November 2026| Volume 05 | Issue 04 | Pages 232-245

Future Technology
Open Access Journal
ISSN 2832-0379

Journal homepage: https://fupubco.com/futech

https://doi.org/10.55670/fpll.futech.5.4.19

Future Publishing LLC

Article

AI-driven multi-objective optimization of CNC machining parameters for enhanced surface finish and tool life

Arvinder Singh Channi[1]*, Altaf Hanifbhai Nalbandh[2], Amol R Madane[3], Vikas Gulati[4], Gadde Raghu Babu[5], Abhijit Chandratreya[6], Deepak Tripathi[7]

[1]Guru Kashi University, Talwandi Sabo, Bathinda (Punjab), India
[2]Government Polytechnic Porbandar, India
[3]Tata Motors Limited, Pune, India
[4]Lovely professional University, Jalandhar, Punjab, India
[5]Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering & Technology, Bachupally, Hyderabad-500118, Telangana, India
[6]Amity L Indira University, Pune, India
[7]Department of CSE-AI&ML, Marwadi University, Rajkot (Gujarat) – 360003, India

ARTICLE INFO

Article history:
Received 29 March 2026
Received in revised form
27 August 2026
Accepted 15 September 2026

Keywords:
CNC machining optimization,
Surface roughness prediction,
Tool-health monitoring, NSGA-II–TOPSIS

*Corresponding author
Email address:
arvindarsinghchanni@gmail.com

DOI: 10.55670/fpll.futech.5.4.19

ABSTRACT

AI-driven CNC machining optimization requires balancing surface quality and productivity while avoiding unsupported integration of independently collected machining data. This study proposes a leakage-controlled, surrogate-assisted framework using two open datasets: a 27-condition turning surface-quality dataset and a 968-observation milling tool-life dataset comprising 14 tools. The datasets were analyzed independently, with no row-wise fusion or DOC–ADOC mapping. Cutting speed, feed rate, and depth of cut were used as controllable turning variables, while milling sensor and metadata features were used separately for remaining useful life (RUL) prediction. The selected Ridge surface model achieved LOOCV MAE = 0.1077, RMSE = 0.1383, and $R^2 = -0.0740$, while nested validation yielded $R^2 = -0.4692$, indicating limited out-of-sample generalization. The no-cycle RUL model achieved MAE = 23.44 cycles, RMSE = 31.39 cycles, and $R^2 = 0.2383$under leave-one-tool-out validation. NSGA-II significantly outperformed equal-budget random search in normalized hypervolume (0.7711 vs. 0.7421). Because the surrogate surface was statistically limited, continuous optimization was treated as exploratory. Final selection was therefore anchored to measured Pareto-efficient turning conditions. Equal-weight TOPSIS selected $V_c = 630$, $f = 0.16$, and $a_p = 0.60$, with measured $R_a = 0.4263$and $MRR_{\text{proxy}} = 60.48$, as the balanced practical recommendation.

1. Introduction

CNC machining parameter selection is a multi-objective decision problem because cutting speed, feed rate, and depth of cut influence surface finish, productivity, machining cost, and process reliability. Recent data-driven approaches have demonstrated the potential of artificial intelligence for machining optimization involving competing quality and productivity objectives [1]. Tool-condition monitoring is also important because tool degradation affects dimensional accuracy, process stability, and downtime [2]. In parallel, surface-roughness prediction has become an important AI application for reducing rework and improving manufacturing quality [3,4]. Despite these advances, several methodological issues remain. Many machining studies address surface roughness, productivity, or tool condition separately, while studies that combine surface quality and tool wear commonly rely on measurements from the same experimental system [5]. In open-data research, however, surface-quality and tool-life datasets may originate from different machines, materials, tools, and machining processes. Direct row-wise fusion of such datasets can therefore imply physical relationships that were never experimentally observed. A further concern is that sensor-derived variables are sometimes treated as optimization inputs even though they are measured during machining and are not controllable at the pre-machining parameter-selection stage.

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
CNC	Computer Numerical Control
CV	Cross-Validation
DOC	Depth of Cut
ADOC	Axial Depth of Cut
GPR	Gaussian Process Regression
HV	Hypervolume
LOOCV	Leave-One-Out Cross-Validation
LOGO	Leave-One-Group-Out
ML	Machine Learning
MRR	Material Removal Rate
MRR_{proxy}	Material Removal Rate Proxy
NSGA-II	Non-dominated Sorting Genetic Algorithm II
OAT	One-at-a-Time
OOD	Out-of-Distribution
RF	Random Forest
RUL	Remaining Useful Life
SHAP	SHapley Additive exPlanations
SVR	Support Vector Regression
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution

To address these issues, this study develops a leakage-controlled surrogate-assisted CNC decision-support framework using two independent open datasets. The turning dataset is used for surface-roughness prediction and productivity optimization, whereas the milling dataset is analyzed independently for remaining useful life (RUL) and tool-life monitoring. No row-wise dataset fusion or DOC–ADOC mapping is performed, and milling-derived tool-health quantities are not imposed as turning optimization objectives or constraints. NSGA-II explores the trade-off between predicted surface roughness and the productivity proxy, while TOPSIS selects the compromise solution [6]. Because strict validation showed limited generalization of the surface surrogate, continuous optimization is treated as exploratory, and the final practical recommendation is anchored to experimentally observed Pareto-efficient turning conditions. This revised framing is consistent with the updated Methodology and Results sections. Compared with previous multi-objective CNC studies [1,5,7,8], the present framework introduces four distinguishing features: the use of two independent open machining datasets comprising 27 turning conditions and 968 milling observations from 14 tools; explicit avoidance of observation-level cross-dataset fusion; restriction of optimization to controllable pre-machining variables; and reproducible evaluation of NSGA-II over 30 seeds against an equal-budget random-search baseline of 20,000 evaluations per seed. The objectives of this study are:

- To develop leakage-controlled AI models for surface roughness prediction and independent RUL-based tool-life monitoring using open machining datasets.
- To generate dataset-bounded Pareto solutions using NSGA-II with only controllable turning variables: cutting speed, feed rate, and depth of cut.
- To identify a practical machining recommendation using measured Pareto alternatives, TOPSIS weight-sensitivity analysis, robustness assessment, and experimental design-support verification.

2. Literature review

Tool-life and remaining useful life prediction are central to intelligent machining because tool degradation affects dimensional accuracy, surface integrity, downtime, and production cost. In-process tool-wear prediction using machine-learning techniques and force analysis has shown that monitored force-related signals can provide useful information about tool condition during machining [9]. Broader studies on milling tool-condition monitoring have further identified vibration, current, force, acoustic emission, and related sensor features as common indirect indicators of tool wear and degradation [10]. In turning operations, indirect tool-condition monitoring and decision-making methods have also been critically reviewed, emphasizing the need to robustly interpret sensor information before using it for machining decisions [11]. Surface roughness prediction is another major research direction because Ra, Rq, and Rz directly represent machined-part quality. Roughness is influenced by feed rate, cutting speed, depth of cut, vibration, force, tool geometry, and material response. In ideal turning, the geometric roughness tendency is often associated with R_a≈f^2/(32r_e), where r_e is the tool nose radius. However, real machining responses are nonlinear and sensor-dependent. Therefore, machine-learning models using sensor signals have been applied to improve roughness prediction beyond simple analytical approximations [12].

CNC machining parameter optimization is naturally multi-objective because surface finish, tool life, productivity, and cost often conflict. Evolutionary algorithms are widely used because they can search nonlinear decision spaces without requiring explicit analytical gradients. Recent work on micro-milling demonstrated that improved NSGA-II can effectively optimize cutting parameters for competing machining objectives [13]. Similarly, improved NSGA-II has been applied in hard turning to optimize process parameters under multiple performance criteria [14]. Existing studies confirm the value of sensor-based monitoring for tool wear and degradation using force, vibration, current, and related signals [9-11]. Surface-roughness prediction has also benefited from machine-learning and sensor-based approaches beyond simplified analytical relations [12]. Similarly, NSGA-II has been successfully applied to multi-objective machining optimization [7,8]. However, a key limitation remains when independently collected machining datasets are combined without preserving differences in machine, process, material, and variable meaning. Sensor-derived features may also be incorrectly treated as pre-machining decision variables. Therefore, a framework is needed that keeps independent datasets separate, restricts optimization to controllable variables, and evaluates predictive uncertainty before recommending parameters. The present study addresses this gap through leakage-controlled modeling, independent turning and milling branches, NSGA-II–TOPSIS decision support, and experimental design-support assessment. Although these studies demonstrate progress in tool-condition monitoring, roughness prediction, and multi-objective optimization, several gaps remain. Existing approaches often focus on either surface finish or tool life separately, use sensor variables as if they were pre-machining decision variables, or omit robustness and out-of-distribution screening. This study addresses these gaps by developing a leakage-controlled surrogate-assisted framework that combines independent open datasets, explainable AI, NSGA-II optimization, TOPSIS ranking, robustness testing, and Mahalanobis OOD screening.

