Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal, 2026, 20(10), 407–423
https://doi.org/10.12913/22998624/225411
ISSN 2299-8624, License CC-BY 4.0

Received: 2026.06.15
Accepted: 2026.07.23
Published: 2026.09.01

Physics-informed surrogate modeling for reliability-based robust design optimization of the external load factor in bolted joints under uncertainty

Van Thuy Tran[1]

[1] Faculty of Engineering-Technology, Pham Van Dong University, Quang Ngai Province, Vietnam
E-mail: tvthuy@pdu.edu.vn

ABSTRACT

The external load factor (χ) is a key parameter affecting the performance and reliability of bolted joints. This study presents an integrated framework combining experimental measurement, physics-informed surrogate modeling, and reliability-based robust design optimization (RBRDO) to analyze and optimize χ under uncertainty. A structured experimental design is employed to investigate the effects of friction coefficient (μ), surface roughness (Ra), and material hardness (HB). Results indicate significant nonlinear interactions, with material hardness exerting the strongest influence. A physics-informed surrogate model is then developed to provide accurate and physically consistent predictions. The optimization framework incorporates the expected value, variance, and Conditional Value-at-Risk (CVaR) to balance performance, variability, and risk. The optimal design, characterized by higher friction, lower surface roughness, and greater material hardness, reduces the expected external load factor to approximately 0.24 while satisfying the reliability requirement $P(0.2 < \chi < 0.3) \approx 0.9993$. The proposed framework offers an efficient and robust methodology for mechanical design under uncertainty.

Keywords: physics-informed modeling, robust design optimization, bolted joints, external load factor, Monte Carlo simulation, surrogate modeling.

INTRODUCTION

Bolted joints are among the most fundamental mechanical elements used in engineering systems, including automotive, aerospace, and heavy machinery, due to their simplicity, reliability, and ease of assembly [1]. Their structural performance is primarily governed by the distribution of external loads between the bolt and the clamped members, commonly characterized by the external load factor (χ) [2, 3]. This parameter plays a critical role in determining joint stiffness, load transfer behavior, and fatigue resistance, and is therefore essential for ensuring structural integrity and long-term reliability [4-6].

In practical applications, however, the external load factor cannot be treated as a deterministic quantity. It is strongly influenced by multiple sources of uncertainty, including friction conditions, surface roughness, and material properties [6, 7]. These factors exhibit inherent variability arising from manufacturing tolerances, assembly conditions, and operational environments. Since χ is fundamentally governed by a stiffness ratio between the bolt and the clamped members, uncertainty propagates through the system in a highly nonlinear and coupled manner [3, 8, 9]. As a result, deterministic modeling approaches are insufficient for accurately capturing system behavior and ensuring reliability under realistic conditions.

To address this challenge, surrogate modeling techniques such as Response Surface Methodology (RSM), Kriging (Gaussian process regression), and machine learning models have been widely adopted to approximate complex system responses with reduced computational cost [10–13]. While efficient, these approaches remain largely data-driven and lack explicit representation of underlying physical mechanisms. Consequently, their accuracy depends strongly on data quality and coverage, often leading to poor generalization, especially in extrapolation regions. More critically, they may produce physically inconsistent predictions, which is unacceptable in reliability-based design requiring accurate estimation of low-probability events.

Recent advances in physics-informed machine learning, particularly physics-informed neural networks (PINNs), have attempted to address these limitations by incorporating governing equations into the learning process [14–17]. These methods typically enforce physical consistency through penalty-based constraints or residual minimization. While effective for differential equations, they are less suitable for algebraic mechanical systems such as bolted joints. The reliance on soft constraints introduces additional hyperparameters and does not ensure strict adherence to problem-specific physics. In particular, the stiffness-governed load transfer mechanism is not explicitly embedded, limiting interpretability and extrapolation reliability.

In parallel, reliability-based design optimization (RBDO) has been extensively developed to account for uncertainty in engineering systems [18–21]. By incorporating probabilistic constraints, RBDO ensures that performance requirements are satisfied with a specified level of reliability. Extensions toward robust design further consider variability in system response. However, most existing frameworks focus primarily on mean performance and variance, without explicitly addressing the influence of extreme responses. In many mechanical systems, including bolted joints, rare but severe loading conditions are directly associated with fatigue failure and structural degradation. Risk-sensitive measures such as conditional value-at-risk (CVaR) provide a more comprehensive characterization of such tail behavior [22, 23]. Despite their advantages, the integration of CVaR into the design optimization of bolted joints remains largely unexplored.

From the above discussion, three fundamental limitations can be identified in the current state of the art: (i) existing surrogate models lack inherent physical consistency and may violate fundamental mechanical constraints, particularly in extrapolation regions; (ii) current physics-informed approaches rely on penalty-based formulations and do not structurally embed problem-specific mechanics, such as stiffness-governed load transfer, into the model; (iii) existing reliability-based design frameworks rarely incorporate explicit control of tail risk, limiting their effectiveness in safety-critical applications.

