Journal of Manufacturing Systems 85 (2026) 419–440

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

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

Technical paper

A comprehensive framework for computationally efficient system-level design optimization of machine tools

Deniz Bilgili [a,b,*], Erhan Budak [c], Jasmin Jelovica [b]

[a] Natural Controls, 2 Bloor Street East, Suite 3500, Toronto, ON M4W 1A8, Canada
[b] Department of Mechanical Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[c] Manufacturing Research Laboratory, Sabanci University, Tuzla, Istanbul 34956, Turkey

ARTICLE INFO

Keywords:
Machine tools
Optimization
Surrogate modeling
Machine learning
Multi-objective

ABSTRACT

Mass reduction of machine tool components is a crucial task that can improve performance, accuracy, and energy efficiency. System-level optimization, where multiple components are simultaneously optimized, is noted in the literature as a challenging necessity for complete performance improvement of machine tools. Existing methods focus on optimizing the machine tool components individually, neglecting the critical effects of simultaneous modification on the machine performance and thorough exploration of the design space. To the authors' knowledge, for the first time in the literature, this paper presents a comprehensive framework for system-level machine tool design optimization considering the most significant multi-objective performance indicators for the machining process. Static and dynamic stiffness, thermal and dynamic stability, and fatigue life are evaluated as performance indicators using a multi-objective finite element response set that includes coupled thermal-structural, modal, and frequency response analyses. A minimal parameter set approach is proposed which uses the linear guide joints to minimize the number of design variables, addressing the challenge of increased computational cost in system-level modeling. Machine responses during optimization iterations are predicted by a machine learning model trained on the machine tool's multi-objective finite element response set, achieving higher accuracy than commonly used polynomial-based methods. A constraint relaxation method is proposed that permits limited degradation relative to the base design, yielding designs that substantially outperform those obtained from unconstrained optimization while avoiding over-constraining. Up to 20 % mass reduction is achieved across the machine tool components while the performance indicators are either improved or maintained with negligible degradation.

1. Introduction

Reducing the mass of the moving components improves the machine tool's dynamic performance, positioning accuracy, and energy efficiency [1,2]. However, achieving substantial amounts of mass reduction can be challenging as machine tools experience significant structural and thermal loads during the machining process [3,4]. The structural stress and deformations of the machine tool are mainly caused by gravity, thermal expansion of machine tool components, and dynamic machining process loads due to motion, vibrations, and cutting forces. Mass reduction under structural and thermal process loads has a substantial effect on the machine tool's stiffness, stability, and fatigue life. To achieve significant mass reduction across the machine tool while maintaining these, a system-level multi-objective design optimization method capable of simultaneously modifying multiple components is necessary.

As a computationally efficient alternative to topology optimization, parameter optimization involves selecting machine dimensions such as lengths, thicknesses, or diameters as design variables to be added to the design space while all other dimensions are added to the non-design space [5]. Therefore, only the machine dimensions selected as variables are modified during optimization while the other dimensions are kept constant. Once the design variables are selected, to reduce the computational expense of design evaluations, a surrogate model is created to predict the response of the machine for new values of the variables [6]. The surrogate model is typically created using a set of value-response pairs called samples. Samples can be obtained through simulation, physical testing, or both. Using the generated samples, a

* Corresponding author at: Natural Controls, 2 Bloor Street East, Suite 3500, Toronto, ON M4W 1A8, Canada.
E-mail address: deniz.bilgili@naturalcontrols.ca (D. Bilgili).

https://doi.org/10.1016/j.jmsy.2026.02.005
Received 4 November 2025; Received in revised form 11 January 2026; Accepted 3 February 2026
Available online 10 February 2026
0278-6125/© 2026 The Author(s). Published by Elsevier Ltd on behalf of The Society of Manufacturing Engineers. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

surrogate model of the machine tool can then be created using methods such as response surface method or machine learning [7].

Research on parameter-based optimization of machine tools exists in the machining literature. Existing methods commonly isolate a single component of the machine tool for optimization while the remaining components are excluded [2,8,9]. The design variables are selected from component-specific features which may not be applicable to other components. Furthermore, the mechanical joints connecting machine tool components, which heavily influence the machine's overall static and dynamic [10,11] as well as thermal [12] behavior, are neglected. Therefore, existing component-level methods fail to provide a system-level analysis capturing the coupled structural and thermal interactions between the machine tool components. On the other hand, Wu et al. [13] included multiple machine tool components in optimization by using a combination of component-specific principal and local design features selected on each component as design variables. Although multiple components contribute to the objective function in [13], mass reduction is driven only by the maximum static deflection of individual components, ignoring other crucial performance indicators such as thermal and dynamic stability, natural frequencies, and durability. Furthermore, the load case is simplified to only include a single unit-load applied at the spindle in the horizontal direction. This fails to capture the effects of the three-dimensional loading by the cutting force and neglects the coupled thermal-structural load case by gravity, the cutting force, and the heat generated at the spindle bearings. Hence, previous studies fall short of achieving a system-level effect in optimization driven by a multi-objective response set of the machine tool's most important performance indicators.

To the best of authors' knowledge, this study presents the first comprehensive framework addressing the challenge of system-level design optimization of machine tools where the components are simultaneously modified to achieve substantial mass reduction while improving or maintaining critical performance indicators. Instead of using design-specific variables as commonly done in the literature, which would rapidly increase the number of variables and computational cost in system-level modeling, the linear guides, which are widely used in machine tools, are used for parameterization. Consequently, substantial mass reduction is achieved across the machine tool with limited computational cost. Important performance indicators including static and dynamic stiffness, thermal and dynamic stability, and fatigue life, which are neglected in existing multi-component methods, are used to guide optimization to ensure the performance is improved or maintained. Focus is placed on limiting the computational cost of optimization to make the method computationally accessible. Machine learning (ML) algorithms are used for surrogate modeling to guide optimization iterations with greater accuracy in predicting the machine's responses than the widely used response surface-based methods. A global optimization algorithm is employed to ensure the discovery of optimal trade-off designs. The paper is structured as follows: Section 2 introduces a parameter selection method taking advantage of the linear guides to limit the number of design variables and increase computational efficiency. Section 3 discusses FEA modeling for ML training sample generation. An energy-based order reduction approach for frequency response functions is introduced to include dynamic stability indicators in design space exploration with reduced computational cost. Section 4 demonstrates a surrogate modeling strategy using a minimal number of training samples for the ML model. Details of machine design sample generation for training and automatic hyperparameter tuning are shared. Section 5 explains the integration of the surrogate model with optimization algorithms and presents the optimization results. Section 6 presents a comparison of the base and optimized designs and demonstrates performance improvements. Section 7 gives the concluding remarks and motivation for future research.

2. System-level parametric modeling of machine tools

The dimensions of the machine tool selected as modifiable parameters for optimization are called design variables. Design variables are used as inputs to the response model of the machine tool to predict the change in the machine's response due to design modifications during optimization. It is necessary that the number of design variables balance computational complexity and accuracy. While an increased number of design variables can potentially enable greater improvements in performance, it also increases computational cost as more design samples are necessary to generate the response model [14]. Several design variables are used for each machine tool component in existing methods, which rapidly increases the total variable count and the computational cost when applied to all components across the machine tool for system-level optimization. In addition, to enable repeatability and applicability to other machines, the selection of design variables should be based on common dimensions rather than specifics. To address these challenges, this section presents a minimal parameter set approach to enable substantial design modifications across the machine tool with a limited number of design variables.

To achieve mass reduction throughout the machine tool with reduced computational cost, a minimal parameter set approach is proposed which prioritizes the fundamental features of the machine tool. The following rules are applied in design variable selection:

i. The distances between the linear guide rails must be selected as design variables.
ii. The wall thicknesses of the main components must be selected as design variables.

As shown in Fig. 1, the distance L between the linear guide rails largely determines the width or height as well as the cross section of both the static and the moving component in a linear guide joint. Therefore, with the first rule, a single design variable can have significant control over the overall design of two machine tool components simultaneously. On the other hand, although the wall thickness T only affects a single component, it has a strong effect on the component's mass, rigidity, and durability, which determine the component's static and dynamic response. Therefore, the second rule enables control over multiple aspects of a component's response with only a single design variable, helping limit the computational cost. The combined effect of the first and second rules on the cross-section of machine tool components is observed in Fig. 1**a**. As seen, by simultaneous modification of variables L and T, it is possible to achieve significant cross-section modifications with a limited number of design variables, which enables system-level optimization with reduced computational complexity.

The application of the minimal parameter set on the example machine tool is demonstrated in Fig. 2. As seen, the first and second rules are applied to select the linear guide distances and wall thicknesses of the main components as design variables which cover the entire machine with only seven variables: The linear guide distance of the bed (*Bed_L*) and the column (*Col_L*), the wall thickness of the bed (*Bed_T*), column (*Col_T*), X-axis carriage (*XCar_T*), Y-axis carriage (*YCar_T*), and Z-axis carriage (*ZCar_T*). However, certain machine tool design features which would normally be selected as variables are excluded due to assumed physical constraints. As seen in Fig. 2, the linear guide distance of the Y-axis carriage is not selected as a variable, which is due to the narrow spacing of the rails not allowing further length reduction. Similarly, the distance between the top and the middle rails of the column's linear guides is constrained to have the constant ratio of 1.22 to the distance between the middle and bottom rails to ensure appropriate physical proportions.

The proposed minimal parameter set approach can be adapted to product-related requirements not directly related to design and performance optimization, such as the axis motion limits and the table's width

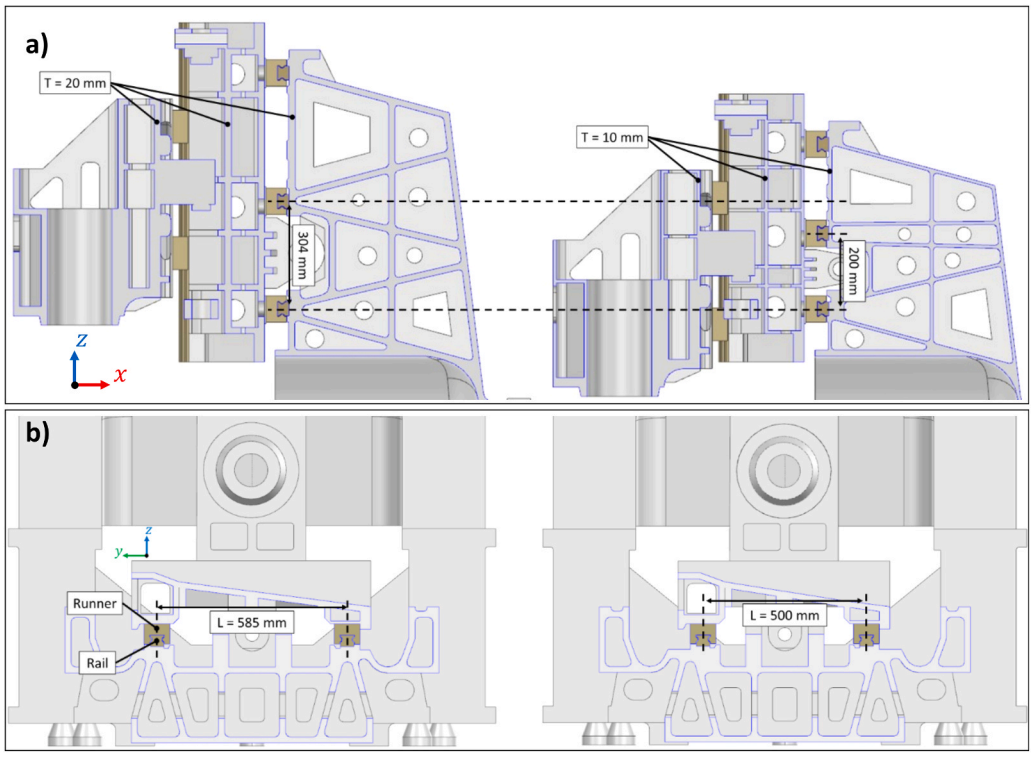


Fig. 1. The effect of linear guide distance (L) and wall thickness (T) on the height, width, and cross section of machine tool components: a) with varied height, b) with varied width.

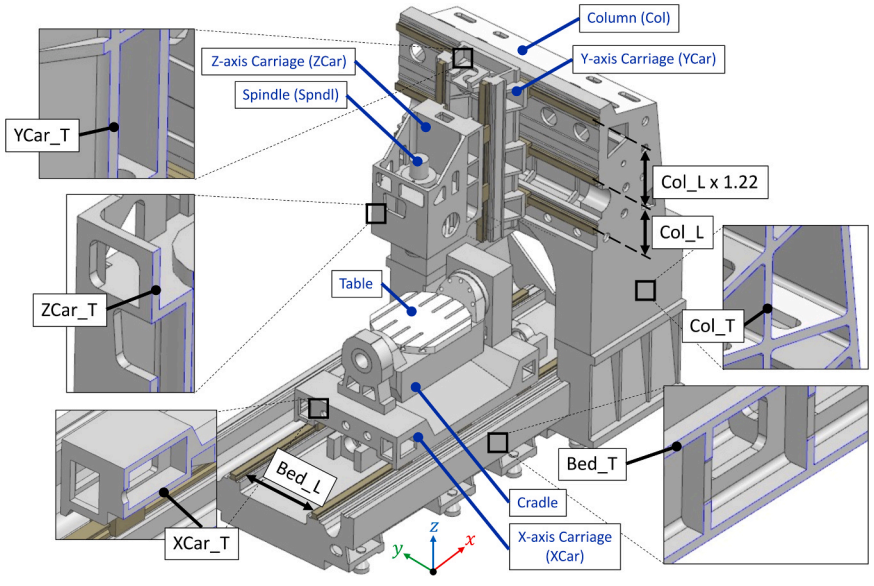


Fig. 2. The components (blue) and design variables (black) of the machine tool. The zoomed views show cross-sections.

and length which determine the maximum work volume of the machine. In this study, it is assumed that the machine tool's maximum work volume is a product-related requirement and must not be reduced after optimization. Hence, as also seen in Fig. 1**b**, while the width of the X-axis carriage is completely controlled by the linear guide distance *Bed_L*, only the inner width of the bed structure is changed while the outer width is kept fixed. This is due to the assumption that reducing the outer width of the bed would require similarly reducing the outer width of the column structure, in which case it may not be possible to provide the original design's Y-axis motion range. Hence, the design variable *Bed_L* is not allowed to modify the outer widths of the bed and the column. On the other hand, it is observed that the Z-axis motion range of the original design can be maintained for $Col\_L > 200mm$. While this shows the adaptability of the proposed approach, it also shows that product-related requirements may act as constraints or modifiers on the proposed approach. Such product-related considerations must inform the parameterization step to ensure that the design space, through adjustment of the design variable limits, is aligned with product goals. This further implies that, for cases with exceptionally challenging product requirements, the design space formed by the minimal parameter set approach may be significantly limited, reducing the level of attainable optimization.

3. Finite element modeling for multi-objective optimization

Existing multi-component optimization methods use simplified load cases where only a downward unit force is applied on the spindle nose [13], neglecting the complex loading by gravity, the thermal expansion of the components [4], and the cutting forces. Furthermore, important performance indicators including static and dynamic stiffness, thermal and dynamic stability, and fatigue life are not used to guide optimization as multiple components are simultaneously modified. To accurately guide optimization under complex process loads while ensuring the optimized designs improve or maintain the multi-objective performance indicators, three types of FEA simulations are performed:

i. Coupled thermal-structural analysis under gravity, cutting forces, and heat generated at the spindle bearings,
ii. Modal analysis for the first four natural frequencies of the machine tool,
iii. Spindle nose frequency response function (FRF) calculations.