3. Methodology

This study presents a software-only, surrogate-assisted framework for CNC machining optimization using two independent open datasets. It includes preprocessing, feature extraction, leakage-controlled modeling, explainable AI, NSGA-II optimization, TOPSIS decision-making, robustness analysis, and experimental design-support assessment. The turning and milling datasets were modeled independently and were not physically merged or coupled within the optimization problem.

3.1 Overall research framework

The framework included data acquisition, preprocessing, predictive modeling, explainability, Pareto optimization, TOPSIS ranking, and robustness assessment. The turning dataset was used for surface-roughness and productivity optimization, whereas the milling dataset was used independently for RUL and tool-life monitoring. Tool-health predictions were not imposed as constraints on the turning optimizer. Because the surface surrogate showed limited generalization, continuous NSGA-II solutions were treated as exploratory, while the final practical recommendation was selected from experimentally observed Pareto-optimal turning conditions. Table 1 summarizes the complete experimental and reproducibility configuration used in this study. The complete computational workflow, from independent dataset acquisition to measured Pareto-based final parameter selection, is illustrated in Figure 1.

3.2 Dataset description and acquisition

Two open machining datasets were used: a milling tool-life dataset [13] and a turning surface-roughness dataset [14]. The final analysis contained 968 milling observations from 14 tools and 27 turning conditions, forming a complete 3×3×3design. The datasets were obtained from different machining processes and experimental settings. Therefore, no row-wise merging or turning-DOC-to-milling-ADOC mapping was performed.

3.3 Data preprocessing and feature extraction

For the turning dataset, R_a was the primary response, while speed, feed, and depth of cut were retained as controllable predictors. For the milling dataset, RUL was used as the primary prognostic target. The normalized health/life index and life-stage classes were retained only as auxiliary outcomes. Tool cycle index was excluded from the primary RUL model and used only for ablation analysis.

3.4 Independent dataset handling and safe integration

The two datasets were not harmonized at the observation or objective-function level. Instead, they formed two independent analytical branches: turning for R_a–MRR_{proxy} decision support and milling for RUL/tool-life monitoring. Milling RUL or health-index predictions were therefore not used as turning optimization objectives or feasibility constraints.

3.5 Feature selection and leakage control

Speed, feed, and depth of cut were the only optimization variables. Sensor-derived features were used only for prediction, monitoring, or explainability. Feature selection was performed inside the training folds to avoid leakage. For RUL, the primary configuration used sensor and milling metadata without cycle information.

3.6 AI model development and validation

Surface models included Dummy Mean, Ridge, response-surface/Ridge, SVR, GPR, tree-based models, and gradient boosting. LOOCV, repeated five-fold cross-validation, nested model-selection validation, and 1,000 bootstrap resamples were used to quantify generalization and uncertainty because the surface dataset contained only 27 conditions. Grouped bootstrap resampling with 1,000 repetitions was additionally used to quantify uncertainty in unseen-tool performance. For RUL prediction, nested feature selection with Extra Trees was evaluated using outer leave-one-tool-out validation and inner grouped cross-validation. Cycle-only and cycle-included models were retained as ablations.

3.7 Explainable AI, optimization, and decision-making

Grouped permutation importance was used as the primary RUL explainability method, with SHAP retained as supplementary descriptive interpretation. NSGA-II was used to minimize predicted R_aand maximize MRR_proxy=V_c fa_p. Tool health was excluded from the optimization objectives and constraints. The final NSGA-II configuration used a population size of 100, 200 generations, 30 independent seeds, and 20,000 evaluations per seed. Random search received the same 20,000-evaluation budget per seed and was evaluated independently rather than pooled with the NSGA-II Pareto candidates. Hypervolume was calculated in normalized two-objective minimization space using the fixed reference point [1.1,1.1]. TOPSIS used min–max normalization, with R_atreated as a cost criterion and MRR_proxy as a benefit criterion. Weight sensitivity was evaluated for w_(R_a)=0.05,0.10,...,0.95, with w_MRR=1-w_(R_a), and ranking stability was assessed using Spearman and Kendall correlations. TOPSIS was also applied directly to the measured Pareto alternatives for final practical selection.

3.8 Robustness, data-support assessment, and reproducibility

Robustness was examined using bootstrap uncertainty, repeated optimizer runs, equal-budget random-search comparison, TOPSIS weight sensitivity, and one-at-a-time parameter perturbation. Experimental support was assessed primarily using the complete factorial design and nearest-observed-condition distance. The normalized factorial coverage radius was 0.481611, and candidates with a nearest-observation distance not exceeding this value were classified as interpolation-supported.

Table 1. Dataset structure, variables, model settings, robustness tests, and reproducibility parameters used in the study

Category	Item	Final configuration	Interpretation
Dataset 1	Milling tool-life dataset [13]	968 observations; 14 tools	Independent RUL/tool-life analysis
Dataset 2	Turning surface dataset [14]	27 conditions; 3 × 3 × 3factorial design	Surface prediction and parameter optimization
Dataset integration	Cross-dataset relation	No row-wise merge; no DOC–ADOC mapping	Turning and milling analyzed independently
Surface target	Primary response	R_a	Surface-quality objective
Tool-life target	Primary response	RUL	Independent prognostic target
Decision vector	$x = [V_c, f, a_p]$	Cutting speed, feed rate, depth of cut	Controllable variables only
Bounds	Search space	$V_c = 400–1000$; $f = 0.05–0.16$; $a_p = 0.20–0.60$	Dataset-bounded optimization
Surface validation	Validation protocol	LOOCV, repeated 5-fold CV, nested validation	Controls overfitting for $n = 27$
RUL validation	Validation protocol	Outer LOGO + inner grouped CV	Unseen-tool evaluation
RUL predictors	Primary configuration	Sensor + milling metadata; cycle excluded	Cycle retained only for ablation
Optimizer	NSGA-II	Population = 100; generations = 200; 30 seeds	Final executed settings
Computational budget	NSGA-II vs. random search	20,000 evaluations/seed each	Fair equal-budget comparison
Objectives	Main formulation	Minimize predicted R_a; maximize MRR_{proxy}	No HI constraint
Productivity	Relative productivity proxy	$MRR_{\text{proxy}} = V_c f a_p$	Not dimensional turning MRR
Hypervolume	Optimizer performance	Normalized exact 2-D HV; reference point = [1.1, 1.1]	Consistent optimizer comparison
TOPSIS	Decision analysis	Min–max normalization; $w_{R_a} = 0.05–0.95$, $w_{MRR} = 1 - w_{R_a}$	Weight-sensitivity analysis
Rank stability	Statistics	Spearman and Kendall correlations	Quantifies ranking sensitivity
Support assessment	Primary method	Factorial design support + nearest-observed distance; coverage radius = 0.481611	Interpolation-domain verification
Mahalanobis	Secondary method	Standardized V_c, f, a_p+ Ledoit–Wolf covariance	Centrality diagnostic only; no rejection threshold
Practical recommendation	Observed Pareto solution	$V_c = 630$, $f = 0.16$, $a_p = 0.60$; measured $R_a = 0.4263$; $MRR_{\text{proxy}} = 60.48$	Equal-weight TOPSIS Rank 1
Robustness	OAT analysis	±5% and ±10% perturbations	Clipping at bounds reported explicitly
Uncertainty	Bootstrap analysis	1,000 metric/group resamples; 500 final-model refits	95% uncertainty assessment

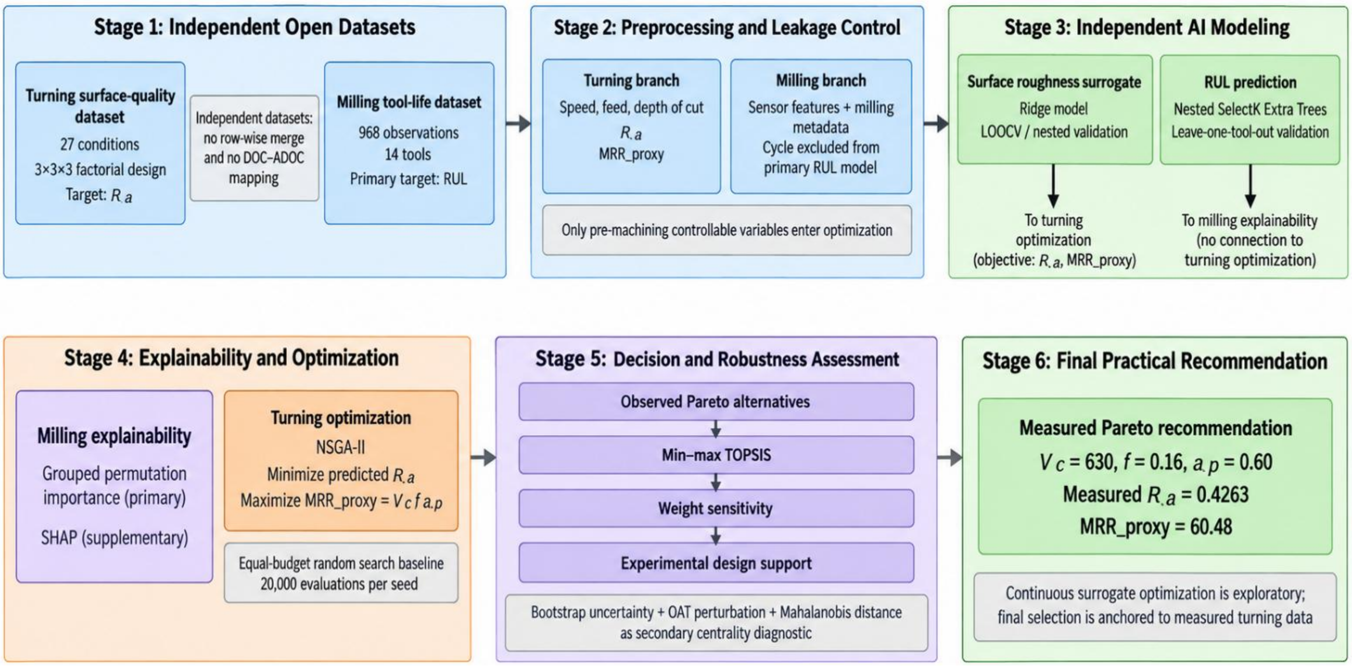

Figure 1. Overall workflow of the proposed AI-driven surrogate-assisted CNC machining optimization framework

