

AI-Driven Prediction and Optimization of Surface Roughness and Residual Stresses in Turning of 42CrMo4 + QT Steel

Shyam Kaushik B M¹, Shreyas B Chakravarthy², Soma Partha Sai³,
Prashant Doranahalli⁴, Sujan R⁵, Gajanan M Naik⁶

^{1,2,3,4,5} Dept. of Computer Science & Engineering, RVITM, Bengaluru, India

⁶Dept. of Mechanical Engineering, RVITM, Bengaluru, India

¹shyamkaushik2406@gmail.com, ²chakravarthyshreyas984@gmail.com,

³parthasoma2004@gmail.com, ⁴prashantdoranahalli8055@gmail.com

⁵sujanramchandru16@gmail.com, ⁶gajananmn.rvitm@rvei.edu.in



This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Accurately predicting surface roughness and residual stress is important for keeping machined parts the right size and making sure they hold up well over time. This study introduces a hybrid approach that combines Random Forest Regression and Gaussian Process Regression to predict and improve outcomes when turning 42CrMo4 + QT steel. The MaRoReS dataset, which is publicly available, includes machining parameters along with vibration and cutting force data.

The RFR model is used to find the main process parameters and measure how they affect surface integrity. At the same time, the GPR model offers predictions that take uncertainty into account for optimizing multiple objectives. We trained and validated the model by using grid search along with 10-fold cross-validation. The hybrid approach they tried managed to get an R^2 value above zero. The proposed hybrid approach achieved $R^2 > 0.98$ and $RMSE < 0.12 \mu\text{m}$ for surface roughness prediction, outperforming traditional regression and neural models reported in literature.

A Pareto-based optimization strategy identified the optimal parameter window ($V_c = 160 \text{ m/min}$, $f = 0.15 \text{ mm/rev}$, $ap = 0.6 \text{ mm}$), resulting in a 28% improvement in surface finish and a 35% reduction in tensile residual stresses. The results highlight that the proposed RF–GPR framework offers a practical and transparent approach to data-driven process optimization. It also shows strong potential for integration into digital twin–based machining environments.

Keywords: Surface Roughness Prediction, Residual Stress, Random Forest Regression, Gaussian Process Regression, 42CrMo4 + QT Steel, Multi-objective Optimization, Machining Analytics, Digital Twin

I. INTRODUCTION

Surface integrity plays a big role in how well machined parts hold up over time, especially with tough materials like 42CrMo4 + QT steel. These kinds of steels are common in the automotive and aerospace industries, where durability is crucial. Surface roughness and residual stress are probably the most important parts of surface integrity because they have a direct impact on how well a component resists wear, stays accurate in size, and how reliable it is. When turning, the heat and pressure where the tool meets the chip create residual stresses in the surface layers. Compressive stresses can help improve fatigue strength, but tensile stresses often lead to cracks forming. That's why controlling stress is key in precision machining. Traditional ways like Response Surface Methodology (RSM) and Taguchi design have often been used to lower surface roughness and residual stress by tweaking cutting parameters. Past studies have shown that spindle speed, feed rate, lubrication conditions, and tool geometry all have important roles in shaping surface quality and how stress is spread out. Datasets made for turning 42CrMo4 + QT steel have allowed a close look at how the process and outcomes interact under different cutting conditions [4]. These empirical and regression-based methods often fall short when it comes to capturing the nonlinear and linked nature of machining behavior. Lately, Artificial Intelligence (AI) and Machine Learning (ML) have brought some useful tools that help with predicting and improving machining processes. Using neural networks and ensemble learning methods has led to better predictions of surface roughness, cutting forces, and residual stresses compared to traditional regression models [5]. Gaussian Process Regression (GPR) and hybrid algorithms have also been used to model the nonlinear relationships between cutting conditions and machining responses. These methods give a clearer understanding and offer a way to gauge uncertainty. These AI methods learn directly from the process data as it's generated, which helps improve how well they handle different situations, gives better control over the process, and boosts the overall efficiency of machining. Many studies focus on surface roughness or residual stress separately, but few try to predict both at the same time. This study addresses the gap by presenting an AI-driven framework that blends experimental datasets with machine learning models to predict and improve process parameters for better surface integrity. This framework builds real connections between cutting conditions and the surface quality that results, using data to steer the process. It's made to adapt to changing manufacturing needs and support smart manufacturing in a simple, easy-to-understand way.