A reference FEA model, shown in Fig. 3, is shared between the three simulation cases. The machine tool is modeled in its weakest pose with the Y-axis carriage at the center, $Y = 0$ mm, and the Z-axis carriage at its highest position, $Z = +140$ mm. X, A, and C axes are also placed at their zero positions. The machine is fully fixed to the ground through its ground supports using RBE2 rigid elements as shown in Fig. 3-**A**. To enable heat conduction between the components, the sliding and rotating interfaces are connected with one-dimensional conductive, elastic, cylindrical bar elements shown in Fig. 3-**B**. For linear guide rail-runner couples, a bar element connects a rail node to a runner node. For rotary interfaces, a bar element connects a node on the bearing inner ring surface to a node on the bearing outer ring surface. The actual bearing rings are not included in the model. Connectivity between the static components, namely the bed and the column, is achieved through node connectivity where FE nodes are shared between the FE elements of the bed and the column as shown in Fig. 3-**C**, which establishes a flexible connection without artificial stiffening. For displacement and frequency response calculations, RBE2 rigid elements are created on the spindle nose and the table center as shown in Fig. 3-**D** and **E**. The machine tool is meshed with quadratic tetrahedrons to avoid artificial stiffening of the components. A primary mesh size of 25 mm is used for most regions, with 5 mm applied at high-stress locations to ensure result convergence. Young's modulus, density, thermal expansion coefficient, and Poisson's ratio for the cast iron machine tool components are taken as 124000 MPa, 7.3×10^{-9} tonne/mm^3, 1.04×10^{-5} ℃$^{-1}$, and 0.3, respectively.

The outputs of the FEA simulations performed for the machine tool, referred to as responses, are given in Table 1. The maximum principal stresses in the components are calculated to estimate the effect of optimization on the machine's fatigue life. The first four natural frequencies are calculated to include the effect of dynamic stiffness in optimization. Spindle nose FRFs are calculated to capture the machine tool's contribution to dynamic stability. As discussed in later sections, along with the design variables in Fig. 2, the FEA responses in Table 1 are used to train an ML model as a computationally efficient surrogate for the original FEA model of the machine tool. Therefore, the FEA model and the surrogate ML model share the same inputs (design variables, Fig. 2) and outputs (responses, Table 1).

3.1. Coupled thermal-structural analysis

For the coupled thermal-structural analysis, the cutting force projected to the spindle nose is applied on the master node of the RBE2 rigid element in Fig. 3-**D**. With an assumed feed in the +X direction of the machine using a helical endmill with climb milling, the X, Y, and Z components of the cutting force in the spindle nose frame are applied as −1500 N, 1000 N, and −2000 N, respectively. The cutting force is also applied at the center of the table (Fig. 3-**E**) in the opposite direction of the spindle nose forces with the same amplitudes. A combined process load is applied including gravity, thermal loads, and cutting forces. Gravity is applied in the -Z direction of the machine as 9810 mm/s^2. The machine tool is initially assumed to be at thermal equilibrium at 21 ℃, which is applied as the reference temperature to all FE nodes. The

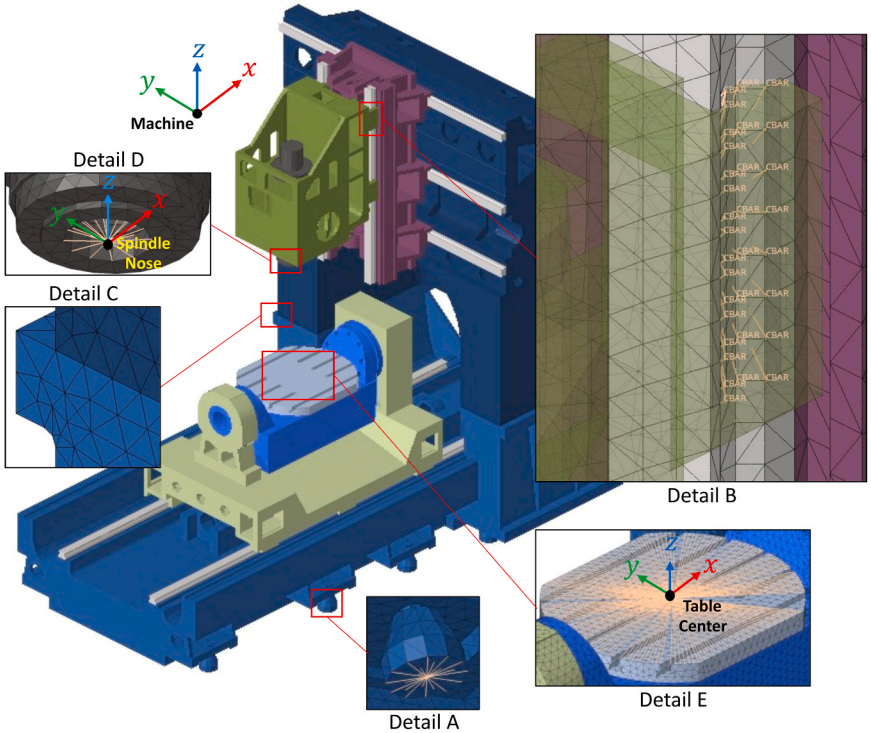


Fig. 3. The reference FEA model for the coupled thermal-structural, modal, and frequency response analyses of the machine tool.

Table 1
FEA outputs (responses) included in ML model training and parameter optimization.

Load Case	Response	Res. ID	Detail	Unit
None	*Mass*	R1	Total mass of the machine tool	tonne
Coupled Thermal-Structural	*SN_Disp*	R2	Max. displacement at spindle nose	mm
	TB_Disp	R3	Max. displacement at table center	mm
	Stress_Bed	R4	Max. Principal Stress in Bed	MPa
	Stress_Col	R5	Max. Principal Stress in Column	MPa
	Stress_XCar	R6	Max. Principal Stress in X-axis Carriage	MPa
	Stress_YCar	R7	Max. Principal Stress in Y-axis Carriage	MPa
	Stress_ZCar	R8	Max. Principal Stress in Z-axis Carriage	MPa
	Stress_Spndl	R9	Max. Principal Stress in Spindle	MPa
Modal	*Mode1*	R10	First natural frequency of the assembled machine	Hz
	Mode2	R11	Second natural frequency of the assembled machine	Hz
	Mode3	R12	Third natural frequency of the assembled machine	Hz
	Mode4	R13	Fourth natural frequency of the assembled machine	Hz
FRF	*EnergyXX*	R14	The square of the L_2 (Euclidean) norm of the spindle nose FRF (XX) signal	$(mm/N)^2$
	EnergyYY	R15	The square of the L_2 (Euclidean) norm of the spindle nose FRF (YY) signal	$(mm/N)^2$

steady-state temperatures of the lower and upper spindle bearing locations are assumed to be 55 ℃ and 45 ℃, respectively, based on the results by Yuksel et al. [4]. The spindle bearing steady-state temperatures are applied as temperature boundary conditions to the surface nodes at the outer and inner bearing ring locations.

The contributions of the thermal and structural loads to the combined process load are identified by isolating each load case. The calculations are performed for 100 machine design samples to capture the distribution of thermal and structural load contribution over the design space. As seen in Fig. 4**a**, while the thermal deflection makes up a significant portion of the spindle nose deflection *SN_Disp* due to the combined load, its variance over the design samples is not substantial. On the other hand, as seen in Fig. 4**b**, **c**, and **d**, the maximum principal stress is dominated by thermal stresses. Hence, using the thermal deformations of the spindle nose or the structural stresses of the machine tool components in individual optimization objectives would not yield meaningful results. Thus, to limit the number of optimization objectives and increase computational efficiency, the combined load case is used instead of individual thermal and structural load cases. Although transient thermo-mechanical effects are neglected in this study, the decision to use combined static thermal-structural loads is consistent with existing transient load modeling approaches [15], which could be integrated with the combined static load component to account for transient effects in system-level optimization.

3.2. Modal analysis

Modal analysis is conducted using the same FEA model along with the boundary conditions in Fig. 3-**A**. The first four elastic vibration modes of the machine tool are calculated, which are shown in Appendix A for the base design. As seen, the first four modes capture the overall

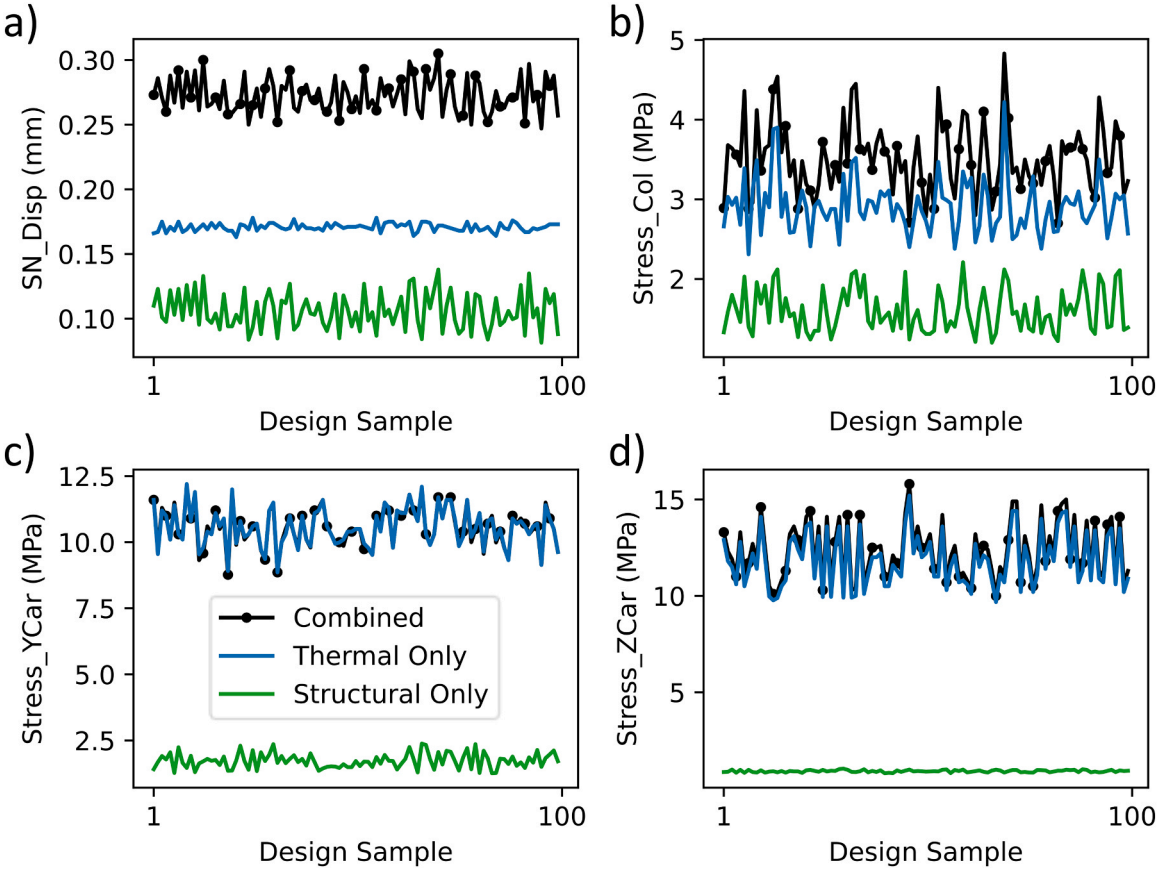


Fig. 4. The contribution of thermal and structural load cases to the combined load case.

dynamic stiffness of the machine tool in all bending and torsional directions which will be sufficient for optimizing the machine tool's dynamic stiffness. Thus, only the first four modes are included from modal analysis as responses to limit the computational cost.

3.3. Frequency response analysis

The spindle nose FRF of the machine tool is calculated in the X and Y directions of the spindle nose frame shown in Fig. 3-**D** which are parallel to the machine frame's X and Y directions. The tool tip FRF directly influences the chatter stability of the machining process [16], and the spindle nose FRF is a significant contributor to the resulting tool tip FRF [17]. Thus, the spindle nose FRF can serve as a common indicator of the machine's overall dynamic stability before the contribution of a specific tool is added. However, using the spindle nose FRF directly for ML training would significantly increase the computational cost as discrete FRFs are represented by many points. Therefore, a reduced-order representation for the spindle nose FRF as a single scalar is proposed to maximize computational efficiency.

The single degree of freedom (SDOF) representation for the milling of a flexible workpiece and a relatively rigid tool is presented in [16] where the stiffness and damping of the workpiece are represented as discrete, lumped parameters. As shown in Fig. 5, a similar lumped parameter representation is adopted for the machine tool's stiffness and damping. By assuming a rigid tool and workpiece, design optimization exclusively targets the elastic properties of the machine tool without the contribution of a specific tool or workpiece. The critical axial depth of cut a_{lim} is obtained using the single degree-of-freedom expression [16]:

$$a_{lim} = \frac{1}{\frac{N}{2\pi}\alpha_{jj}K_t \mathrm{Re}\left[G_{jj}(i\omega_c)\right]}, \quad j \in \{X, Y\} \tag{1}$$

where N is the number of flutes, ω_c is the frequency of chatter vibrations in rad/s, K_t is the tangential cutting coefficient, and α_{jj} and $\mathrm{Re}\left[G_{jj}(i\omega_c)\right]$ are the directional coefficient and the real part of the FRF calculated at the point of tool-workpiece engagement in the direction j, respectively. The directional coefficient α_{jj} can be positive or negative depending on the angle of engagement between the tool and the workpiece, the milling direction (up or down milling), and the radial cutting coefficient [16]. As a_{lim} is a positive number, the sign of $\mathrm{Re}\left[G_{jj}(i\omega_c)\right]$ must be the same as that of α_{jj}. It is seen in Eq. (1) that the critical depth of cut is inversely proportional to the multiplication of α_{jj} and $\mathrm{Re}\left[G_{jj}(i\omega_c)\right]$. This indicates that the dynamic stability of the machine tool increases as the real part of the tool tip FRF $\mathrm{Re}\left[G_{jj}(i\omega_c)\right]$ approaches zero.

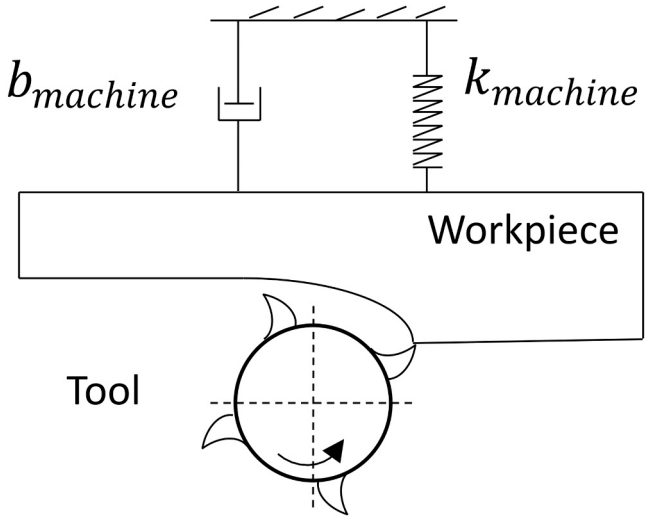


Fig. 5. SDOF representation of a flexible machine tool which has an equivalent damping and stiffness of $b_{machine}$ and $k_{machine}$, milling a rigid workpiece with a rigid tool.