Mahalanobis distance was retained only as a secondary centrality diagnostic using standardized speed, feed, and depth of cut with Ledoit–Wolf covariance estimation. The previous Mahalanobis threshold of 2.137470 and the 0.95 safety-margin rule were removed. For OAT analysis, requested and realized perturbations were reported separately when parameter changes were clipped at experimental bounds. These values are reported in Table 1 so that the computational study can be reproduced or extended under the same assumptions.

4. Mathematical Modeling

This section formulates the computational basis of the proposed CNC optimization framework. It defines the machining decision variables, surface roughness surrogate, RUL and auxiliary normalized-life models, productivity proxy, multi-objective optimization problem, Pareto dominance, TOPSIS ranking, and data-support assessment. All equations directly support the implemented workflow and ensure dataset-bounded and leakage-controlled parameter selection. Because the surface-finish and tool-life datasets were obtained from independent turning and milling processes, the tool-life model was analyzed separately and was not imposed as an optimization objective or constraint.

4.1 Legal admissibility and procedural safeguards

The optimization problem was formulated using only controllable machining parameters available before machining. This leakage-controlled design ensured that the optimizer did not use sensor responses or target-derived variables as direct decision variables. The machining decision vector was defined as:

$$x = \left[V_c, f, a_p\right]^T \tag{1}$$

where V_c is cutting speed, f is feed rate, and a_p is depth of cut. Based on the observed surface-finish dataset, the search space was restricted to:

$$\begin{aligned} 400 &\le V_c \le 1000 \\ 0.05 &\le f \le 0.16 \\ 0.20 &\le a_p \le 0.60 \end{aligned} \tag{2}$$

These bounds were used to prevent extrapolation beyond the available data-supported machining region.

4.2 Surface roughness surrogate model

The primary surface-quality response was average surface roughness, R_a. A parameter-only AI surrogate was used for optimization because it depends only on pre-machining controllable variables. The surrogate mapping was expressed as:

$$\widehat{R_a} = g(x) \tag{3}$$

where $g(\cdot)$ denotes the trained surface roughness prediction model. The surface-quality objective was therefore formulated as:

$$f_1(x) = \widehat{R_a}(x) \tag{4}$$

Although sensor-enhanced models were trained for comparison and explainability, they were not used as optimization decision models because vibration, force, and other sensor features are measured during machining and are not direct pre-machining control variables. Because the validated surface surrogate showed limited generalization, continuous optimization was treated as exploratory, while the final practical recommendation was additionally supported using experimentally observed Pareto alternatives.

4.3 Tool-health and remaining useful life model

The milling tool-life dataset was processed at the tool-cycle level to estimate remaining useful life. For tool t at cycle i, RUL was defined as:

$$RUL_{i,t} = N_{\mathrm{end},t} - N_{i,t} \tag{5}$$

where $N_{\mathrm{end},t}$ is the terminal recorded cycle for tool t, and $N_{i,t}$ is the current cycle. A normalized remaining-life index was defined as:

$$HI_{i,t} = \frac{RUL_{i,t}}{\max(RUL_t)} \tag{6}$$

The primary prognostic model was expressed as:

$$\widehat{RUL} = h(z) \tag{7}$$

where $h(\cdot)$ denotes the trained RUL model and z contains milling sensor and metadata features. Cycle progression was excluded from the primary RUL model and retained only for ablation analysis. The normalized HI was treated as an auxiliary remaining-life indicator rather than an independent physical wear measurement.
Since the milling and turning datasets were independently collected, neither HI nor RUL was used as a feasibility constraint in the turning optimization.

4.4 Material removal rate

Productivity was represented using a material-removal-rate proxy, because complete workpiece-geometry information required for dimensional turning MRR was unavailable. The proxy was computed as:

$$MRR_{\mathrm{proxy}} = V_c f a_p \tag{8}$$

Because a higher value represents greater relative productivity,

$$\max MRR_{\mathrm{proxy}} \tag{9}$$

For compatibility with minimization-based multi-objective optimization, this was expressed as:

$$f_2(x) = -MRR_{\mathrm{proxy}}(x) \tag{10}$$

Thus, MRR_{proxy} was interpreted as a relative productivity indicator rather than absolute volumetric turning MRR because complete workpiece-geometry information was unavailable.

4.5 Multi-objective optimization formulation

The main optimization formulation treated surface quality and productivity as the two active objectives. Tool health was not included as a feasibility constraint because it was obtained from an independent milling dataset. The optimization problem was therefore defined as:

$$\min_x F(x) = \left[\widehat{R_a}, (x) - MRR_{\mathrm{proxy}}(x)\right] \tag{11}$$

subject to

$$x_{\min} \le x \le x_{\max} \tag{12}$$

where $x_{\min}$ and $x_{\max}$ denote the lower and upper bounds defined in Eq. (2). HI was excluded from the optimization objectives and constraints because the independent milling and turning datasets do not provide experimentally supported cross-process coupling between tool health and the turning decision variables.

4.6 NSGA-II and Pareto dominance

The multi-objective search was solved using NSGA-II. A candidate solution x^A was considered to dominate another candidate x^B if it was no worse in all objectives.

$$f_k(x^A) \le f_k(x^B), \forall k \tag{15}$$

and strictly better in at least one objective.

$$\exists k: f_k(x^A) < f_k(x^B) \tag{16}$$

NSGA-II ranked solutions using non-dominated sorting, crowding-distance preservation, and elitist selection. The resulting non-dominated solutions represented the estimated trade-off between predicted surface roughness and MRR_{proxy}.

The final implementation used a population size of 100, 200 generations, and 30 independent seeds, corresponding to 20,000 objective evaluations per seed. Random search used the same 20,000-evaluation budget per seed and was evaluated independently as a baseline. Hypervolume was calculated in normalized two-objective minimization space using the fixed reference point [1.1, 1.1]. A paired one-sided Wilcoxon signed-rank test across the 30 matched seeds was used to assess whether NSGA-II achieved higher hypervolume than equal-budget random search.

4.7 TOPSIS ranking

After Pareto optimization, TOPSIS was applied to identify practical compromise solutions. The final implementation used min-max normalization. For the cost criterion R_a,

$$r_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \tag{15}$$

whereas for the benefit criterion MRR_{proxy},

$$r_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \tag{16}$$

The weighted normalized matrix was calculated as:

$$v_{ij} = w_j r_{ij} \tag{17}$$

The ideal best and worst solutions were defined as:

$$A^+ = \left\{\max_i v_{ij}\right\}, A^- = \left\{\min_i v_{ij}\right\} \tag{18}$$

The separation distances were

$$S_i^+ = \sqrt{\sum_j (v_{ij} - A_j^+)^2},\ S_i^- = \sqrt{\sum_j (v_{ij} - A_j^-)^2} \tag{19}$$

The TOPSIS closeness coefficient was calculated as:

$$C_i = \frac{S_i^-}{S_i^+ + S_i^-} \tag{20}$$

A larger C_i represents a more preferred alternative. R_a was treated as the cost criterion and MRR_{proxy} as the benefit criterion. Equal weighting used 0.50/0.50, while weight sensitivity was evaluated for $w_{R_a} = 0.05$ to 0.95 in increments of 0.05. Spearman and Kendall rank correlations were used to quantify ranking sensitivity.

TOPSIS was also applied directly to the experimentally observed Pareto alternatives using measured R_a, which formed the basis of the final practical recommendation.

4.8 Data-support and Mahalanobis assessment

Because continuous optimization can generate intermediate parameter combinations, experimental design support was used as the primary criterion for verifying interpolation validity. Each decision variable was normalized within its observed range, and the distance to the nearest experimental condition was calculated as:

$$d_{\min}(x) = \min_j \|\tilde{x} - \tilde{x}^{(j)}\|_2 \tag{21}$$

For the complete factorial design, the corresponding coverage radius was

$$d_{\text{support}} = 0.481611 \tag{22}$$

Candidates within the experimental bounds and within this coverage radius were considered interpolation supported.