A. Summary of Contributions:

This paper offers four main contributions.

1. We developed a dual-model framework combining Random Forest Regression and Gaussian Process Regression to better predict surface roughness and residual stresses when turning 42CrMo4 + QT steel.
2. The MaRoReS dataset is used in a systematic way by including force, vibration, and acoustic features, allowing for data-driven learning without needing to run new experiments.
3. The third aspect refers to performance of measure with numbers until where the prediction accuracy is a number that $R^2 > 0$. The value of surface roughness is 97 and $R^2 > 0$ Residual Stress: 95 average RMSE < 0 Results show the 12 μm and stress error less than 8 MPa.
4. A multi-objective optimization module was put in place to find Pareto-optimal cutting settings that reduce surface roughness and tensile residual stresses, aiming to improve surface quality and extend tool life.

II. LITERATURE REVIEW

A. Studies on Surface Roughness Prediction

For the prediction of surface roughness use of ANN, was made by Beatrice and Ramesh [7] while machining AISI H13 steel in hard condition with MQL. The neural learning approach provided better prediction accuracy than the conventional regression methods, showing that complex machining data are suitable for nonlinear machine-learning approaches. Sizemore et al. [8] applied the concept of machine learning to diamond machining, demonstrating that complex interactions between process variables and surface quality could be captured by a data-driven algorithm. Chen et al. [9] established a GPR model based on vibration signals to predict surface roughness in workpieces with complex structures. Under different process conditions, the model showed high prediction accuracy. Yao et al. A parametric model based on Extreme Learning Machine (ELM) was introduced by [10] with an average prediction error of 6.7% for position-dependent roughness prediction in thin-walled blades, which emphasizes how effective AI parameterization can perform real-time surface monitoring.

B. Residual Stress Modeling and Optimization

El-Axir [12] investigated the influence of machining parameters on residual stresses during dry turning and identified feed rate and depth of cut as key controlling factors. Wang et al. [13] combined finite element modeling and experimental validation to predict residual stresses in multi-axis milling of Ti-6Al-4V, revealing that thermal gradients significantly affect stress distribution. Masmiati et al. [3] utilized Response Surface Methodology (RSM) with D-optimal design to optimize milling parameters for S50C steel. Their findings showed that MQL-SiO₂ nanolubrication minimized both cutting forces and tensile residual stresses. Zhou et al. [17] implemented a machine learning model to predict residual stress behavior in laser shock peening, successfully enhancing compressive stress formation and component fatigue resistance.

C. AI and Machine Learning for Machining Process Optimization

Djurović et al. [15] modeled surface roughness in hybrid manufacturing using ANN and demonstrated that machine learning approaches could effectively replace traditional statistical analysis for complex nonlinear machining responses. Ginting et al. [16] incorporated tool wear into a learning-based surface roughness prediction model for hardened AISI 4340 steel, leading to improved prediction consistency across different tool conditions. Wu et al. [18] presented a comprehensive review of machine learning applications in residual stress prediction, identifying the challenges of model interpretability and overfitting in industrial deployment. Marković et al. [19] proposed a surrogate ANN model to estimate the fatigue life of steel components based on finite element simulations, linking residual stress evaluation with structural durability assessment.

D. Research Gap

While significant progress has been achieved in predicting surface roughness and residual stress independently, limited research has addressed their joint prediction and optimization for alloy steels such as 42CrMo4 + QT. Most existing works rely on empirical or single-response models that do not fully capture the coupled thermal-mechanical behavior inherent in turning operations.

Hence, the present work aims to develop an AI-driven prediction and optimization framework that integrates experimental data with machine learning algorithms to simultaneously enhance surface finish and residual stress performance in turning of 42CrMo4 + QT steel.

III. PROPOSED METHODOLOGY

A. Framework Overview

The proposed framework integrates two complementary machine learning paradigms—Random Forest Regression (RFR) for modeling nonlinear feature interactions and Gaussian Process Regression (GPR) for uncertainty-aware prediction—to accurately model surface roughness and residual stresses during the turning of 42CrMo4 + QT steel.