Receptance coupling shows that the spindle nose FRF can have a varying effect on the tool tip FRF $\mathrm{Re}\left[G_{jj}(i\omega_c)\right]$ depending on the dynamic interactions between the machine tool, spindle, tool holder, and the tool [18,19]. As a computationally efficient performance indicator for the machine tool's dynamic stability that can be adjusted for varying dynamics and design goals, the concept of FRF energy is introduced as a reduced-order representation for the spindle nose FRF. This study proposes minimizing the FRF energy as a baseline to bring the real part of the tool tip FRF closer to zero, increase the machine tool's dynamic rigidity, and improve process stability. However, depending on the dynamic interactions between the components, stability may also be increased by maximizing the FRF energy or by targeting a given FRF energy value.

The L_2 norm ε of the real part of the discrete spindle nose FRF $G[k]$ is calculated in Eq. (2) to reduce the discrete FRF curve containing k amplitude-frequency pairs to a single scalar. As the square of the L_2 norm of a signal represents the energy of the signal [20], the energy of the real part of the spindle nose FRF, E, is calculated as the square of ε in Eq. (2). Although E does not directly represent the energy of the entire spindle nose FRF $G[k]$, it is used as a response due to its strong relevance to the real part of the spindle nose FRF which has a direct influence on dynamic stability [16]. The energies E_{XX} and E_{YY} calculated in orthogonal directions X and Y are denoted *EnergyXX* and *EnergyYY*, respectively, in Table 1.

$$\begin{aligned} &\varepsilon_{jj} = \sqrt{\sum_k \left(\mathrm{Re}\left[G_{jj}[k]\right]\right)^2} \\ &\rightarrow \varepsilon_{jj}^2 = E_{jj} = \sum_k \left(\mathrm{Re}\left[G_{jj}[k]\right]\right)^2, \; j \in \{X, Y\} \end{aligned} \tag{2}$$

4. Surrogate modeling of machine tools

Surrogate models are widely used as computationally efficient alternatives to full-order FEA models in parameter-based optimization of machine tools. Existing studies dominantly use polynomial-based response surface methods (RSMs) in which the machine tool's multi-objective response is predicted by a polynomial [2,8,9,13]. The polynomial's coefficients are identified through a learning process in which the error between the polynomial's predictions and the original FEA results are iteratively minimized. This section proposes a machine learning-based surrogate modeling approach exceeding the accuracy of polynomial-based RSMs in predicting multi-objective machine response while limiting the number of training samples to maximize computational efficiency.

4.1. Sampling machine tool design variants

While the surrogate ML model itself is computationally highly efficient, generating machine tool samples can be computationally demanding depending on the number of samples generated. As explained in Section 3, each sample requires four FEA calculations to be performed (i.e., coupled thermal-structural stress analysis, modal analysis, and spindle nose FRF calculation in X and Y directions). Therefore, it is crucial to minimize the number of samples while maintaining training performance and ML prediction accuracy.

The design space is defined based on the lower and upper limits for each design variable as given in Table 2. As the primary goal of design optimization is mass reduction, the upper limits are selected to be equal to the base design of the machine, while the lower limits are based on reasonable approximations to ensure feasibility and manufacturability. The maximum possible mass reduction in Table 2 shows that the selected limits may enable a significant amount of mass reduction throughout the machine tool. For the variables *Bed_L* and *Col_L*, a

Table 2
The lower and upper limits for the optimization design variables. The upper limits are selected following the base design.

Component	Design Parameter	Variable	Lower Limit (mm)	Upper Limit (mm)	Mass at Lower Limit (tonne)	Mass at Upper Limit (tonne)	Max. Mass Reduction (tonne)
Bed	Linear guide distance	*Bed_L*	400	585	2.523	3.316	0.793
Bed	Wall thickness	*Bed_T*	10	20			
Column	Linear guide distance	*Col_L*	200	304	1.591	2.730	1.139
Column	Wall thickness	*Col_T*	10	20			
X-axis Carriage	Wall thickness	*XCar_T*	10	20	0.471	0.652	0.181
Y-axis Carriage	Wall thickness	*YCar_T*	10	20	0.309	0.458	0.148
Z-axis Carriage	Wall thickness	*ZCar_T*	10	20	0.295	0.378	0.084
Total							2.345

geometrically consistent cross section cannot be found if these variables are set to a value below the lower limit. For the five wall thicknesses, the lower limit of 10 mm is selected to ensure the optimized components can be manufactured through casting.

The trained ML model is required to produce accurate predictions for the machine tool's thermal and mechanical response within the selected lower and upper limits. Thus, the design samples must be distributed evenly over the seven-dimensional design space defined by the design variables. A sampling method that covers the entire design space is said to be space-filling. In this study, the Latin Hypercube Sampling (LHS) method, specifically the maximin LHS [21], is used to generate a space-filling design sample set. Based on Table 2, a sample $\boldsymbol{u}$ in the machine tool's seven-dimensional design space is represented as:

$$\boldsymbol{u} = (u_{Bed\_L},\ u_{Bed\_T},\ u_{Col\_L},\ u_{Col\_T},\ u_{XCar\_T},\ u_{YCar\_T},\ u_{ZCar\_T})$$

$$u_{Bed\_L} \in [400,\ 585],\ u_{Bed\_T} \in [10,\ 20]$$
$$u_{Col\_L} \in [200,\ 304],\ u_{Col\_T} \in [10,\ 20]$$

$$u_{XCar\_T} \in [10,\ 20],\ u_{YCar\_T} \in [10,\ 20],\ u_{ZCar\_T} \in [10,\ 20] \tag{3}$$

The Euclidean distance between any two design samples $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ is calculated as:

$$d(\boldsymbol{u}_i,\ \boldsymbol{u}_j) = \sqrt{\sum_{\substack{All\ variables \\ (Table\ 2)}} |\boldsymbol{u}_i - \boldsymbol{u}_j|^2} \tag{4}$$

In a sample set with n samples, the minimum pairwise Euclidean distance is calculated as:

$$\min_{1 \le i < j \le n} d(\boldsymbol{u}_i,\ \boldsymbol{u}_j) \tag{5}$$

The maximin algorithm maximizes Eq. (5) [21] which yields a space-filling sample set with well-distributed samples, covering the entirety of the design space.

The full sample set obtained by LHS is split into training and test sets in an 80:20 ratio. The samples in the test set are not used in training and remain unknown to the trained ML model. While the test set also needs to be space-filling to test the trained ML model's prediction accuracy throughout the design space, it also needs to be dissimilar to the training set. As the training process fits the ML model's hyperparameters to the training set, a test set whose samples are similar to the training samples would overestimate the ML model's prediction accuracy. It is common to obtain the test set by randomly shuffling the full sample set and taking a random slice. Although this approach usually yields sufficient dispersion of samples over the design space, a larger number of samples is needed to ensure both the training and test sets are evenly distributed.

To generate space-filling training and test sets with a limited number of samples for increased computational efficiency, a two-step sampling approach is proposed by coupling the maximin LHS method with the diversity subsampling method by Shang et al. [22]. The diversity subsampling method creates a space-filling subsample of a discrete sample set without requiring knowledge of its probability distribution. Therefore, the diversity subsampling method is suitable for identifying a space-filling subset of a given machine tool design sample set. As seen in Fig. 6, an initial space-filling set of design samples is generated from the continuous design space using the maximin LHS method. Subsequently, the diversity subsampling method is applied, identifying the test set as a subset of the initial set that is also space-filling in the original design space. While it is possible to apply diversity subsampling to select the training samples first, it is observed that the resulting test set achieves better dispersion over the design space if diversity subsampling is applied to draw the test samples as shown in Fig. 6**b**. Fig. 6**a** shows that the two-step sampling method produces training and test sets achieving satisfactory dispersion over the design space with minimal overlap using only 100 design samples. It is also seen that a ~~mostly~~ nearly uniform distribution is achieved in each design variable as seen in the histograms along the diagonal. Section 4.2 presents the accuracy of three ML models and shows the effect of sample set size on the accuracy of the best-performing model.

4.2. Machine learning model training with automatic hyperparameter tuning

The ML model is trained using supervised learning to serve as a computationally efficient surrogate for the original FEA model of the machine tool. The training set, shown in Fig. 6**a**, consists of 80 design samples. The seven design variables given in Table 2 are selected as features (i.e., inputs) for the ML model, while the 15 FEA results given in Table 1 are selected as responses (i.e., outputs). To prevent bias in the trained ML model's predictions, the features and responses are normalized to [-1, 1] based on the minimum and maximum values of the training set. During training, the ML model's hyperparameters are fit to the training samples which cannot guarantee that the trained ML model will achieve high prediction accuracy with unseen samples after training. On the other hand, the test set is not included in the training

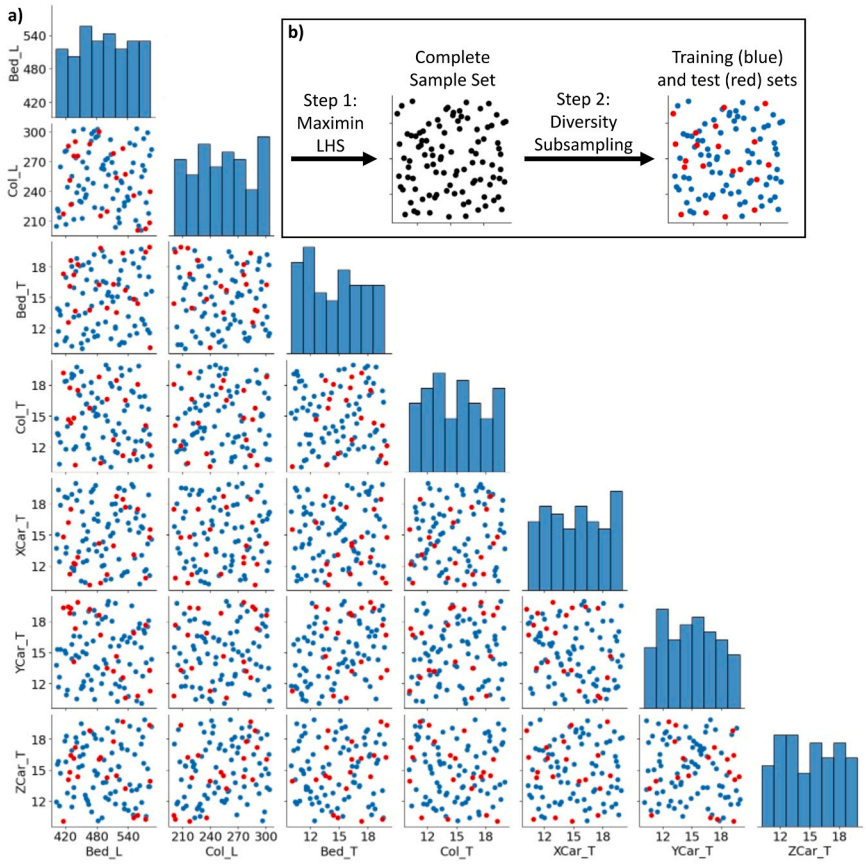


Fig. 6. a) The space-filling training and test sets obtained using b) the two-step sampling approach. 80 training and 20 test samples are shown.

and hyperparameter tuning process but used only once to evaluate the trained model's success at generalizing to unseen data.

To ensure that the trained model achieves high accuracy and generalizes well to unseen data, 10-fold cross-validation (CV) [23] is employed, while Bayesian optimization [24] is used to automatically tune the hyperparameters to minimize the prediction error. For a given hyperparameter set θ, feature vector x_i, and response vector y_i, the validation loss $L_{val}^{(i)}(\theta)$ representing the ML model's mean squared prediction error on the i-th fold is calculated as:

$$L_{val}^{(i)}(\theta) = \frac{1}{N}\sum_{j=1}^{N}\left(h_\theta\left(x_j^{(i)}\right) - y_j^{(i)}\right)^2, \quad 1 \le i \le k = 10. \tag{6}$$

where h_θ is the trained ML model and $N = 80/10 = 8$ for a training set with 80 samples. Then, the objective function $f(\theta)$ for Bayesian optimization of the hyperparameters θ is defined as the average validation loss:

$$f(\theta) = \frac{1}{k}\sum_{i=1}^{k} L_{val}^{(i)}(\theta) \tag{7}$$

which is minimized to automatically identify the fine-tuned values for the hyperparameters θ.

Three ML algorithms, polynomial regression (PR), Gaussian process regression (GPR), and extreme gradient boosting (XGB), are compared to select the best performing algorithm. Various hyperparameters are selected for automatic tuning using Bayesian optimization which are explained in Appendix B. The Python libraries scikit-learn [25] and xgboost [26] are used to train and tune the ML models. Hyperparameters whose range spans multiple orders of magnitude are sampled with logarithmic uniform sampling while linear uniform sampling is used for others. The average training and test errors of the three algorithms are given in Table 3 where bold indicates the best performing algorithm for a given objective. The low training errors indicate that none of the algorithms underfit the training data, which means a good level of learning is achieved. The small difference between the training and test errors indicate that the algorithms do not overfit to the training data and perform well in predicting the test samples, indicating their overall success in generalizing to unseen data. However, it is seen that PR

Table 3
Training and test performances of various machine learning algorithms. Bold indicates the best performing algorithm for a given response.

	Average Error (%)					
	Training			**Test**		
Response	PR	GPR	XGB	PR	GPR	XGB
Mass	0.13	**0.06**	0.36	0.14	**0.06**	1.04
SN_Disp	0.70	0.43	**0.27**	0.78	**0.45**	0.78
TB_Disp	3.49	1.93	**0.86**	5.60	2.72	**1.72**
Stress_Bed	5.88	3.30	**1.29**	8.51	4.00	**2.88**
Stress_Col	4.07	3.18	**1.34**	4.13	4.56	**3.64**
Stress_XCar	5.99	4.03	**1.94**	6.73	5.41	**4.56**
Stress_YCar	5.41	4.47	**1.73**	5.91	6.08	**5.66**
Stress_ZCar	2.83	2.29	**0.95**	**2.70**	2.73	2.87
Stress_Spndl	0.97	0.81	**0.39**	**0.68**	1.05	1.01
Mode1	0.81	**0.44**	0.63	1.08	**0.58**	1.54
Mode2	0.69	**0.34**	0.58	0.72	**0.42**	1.52
Mode3	1.80	1.10	**0.49**	2.41	1.25	**0.79**
Mode4	0.71	0.40	**0.38**	1.13	**0.58**	0.91
EnergyXX	7.75	4.26	**1.77**	11.59	5.92	**4.79**
EnergyYY	3.91	2.24	**1.32**	5.41	**2.77**	3.06
Average	3.01	1.95	**0.95**	3.83	2.57	**2.45**

achieved a lower training and test accuracy than the other two algorithms.