Mahalanobis distance was additionally calculated as:

$$D_M(z) = \sqrt{(z-\mu)^T \Sigma^{-1} (z-\mu)} \tag{23}$$

where z contains the standardized cutting speed, feed rate, and depth of cut, and Σ was estimated using the Ledoit-Wolf covariance estimator.

Mahalanobis distance was retained only as a secondary centrality diagnostic. The previous threshold of 2.137470 and the 0.95 safety-margin rejection rule were therefore removed from the final formulation. Final candidate support was determined by the experimentally defined interpolation-support criterion rather than an arbitrary Mahalanobis rejection threshold.

5. Results

5.1 Dataset and preprocessing results

The preprocessing stage produced two independent analytical spaces. The turning surface-quality dataset contained 27 complete machining conditions, representing a $3 \times 3 \times 3$ factorial combination of cutting speed, feed rate, and depth of cut. R_a was retained as the primary surface-quality response. The milling tool-life dataset contained 968 observations from 14 tools, with RUL treated as the primary prognostic target. The normalized remaining-life index and life-stage classes were retained as auxiliary outputs. No row-wise integration of the turning and milling observations was performed, and turning depth of cut was not mapped to milling ADOC. Thus, surface-quality optimization and milling tool-life prediction remained independent analytical branches. The resulting dataset structure, primary targets, grouping strategy, and analytical roles are summarized in Table 2.

5.2 Baseline machining analysis

Baseline analysis was conducted to examine the relationship between controllable turning parameters and surface roughness. The correlations were generally weak. Depth of cut showed the largest Pearson association with $R_a (r = 0.3707)$, followed by $MRR_{\text{proxy}} (r = 0.3315)$. However, none of the tested associations reached statistical significance at $p < 0.05$. The Pearson coefficients, 95% confidence intervals, p-values, and Spearman correlations for the controllable variables are reported in Table 3. These results also provide an empirical check on the classical turning tendency $R_a \propto f^2/(32 r_e)$. Although the analytical expression suggests a feed-squared relationship, the present dataset did not show statistically significant evidence for this tendency ($r = 0.0721$, $p = 0.7210$). The absence of tool nose-radius information prevented a complete dimensional comparison.

Table 2. Dataset structure used for final model development

Dataset	Observations	Primary target	Group structure	Main use
Turning surface-quality	27	R_a	$3 \times 3 \times 3$ machining conditions	Surface prediction and parameter decision support
Milling tool-life	968	RUL	14 tools	Independent tool-life prognostics

Table 3. Statistical relationships between controllable variables and R_a

Variable	Pearson r	95% CI	p-value	Spearman ρ	Significant
Speed	0.1509	−0.2431 to 0.5021	0.4525	0.2562	No
Feed	0.0606	−0.3269 to 0.4307	0.7639	−0.0349	No
Feed^2	0.0721	−0.3166 to 0.4400	0.7210	−0.0349	No
Depth of cut	0.3707	−0.0108 to 0.6580	0.0570	0.2679	No
MRR_{proxy}	0.3315	−0.0555 to 0.6319	0.0912	0.2315	No

Table 4. Surface-roughness performance under strict validation

Assessment	Model/procedure	MAE	RMSE	R^2	Interpretation
Fixed parameter model	Ridge	0.1077	0.1383	−0.0740	Leakage-controlled parameter model
Nested model selection	Outer LOOCV + inner CV	0.1213	0.1617	−0.4692	Conservative generalization estimate
Reference baseline	DummyMean	0.1071	0.1385	−0.0784	Mean-response reference

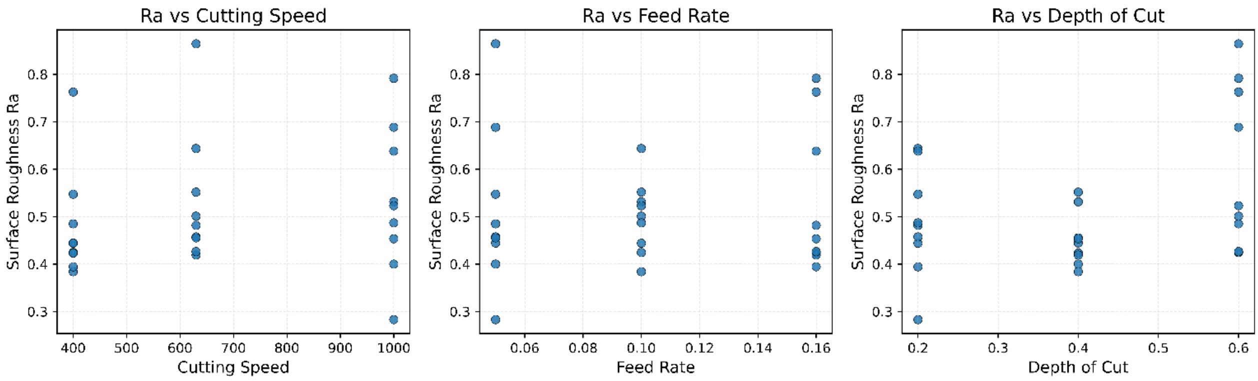

Figure 2. Relationship between measured R_a and cutting speed, feed rate, and depth of cut

The measured surface-roughness responses across cutting speed, feed rate, and depth of cut are visualized in Figure 2.

5.3 Surface roughness model results

Surface roughness was evaluated using parameter-only and comparative models under strict cross-validation. The selected fixed parameter-only Ridge model produced LOOCV MAE = 0.1077, RMSE = 0.1383, and $R^2 = -0.0740$. The nested model-selection procedure, which additionally accounts for model-selection uncertainty, produced MAE = 0.1213, RMSE = 0.1617, and $R^2 = -0.4692$.

For reference, the DummyMean predictor produced MAE = 0.1071, RMSE = 0.1385, and $R^2 = -0.0784$. The comparative performance of the selected surface model, nested model-selection procedure, and mean-response baseline is summarized in Table 4. Bootstrap assessment of the fixed Ridge model produced an RMSE interval of approximately 0.0957–0.1791 and an R^2interval of approximately −0.4498 to −0.0441. The nested procedure did not significantly reduce prediction error relative to the mean predictor ($p = 0.8901$).

The negative R^2values indicate that the limited 27-condition dataset does not support strong out-of-sample prediction of $R_a$from the three controllable parameters alone. Importantly, this result was used to define the appropriate scope of the model rather than overstate its accuracy**.** Continuous surrogate optimization was therefore retained only as an exploratory screening mechanism. To limit dependence on the weak surrogate, final parameter selection was subsequently anchored to experimentally observed Pareto-efficient turning conditions using measured R_a. Thus, the practical recommendation does not rely solely on the continuous model prediction. The relationship between observed and LOOCV-predicted surface roughness is illustrated in Figure 3.

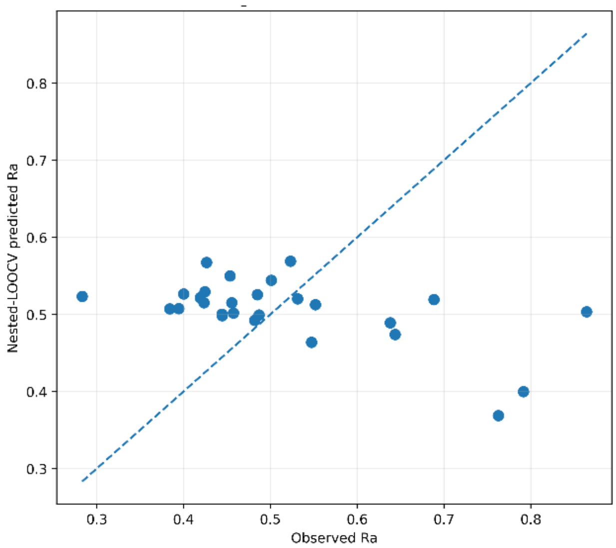

Figure 3. Observed versus LOOCV-predicted $R_a$for the selected parameter-only surface model

5.4 Tool-life model results

RUL prediction was treated as the primary prognostic task for the milling dataset. The final model combined sensor information with milling metadata while excluding cycle progression from the primary predictor set. Nested feature selection with Extra Trees was evaluated using leave-one-tool-out validation. The model achieved MAE = 23.44 cycles, RMSE = 31.39 cycles, $R^2 = 0.2383$, and NRMSE = 21.07%**.** The grouped-bootstrap 95% interval for $R^2$was approximately **−**0.0599 to 0.5660**,** demonstrating positive pooled predictive skill but also substantial between-tool heterogeneity.

The auxiliary normalized remaining-life index achieved $R^2 = 0.5083$. Because this variable is derived from normalized RUL, it was interpreted as an auxiliary prognostic representation rather than an independent physical wear measurement. The primary RUL, normalized-life, life-stage, and cycle-ablation results are summarized in Table 5.