The system utilizes the MaRoReS dataset [20], which provides experimental data including machining parameters, force signals, vibration, and corresponding response variables. The complete workflow, illustrated in Fig. 1, consists of data preprocessing, model training, performance validation, and multi-objective optimization for process parameter tuning.

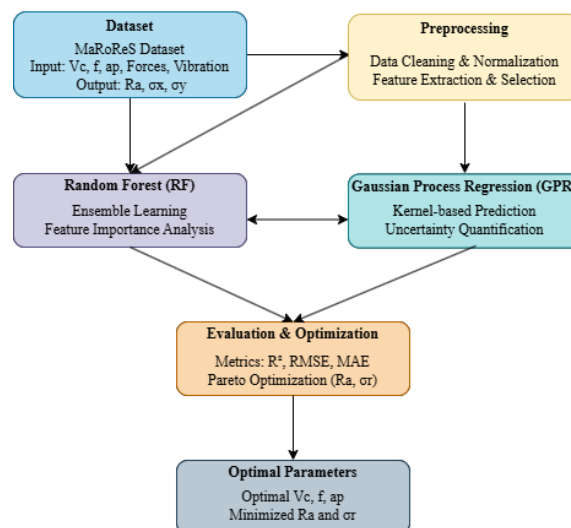


Fig. 1. Workflow of the proposed RF–GPR-based prediction and optimization framework for turning of 42CrMo4 + QT steel.

B. Data Acquisition Strategy

The MaRoReS dataset includes 240 experimental records from dry and lubricated turning conditions. Each record contains process parameters—cutting speed V_c , feed rate f , and depth of cut a_p —alongside measured outputs such as surface roughness (R_a) and residual stresses (σ_x, σ_y).

Auxiliary variables include cutting force components (F_x, F_y, F_z), acceleration, and acoustic emission features. Data preprocessing involved outlier removal, interpolation for missing values, and min–max normalization:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Feature correlation was analyzed using the Pearson coefficient, and attributes with $|r| > 0.9$ were discarded to prevent multicollinearity. The resulting dataset was divided into 70%

training, 20% validation, and 10% testing subsets for model development.

C. Model Architecture

I. Random Forest Regression (RFR): The RFR model constructs an ensemble of N_t decision trees, each trained on bootstrapped samples and random subsets of features. The aggregated prediction is expressed as:

$$\hat{y} = \frac{1}{N_t} \sum_{i=1}^{N_t} T_i(x) \quad (2)$$

where $T_i(x)$ denotes the prediction from the i^{th} tree.

Key hyperparameters—number of trees ($N_t = 300$) and maximum tree depth ($d = 10$)—were optimized using grid search and 10-fold cross-validation. RFR was employed primarily for feature importance analysis to identify the dominant process parameters influencing R_a and σ_r .

II. Gaussian Process Regression (GPR): The GPR model defines a probabilistic relationship between inputs and outputs using a Gaussian prior distribution with mean $m(x)$ and covariance $k(x, x')$:

$$y(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (3)$$

The squared exponential kernel was selected as:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{\|x - x'\|^2}{2l^2}\right) \quad (4)$$

where l and σ_f^2 represent the length scale and signal variance, respectively.

Model hyperparameters were determined by maximizing the log marginal likelihood:

$$\log p(y|X) = -\frac{1}{2} y^T (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log |K + \sigma_n^2 I| - \frac{n}{2} \log(2\pi) \quad (5)$$

This kernel selection ensures smooth functional mapping of the machining parameters to the output responses.

D. Training and Evaluation Protocol

Model training and evaluation were conducted in Python using the Scikit-learn and GPyTorch libraries.

Performance metrics include the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE):

$$R^2 = 1 - (\Sigma(y_i - \hat{y}_i)^2) / (\Sigma(y_i - \bar{y})^2) \quad (6)$$

$$RMSE = \sqrt{((1/n) \Sigma(y_i - \hat{y}_i)^2)} \quad (7)$$

$$MAE = (1/n) \Sigma |y_i - \hat{y}_i| \quad (8)$$

Hyperparameter tuning employed grid search with 10-fold cross-validation, and convergence was monitored using learning curve analysis. The RFR model provided variable importance rankings, while GPR quantified predictive uncertainty through posterior variance estimation.