GPR and XGB achieved overall similar test performances, although XGB achieved a slightly higher test accuracy than GPR. An advantage of GPR over XGB is its ability to provide a prediction together with an associated uncertainty, indicating its level of confidence in each output. Hence, GPR is commonly used for surrogate-assisted optimization where a posterior mean and variance provided by the surrogate are used to decide whether to accept the surrogate prediction or run the expensive evaluation (e.g., with the FEA model) [27]. While this approach can be used within the proposed system-level framework, for maximized computational efficiency, this study aims to avoid expensive FEA evaluations within the optimization iterations. Therefore, focus is placed on maximizing the prediction accuracy of the surrogate model as achieved through model training using the proposed two-step sampling approach with successful coverage of the design space. Accordingly, XGB is selected as the algorithm for the machine tool surrogate model. The complete training and test results of XGB are given in Appendix C where most test samples are predicted with less than 10 % error. Finally, a convergence study is carried out to analyze the effect of training sample count on the average test errors to identify the minimum number of training samples to achieve sufficient accuracy. As seen in Fig. 7, there is a sharp decrease in the average test errors between 40 and 80 training samples. However, test accuracy only marginally increases after 80 training samples. Hence, 80 training and 20 test samples are used to train and test the XGB model to limit the computational cost.

5. Global design optimization of machine tools

In a complex problem such as machine tool optimization, the objective functions can have multiple troughs and peaks. This creates a challenge for gradient-based optimization algorithms in identifying optimal designs due to their tendency to converge to local optima or saddle points [28]. On the other hand, evolutionary algorithms use a mechanism of representation, selection, crossover, and mutation to search the design space for optimal designs [29]. In representation, the population of solutions from the current iteration are encoded into a manipulable form such as binary or real-valued vectors while preserving the problem structure. Then, parent solutions are probabilistically selected from the candidates based on a calculated fitness value. Finally, new solutions are produced from the parent solutions through crossover and mutation. Crossover is an operator which mixes the represented components of parent solutions, while mutation injects variations into the represented components of individual parents. Hence, evolutionary algorithms are more robust against differences in the starting points or the gradient of the objective functions.

Among the multi-objective evolutionary optimization algorithms from the Pymoo Python library [30] compared in this study, NSGA-III is selected for optimization due to performing better than the others–C-TAEA and SMS-EMOA–in terms of optimization performance. The comparison results are shown in Appendix D, and a detailed discussion is given in Section 5.3. Instead of a single optimal design, NSGA-III identifies multiple designs which are referred to as Pareto-optimal designs. A Pareto-optimal machine tool design cannot be improved in any of its responses without worsening at least one other response. Compared to non-Pareto-optimal designs, Pareto-optimal designs are better in at least one response and as good in all other responses. Identifying a set of Pareto-optimal designs instead of a single optimal design provides insight into the trade-offs among conflicting objectives while allowing selection among alternatives based on specific preferences or constraints. This section discusses evolution-based optimization of machine tools using the trained ML model from Section 4 as the surrogate model for rapid evaluation of the objective functions.

5.1. Multi-objective optimization without constraints

The optimization objectives defined for the machine responses are given in Table 4. Since NSGA-III iteratively minimizes all objectives F1–15, the minimization or maximization of the responses are determined by the signs used in the objective definitions. The minimization objectives F1–9, F14, F15 are defined as equal to the responses R1–9, 14, 15, while the maximization objectives F10–13 are defined by inverting the signs of the responses R10–13. As seen, no constraints are defined on the objectives, allowing NSGA-III to follow the objectives without trying to achieve any specific conditions for the machine's responses. Hypervolume (HV), which is the volume of the objective space dominated by the non-dominated solutions, is used as a numerical measure of the

Table 4
The optimization objectives (obj.) defined for the machine tool responses (res.) in unconstrained optimization. MIN. and MAX. indicate minimization and maximization, respectively.

Response	Res. ID	Obj. Definition	Obj. Description	Constraint
Mass	R1	F1 = R1	MIN. F1 (MIN. R1)	-
SN_Disp	R2	F2 = R2	MIN. F2 (MIN. R2)	-
TB_Disp	R3	F3 = R3	MIN. F3 (MIN. R3)	-
Stress_Bed	R4	F4 = R4	MIN. F4 (MIN. R4)	-
Stress_Col	R5	F5 = R5	MIN. F5 (MIN. R5)	-
Stress_XCar	R6	F6 = R6	MIN. F6 (MIN. R6)	-
Stress_YCar	R7	F7 = R7	MIN. F7 (MIN. R7)	-
Stress_ZCar	R8	F8 = R8	MIN. F8 (MIN. R8)	-
Stress_Spndl	R9	F9 = R9	MIN. F9 (MIN. R9)	-
Mode1	R10	F10 = -R10	MIN. F10 (MAX. R10)	-
Mode2	R11	F11 = -R11	MIN. F11 (MAX. R11)	-
Mode3	R12	F12 = -R12	MIN. F12 (MAX. R12)	-
Mode4	R13	F13 = -R13	MIN. F13 (MAX. R13)	-
EnergyXX	R14	F14 = R14	MIN. F14 (MIN. R14)	-
EnergyYY	R15	F15 = R15	MIN. F15 (MIN. R15)	-

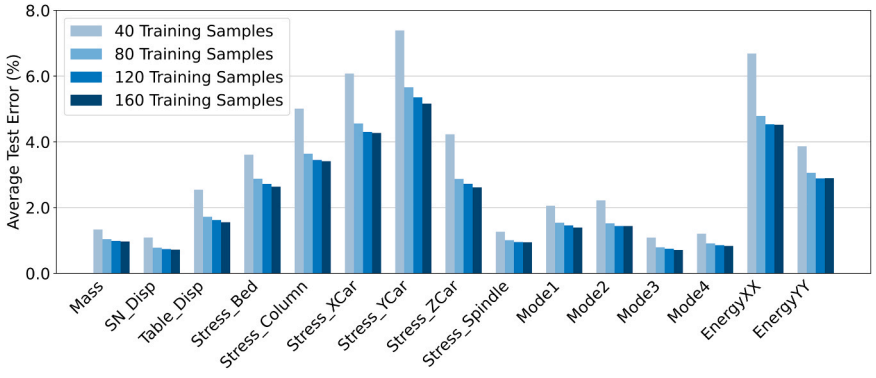


Fig. 7. The effect of the number of training samples on XGB's average test error.

success of multi-objective optimization [31]. The objective space is normalized with respect to the utopian point to quantify the optimization success relative to an idealized, unattainable upper bound of the objective space. With normalization, the values of HV remain between zero and one where values closer to one indicate greater optimization success. The utopian point P_U is generated as the set containing the minimum objective values achieved as:

$$P_U = (\min\ F_1,\ \min\ F_2,\ \ldots,\ \min\ F_{15}) \tag{8}$$

Then, the maximum possible HV is calculated as:

$$HV_{\max} = \prod_{i=1}^{15} (1.35 - P_U[i]) \tag{9}$$

where 1.35 represents the lower bound reference selected based on the estimated Nadir point. Finally, to eliminate the effect of random initialization, optimization is repeated for 30 runs with the same parameters where Eqs. (8) and (9) are calculated for each run. The median HV value across 30 runs for each generation is reported as the median normalized HV as shown in Fig. 8. As seen, although HV converges over generations, the low normalized HV values indicate that only a small portion of the objective space is dominated, indicating low optimization success. This can also be seen in Table 5 which gives the responses of the non-dominated designs, whose dimensions are given in Appendix E-I, in comparison to those of the base design. The down arrows (↓) indicate the worsened responses (i.e., greater than the base response in minimized objectives, lower than the base response in maximized objectives), while the up arrows (↑) represent the improved responses. As seen, optimization success remains low for most of the responses. The next section presents a novel constrained optimization method which significantly improves the performance of the identified non-dominated designs.

5.2. Multi-objective optimization with constraints

Inequality constraints are defined to improve the optimization performance for the worsened responses R2–5, R8, R10–14, and R15. An inequality constraint G defined on the response r specifies the feasibility conditions used by NSGA-III where a candidate design d is feasible if:

$$G_r^{(d)} \leq 0 \tag{10}$$

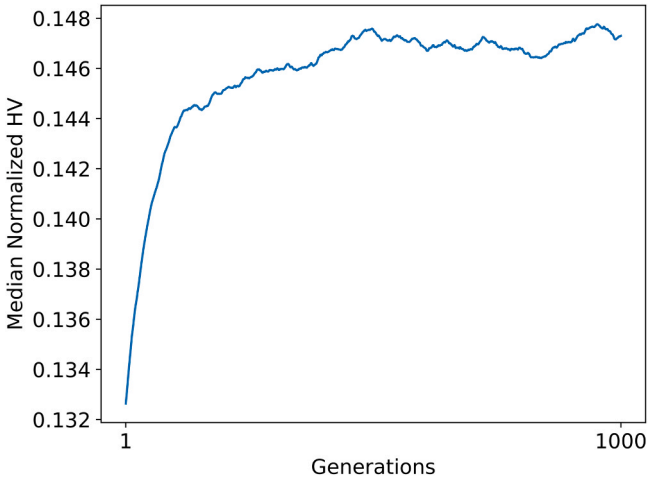


Fig. 8. Median normalized hypervolume (HV) for the unconstrained case.

which leads to the feasible designs satisfying the constraints getting ranked higher than the infeasible designs in the selection phase. As a result, the infeasible results are dominated by the feasible results and are not included in the non-dominated output of NSGA-III. Initially, the constraints are defined to force the non-dominating designs to perform at least as good as the base design, as commonly performed in existing methods [2,8,9,13], in the responses R2–5, R8, R10–14, and R15 by enforcing:

$$G_r^{(d)} = g_r^{(d)} - R_r^{(Base)} \leq 0 \tag{11}$$

for minimization objectives (i.e., R2–5, R8, R14, R15), and,

$$G_r^{(d)} = R_r^{(Base)} - g_r^{(d)} \leq 0 \tag{12}$$

for maximization objectives (i.e., R10–13) where $g_r^{(d)}$ and $R_r^{(Base)}$ are the responses of a candidate design d and the base design, respectively. However, no feasible designs are initially identified, indicating that the optimization problem becomes overly constrained due to the addition of Eqs. (11) and (12). To balance the severity of the inequality constraints, a relaxation factor C is introduced as:

$$G_r^{(d)} = g_r^{(d)} - (1 + C_r)\, R_r^{(Base)} \leq 0 \tag{13}$$

and,

$$G_r^{(d)} = (1 - C_r) \quad R_r^{(Base)} - g_r^{(d)} \leq 0 \tag{14}$$

The relaxation factor allows a small amount of deviation from the goal of achieving the same performance of the base design. As given by Eq. (13) for minimization objectives, it allows exceeding the base design's responses by C_r in percentage of the response r. Similarly, Eq. (14) gives that the relaxation factor allows undershooting the base design's response by a defined percentage C_r for the response r.

As shown in Table 6, a small relaxation factor of 5 % (i.e., $C_r = 0.05$) is applied for nine of the worsened responses while a factor of 25 % is used for the FRF energies R14 and R15 as they represent squared properties. The resulting non-dominated set of Pareto-optimal designs aggregated from 30 NSGA-III runs and their corresponding responses predicted by the surrogate ML model are given in Appendix E-II and Table 7, respectively. As seen in Table 7, varying amounts of reduction are achieved in the design variables with many designs showing significant reductions particularly in the wall thicknesses of the moving components. Furthermore, it is observed in Table 7 that the 11 responses, which previously performed worse than the base design, are significantly improved through the application of relaxed constraints. The first three natural frequencies that were previously lower than those of the base design now surpass the base design as marked by the up arrows. The angled arrows (↘) mark the responses which performed worse than the base design while remaining within the ranges defined by the relaxed constraints where the cell values indicate the signed percent change with respect to the corresponding response of the base design. As seen, most of the resulting non-dominated designs use only a portion of the allowed relaxation percentages demonstrating that the proposed constraint relaxation method acts as a general guideline rather than a strict boundary. Therefore, the proposed constraint relaxation method enables the discovery of improved designs that could be missed in unconstrained optimization or traditional optimization with strict constraints, as commonly observed in existing approaches.

To score the non-dominated designs by prioritizing the dimensional reduction achieved in the design variables relating to the moving components, namely, *XCar_T*, *YCar_T*, and *ZCar_T*, a variable, S_{MC} (MC: moving components), is introduced. For a non-dominated design d, S_{MC}

Table 5
The responses of the non-dominated designs in comparison to the base design for unconstrained optimization. The down arrows (↓) indicate the worsened responses (i.e., greater than the base response in minimized objectives, lower than the base response in maximized objectives), while the up arrows (↑) represent the improved responses. The cell values show the signed percent change with respect to the base design.

Design	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15
1	↑	↓3.54	↑	↑	↓28.32	↑	↑	↓43.53	↑	↓-26.20	↓-15.20	↓-26.04	↓-19.51	↓33.22	↓74.47
2	↑	↓1.56	↓6.27	↓14.22	↑	↓46.53	↑	↓5.04	↑	↓-19.38	↓-8.32	↓-23.47	↓-16.80	↓5.07	↓18.97
3	↑	↓3.08	↓11.88	↓18.55	↑	↓43.43	↑	↓3.90	↑	↓-20.52	↓-12.74	↓-20.90	↓-13.34	↓10.26	↓30.75
4	↑	↓0.13	↓1.03	↓1.62	↑	↑	↓2.08	↓35.03	↑	↓-26.48	↓-17.48	↓-25.55	↓-19.94	↓15.66	↓28.23
5	↑	↓1.89	↓13.02	↓31.46	↓1.19	↑	↑	↓31.45	↑	↓-16.33	↓-9.50	↓-21.20	↓-15.07	↓20.81	↓36.90
6	↑	↓3.08	↑	↑	↓7.79	↑	↑	↓31.21	↑	↓-28.45	↓-19.68	↓-25.67	↓-19.10	↓25.54	↓66.87
7	↑	↓4.17	↑	↑	↓8.29	↑	↑	↓30.24	↑	↓-28.44	↓-19.47	↓-25.59	↓-18.74	↓29.76	↓67.26
8	↑	↓3.70	↓10.58	↓16.32	↑	↑	↓5.31	↓11.15	↑	↓-21.35	↓-13.49	↓-21.10	↓-13.89	↓48.80	↓44.41
9	↑	↓1.36	↓0.29	↑	↑	↑	↓0.58	↓6.18	↑	↓-26.62	↓-17.48	↓-24.54	↓-18.57	↓13.63	↓31.66
10	↑	↓3.41	↑	↑	↓10.17	↑	↑	↓5.09	↑	↓-26.93	↓-18.36	↓-24.65	↓-18.15	↓44.33	↓46.69
11	↑	↓7.40	↓5.69	↓10.49	↓29.68	↑	↑	↓3.96	↑	↓-23.23	↓-15.78	↓-21.28	↓-13.65	↓31.09	↓87.00
12	↑	↓1.56	↓1.15	↓1.77	↑	↑	↑	↓6.50	↑	↓-25.98	↓-16.91	↓-24.54	↓-18.58	↓5.24	↓25.57
13	↑	↓6.18	↑	↑	↓39.16	↑	↑	↓2.58	↑	↓-25.38	↓-14.57	↓-24.78	↓-17.81	↓21.25	↓72.27
14	↑	↓5.52	↓10.31	↓17.24	↓17.20	↑	↑	↓2.54	↑	↓-18.02	↓↓-8.17	↓-21.23	↓-13.50	↓11.54	↓35.11
15	↑	↓1.50	↓3.67	↓9.37	↓7.74	↑	↑	↓11.44	↑	↓-20.40	↓-9.47	↓-24.44	↓-18.12	↓7.10	↓17.96
16	↑	↓3.33	↓10.68	↓16.09	↓6.87	↑	↑	↓5.13	↑	↓-20.40	↓-12.45	↓-21.73	↓-14.03	↓9.20	↓29.93
17	↑	↓4.57	↓8.80	↓15.73	↓14.92	↑	↓3.56	↓2.44	↑	↓-21.28	↓-13.33	↓-21.37	↓-14.07	↓4.02	↓36.04

Table 6
The optimization objectives (obj.) and constraints defined for the machine tool responses (res.) in constrained optimization with relaxation factors. MIN. indicates minimization.