The no-cycle sensor model also produced significantly lower per-tool RUL error than the cycle-only comparator ($p = 0.0101$), demonstrating that the sensor information contributed prognostic value beyond simple life progression. Adding cycle information to the sensor configuration did not provide a significant additional reduction in error. Accordingly, the RUL model is interpreted as demonstrating moderate pooled unseen-tool prognostic capability rather than universal tool-life prediction. This interpretation is consistent with the strict LOGO evaluation and the observed between-tool uncertainty. The observed and leave-one-tool-out predicted RUL values are compared in Figure 4. The class-wise prediction pattern of the no-cycle normalized life-stage model is shown by the confusion matrix in Figure 5.

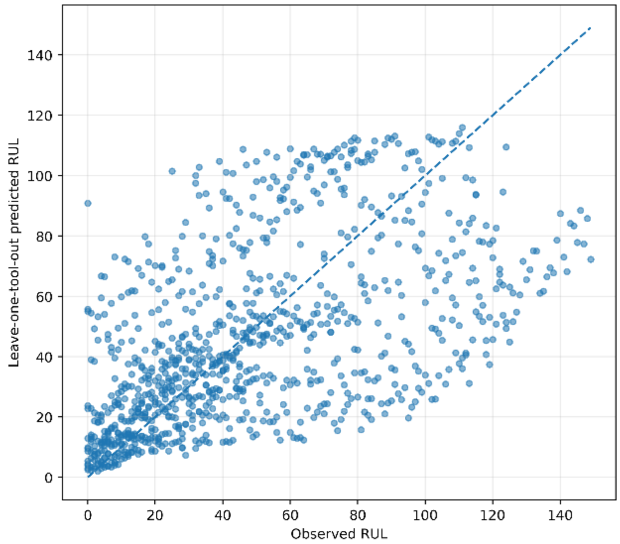

Figure 4. Actual versus predicted RUL under leave-one-tool-out validation

Table 5. Final RUL and auxiliary life-stage performance

Task	Feature configuration	Validation	Performance	Interpretation
Primary RUL regression	Sensor + metadata, no cycle	LOGO	MAE = 23.44; RMSE = 31.39; $R^2 = 0.2383$	Moderate unseen-tool predictive skill
Normalized-life index	Sensor + metadata, no cycle	LOGO	$R^2 = 0.5083$; RMSE = 0.2045	Auxiliary normalized-life prediction
Life-stage classification	Sensor + metadata, no cycle	LOGO	Accuracy = 55.06%; balanced accuracy = 56.61%; macro-F1 = 0.5576	Secondary life-stage analysis
Cycle-included sensitivity	Sensor + metadata + cycle	LOGO	Accuracy = 69.01%; balanced accuracy = 70.40%; macro-F1 = 0.6979	Ablation only

5.5 Explainable AI results

Explainability analysis was focused on the primary RUL model. Grouped permutation importance under held-out-tool evaluation was used as the principal interpretation method because it remains consistent with the leakage-controlled validation design. SHAP analysis of the full-data model was retained as supplementary descriptive evidence. The permutation analysis showed that current-related features, particularly driving-axis current mean and variability descriptors, were among the most influential predictors of unseen-tool RUL. This indicates that electrical-load behavior contained useful degradation information beyond simple cycle progression. The overlap between grouped permutation importance and descriptive SHAP rankings further indicated that several current-related variables were consistently identified by both methods. However, SHAP rankings were interpreted descriptively rather than as independent validation evidence. Grouped permutation importance identified current-related descriptors as the dominant contributors to unseen-tool RUL prediction. The six highest-ranked features are summarized in Table 6.

The three highest-ranked features are mean current measurements from the driving X-axis across the three electrical phases, indicating that electrical-load behavior provided the strongest permutation-based prognostic information. Figure 6 presents the grouped permutation-importance ranking, while Figure 7 provides supplementary SHAP-based interpretation of the final no-cycle RUL model.

5.6 Pareto optimization results

NSGA-II optimization was conducted within the turning parameter bounds using a population of 100, 200 generations, and 30 independent seeds, resulting in 20,000 objective evaluations per seed. Random search was evaluated independently using exactly the same number of evaluations per seed. Unlike the previous implementation, random-search candidates were not pooled with the NSGA-II candidates for Pareto ranking or TOPSIS selection. Random search served only as an equal-budget benchmarking procedure. The median exact normalized two-dimensional hypervolume was

$$HV_{\text{NSGA-II}} = 0.77110$$

compared to

$$HV_{\text{Random}} = 0.74212.$$

The median difference was 0.02897, and a paired one-sided Wilcoxon test confirmed significantly higher hypervolume for NSGA-II ($p = 9.31 \times 10^{-10}$). The equal-budget performance comparison between NSGA-II and random search is summarized in Table 7. The result demonstrates that NSGA-II searched the fitted roughness–productivity objective space more effectively than equal-budget random sampling.

Table 6. Leading grouped permutation-importance features for RUL prediction should be reported as ranked features directly

Rank	Feature	Mean RMSE increase	Positive folds
1	Driving-axis X current mean, L1	1.6930	78.6%
2	Driving-axis X current mean, L2	1.6921	92.9%
3	Driving-axis X current mean, L3	1.3828	71.4%
4	Driving-axis X current std., L2	0.5445	71.4%
5	Spindle current mean, L2	0.5175	57.1%
6	Driving-axis X current std., L1	0.4290	50.0%

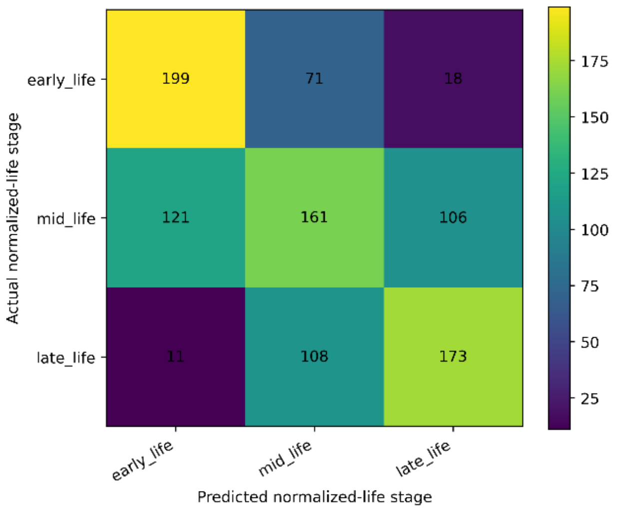

Figure 5. Confusion matrix for the three-class no-cycle normalized life-stage model

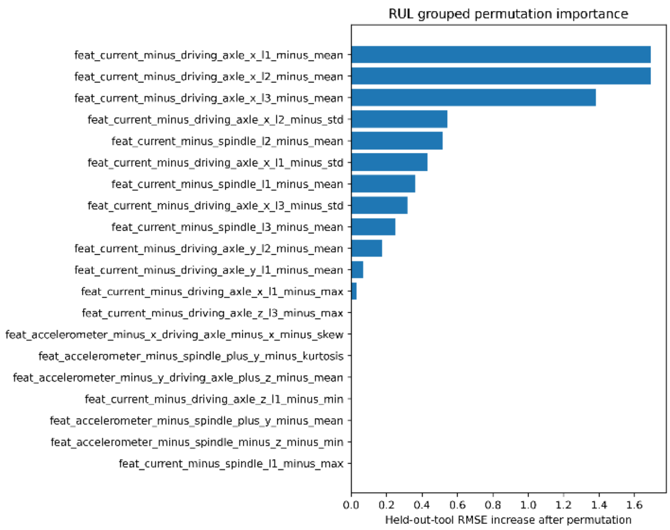

Figure 6. Grouped permutation importance for the primary no-cycle RUL model

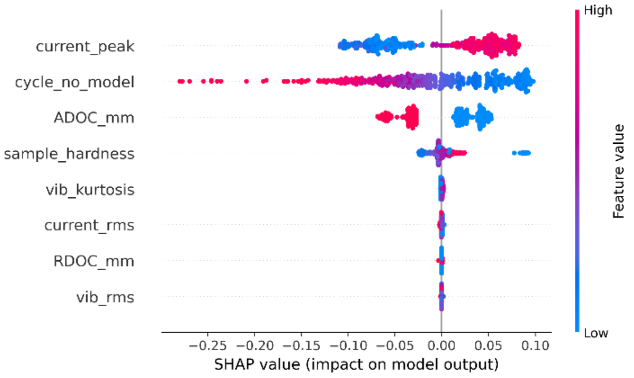

Figure 7. SHAP summary for the final full-data no-cycle RUL model

It should not be interpreted as independent physical validation of the continuous machining settings because the surface surrogate itself showed limited predictive generalization. The fitted roughness–productivity trade-off and the seed-wise hypervolume comparison are presented in Figure 8.