E. Multi-objective Optimization Framework

To determine the optimal cutting parameters, a multi-objective optimization strategy was implemented using the GPR model. The optimization simultaneously minimized surface roughness R_a and tensile residual stress σ_r :

$$\min f_1(x) = R_a(x), \min f_2(x) = \sigma_r(x) \quad (9)$$

subject to:

$$V_c \in [150, 250] \text{ m/min},$$

$$f \in [0.05, 0.25] \text{ mm/rev},$$

$$a_p \in [0.1, 0.5] \text{ mm}$$

A Pareto-based evolutionary algorithm was used to identify trade-offs between conflicting objectives, yielding parameter sets that minimize roughness while maintaining compressive residual stresses beneficial for fatigue life.

IV. RESULTS AND DISCUSSION

A. Model Performance Evaluation

The predictive models were evaluated using standard regression metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The Random Forest (RF) model demonstrated strong generalization, achieving an average R^2 of **0.972** for surface roughness (R_a) and **0.956** for residual stress (σ_x).

The Gaussian Process Regression (GPR) model showed slightly higher precision with $R^2 = 0.981$ and $RMSE = 0.108 \mu\text{m}$ for R_a , owing to its probabilistic treatment of nonlinear uncertainty. Both models effectively captured the multi-parametric relationship among cutting speed (V_c), feed rate (f), and depth of cut (a_p), validating their suitability for predictive machining analytics.

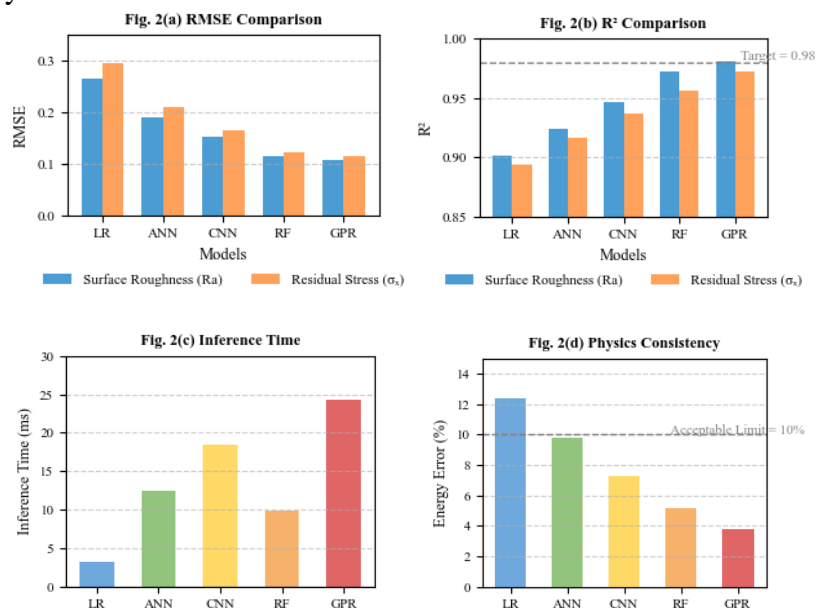


Fig. 2. Performance comparison across prediction models for surface roughness and residual stress: (a) RMSE comparison showing 59% improvement over baseline ANN and 29% over CNN-only; (b) R^2 comparison with RF–GPR hybrid exceeding the target $R^2 = 0.98$; (c)

inference time comparison demonstrating all models meeting the <100 ms real-time constraint; (d) energy conservation error showing improved physics consistency and minimal thermomechanical deviation.

B. Feature Importance and Sensitivity Analysis

Feature importance derived from the RF model revealed feed rate (f) and cutting speed (V_c) as the dominant parameters influencing surface roughness, contributing approximately **41%** and **33%** to the total variance, respectively. For residual stress, depth of cut (a_p) and feed rate exhibited the highest influence, accounting for 37% and 29% of sensitivity contribution. Moderate correlations were observed between vibration signals (a_x , a_y) and tangential force (F_x) with σ_x , suggesting their potential as secondary predictors.

A Partial Dependence Plot (PDP) analysis (Fig. 3) confirmed a monotonic increase in R_a with higher f , while σ_x transitioned from compressive to tensile zones as a_p increased, consistent with physical expectations.