Response	Res. ID	Obj. Definition	Obj. Description	Constraint
Mass	R1	F1 = R1	MIN. F1 (MIN. R1)	-
SN_Disp	R2	-	-	≤ 105 % of base
TB_Disp	R3	-	-	≤ 105 % of base
Stress_Bed	R4	-	-	≤ 105 % of base
Stress_Col	R5	-	-	≤ 105 % of base
Stress_XCar	R6	F6 = R6	MIN. F6 (MIN. R6)	-
Stress_YCar	R7	F7 = R7	MIN. F7 (MIN. R7)	-
Stress_ZCar	R8	-	-	≤ 105 % of base
Stress_Spndl	R9	F9 = R9	MIN. F9 (MIN. R9)	-
*Mode*1	R10	-	-	95 % of base ≥
*Mode*2	R11	-	-	95 % of base ≥
*Mode*3	R12	-	-	95 % of base ≥
*Mode*4	R13	-	-	95 % of base ≥
EnergyXX	R14	-	-	≤ 125 % of base
EnergyYY	R15	-	-	≤ 125 % of base

is calculated as:

$$S_{MC}^{(d)} = 2 - 2\left(\frac{XCar\_T^{(d)}}{XCar\_T^{(Base)}} + \frac{YCar\_T^{(d)}}{YCar\_T^{(Base)}} + \frac{YCar\_T^{(d)}}{YCar\_T^{(Base)}}\right) \Big/ 3 \tag{15}$$

which is between zero and one where values closer to one indicate designs with lighter moving components. Another variable, S_{BDD} (BDD: base design domination), is defined to score the non-dominated designs based on the number of the base design's responses they dominated as marked by the up, angled, and down arrows in Table 7. S_{BDD} is defined as:

$$S_{BDD}^{(d)} = n_{up} - (n_{down} + \max(C)\, n_{angled}) \tag{16}$$

where $\max(C)$ is the maximum relaxation factor used for any of the constraints, which is 0.25 for this problem and can be adjusted for other problems. Then, a final rating variable S_F is defined as the multiplication of variables S_{MC} and S_{BDD} for a given design d as:

Table 7
The responses of the non-dominated designs in comparison to the base design for constrained optimization, sorted in final design score S_F. The up arrows (↑) show improved responses; down arrows (↓) show worsened responses outside constraints; angled arrows (↘) show worsened responses within relaxed constraints. The cell values show the signed percent change with respect to the base design.

Design	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	S_{MC}	S_{BDD}	S_F
19	↑	↘1.3	↘3.4	↘3.8	↑	↓37.5	↑	↘5.0	↑	↑	↑	↑	↘-3.5	↘8.4	↘18.6	0.25	4.25	1.06
20	↑	↘1.3	↘3.4	↘3.8	↑	↓37.2	↑	↘4.9	↑	↑	↑	↑	↘-3.4	↘8.1	↘18.4	0.25	4.25	1.04
11	↑	↘1.5	↘3.3	↘2.3	↑	↓38.9	↑	↘4.6	↑	↑	↑	↑	↘-3.4	↘5.3	↘22.4	0.24	4.25	1.00
14	↑	↘1.2	↘3.6	↘1.9	↘2.6	↓50.3	↑	↘5.1	↑	↑	↑	↑	↘-3.7	↘12.0	↘23.7	0.30	3.00	0.89
16	↑	↘1.5	↘4.9	↘4.6	↘4.6	↓81.5	↑	↘4.9	↑	↑	↑	↑	↘-3.9	↘8.6	↘24.4	0.28	3.00	0.84
18	↑	↘1.4	↘4.1	↘0.4	↘3.2	↓81.3	↑	↘4.4	↑	↑	↑	↑	↘-3.7	↘8.3	↘24.1	0.27	3.00	0.82
8	↑	↘1.3	↘2.2	↘1.5	↘4.0	↓44.8	↑	↘4.9	↑	↑	↑	↑	↘-3.6	↘9.1	↘20.7	0.25	3.00	0.75
4	↑	↘1.6	↘3.6	↘0.4	↘4.0	↓53.6	↑	↘4.9	↑	↑	↑	↑	↘-3.6	↘6.8	↘24.6	0.25	3.00	0.75
10	↑	↘2.1	↘1.5	↘0.4	↘4.0	↓19.2	↑	↘4.7	↑	↑	↑	↑	↘-3.6	↘4.8	↘23.2	0.19	3.00	0.58
6	↑	↘2.2	↘0.5	↘1.6	↘4.6	↑	↑	↘3.6	↑	↑	↑	↑	↘-2.9	↘5.9	↘23.9	0.12	5.00	0.58
13	↑	↘2.0	↘1.9	↘3.9	↘2.2	↑	↑	↘4.0	↑	↑	↑	↑	↘-2.9	↘5.8	↘18.0	0.11	5.00	0.56
3	↑	↘2.3	↘4.8	↘3.9	↘4.4	↓81.4	↓4.8	↘4.3	↑	↑	↑	↑	↘-4.0	↑	↘22.9	0.24	2.25	0.54
5	↑	↘2.1	↘3.0	↘3.8	↘4.6	↓22.2	↑	↘3.5	↑	↑	↑	↑	↘-3.5	↘4.6	↘17.1	0.17	3.00	0.52
2	↑	↘2.1	↘0.9	↘2.9	↘3.5	↑	↑	↘4.7	↑	↑	↑	↑	↘-2.8	↑	↘17.2	0.08	6.25	0.50
7	↑	↘2.2	↘0.7	↘2.3	↘2.0	↑	↓4.6	↘4.5	↑	↑	↑	↑	↘-2.9	↘7.1	↘24.2	0.15	3.00	0.44
15	↑	↘2.3	↘1.3	↘3.8	↘2.2	↑	↓4.1	↘4.9	↑	↑	↑	↑	↘-2.9	↘0.8	↘23.7	0.09	3.00	0.28
17	↑	↘2.1	↘2.0	↘2.3	↘3.3	↓46.8	↓4.0	↘4.3	↑	↑	↑	↑	↘-3.8	↘6.8	↘24.6	0.28	1.00	0.28
9	↑	↘1.9	↘3.5	↘2.8	↘3.6	↓49.8	↓1.3	↘4.5	↑	↑	↑	↑	↘-3.8	↘7.1	↘22.6	0.28	1.00	0.28
12	↑	↘2.2	↘1.8	↘2.2	↘3.0	↓22.9	↓5.2	↘5.0	↑	↑	↑	↑	↘-3.5	↘6.2	↘23.2	0.25	1.00	0.25
1	↑	↘2.0	↘3.0	↘3.8	↘1.6	↓21.8	↓0.7	↘4.6	↑	↑	↑	↑	↘-3.5	↘7.1	↘20.0	0.23	1.00	0.23

$$S_F^{(d)} = S_{MC}^{(d)} S_{BDD}^{(d)} \tag{17}$$

As observed in Table 7, Eq. (17) produces a higher rating for non-dominated designs that achieve greater amounts of dimensional reduction and exceed the performance of the base design's responses. The rating behavior can be adjusted by modifying Eqs. (15) and (16). For example, Design 20 rates higher than Design 2 with the current implementation of Eqs. (15) – (17), which is largely due to the low S_{MC} score Design 2 received. It is seen in Appendix E-II that this is primarily because of the low amount of reduction achieved on *XCar_T*. On the other hand, Table 7 shows that Design 2 received a higher S_{BDD} score than Design 20, demonstrating its dominance over more of the base design's responses. Another optimization case could focus on optimizing only the tool's kinematic chain which includes the components column, Y-axis carriage, and Z-axis carriage. This would leave X-axis carriage out as it belongs to the table's kinematic chain. In this case, Design 2 would receive a higher S_{MC} score and an S_F rating than Design 20. Therefore, the adjustable behavior of the proposed design rating method can adapt to various optimization problems with varying priorities.

5.3. The effect of algorithms and algorithm parameters

Crossover and mutation can significantly impact the diversity of identified optimal designs in evolutionary algorithms [31]. The population size and the number of generations are also important which influence the extent of an evolutionary algorithm's exploration of the objective space. Additionally, the choice of evolutionary algorithm can substantially impact the identified optimal designs. This section presents a set of comprehensive comparisons to demonstrate the effect of different evolutionary algorithms as well as their parameters on the median normalized HV as an indicator of overall optimization success. As before, the median normalized HV from 30 runs is used to eliminate the effects of random initialization. All cases are run for 200 generations to ensure HV convergence is achieved.

The effect of population size is shown in Appendix D-I. Three partition counts, 6, 9, and 12, are selected for Das Dennis partitioning [32] which correspond to the population sizes 84, 220, and 455, respectively. As seen, increased population size results in increased HV, indicating a more thorough exploration of the objective space. The effect of evolutionary parameters crossover and mutation is demonstrated in Appendix D-II and III for two population sizes. In both cases, HV is observed to vary minimally with changing values of crossover probability *COprob*, crossover distribution index *COeta*, and mutation distribution index *MUeta*. Despite the minimal effect, *COprob*, *COeta*, and *MUeta* are set to 0.9, 5, and 5, respectively, to promote diversity among the solutions and encourage design exploration. The effect of the optimization algorithm is demonstrated in Appendix D-IV and V by comparing the median normalized HV values of the NSGA-III, C-TAEA, and SMS-EMOA algorithms. The same set of values are used for *COprob*, *COeta*, and *MUeta* to encourage exploration as before while the default values are used for any remaining parameters. While the details of evolutionary algorithms C-TAEA and SMS-EMOA are beyond the scope of this paper, C-TAEA is a constrained multi-objective algorithm balancing convergence, diversity, and feasibility, and SMS-EMOA is another multi-objective algorithm that selects solutions by maximizing the dominated hypervolume. As seen in Appendix D-IV and V, NSGA-III performed the best among the three algorithms.

The number of fitness evaluations performed by NSGA-III is equal to the multiplication of the population size and the number of generations. To further reduce the computational cost of NSGA-III, the effect of varying the population size and the number of generations while maintaining the total number of fitness evaluations is investigated based on the results in Appendix D-I. The final HV value is defined as the average of the last 10 generations. Then, the number of generations required to reach 98 % of the reference HV, called the convergence HV, is defined as the convergence generations. For the low-population case (Appendix D-I, blue), the convergence HV and the number of convergence generations are calculated as 0.75 and 152, respectively. The total number of fitness evaluations for the low-population case is then calculated as $84 \times 152 = 12768$, which is used as the reference for the middle- and high-population cases (Appendix D-I, orange and green). The number of convergence generations for the middle- and high-population cases are calculated as $12768/220 \cong 58$ and $12768/455 \cong 28$, respectively. The constrained NSGA-III procedure with relaxed constraints is performed for the low-, middle-, and high-population cases by limiting the number of generations to the corresponding convergence generations of each case. The resulting median normalized HV values and the computation times are given in Fig. 9. As seen, although a negligible difference is observed in HV, increasing the number of populations while decreasing the number of generations significantly reduces computation time. This can be attributed to the reduced number of generations decreasing the number of sequential calculations while HV is maintained due to the increased population size. Therefore, for increased computational efficiency, the primary modification should be to reduce the number of generations rather than the population size.

6. Comparison of the base and optimized designs

Although the surrogate ML model is used to efficiently guide optimization, final design validation is based on FEA results for accuracy. The optimized designs 2 and 20, whose dimensions are shown in Appendix E-II, are applied the same load cases shown in Table 1 to obtain the corresponding responses. The mass comparison between the base and optimal designs is given in Table 8. As seen, significant amounts of mass reduction up to nearly 20 % is achieved across the machine tool components despite using a small number of design variables. The

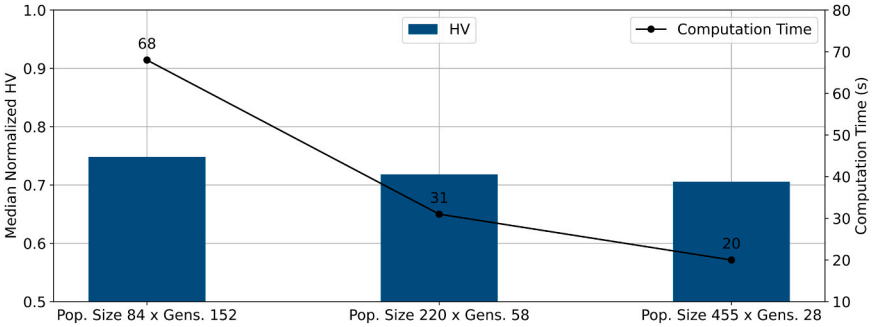


Fig. 9. The effect of population size and the number of generations on median normalized hypervolume (HV) and total computation time under constant fitness evaluation count (population size × number of generations).

Table 8
Mass comparison between the base and optimized designs. The masses of any rails and runners are not included.

Component	Mass (tonne)			Difference (tonne)		Difference (%)	
	Base	Design 2	Design 20	Design 2	Design 20	Design 2	Design 20
Bed	3.316	3.375	3.390	0.059	0.074	+ 1.8	+ 2.2
Column	2.730	2.431	2.630	-0.299	-0.100	-10.9	-3.7
X-axis Carriage	0.652	0.617	0.523	-0.035	-0.129	-5.3	-19.8
Y-axis Carriage	0.458	0.388	0.387	-0.070	-0.071	-15.3	-15.5
Z-axis Carriage	0.378	0.369	0.357	-0.009	-0.021	-2.5	-5.5

Table 9
Machine response comparison between the base and optimized designs. The up arrows (↑) show improved responses; down arrows (↓) show worsened responses outside constraints; angled arrows (↘) show worsened responses within relaxed constraints.

Response	Units	Design			Change from Base (%)	
		Base	Design 2	Design 20	Design 2	Design 20
Mass	tonne	7.985	↑7.631	↑7.738	-4.43	-3.09
SN_Disp	mm	0.251	↘0.256	↑0.251	+ 1.90	+ 0.12
TB_Disp	mm	0.022	↑0.022	↘0.022	-2.06	+ 1.96
Stress_Bed	MPa	9.92	↑9.77	↘10.15	-1.49	+ 2.31
Stress_Col	MPa	3.18	↑3.01	↑3.10	-5.29	-2.57
Stress_XCar	MPa	0.93	↑0.68	↓1.26	-26.63	+ 36.35
Stress_YCar	MPa	10.20	↓10.96	↓10.34	+ 7.49	+ 1.41
Stress_ZCar	MPa	10.20	↓10.80	↓11.44	+ 5.84	+ 12.14
Stress_Spndl	MPa	20.00	↑19.90	↑19.68	-0.50	-1.61
Mode1	Hz	37.0	↑39.1	↑38.6	+ 5.74	+ 4.30
Mode2	Hz	37.3	↑39.9	↑39.2	+ 6.96	+ 4.92
Mode3	Hz	92.5	↑93.9	↑93.1	+ 1.44	+ 0.63
Mode4	Hz	100.0	↘98.7	↘99.2	-1.35	-0.83
EnergyXX	10^{-6} $(mm/N)^2$	0.28	↑0.27	↑0.26	-3.74	-8.78
EnergyYY	10^{-6} $(mm/N)^2$	1.76	↘2.09	↘1.92	+ 19.02	+ 9.31

diversity of the two Pareto-optimal designs, Design 2 and Design 20, is evident in that Design 20 focuses on the moving components while Design 2 balances focus between the static and moving components. The diversity of optimal results enables selection based on specific priorities, highlighting the advantage of using NSGA-III and a Pareto front-based optimization approach over solving for a single optimal solution.

