. Liu, H. Su, Machine learning aided prediction of martensite transformation temperature of NiTi-based shape memory alloy, Mater. Today Commun. 41 (2024) 110720, https://doi.org/10.1016/j.mtcomm.2024.110720.
[25] S. Shoar, N. Chileshe, J.D. Edwards, Machine learning-aided engineering services' cost overruns prediction in high-rise residential building projects: application of

random forest regression, J. Build. Eng. 50 (2022) 104102, https://doi.org/10.1016/j.jobe.2022.104102.

[26] A. Bustillo, G. Urbikain, J.M. Perez, O.M. Pereira, L.N. Lopez de Lacalle, Smart optimization of a friction-drilling process based on boosting ensembles, J. Manuf. Syst. 48 (2018) 108–121, https://doi.org/10.1016/j.jmsy.2018.06.004.

[27] S.-M. Lee, H.-M. Chow, B.-H. Yan, Friction drilling of IN-713LC cast superalloy, Mater. Manuf. Process. 22 (2007) 893–897, https://doi.org/10.1080/10426910701451697.

[28] L. Ozler, N. Dogru, An experimental investigation of hole geometry in friction drilling, Mater. Manuf. Process. 28 (2013) 470–475, https://doi.org/10.1080/10426914.2012.746699.

[29] S.F. Miller, J. Tao, A.J. Shih, Friction drilling of cast metals, Int. J. Mach. Tools Manuf. 46 (2006) 1526–1535, https://doi.org/10.1016/j.ijmachtools.2005.09.003.

[30] Y. Ao, H. Li, L. Zhu, S. Ali, Z. Yang, The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling, J. Pet. Sci. Eng. 174 (2019) 776–789, https://doi.org/10.1016/j.petrol.2018.11.067.

[31] D.A. Otchere, T.O.A. Ganat, J.O. Ojero, B.N. Tackie-Otoo, M.Y. Taki, Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions, J. Pet. Sci. Eng. 208 (2022) 109244, https://doi.org/10.1016/j.petrol.2021.109244.

[32] M.A. Davari, A. Kadkhodaie, Comprehensive input models and machine learning methods to improve permeability prediction, Sci. Rep. 14 (2024) 22087, https://doi.org/10.1038/s41598-024-73846-2.

[33] L.W. Rizkallah, Enhancing the performance of gradient boosting trees on regression problems, J. Big Data 12 (2025) 35, https://doi.org/10.1186/s40537-025-01071-3.

[34] Anaconda n.d. https://www.anaconda.com.

[35] H. Zheng, Y. Wu, A XGBoost model with weather similarity analysis and feature engineering for short-term wind power forecasting, Appl. Sci. 9 (2019) 3019, https://doi.org/10.3390/app9153019.

[36] M. Khedr, W. Li, N. Min, W. Abd-Elaziem, Strengthening contributions of mechanical twinning and dislocations to the flow stress of hadfield high-manganese steel: quantitative analysis, J. Mater. Eng. Perform. 32 (2023) 501–511, https://doi.org/10.1007/s11665-022-07113-8.

[37] J.D. Skovron, R. Rohan Prasad, D. Ulutan, L. Mears, D. Detwiler, D. Paolini, et al., Effect of thermal assistance on the joint quality of Al6063-T5A during flow drill screwdriving, J. Manuf. Sci. Eng. 137 (2015), https://doi.org/10.1115/1.4031242.

[38] H. Vanhove, E. Ozden, J.R. Duflou, An experimental study on bushing formation during friction drilling of titanium grade 2 for medical applications, J. Manuf. Mater. Process. 7 (2023) 220, https://doi.org/10.3390/jmmp7060220.

[39] S.B. Nadkarni, G.S. Vijay, R.C. Kamath. Comparative Study of Random Forest and Gradient Boosting Algorithms to Predict Airfoil Self-Noise. RAiSE-2023, MDPI, Basel Switzerland, 2023, p. 24. ⟨http://doi.org/10.3390/engproc2023059024⟩.

[40] C. Fang, S. Yu, H. Peng, X. Deng, J. WangIonic Effect of NaCl and KCl on the Flotation of Diaspore and Kaolinite Using Sodium Oleate as Collector, 2020, p. 40–46. ⟨http://doiorg/10.1007/978-3-030-36408-3_6⟩.

[41] N. Hamzawy, T.S. Mahmoud, I. El-Mahallawi, T. Khalifa, M. Khedr, Optimization of thermal drilling parameters of 6082 Al-alloy based on response surface methodology, Arab. J. Sci. Eng. 48 (2023) 12001–12014, https://doi.org/10.1007/s13369-023-07628-9.