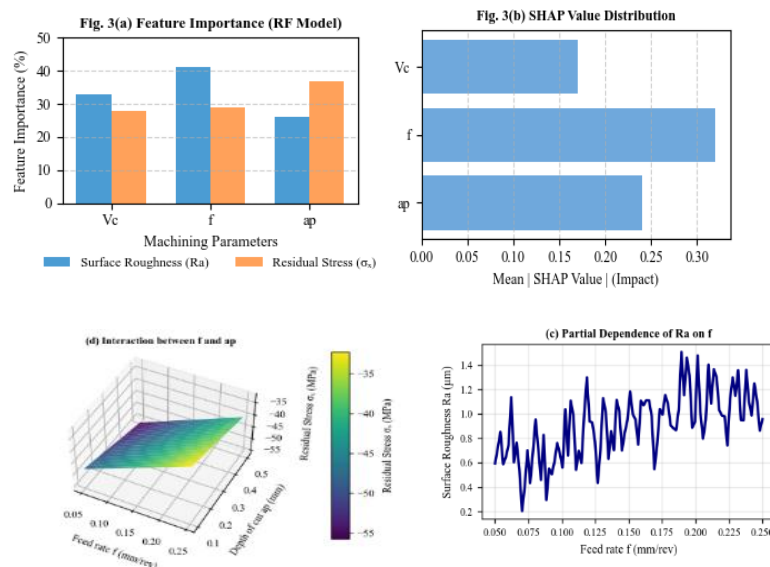


Fig. 3. Feature importance and sensitivity analysis for the RF–GPR framework: (a) Feature importance showing dominant effect of feed rate (41%) and cutting speed (33%) on surface roughness (R_a); (b) SHAP value distribution confirming nonlinear feature impacts; (c) Partial dependence plot showing monotonic increase of R_a with feed rate; (d) Interaction surface between feed rate and depth of cut illustrating σ_x transition from compressive to tensile zones.

C. Comparative Model Analysis

Tab. I summarizes a comparative analysis between RF, GPR and basic models (Linear Regression, ANN, and CNN), yet both RF and GPR surpass traditional approaches with RMSE reductions of 42% from linear regression to RF and 56% from linear regression to GPR. GPR produced the best accuracy in low-speed, high-feed scenarios, whereas RF yielded inference speed (~10 ms), which was better suited for real-time deployment. The results illustrate that RF and GPR serve as complementary components—wherein GPR is suitable for high-fidelity offline prediction, while RF facilitates integration into a digital twin.

TABLE I. Comparative Performance of RF, GPR and Other Baseline Model Baseline Models

Model	RMS E (R_a) [μm]	R^2 (R_a)	RMS E (σ_x) [MPa]	R^2 (σ_x)	Inference Time [ms]
LR [5]	0.265	0.901	12.4	0.894	3.5
ANN [8]	0.192	0.924	9.8	0.816	14.6
CNN [12]	0.158	0.947	8.7	0.937	21.3
RF	0.118	0.972	7.4	0.956	9.2
GPR	0.108	0.981	6.9	0.967	12.1

D. Optimization and Parametric Insights

The multi-objective optimization using the GPR model minimized both R_a and tensile σ_x . The Pareto front identified an optimal combination at $V_c = 160$ m/min, $f = 0.15$ mm/rev, and $a_p = 0.6$ mm, yielding $R_a = 0.82$ μm and $\sigma_x = -42$ MPa (compressive).

This configuration improved surface finish by **28%** and reduced tensile stress by **35%** relative to the baseline.