An interesting finding is that, despite limiting the design variables to remain below the base machine's dimensions (Table 2), it is seen in Table 8 that a small amount of mass is added to the bed in both optimal designs. The bed acts as the foundation of the machine tool, supporting the bending and torsional deformations of the column. The small amount of mass added to the bed balances the substantial mass reduction in the other components, preventing a significant loss of performance in the machine responses. A system-level balancing of added and removed mass may not be possible with existing methods as only certain parts of the machine are enabled for optimization at a time. These results demonstrate the proposed method's advantage over existing methods in thoroughly exploring the design space to find the most optimal designs.

The response comparison between the base and the two selected optimal designs is given in Table 9. Most of the machine's responses achieved better performance than the base design while most of the responses that performed worse closely approach the base design's performance. For example, as observed in Table 9, performance degradation in the table displacement, maximum principal stress in bed, and the fourth natural frequency of Design 20 remain under 2.31 %. The contribution of the thermal load to the displacements *SN_Disp* for the preferred optimized design 20 is calculated as 0.172 mm in comparison to the base design's 0.168 mm, indicating only a 2.3 % degradation in thermal stability. The thermal contribution to the table displacements *TB_Disp* is observed to be nearly zero for both the base and optimized designs. In terms of the spindle nose FRF energies as an indicator of dynamic stability of the machine, it is observed that both optimal designs' point FRFs in the X direction of the machine exceed the performance of the base design. In the Y direction, the optimal designs Design 2 and Design 20 experienced a performance reduction of 19.02 % and 9.31 % with respect to the base design, respectively. The most notable performance reduction is observed in the maximum principal stresses in the X, Y, and Z-axis carriages, which necessitates a thorough high-cycle fatigue life analysis to interpret as given in Section 6.1. A detailed process stability analysis, as given in Section 6.2, is also necessary to verify the validity of the proposed energy-based stability optimization approach.

6.1. Fatigue life comparison of the base and optimal designs

For cast iron containing flake graphite (GJL), the fatigue limit can be assumed to be between 25 – 35 % of its ultimate tensile strength (UTS) [33]. The 300-series cast iron such as FC300 which is widely used in machine tool structural components has a UTS of 300 MPa. Hence, a conservative fatigue limit for the machine tool material can be taken as 75 MPa. The machine tool's cast iron components achieve infinite fatigue life if the amplitude of the dynamic stresses due to the dynamic cutting forces remains below the fatigue limit σ_l of 75 MPa. On the other hand, mean tensile stress reduces fatigue life and must be included in high-cycle fatigue calculations. The stress amplitude can be corrected to include the mean stress effect using the Goodman model [34]:

Table 10
The maximum values of the assumed FEA-test correlation factor for infinite high-cycle fatigue life. The values are unitless and rounded to the nearest integer.

Design	Stress_Bed	Stress_Col	Stress_XCar	Stress_YCar	Stress_ZCar	Stress_Spndl
Base	28	93	332	30	30	15
Design 2	28	98	507	28	28	14
Design 20	27	92	228	29	26	15

$$\sigma_{a,eff} = \frac{\sigma_a}{1 - \frac{\sigma_m}{\sigma_{UTS}}}, \quad \sigma_{a,eff} \leq \sigma_l \tag{18}$$

where σ_m, σ_a, $\sigma_{a,eff}$, and σ_l are the mean, the amplitude, the effective amplitude of the maximum principal stress, and the fatigue limit, respectively. To simulate the dynamic loading from oscillating cutting forces, the X, Y, Z force components reported in Section 2 are taken as the maximum point of the cutting force while a coefficient of 25 % is applied to all three force components to apply the minimum cutting force components as −375 N, 250 N, and −500 N in the X, Y, and Z directions. The coupled thermal-structural FEA procedure is applied to the base and optimal designs using the minimum cutting force components. Then, σ_a and σ_m are calculated as:

$$\sigma_a = (\sigma_{max} - \sigma_{min})/2 \tag{19}$$

$$\sigma_m = (\sigma_{max} + \sigma_{min})/2 \tag{20}$$

where σ_{max} and σ_{min} are the maximum principal stresses calculated using the maximum and minimum cutting forces, respectively.

σ_{min}, σ_{max}, σ_a, σ_m, and $\sigma_{a,eff}$ for the base and optimal designs are given in Appendix F. Although all effective stress amplitudes are well below the fatigue limit of 75 MPa for FC300 cast iron, an FEA-test correlation factor $k = \sigma_a^{(Test)}/\sigma_a^{(FEA)} = \sigma_m^{(Test)}/\sigma_m^{(FEA)}$ is considered to account for potential simulation errors in stress calculations. Applying the correlation factor to σ_a and σ_m, the effective stress amplitude $\sigma_{a,eff}$ is calculated as:

$$\sigma_{a,eff,corr} = \frac{k\sigma_a}{1 - \frac{k\sigma_m}{\sigma_{UTS}}} \tag{21}$$

Then, the correlation factor k is obtained as:

$$k = \frac{\sigma_{a,eff,corr}}{\sigma_a + \sigma_{a,eff,corr}\frac{\sigma_m}{\sigma_{UTS}}} \tag{22}$$

Finally, the maximum value of k for infinite fatigue life can be calculated by substituting $\sigma_{a,eff,corr} = \sigma_l$ in Eq. (22):

$$k_{max} = \frac{\sigma_l}{\sigma_a + \sigma_l\frac{\sigma_m}{\sigma_{UTS}}} \tag{23}$$

which yields, with $\sigma_l = 75 \quad MPa$ and $\sigma_{UTS} = 300 \quad MPa$:

$$k_{max} = \frac{75}{\sigma_a + 0.25\sigma_m} \tag{24}$$

The maximum FEA-test correlation factors k_{max} for infinite life calculated through Eq. (24) are given in Table 10. The minimum correlation factor for the stresses exceeding those of the base design, namely, *Stress_XCar*, *Stress_YCar*, and *Stress_ZCar*, is observed to be 26. Then, the corresponding percent simulation error ε_{max} for $k_{max} = 26$ is calculated as:

$$\varepsilon_{max} = \frac{|1 - k_{max}|}{k_{max}} \times 100 = 96.15\% \tag{25}$$

which demonstrates that the base and optimal designs would have

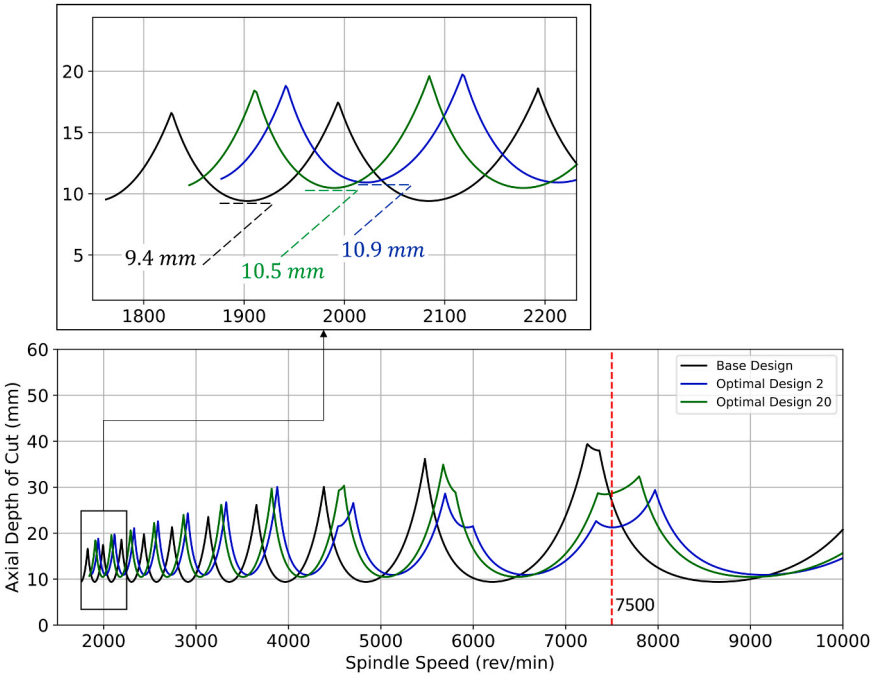


Fig. 10. Stability diagrams of the base and optimal designs calculated using FRFs from FEA. The black, blue, and green dashed lines indicate the absolute stability limits.

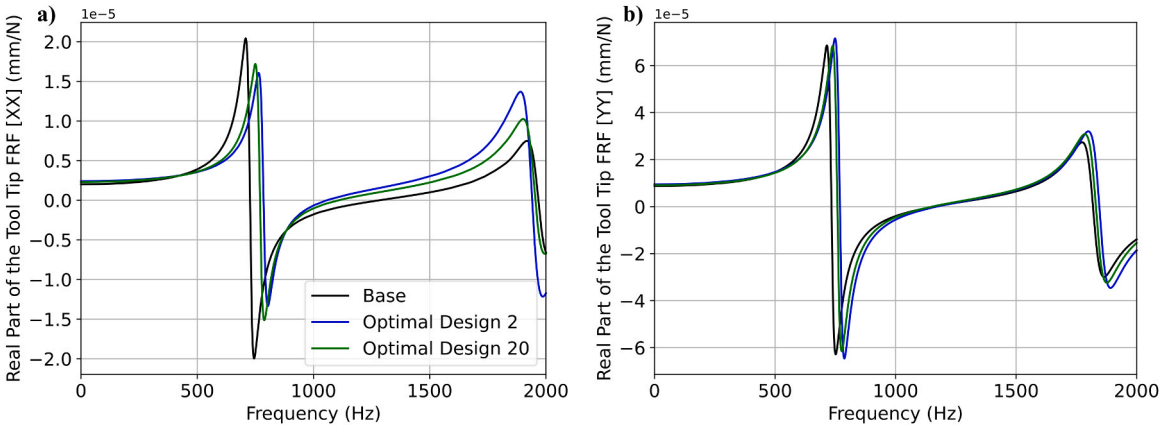


Fig. 11. The real parts of the tool tip FRFs in the I) X and II) Y directions used to calculate the stability diagrams.

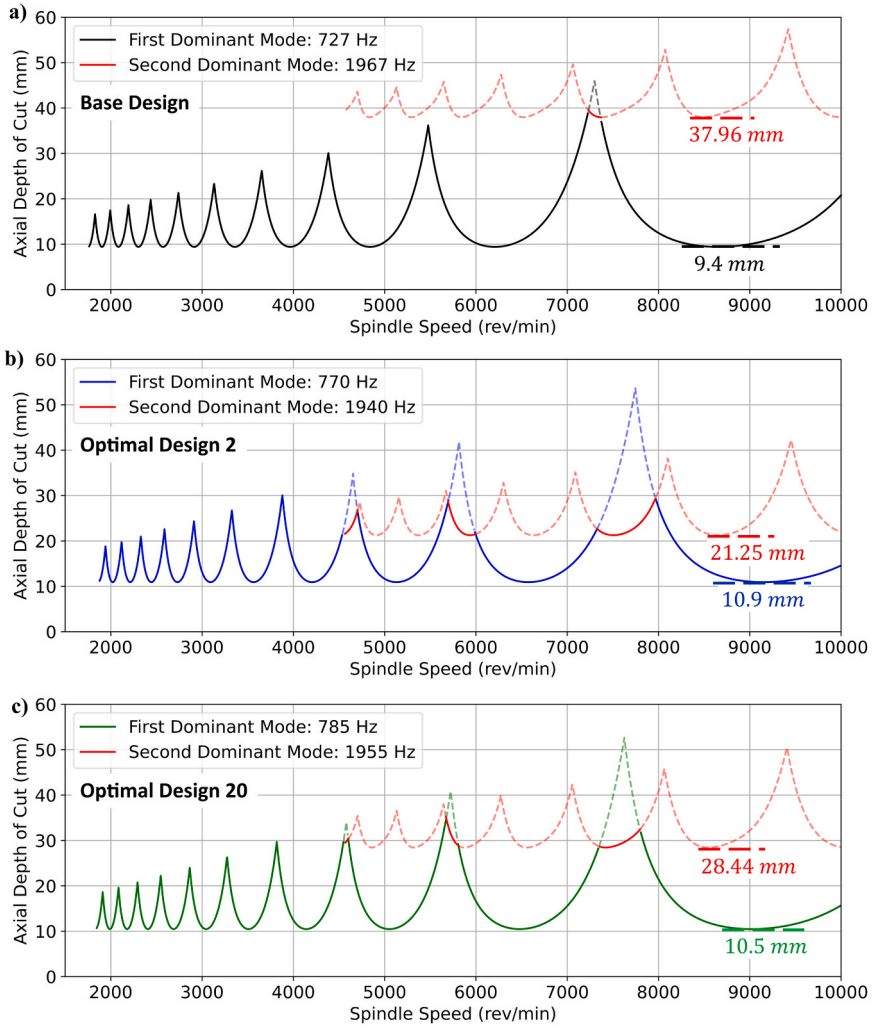


Fig. 12. Contributions of the first and second dominant modes of the tool+holder+machine assembly to the stability diagrams. The dashed and faded segments indicate trimmed regions.

infinite fatigue life even if the FEA results in Appendix F underpredicted the actual stresses by up to 96.15 %. In the absence of high-cycle fatigue life test results, as in this study, a probabilistic model can be used to estimate the probability of the FEA errors exceeding 96.15 % (i.e., probability of $k_{max} \geq 26$). As given in [35], a mean $m = 1.0$ and coefficient of variability $v = 0.1$ can be used to create FEA model uncertainty model based on log-normal distribution as:

$$\sigma_{std} = \sqrt{\ln(1+v^2)} = 0.099751$$

$$\mu = \ln m - \sigma_{std}^2/2 = -0.004975 \tag{26}$$

where σ_{std} and μ are the standard deviation and the mean of the log-

normal distribution. Then, the probability of k_{max} exceeding 26, which would invalidate the condition that the base and optimal designs achieve infinite fatigue life, is calculated as:

$$P(k_{max} \geq 26) = 1 - \Phi\left(\frac{\ln 26 - \mu}{\sigma_{std}}\right) \approx 5.26 \times 10^{-235} \approx 0 \quad (27)$$

where Φ is the standard normal cumulative distribution function. It is inferred from Eq. (27) that the base and optimal designs achieve infinite fatigue life with near certainty as there is nearly a zero probability that the assumed FEA-test correlation factor k_{max} exceeds 26. Therefore, despite experiencing slightly increased stresses compared to the base design (Table 9 and Appendix F), the optimal designs Design 2 and Design 20 are effectively equivalent to the base design in terms of high-cycle fatigue life while being significantly lighter.