5.7 TOPSIS and final parameter selection

Because the surface surrogate showed limited out-of-sample predictive capability, practical parameter selection was based on the experimentally observed Pareto-efficient turning alternatives rather than relying solely on a continuous surrogate optimum. Under equal TOPSIS weighting, $w_{R_a} = 0.50, w_{MRR} = 0.50$, the highest-ranked measured condition was $V_c = 630$, $f = 0.16$, $a_p = 0.60$, with R_a=*0.4263*, and $MRR_{\text{proxy}} = 60.48$.

Its TOPSIS score was approximately 0.6498, giving it Rank 1 under the balanced equal-weight scenario. The TOPSIS weight sweep showed that candidate preference changed according to the relative importance assigned to surface quality and productivity. The selected balanced recommendation achieved Rank 1 in 4 of the 19 weighting scenarios (21.1%) and remained within the top three in 16 of 19 scenarios (84.2%). Thus, the recommendation should not be described as universally invariant to weighting. Instead, it represents a balanced compromise that remains highly ranked over most of the examined preference range.

Strong productivity weighting favored the maximum-productivity conditions, whereas strong surface-quality weighting shifted the preference toward low-R_a alternatives. Spearman and Kendall rank analyses similarly confirmed that the ordering of candidates changed as criterion priorities changed. The measured-versus-surrogate audit further justified the empirical selection strategy. At the selected condition, measured $R_a = 0.4263$, whereas the fitted surrogate predicted approximately 0.5292, corresponding to an overprediction of approximately 24.1%. At the experimentally observed maximum-productivity condition $(1000, 0.16, 0.60)$, measured $R_a = 0.7917$, whereas the surrogate predicted approximately 0.5357, corresponding to an underprediction of approximately 32.3%.

These discrepancies provide quantitative evidence for treating continuous surrogate solutions as exploratory and using measured Pareto observations for the final practical recommendation. The equal-weight TOPSIS ranking of the seven experimentally observed Pareto-efficient turning alternatives is reported in Table 8. With $w_{R_a} = w_{MRR} = 0.50$, the condition $V_c = 630$, $f = 0.16$, and $a_p = 0.60$achieved the highest TOPSIS score of 0.6498 and was therefore selected as the balanced practical recommendation. Complete rankings across all 19 weighting scenarios are provided in Table 8. Figure 9 demonstrates multi-panel final decision analysis.

5.8 Robustness and reliability results

Robustness assessment included surface-model uncertainty, RUL uncertainty, optimizer repeatability, TOPSIS sensitivity, local perturbation analysis, and experimental-domain support. The primary design-support analysis showed that all 1,252 pooled NSGA-II Pareto candidates remained inside the interpolation-supported factorial region. The maximum normalized nearest-observation distance was 0.3159, which was below the factorial coverage radius of 0.4816. Consequently, the support assessment accepted 1,252 of 1,252 candidates, with no candidate rejected. This result indicates that the search remained within the experimentally covered interpolation region. The support test should therefore be interpreted as confirmatory rather than selective. Mahalanobis distance was retained as a secondary measure of multivariate centrality.

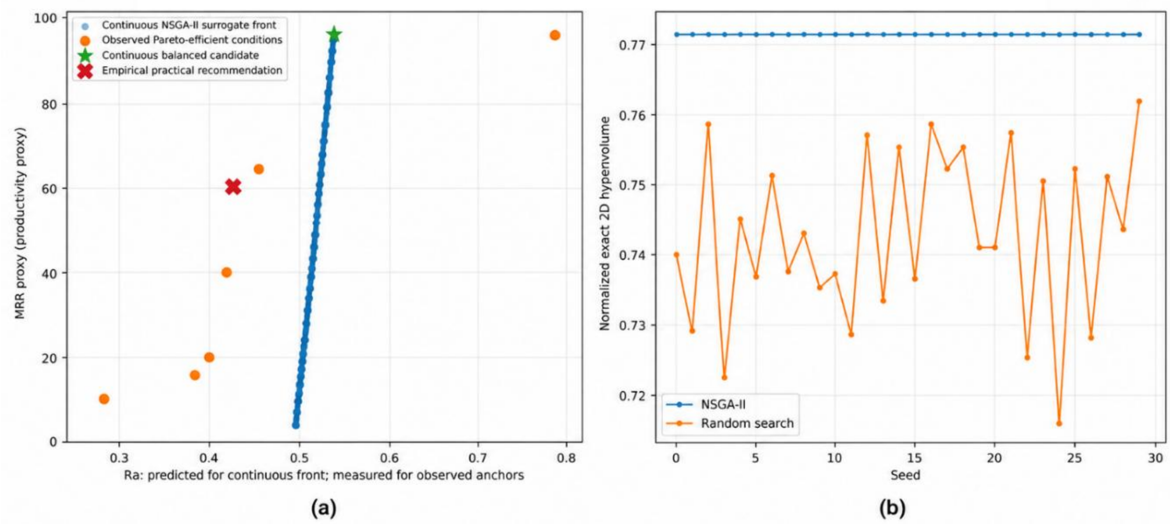

Figure 8. Multi-panel optimization performance: (a) predicted R_a–MRR_{proxy} tradeoff and final candidate structure; (b) equal-budget hypervolume comparison between NSGA-II and random search

Table 7. NSGA-II versus equal-budget random-search performance

Measure	NSGA-II	Random search
Seeds	30	30
Evaluations per seed	20,000	20,000
Median hypervolume	0.77110	0.74212
Hypervolume difference	0.02897	—
Paired Wilcoxon p	9.31×10^{-10}	—
Pooled NSGA-II nondominated candidates	1,252	—

Table 8. Equal-weight TOPSIS ranking of experimentally observed Pareto-efficient turning alternatives

Rank	V_c	f	a_p	Measured R_a	MRR_{proxy}	TOPSIS score
1	630	0.16	0.60	0.4263	60.48	0.6498
2	1000	0.16	0.40	0.4533	64.00	0.6463
3	630	0.16	0.40	0.4190	40.32	0.5372
4*	1000	0.05	0.20	0.2830	10.00	0.5000
4*	1000	0.16	0.60	0.7917	96.00	0.5000
6	1000	0.05	0.40	0.4000	20.00	0.4603
7	400	0.10	0.40	0.3840	16.00	0.4582

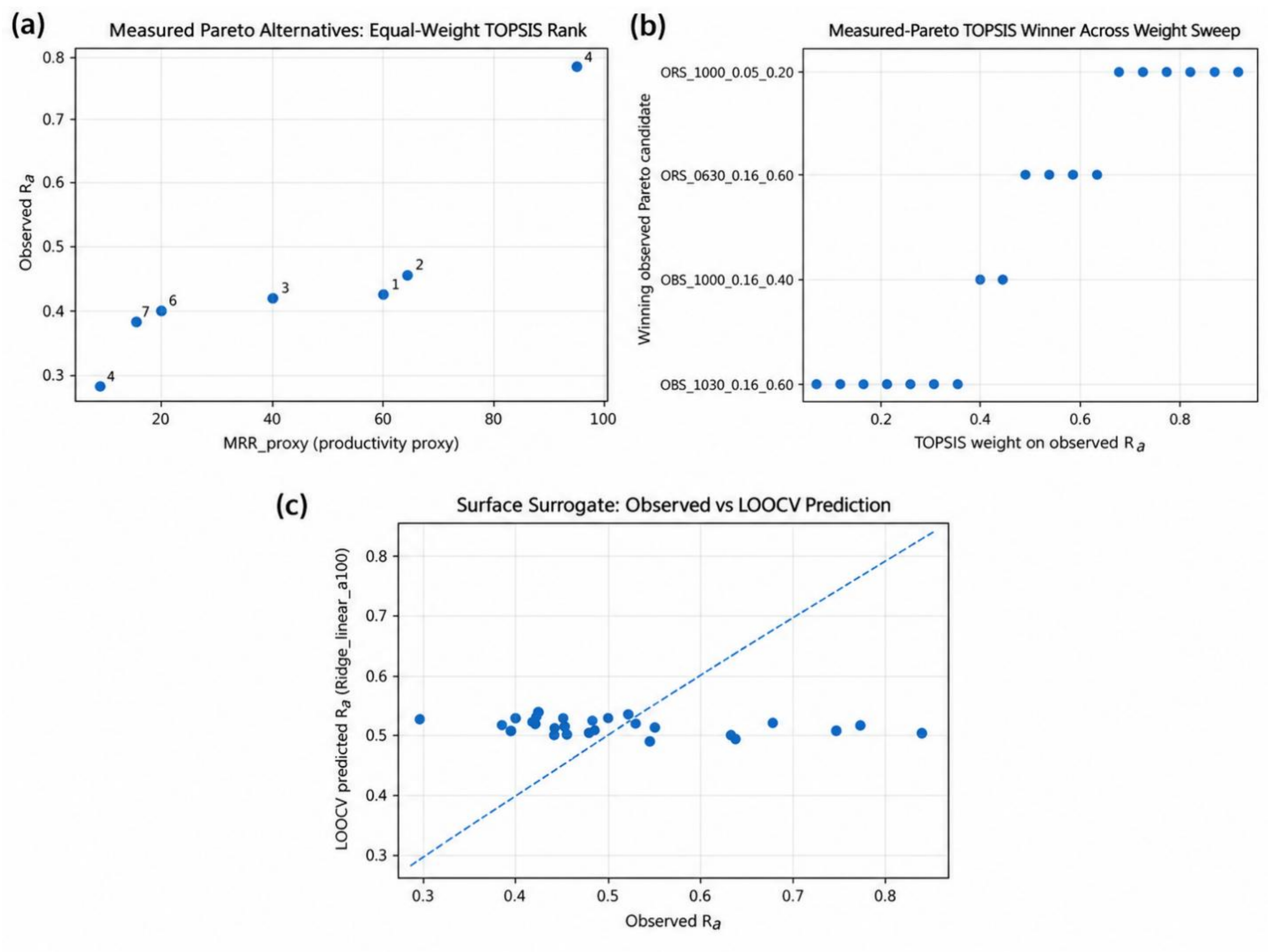

Figure 9. Multi-panel final decision analysis: (a) equal-weight TOPSIS ranking of observed Pareto alternatives; (b) TOPSIS winner/rank sensitivity across $w_{R_a} = 0.05–0.95$; (c) measured versus surrogate R_a comparison