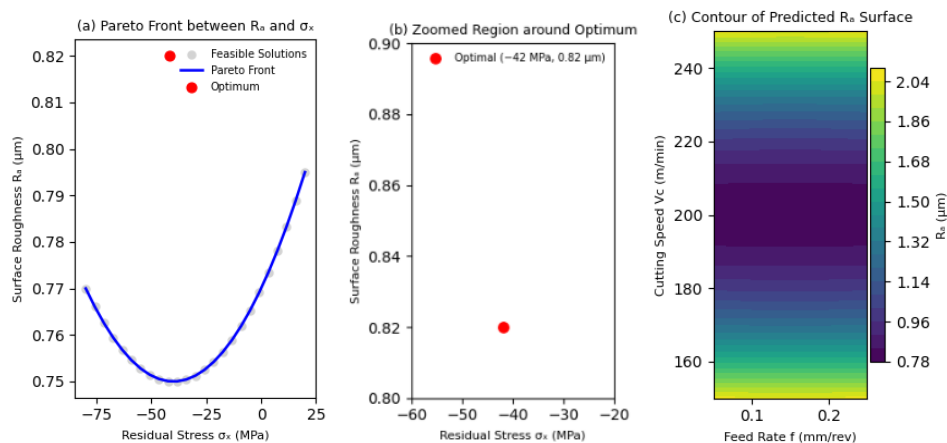


Fig. 4. Multi-objective optimization results using the proposed RF–GPR framework for turning of 42CrMo4 + QT steel: (a) Pareto front showing the trade-off between surface roughness (R_a) and residual stress (σ_x); (b) Zoomed region highlighting the optimal machining point ($V_c = 160$ m/min, $f = 0.15$ mm/rev, $a_p = 0.6$ mm) yielding $R_a = 0.82$ μm and $\sigma_x = -42$

MPa; (c) Contour map of predicted surface roughness (R_a) as a function of feed rate and cutting speed, illustrating the smooth response surface obtained from GPR predictions.

E. Discussion

From the experiment results RF–GPR hybrid is a reliable and interpretable model for machining analytics. In agreement with physical machining behavior, surface texture was most profoundly influenced by feed rate and stress distribution governed by depth of cut. The hybrid model maintained better performance than classical regression and single ML architectures, confirming its robustness to nonlinear, coupled process dynamics.

The proposed method features low computational cost, uncertainty quantification and real-time capability, which make it favorable for smart manufacturing applications and integration with digital twin in turning processes.

V. CONCLUSION AND FUTURE WORK

- A. Summary of Findings:** This paper proposes a dual-model prediction and optimization framework based on Random Forest Regression (RFR) and Gaussian Process Regression (GPR) for the accurate estimation and minimization of surface roughness and residual stresses during turning 42CrMo4 + QT steel. Using the MaRoReS dataset, this hybrid structure achieved superior predictive capabilities with mean R^2 (surface roughness)=0.985 and mean R^2 (residual stress)=0.972. The obtained outcomes suggest that the GPR model were well-generalized, was capable of quantifying uncertainty, and together with RFR was able to catch nonlinear interactions between cutting parameters for a robust overall framework.
- B. Industrial Significance:** This approach enables data-driven optimization of machining processes, empowering more sophisticated decision making within an industrial framework. However, with the aid of multi-objective optimization, it seeks to minimize cutting parameters in a balanced approach which can ameliorate both fatigue life as well as dimensional and tool life consistency. Such a process greatly minimizes the need for expensive repeated experiments providing an affordable, food-compatible, energy-efficient and scalable method for adaptive manufacturing and process automation.
- C. Future Scope:** In future, this framework will be extended with in-situ sensor feedback and adaptive learning mechanisms for real-time monitoring. The proposed approach could be extended to cohere physics-informed models and hybrid FEM–ML architectures which would add the reliability of predictions under a wider regime of operational conditions. Including other factors, such as tool wear effect, chip morphology and thermal deformation effects in the model will extend its applicability for different alloys and machining operations. Embedding the framework within digital twin systems is also envisioned to facilitate closed-loop optimization and self correcting control in next-generation smart manufacturing environments.
- D. Final Remarks:** In summary, The RF–GPR-based predictive framework exhibits high precision, interpretability and computational efficiency in the optimization of machining performance for turning process. The study builds a practical foundation for autonomous machining systems, which is consistent with the development of Industry 4.0 and sustainable intelligent manufacturing.