6.2. Comparison of stability diagrams

The critical axial depths of cut for stable cutting are calculated using the zero-order solution for milling by Budak and Altintas [16]. The tooltip FRFs are calculated in FEA for a 15 mm diameter end-mill with two flutes, and the entry and exit angles of tool-workpiece engagement are taken as 10° and 170°, respectively. The cutting coefficients are taken as $K_t = 1800MPa$ and $K_r = 0.2$. The comparison of the stability diagrams is given in Fig. 10. As seen, both optimal designs achieve greater absolute stability limits, which are calculated as 10.9 mm and 10.5 mm for Design 2 and Design 20, respectively, in comparison to the base design's 9.4 mm. Furthermore, both optimal designs achieve improved dynamic stability below 4000 rev/min, while Design 20 maintains the dynamic stability of the base design in the 4000–6000 rev/min range. Therefore, the optimal design 20 improves the overall dynamic stability of the machine while achieving significant mass reduction across multiple components.

The only speed region where a loss of stability is observed is around 7500 rev/min, which is explained by the contribution of multiple dominant vibration modes observed in the tool tip FRFs to the stability diagrams. The stability diagrams in Fig. 10 are determined entirely by the first two dominant modes present in the tool tip FRFs, shown in Fig. 11, which are different from the machine responses *Mode*1 and *Mode*2. Fig. 11 shows that the optimal designs reduce the vibration amplitudes of the first dominant mode in the X direction while they remain similar to the base design in the Y direction, leading to increased stability contributed by the first dominant mode seen in Fig. 12. However, both optimal designs experience increased vibration amplitudes in the second dominant mode, reducing its stability contribution by pushing the lobes down vertically, which results in the first three stability lobes being trimmed as seen in Fig. 12. The lower FRF amplitudes of the optimal design 20 in comparison to those of the optimal 2 around the second dominant mode results in a smaller vertical shift of the lobes, limiting the stability loss due to trimming. These results demonstrate the consistency and success of the proposed energy-based FRF order reduction method in capturing the complete behavior of the full FRF curves.

It is seen in Fig. 11 that the first dominant mode also dominates the energy of the FRF curves due to containing larger amplitudes which get further amplified when squared in Eq. (2). Therefore, a minimization of the FRF energy as calculated by the proposed FRF order reduction method using Eq. (2) is likely to focus on minimizing the first dominant mode more than the others, as also observed in Fig. 11. The proposed energy-based FRF order reduction method can be used to individually control the stability contributions of different modes by applying Eq. (2) to each dominant mode of the FRFs to minimize or prevent stability loss in cases of multiple dominant modes. The important performance and energy advantages of mass reduction, particularly in the moving components, enabled by the proposed design optimization method come without a reduction of the machine's absolute stability limits. The trimming between the stability lobes is affected by many additional factors including the particular tool and holder used where changes in configurations are expected to result in varied maximum stability limits.

7. Conclusion

This paper presents, for the first time, a comprehensive framework for system-level multi-objective machine tool design optimization enabling simultaneous modification of machine tool components to achieve substantial mass reduction across the entire machine. The results show that the proposed method simultaneously balances mass addition and removal between the components to meet or exceed performance targets, which cannot be achieved by existing single-component methods. A small amount of mass is added to the bed, a stationary component, which enabled substantial mass reductions reaching up to 20 % across the other components while the performance indicators are improved or maintained with negligible degradation. The number of design variables are kept to a minimum by taking advantage of the linear guides to address the challenge of rapidly increasing computational cost due to system-level modeling involving multiple components. The proposed method avoids using machine-specific features as design variables, which is commonly exercised in existing approaches, as it would result in significant accumulation of variables, increasing the computational cost. The minimal design variable set allows the surrogate ML model to be trained with a limited number of samples while achieving good accuracy, minimizing training length, and making the method computationally accessible. With the proposed method, a multi-objective machine learning training and test set with 100 samples took 27 h to generate on a consumer-grade computer with 12 cores and 128 GB RAM.

Future studies could extend this research by integrating recent diffusion-based generative design methods [36] with the minimal parameter set approach as a computationally efficient hybrid. This could enable local optimizations without requiring component-level parameters while maintaining system-level optimizations across the machine tool. Also, the constraint relaxation factors could be interactively and automatically determined through a reverse design approach [37].

CRediT authorship contribution statement

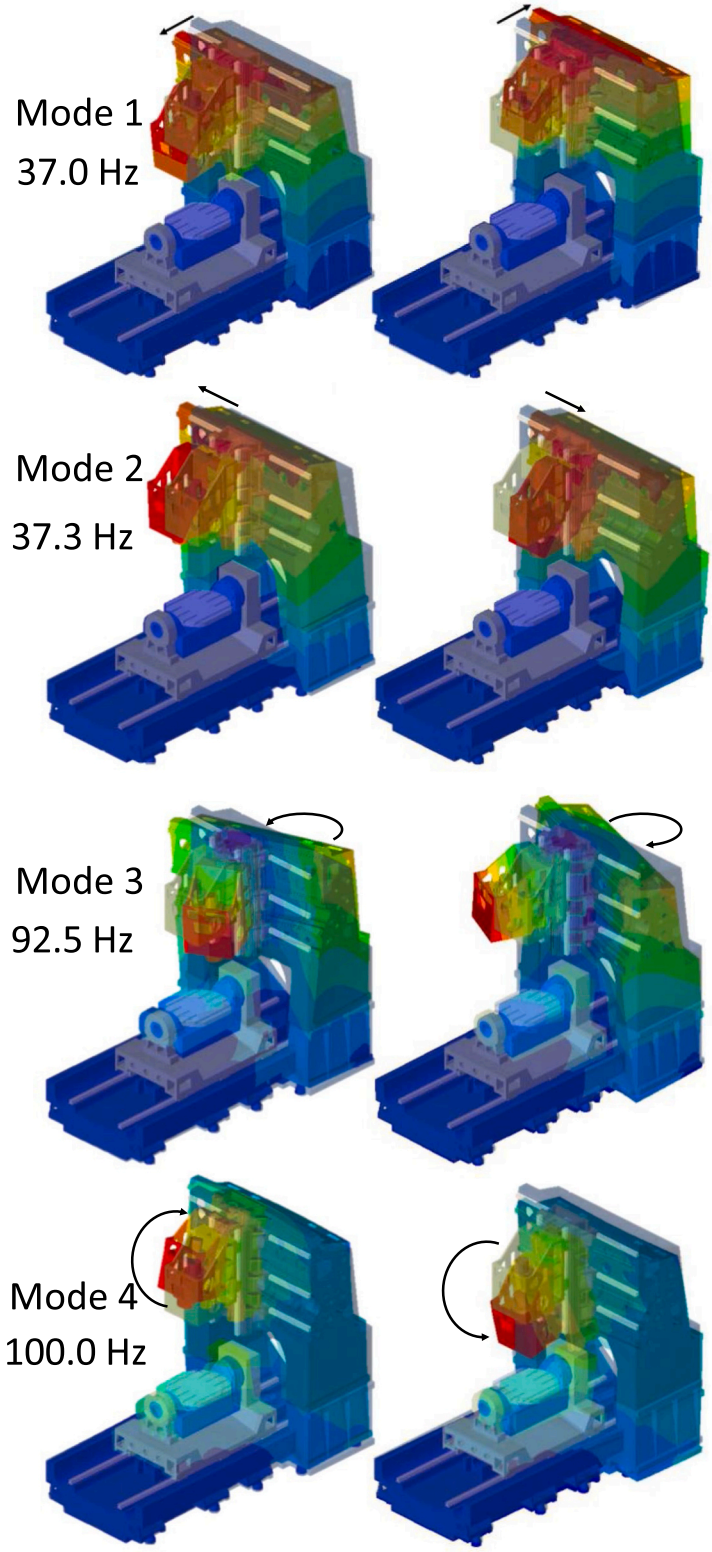
Deniz Bilgili: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jasmin Jelovica:** Writing – review & editing, Supervision, Resources. **Erhan Budak:** Writing – review & editing, Supervision, Resources.

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.

Appendix A

The first four elastic vibration modes of the base machine tool design.



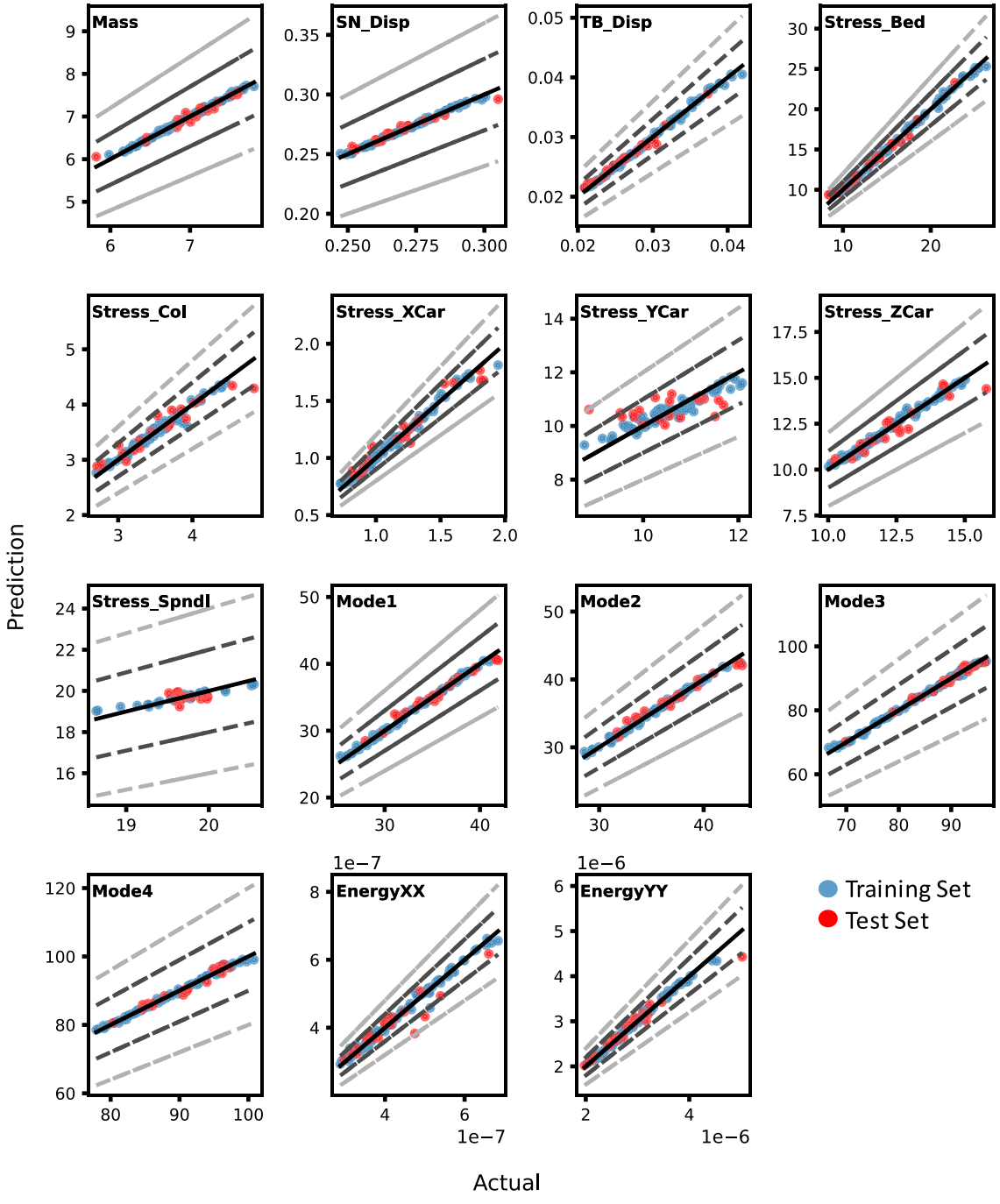
Appendix B

The hyperparameter space for Bayesian optimization-based tuning. The hyperparameters are named as the appear in the Python libraries scikit-learn [19] and xgboost [20].

Architecture	Hyperparameter	Type	Sampling	Min	Max
PR	*polynomialfeatures_degree*	Integer	Uniform	1	6
GPR	*kernel_length_scale*	Real	Log. Uniform	1.00E-07	1.00E+ 07
	alpha	Real	Log. Uniform	1.00E-10	1.00E+ 10
	n_estimators	Integer	Log. Uniform	1	1000
	learning_rate	Real	Log. Uniform	0.001	0.1
	max_*depth*	Integer	Uniform	1	5
	min_*child_weight*	Integer	Uniform	1	5
XGB	*alpha*	Real	Log. Uniform	0.001	10
	lambda	Real	Uniform	1	100
	gamma	Real	Uniform	0.01	0.50
	subsample	Real	Uniform	0.5	1
	colsample_bytree	Real	Uniform	0.5	1

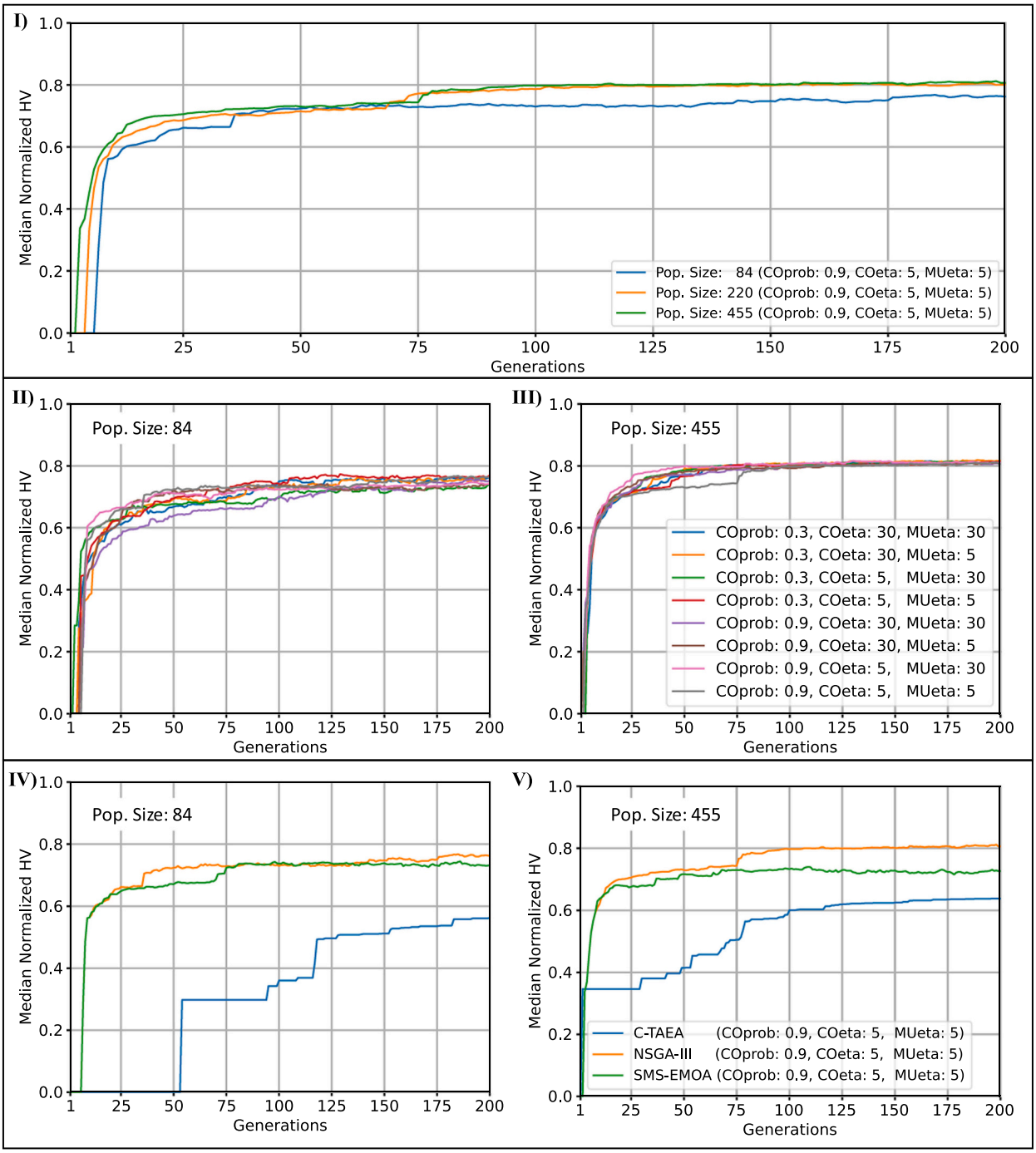
Appendix C

Training and test performances of the trained XGB model. The solid black, dashed black, and dashed gray lines indicate 0 %, 10 %, and 20 % prediction error, respectively.