The maximum observed Mahalanobis distance among the Pareto candidates was approximately 2.1907. No candidate was rejected using the previous threshold of 2.137470, because the final methodology no longer uses Mahalanobis distance as the primary acceptance criterion. One-at-a-time sensitivity analysis evaluated $\pm 5\%$and $\pm 10\%$changes around the selected condition. Positive perturbations in feed and depth of cut exceeded their experimental upper bounds and were consequently clipped. Overall, 4 of the 15 evaluated OAT cases were bound-clipped and requested and realized changes were reported separately.

Surface-model refit uncertainty was additionally evaluated using 500 bootstrap refits. At the final condition, the predicted $R_a$distribution produced an approximate 95% interval of 0.4765–0.5937. The experimentally measured value, $R_a = 0.4263$, was below this interval. This result reinforces that the bootstrap interval represents model-refit uncertainty rather than experimental repeatability or prospective physical validation. A consolidated summary of model uncertainty, optimizer performance, ranking sensitivity, experimental support, and local robustness is provided in Table 9.

Overall, the results show that the framework provides a transparent pre-laboratory decision-support process rather than an experimentally validated global optimizer. NSGA-II demonstrated statistically superior search efficiency on the fitted objective space, the independent RUL branch provided moderate sensor-based prognostic capability, and the use of measured Pareto alternatives reduced reliance on the limited surface surrogate. The final balanced recommendation was therefore $V_c = 630$, $f = 0.16$, and $a_p = 0.60$, with measured $R_a = 0.4263$and $MRR_{\text{proxy}} = 60.48$. The robustness of the optimized solutions was further examined through experimental design-support analysis, Mahalanobis centrality assessment, and local OAT perturbation testing, as shown in Figure 10. The design-support analysis confirmed that the candidate solutions remained within the interpolation-supported experimental region, while Mahalanobis distance was used only as a secondary centrality diagnostic. The OAT analysis further showed the effect of parameter-bound clipping by distinguishing between the requested and actually realized perturbations.

6. Discussion

The findings indicate that the proposed framework should be interpreted as a data-supported pre-laboratory decision-support approach rather than a fully validated predictive optimizer. For surface roughness, the parameter-only Ridge model achieved LOOCV $R^2 = -0.0740$, while nested model-selection validation produced $R^2 = -0.4692$. A negative R^2indicates that the model underperformed mean-response prediction for unseen conditions. Moreover, the nested procedure did not significantly reduce prediction error relative to the mean predictor ($p = 0.8901$). These findings demonstrate that the 27-condition turning dataset provides insufficient evidence for claiming reliable continuous surface-response prediction. Consequently, the NSGA-II-generated continuous solutions were treated as exploratory, and the final decision was anchored to experimentally measured Pareto-efficient conditions. This conservative interpretation is consistent with recent explainable data-driven approaches to machining-response modeling [15]. The independent milling branch provided stronger, although still heterogeneous, prognostic evidence. The no-cycle RUL model achieved $R^2 = 0.2383$, MAE = 23.44 cycles, and RMSE = 31.39 cycles under leave-one-tool-out validation. Its grouped-bootstrap R^2interval of approximately -0.0599 to 0.5660 indicates substantial between-tool variability. Nevertheless, sensor information significantly reduced per-tool RUL error relative to the cycle-only comparator, supporting the use of electrical-load descriptors for condition monitoring. This process-specific interpretation is consistent with physics-informed tool-wear modeling [16].

NSGA-II achieved a median normalized hypervolume of 0.7711 compared with 0.7421 for equal-budget random search, with a median improvement of 0.02897. The paired one-sided Wilcoxon test confirmed significantly better search performance. This demonstrates improved exploration of the fitted roughness–productivity space, consistent with evolutionary machining optimization [17,18], but does not compensate for the limited generalization of the surface surrogate. For practical selection, equal-weight TOPSIS identified $V_c = 630$, $f = 0.16$, and $a_p = 0.60$, with measured $R_a = 0.4263$and $MRR_{\text{proxy}} = 60.48$. The condition remained within the top three in 84.2% of tested weighting scenarios, although it achieved Rank 1 in only 21.1%, demonstrating useful but preference-sensitive stability.

Table 9. Final robustness and reliability summary

Assessment	Main result	Interpretation
Surface nested validation	$R^2 = -0.4692$; $p = 0.8901$vs DummyMean	Limited parameter-only generalization
Primary RUL LOGO	$R^2 = 0.2383$; 95% CI approximately -0.0599 to 0.5660	Moderate pooled unseen-tool capability
NSGA-II benchmark	HV 0.7711 vs 0.7421; $p = 9.31 \times 10^{-10}$	Superior surrogate-space search
TOPSIS sensitivity	elected condition top-3 in 84.2% of weight scenarios	Balanced but preference-sensitive
Design support	1,252/1,252 candidates supported	Search remained within interpolation domain
laximum nearest distance	0.3159 vs radius 0.4816	Within factorial coverage
OAT sensitivity	4/15 cases clipped	Boundary effects explicitly accounted for
Surface bootstrap	95% interval 0.4765–0.5937	Model uncertainty, not physical validation

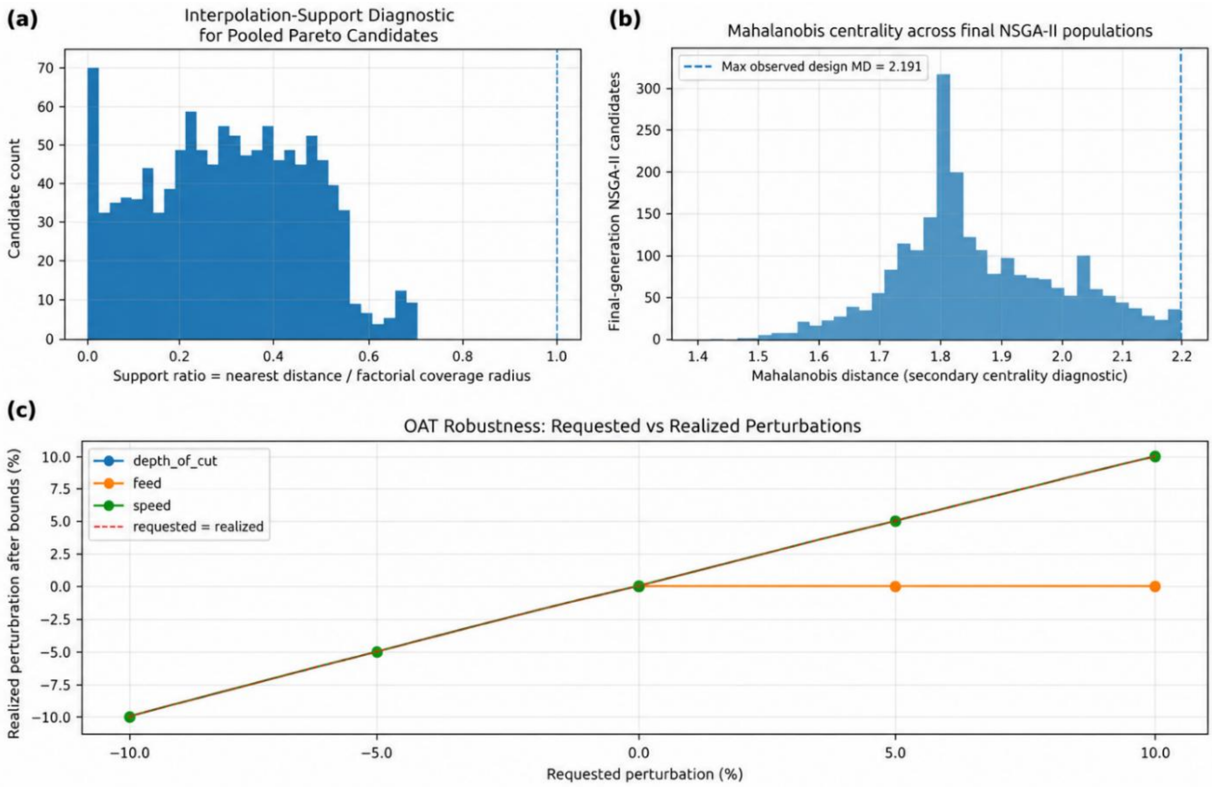

Figure 10. Robustness assessment using (a) design-support ratio, (b) Mahalanobis distance, and (c) OAT perturbation analysis

The separation of sensor-based monitoring from pre-machining optimization is also consistent with multi-sensor prognostic and fused-feature studies [19,20]. The principal limitations are the small 27-condition surface dataset, the use of independent turning and milling datasets, reliance on MRR_{proxy}rather than dimensional MRR, between-tool variability in RUL prediction, and the absence of prospective physical validation. Future work should acquire surface-quality and tool-wear measurements from the same machining system, incorporate force, temperature, chatter, tool geometry, and wear constraints, and compare NSGA-II with MOEA/D, MOPSO, and Bayesian optimization. Releasing the preprocessing, validation, and optimization scripts where licensing permits would further improve reproducibility.