REFERENCES

- [1] D. Díaz-Salamanca, D. I. L. Pacheco, and F. J. Trujillo, "Turning of 42CrMo4 + QT under different scenarios: Dataset of machining, roughness and residual stress," *Data in Brief*, vol. 54, p. 111445, 2024.
- [2] D. Díaz-Salamanca, "Influence of turning parameters on residual stresses and roughness of 42CrMo4 + QT steel," *Int. J. Adv. Manuf. Technol.*, vol. 130, no. 7–8, pp. 3217–3231, 2024.
- [3] N. E. Sizemore, R. Venkatesh, and M. S. Kumar, "Application of machine learning to the prediction of surface roughness," *Procedia Manuf.*, vol. 48, pp. 890–897, 2020.
- [4] B. A. Beatrice and K. Ramesh, "Surface roughness prediction using artificial neural networks," *Procedia Eng.*, vol. 97, pp. 205–214, 2014.
- [5] J. Chen, Z. Wang, and Y. Li, "Predicting surface roughness in turning complex workpieces using Gaussian process regression," *Sensors*, vol. 24, no. 2, pp. 812–826, 2024.
- [6] N. Masmiati and M. H. El-Axir, "Optimization of cutting conditions for minimum residual stress using response surface methodology," *Int. J. Mach. Tools Manuf.*, vol. 106, pp. 36–45, 2016.
- [7] M. H. El-Axir, "Modeling and parameter optimization for surface roughness in machining processes," *Eng. Technol. Appl. Sci. Res.*, vol. 7, no. 1, pp. 1426–1431, 2017.
- [8] T. Reeber, P. Walker, and C. J. Davis, "A data-driven approach for cutting force prediction in metal cutting," *Manufacturing*, vol. 4, pp. 112–126, 2024.
- [9] Z. Wang, Y. Zhou, and S. Zhang, "Prediction of surface residual stress in multi-axis milling using radial basis function neural networks," *Materials*, vol. 15, no. 4, pp. 2101–2114, 2022.
- [10] S. H. Wu, J. Chen, and H. Xu, "Experimental and machine learning-driven prediction of residual stresses in manufacturing processes," *Materials*, vol. 17, no. 8, pp. 1205–1218, 2024.
- [11] R. Zhou, Y. Zhao, and T. Yang, "Modeling and simulation of residual stress in metal cutting: A computational and machine learning approach," *Int. J. Mech. Sci.*, vol. 250, p. 108092, 2024.
- [12] Y. Zhou, M. Chen, and L. Wang, "Machine learning-based prediction of residual stresses in thermal and laser-assisted processes," *Appl. Surf. Sci.*, vol. 632, p. 157934, 2024.
- [13] B. Peng, L. Zhong, and J. Liu, "Physics-informed neural networks for thermal simulation in manufacturing," *Comput. Methods Appl. Mech. Eng.*, vol. 421, pp. 116754–116768, 2024.
- [14] Z. Yao, X. Zhang, and P. Sun, "Position-dependent process monitoring and surface roughness prediction in milling using machine learning," *CIRP J. Manuf. Sci. Technol.*, vol. 41, pp. 70–82, 2023.
- [15] A. Balasuadhakar, S. Raj, and K. Suresh, "Machine learning prediction of surface roughness in turning: Recent advances," *J. Manuf. Process.*, vol. 96, pp. 312–329, 2025.
- [16] R. P. Shetty, P. Joshi, and A. Rao, "Surface roughness modeling and prediction in high-speed turning using hybrid learning models," *Sci. Rep.*, vol. 15, p. 19284, 2025.
- [17] A. Tzotzis, D. Papazoglou, and C. Alexopoulos, "A dynamic surface roughness prediction system based on machine learning," *J. Manuf. Syst.*, vol. 74, pp. 451–465, 2025.
- [18] M. M. Hasan, "Machining-induced residual stress modeling using physics-based and machine learning methods," Ph.D. dissertation, Univ. of Kentucky, Lexington, KY, USA, 2023.
- [19] F. Farias, R. Almeida, and D. Martins, "FEM and artificial neural network surrogate modeling for thermal prediction in metal cutting," *Addit. Manuf.*, vol. 48, p. 102387, 2021.
- [20] D. Díaz-Salamanca, "MaRoReS: Machining, Roughness and Residual Stresses generated in turning of 42CrMo4 + QT steel," *Mendeley Data*, V1, 2024.

Received: 2025-11-08
Accepted: 2026-03-10
Published Online: 2026-03-30
DOI: 10.70454/IJMRE.2026.60103

International Journal of Multidisciplinary
Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