Appendix D

The effect of I) population size, II) III) crossover and mutation parameters, and IV) V) the optimization algorithm on median normalized hypervolume (HV).



Appendix E

The non-dominated designs identified through I) unconstrained optimization and II) optimization with relaxed constraints. All values are in millimeters.

I)

Design	Bed_L	Col_L	Bed_T	Col_T	XCar_T	YCar_T	ZCar_T
Base	585.00	304.00	20.00	20.00	20.00	20.00	20.00
1	483.32	261.13	10.67	19.79	19.73	14.96	10.46
2	476.99	262.00	19.84	18.07	12.06	15.92	17.42
3	475.99	223.06	19.38	17.04	12.27	15.78	17.36
4	517.37	219.77	18.90	19.90	17.27	17.09	12.68
5	489.47	229.39	16.85	16.52	19.68	15.74	12.21
6	488.54	222.49	12.90	19.88	19.90	14.81	13.07
7	487.81	225.10	12.40	18.98	19.22	15.42	13.22
8	551.33	213.34	17.58	16.98	19.80	13.16	16.50
9	488.56	226.82	17.54	19.26	19.38	12.85	17.44
10	462.79	200.10	15.65	19.21	18.57	15.48	17.32
11	475.99	223.06	10.65	17.04	19.57	14.98	17.36
12	511.89	226.99	18.80	19.59	19.53	16.44	17.49
13	488.15	255.79	10.79	19.95	19.27	14.93	19.82
14	488.13	254.06	15.81	16.93	19.92	15.75	19.24
15	488.55	256.13	17.04	18.30	19.77	18.38	16.86
16	489.26	227.54	17.04	16.91	19.91	18.52	17.48
17	546.86	222.95	16.46	17.00	19.71	17.27	19.82

II)

Design	Bed_L	Col_L	Bed_T	Col_T	XCar_T	YCar_T	ZCar_T
Base	585.00	304.00	20.00	20.00	20.00	20.00	20.00
1	489.18	255.68	20.00	19.78	14.62	12.89	18.74
2	489.00	245.37	19.47	18.11	19.98	16.50	18.75
3	521.18	230.22	18.77	17.40	11.08	16.21	18.38
4	531.59	230.25	19.10	17.42	11.88	15.81	17.36
5	474.29	258.25	19.96	19.83	14.60	15.45	19.60
6	476.04	230.25	19.98	17.79	19.29	14.90	18.86
7	489.21	237.11	19.16	18.10	19.68	12.82	18.74
8	486.19	255.30	19.10	17.44	12.22	15.37	17.35
9	489.20	245.51	19.46	17.74	11.99	12.92	18.58
10	453.51	230.35	19.93	17.45	14.91	14.82	18.59
11	522.93	230.25	19.10	19.84	12.54	15.79	17.49
12	488.90	237.95	19.99	18.07	13.69	12.74	18.73
13	488.90	258.25	19.98	19.94	19.29	15.31	18.73
14	489.18	257.37	19.52	17.36	11.94	12.82	17.38
15	521.18	230.08	18.84	17.40	19.60	16.39	18.38
16	489.26	233.50	18.83	17.31	10.73	14.99	17.46
17	492.97	234.11	19.86	17.74	12.03	12.82	18.59
18	488.88	230.24	19.66	17.46	11.03	15.22	17.34
19	489.34	255.54	19.04	19.80	12.60	15.00	17.49
20	482.81	258.28	19.55	19.95	12.61	15.32	17.32

Appendix F

Stresses calculated based on the maximum principal stresses from the minimum and maximum cutting force cases in coupled thermal-structural FEA. All values are in MPa.

Design	Component	Stress_Bed	Stress_Col	Stress_XCar	Stress_YCar	Stress_ZCar	Stress_Spndl
	σ_{min}	9.269	3.152	0.941	10.345	10.298	19.586
	σ_{max}	9.920	3.184	0.925	10.245	10.191	19.955
Base	σ_a	0.326	0.016	-0.008	-0.050	-0.053	0.185
	σ_m	9.595	3.168	0.933	10.295	10.245	19.771
	$\sigma_{a.eff}$	0.337	0.016	-0.008	-0.052	-0.055	0.198
	σ_{min}	9.146	2.989	0.736	11.098	10.905	19.225
	σ_{max}	9.773	3.012	0.679	10.964	10.796	19.899
Design 2	σ_a	0.313	0.011	-0.029	-0.067	-0.055	0.337
	σ_m	9.459	3.000	0.707	11.031	10.851	19.562
	$\sigma_{a.eff}$	0.323	0.011	-0.029	-0.070	-0.057	0.361
Design 20	σ_{min}	9.491	2.998	1.223	10.444	11.418	19.655

(*continued on next page*)

(*continued*)

Design	Component	Stress_Bed	Stress_Col	Stress_XCar	Stress_YCar	Stress_ZCar	Stress_Spndl
	σ_{max}	10.149	3.098	1.261	10.344	11.438	19.678
	σ_a	0.329	0.050	0.019	-0.050	0.010	0.011
	σ_m	9.820	3.048	1.242	10.394	11.428	19.667
	$\sigma_{a.eff}$	0.340	0.051	0.019	-0.052	0.010	0.012

References

[1] Möhring H-C, Müller M, Krieger J, Multhoff J, Plagge C, De Wit J, Misch S. Intelligent lightweight structures for hybrid machine tools. Prod Eng 2020;14(5–6):583–600. https://doi.org/10.1007/s11740-020-00988-3.

[2] Ji Q, Li C, Zhu D, Jin Y, Lv Y, He J. Structural design optimization of moving component in CNC machine tool for energy saving. J Clean Prod 2020;246:118976. https://doi.org/10.1016/j.jclepro.2019.118976.

[3] Mayr J, Jedrzejewski J, Uhlmann E, Alkan Donmez M, Knapp W, Härtig F, Wendt K, Moriwaki T, Shore P, Schmitt R, Brecher C, Würz T, Wegener K. Thermal issues in machine tools. CIRP Ann 2012;61(2):771–91. https://doi.org/10.1016/j.cirp.2012.05.008.

[4] Yuksel E, Budak E, Ozlu E, Oral A, Igrek F, Tosun F. Prediction of thermal growth in a high-speed spindle by considering thermo-mechanical behavior. MM Sci J 2021;2021(3):4526–33. https://doi.org/10.17973/MMSJ.2021_7_2021055.

[5] Li J-B, Zhang T-Y, Li Z-D, Chen L, Tao W-Q. Multi-objective parameter optimization design of tapered-type manifold/variable cross-section microchannel heat sink. Appl Therm Eng 2024;251:123587. https://doi.org/10.1016/j.applthermaleng.2024.123587.

[6] Huang Y, Hong H, Zhou H, Xiang H, Yang J. Real-time surrogate compensation architecture for machine-tool thermal error compensation with high-performance model. Discov Appl Sci 2024;7(1):8. https://doi.org/10.1007/s42452-024-06389-w.

[7] Oh S. Comparison of a response surface method and artificial neural network in predicting the aerodynamic performance of a wind turbine airfoil and its optimization. Appl Sci 2020;10(18):6277. https://doi.org/10.3390/app10186277.

[8] Liu S, Yang C, Chen G, Lin M, Qin J, Li M. Multi-objective optimization design for the table of a CNC machine tool based on bionic structures. Trans Can Soc Mech Eng 2024;48(3):369–92. https://doi.org/10.1139/tcsme-2023-0214.

[9] Liu S. Multi-objective optimization design method for the machine tool's structural parts based on computer-aided engineering. Int J Adv Manuf Technol 2015;78(5–8):1053–65.

[10] Agapiou JS. A methodology to measure joint stiffness parameters for toolholder-spindle interfaces. J Manuf Syst 2005;24(1):13–20. https://doi.org/10.1016/S0278-6125(05)80003-2.

[11] Chen D, Fan J, Zhang F. Dynamic and static characteristics of a hydrostatic spindle for machine tools. J Manuf Syst 2012;31:26–33. https://doi.org/10.1016/j.jmsy.2010.11.006.

[12] Li T-J, Zhang Y-M, Su Y, Zhang K, Wang Y-B, Gao S-Z. Dynamic reliability of thermally deduced positioning precision of ball screw systems based on random moving difference method. J Manuf Syst 2021;61:171–82. https://doi.org/10.1016/j.jmsy.2021.09.005.

[13] Wu B-C, Young G-S, Huang T-Y. Application of a two-level optimization process to conceptual structural design of a machine tool. Int J Mach Tools Manuf 2000;40(6):783–94. https://doi.org/10.1016/S0890-6955(99)00113-3.

[14] Curtis A, Lomax A. Prior information, sampling distributions, and the curse of dimensionality. GEOPHYSICS 2001;66(2):372–8. https://doi.org/10.1190/1.1444928.

[15] Lin C-W, Tu JF, Kamman J. An integrated thermo-mechanical-dynamic model to characterize motorized machine tool spindles during very high speed rotation. Int J Mach Tools Manuf 2003;43(10):1035–50. https://doi.org/10.1016/S0890-6955(03)00091-9.

[16] Budak E, Altintas Y. Analytical prediction of chatter stability in milling—part ii: application of the general formulation to common milling systems. J Dyn Syst Meas Control 1998;120(1):31–6. https://doi.org/10.1115/1.2801318.

[17] Schmitz T, Betters E, Budak E, Yüksel E, Park S, Altintas Y. Review and status of tool tip frequency response function prediction using receptance coupling. Precis Eng 2023;79:60–77. https://doi.org/10.1016/j.precisioneng.2022.09.008.

[18] Duncan GS, Tummond MF, Schmitz TL. An investigation of the dynamic absorber effect in high-speed machining. Int J Mach Tools Manuf 2005;45(4–5):497–507. https://doi.org/10.1016/j.ijmachtools.2004.09.005.

[19] Mohammadi Y, Azvar M, Budak E. Suppressing vibration modes of spindle-holder-tool assembly through FRF modification for enhanced chatter stability. CIRP Ann Manuf Technol 2018;67(1):397–400. https://doi.org/10.1016/j.cirp.2018.03.003.

[20] Hassanzadeh M, Shahrrava B. Linear Version of Parseval's Theorem. IEEE Access 2022;10:27230–41. https://doi.org/10.1109/ACCESS.2022.3157736.

[21] Lin CD, Tang B. Latin hypercubes and space-filling designs. In: Dean InA, Morris M, Stufken J, Bingham D, editors. Handbook of design and analysis of experiments. CRC Press; 2015. p. 593–625.

[22] Shang B, Apley DW, Mehrotra S. Diversity subsampling: custom subsamples from large data sets. INFORMS J Data Sci 2023;2(2):161–82. https://doi.org/10.1287/ijds.2022.00017.

[23] Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Stat Surv 2010;4(e). https://doi.org/10.1214/09-SS054.

[24] Shahriari B, Swersky K, Wang Z, Adams RP, De Freitas N. Taking the human out of the loop: a review of bayesian optimization. Proc IEEE 2016;104(1):148–75. https://doi.org/10.1109/JPROC.2015.2494218.

[25] Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, Niculae V, Prettenhofer P, Gramfort A, Grobler J, Layton R, Vanderplas J, Joly A, Holt B, Varoquaux G. API design for machine learning software: experiences from the scikit-learn project. arXiv 2013. https://doi.org/10.48550/arXiv.1309.0238.

[26] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM; 2016. p. 785–94. https://doi.org/10.1145/2939672.2939785.

[27] Li F, Li Y, Cai X, Gao L. A surrogate-assisted hybrid swarm optimization algorithm for high-dimensional computationally expensive problems. Swarm Evolut Comput 2022;72:101096. https://doi.org/10.1016/j.swevo.2022.101096.

[28] Wang Z, Kong X, Guo J, Zhao B, Xie L, Wu N. An improved adaptive gradient-based optimization algorithm for estimating the parameters of three-parameter Weibull distribution: An application of aero-engine reliability assessment. Reliab Eng Syst Saf 2026;265:111610. https://doi.org/10.1016/j.ress.2025.111610.

[29] Eiben AE, Smith JE. Introduction to Evolutionary Computing. Springer Berlin Heidelberg; 2015. https://doi.org/10.1007/978-3-662-44874-8.

[30] Blank J, Deb K. pymoo: Multi-Objective Optimization in Python. IEEE Access 2020;8:89497–509. https://doi.org/10.1109/ACCESS.2020.2990567.

[31] Samanipour F, Jelovica J. Adaptive repair method for constraint handling in multi-objective genetic algorithm based on relationship between constraints and variables. Appl Soft Comput 2020;90:106143. https://doi.org/10.1016/j.asoc.2020.106143.

[32] Das Indraneel, Dennis JE. Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems (URL) SIAM J Optim March 1998;8(3):631–57. https://doi.org/10.1137/S1052623496307510. ⟨https://doi.org/10.1137/S1052623496307510⟩ (URL).

[33] Willidal Th, Bauer W, Schumacher P. Stress/strain behaviour and fatigue limit of grey cast iron. Materials Science Engineering A 2005;413–414:578–82. https://doi.org/10.1016/j.msea.2005.08.200.

[34] Farhat H. Typical service-induced damages. In Operation, Maintenance, and Repair of Land-Based Gas Turbines. Elsevier; 2021. p. 107–30. https://doi.org/10.1016/B978-0-12-821834-1.00006-X.

[35] Joint Committee on Structural Safety. (2000). Probabilistic model code, Part 3: Material properties. Retrieved from ⟨https://www.jcss-lc.org/publications/jcsspmc/part_iii.pdf⟩.

[36] Leng J, Su X, Liu Z, Zhou L, Chen C, Guo X, Wang Y, Wang R, Zhang C, Liu Q, Chen X, Shen W, Wang L. Diffusion model-driven smart design and manufacturing: Prospects and challenges. J Manuf Syst 2025;82:561–77. https://doi.org/10.1016/j.jmsy.2025.07.011.

[37] Leng J, Xu C, Song X, Liu Q, Chen X, Shen W, Wang L. High-performance manufacturing systems: concepts, performance metrics, enablers, challenges, and research directions. Adv Eng Inform 2025;68:103617. https://doi.org/10.1016/j.aei.2025.103617.