7. Conclusion

This study developed a leakage-controlled, surrogate-assisted framework for multi-objective CNC machining decision support using two independent open datasets. The turning dataset was used for surface-roughness and productivity analysis, whereas the milling dataset was used independently for RUL and tool-life monitoring. No row-wise merging or cross-process DOC–ADOC mapping was performed, and milling-derived tool-health information was not imposed as a turning optimization objective or constraint. Strict validation showed that the parameter-only surface surrogate had limited out-of-sample generalization; consequently, continuous NSGA-II solutions were interpreted as exploratory rather than experimentally validated optima. In contrast, the no-cycle RUL model demonstrated moderate pooled unseen-tool predictive capability under leave-one-tool-out validation. NSGA-II also achieved significantly higher hypervolume than equal-budget random search in the fitted roughness–productivity objective space. To reduce dependence on the weak surface surrogate, final selection was based on experimentally observed Pareto-efficient conditions. Equal-weight TOPSIS identified $V_c = 630$, $f = 0.16$, and $a_p = 0.60$, with measured $R_a = 0.4263$and $MRR_{\text{proxy}} = 60.48$, as the balanced practical recommendation. Experimental design support confirmed that optimized candidates remained within the interpolation-supported region, while Mahalanobis distance was retained only as a secondary diagnostic. Overall, the framework provides transparent pre-laboratory screening and decision support, while prospective same-process physical validation remains necessary before industrial deployment.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically regarding authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with research ethics policies. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere. All survey participants provided informed consent before participating, and the anonymity and confidentiality of all respondents were strictly maintained throughout the research process.

Data availability statement

The manuscript contains all the data. However, additional data will be provided by the corresponding author upon reasonable request.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] P. Wu, Y. He, Y. Li, J. He, X. Liu, and Y. Wang, "Multi-objective optimisation of machining process parameters using deep learning-based data-driven genetic algorithm and TOPSIS," Journal of Manufacturing Systems, vol. 64, pp. 40–52, 2022, doi: 10.1016/j.jmsy.2022.05.016.

[2] G. Serin, B. Sener, A. M. Ozbayoglu, and H. O. Unver, "Review of tool condition monitoring in machining and opportunities for deep learning," The International Journal of Advanced Manufacturing Technology, vol. 109, no. 3–4, pp. 953–974, 2020, doi: 10.1007/s00170-020-05449-w.

[3] H. Yang, H. Zheng, and T. Zhang, "A review of artificial intelligent methods for machined surface roughness prediction," Tribology International, vol. 199, Art. no. 109935, 2024, doi: 10.1016/j.triboint.2024.109935.

[4] L. G. P. de Souza, G. A. V. B. Vasconcelos, L. A. R. Costa, M. B. Francisco, A. P. de Paiva, and J. R. Ferreira, "Roughness prediction using machine learning models in hard turning: An approach to avoid rework and scrap," The International Journal of Advanced Manufacturing Technology, vol. 133, no. 9–10, pp. 4205–4221, 2024, doi: 10.1007/s00170-024-13951-8.

[5] A. T. Nguyen, V. H. Nguyen, T. T. Le, and N. T. Nguyen, "Multiobjective optimization of surface roughness and tool wear in high-speed milling of AA6061 by machine learning and NSGA-II," Advances in Materials Science and Engineering, vol. 2022, Art. no. 5406570, 2022, doi: 10.1155/2022/5406570.

[6] M. Behzadian, S. K. Otaghsara, M. Yazdani, and J. Ignatius, "A state-of-the-art survey of TOPSIS applications," Expert Systems with Applications, vol. 39, no. 17, pp. 13051–13069, 2012, doi: 10.1016/j.eswa.2012.05.056.

[7] X. Lu, Y. Zhang, Z. Sun, H. Gu, C. Jiang, and S. Y. Liang, "Multi-objective optimization of cutting parameters for micro-milling nickel-based superalloy thin-walled parts based on improved NSGA-II algorithm," The International Journal of Advanced Manufacturing Technology, vol. 135, no. 1–2, pp. 775–786, 2024, doi: 10.1007/s00170-024-14478-8.

[8] Z. Zhang, F. Wu, and A. Wu, "Research on multi-objective process parameter optimization method in hard turning based on an improved NSGA-II algorithm," Processes, vol. 12, no. 5, Art. no. 950, 2024, doi: 10.3390/pr12050950.

[9] A. Gouarir, G. Martínez-Arellano, G. Terrazas, P. Benardos, and S. J. P. C. Ratchev, "In-process tool wear prediction system based on machine learning techniques and force analysis," Procedia CIRP, vol. 77, pp. 501–504, 2018, doi: 10.1016/j.procir.2018.08.253.

[10] T. Mohanraj, S. Shankar, R. Rajasekar, N. R. Sakthivel, and A. Pramanik, "Tool condition monitoring techniques in milling process—A review," Journal of Materials Research and Technology, vol. 9, no. 1, pp. 1032–1042, 2020, doi: 10.1016/j.jmrt.2019.10.031.

[11] M. Kuntoğlu, A. Aslan, D. Y. Pimenov, Ü. A. Usca, E. Salur, M. K. Gupta, T. Mikolajczyk, K. Giasin, W. Kapłonek, and S. Sharma, "A review of indirect tool condition monitoring systems and decision-making methods in turning: Critical analysis and trends," Sensors, vol. 21, no. 1, Art. no. 108, 2021, doi: 10.3390/s21010108.

[12] K. Antosz, E. Kozłowski, J. Sęp, and S. Prucnal, "Application of machine learning to the prediction of surface roughness in the milling process on the basis of sensor signals," Materials, vol. 18, no. 1, Art. no. 148, 2025, doi: 10.3390/ma18010148.

[13] X. Lu, Y. Zhang, Z. Sun, H. Gu, C. Jiang, and S. Y. Liang, "Multi-objective optimization of cutting parameters for micro-milling nickel-based superalloy thin-walled parts based on improved NSGA-II algorithm," The International Journal of Advanced Manufacturing Technology, vol. 135, no. 1, pp. 775–786, 2024.

[14] Z. Zhang, F. Wu, and A. Wu, "Research on multi-objective process parameter optimization method in hard turning based on an improved NSGA-II algorithm," Processes, vol. 12, no. 5, Art. no. 950, 2024.

[15] G. Piecuch and T. Żabinski, "A new open dataset from a milling process—data for classification and estimation of tool life," figshare, Dataset, 2025, doi: 10.6084/m9.figshare.28589216.v1. The corresponding published data descriptor is also available as Scientific Data, vol. 12, Art. no. 650, 2025, doi: 10.1038/s41597-025-04923-y.

[16] H. Harigovind, B. B. Nair, and N. R. Sakthivel, "A Sensor based turning dataset for data-driven surface roughness estimation," Mendeley Data, Version 3, 2025, doi: 10.17632/ybyhxv4czd.3.

[17] M. Pasic, D. Marinkovic, D. Lukic, D. Begic-Hajdarevic, A. Zivkovic, M. Milosevic, and K. Muhamedagic, "Prediction and optimization of surface roughness and cutting forces in turning process using ANN, SHAP analysis, and hybrid MCDM method," Applied Sciences, vol. 14, no. 23, Art. no. 11386, 2024, doi: 10.3390/app142311386.

[18] F. Pashmforoush, A. E. Araghizad, and E. Budak, "Tool wear prediction in milling process using physics-informed machine learning and thermo-mechanical force model with monitoring applications," Journal of Manufacturing Systems, vol. 82, pp. 1192–1212, 2025, doi: 10.1016/j.jmsy.2025.08.014.

[19] P. Cinat, G. Gnecco, and M. Paggi, "Multi-scale surface roughness optimization through genetic algorithms," Frontiers in Mechanical Engineering, vol. 6, Art. no. 29, 2020, doi: 10.3389/fmech.2020.00029.

[20] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," Complex Systems, vol. 9, no. 2, pp. 115–148, 1995. https://content.wolfram.com/sites/13/2018/02/09-2-2.pdf