To address these limitations, a mechanics-embedded physics-informed surrogate framework is integrated with reliability-based robust design optimization (RBRDO). The proposed hybrid model directly embeds the stiffness-governed load transfer mechanism, ensuring physically consistent and interpretable predictions while capturing nonlinear behavior.

Building upon this modeling framework, a risk-aware RBRDO formulation is developed by incorporating CVaR as a tail-risk measure [24–26]. This enables simultaneous control of expected performance, variability, and extreme response behavior, providing a more comprehensive and practically relevant design strategy for engineering systems operating under uncertainty.

The proposed approach is validated using experimentally measured data of the external load factor, ensuring strong physical credibility and distinguishing this work from purely simulation-based studies. In addition, the surrogate-based implementation significantly reduces computational cost, enabling efficient uncertainty propagation and large-scale Monte Carlo simulation.

The main contributions of this study can be summarized as follows:
- A mechanics-embedded physics-informed surrogate model that directly integrates stiffness-based load transfer behavior into the model structure, ensuring physical consistency without relying on penalty-based constraints.
- A hybrid physics–data modeling strategy that improves predictive accuracy, enhances extrapolation robustness, and reduces model bias under uncertainty.
- A risk-aware RBRDO framework incorporating CVaR to explicitly control extreme response behavior.
- Experimental validation of the external load factor, providing strong physical grounding and improving the reliability of the proposed approach.
- A computationally efficient framework enabling large-scale uncertainty quantification and optimization.

Overall, this work establishes a unified and generalizable framework for integrating physics-informed modeling with reliability-based robust design optimization in engineering systems characterized by nonlinear behavior, uncertainty propagation, and limited data.

METHODOLOGY

This section presents the methodology adopted in this study. It includes the mechanical modeling and experimental measurement of the external load factor, the uncertainty modeling of the input parameters, and the formulation of the reliability-based robust design optimization framework. These components provide the foundation for the proposed physics-informed surrogate model.

Definition and mechanical modeling of the external load factor

The mechanical behavior of a bolted joint under external loading is governed by the load-sharing mechanism between the bolt and the clamped members. As illustrated in Figure 1, the bolt and the clamped components can be idealized as two elastic elements with stiffnesses (k_b) and (k_c), respectively.

As illustrated in Figure 1, the applied external load F is shared between the bolt and the clamped members. This load-sharing mechanism is quantified by the external load factor, denoted as χ, which represents the fraction of the external load carried by the bolt. The external load factor is defined as:

$$\chi = \frac{\Delta F_b}{F} = \frac{V - V_0}{F} = \frac{k_b}{k_b + k_c} \tag{1}$$

where: F is the applied external load; ΔF_b is the increment in bolt force induced by F; V_0 is the initial tightening force (preload); V is the bolt force under external loading; and k_b and k_c are the stiffnesses of the bolt and the clamped members, respectively. In the experimental context, ΔF_b is determined as $\Delta F_b = V - V_0$.

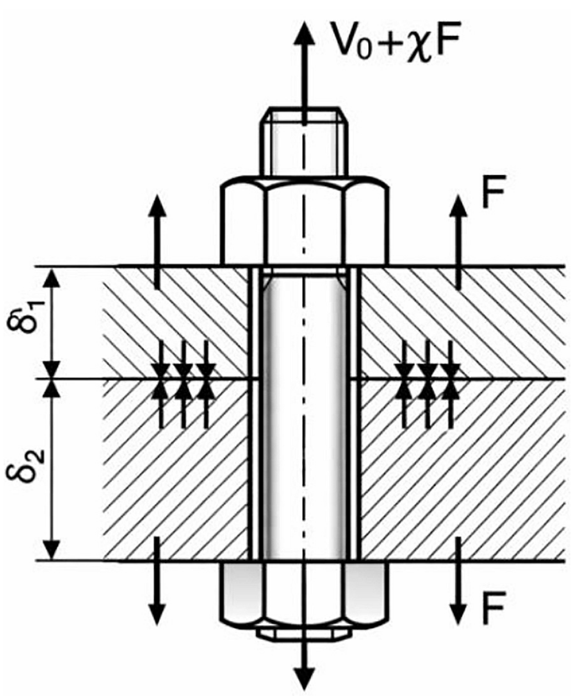

Figure 1. Axial load distribution in a bolted joint

Equation 1 shows that the load distribution depends on the relative stiffness of the bolt and the clamped members. Increasing the bolt stiffness increases the load transferred to the bolt, whereas increasing the clamped-member stiffness reduces it. Since the clamped-member stiffness is governed by contact conditions and material properties, the external load factor is influenced by the friction coefficient, surface roughness, and material hardness.

This stiffness-governed load transfer mechanism forms the physical foundation of the proposed physics-informed surrogate model developed in the subsequent sections.

Experimental setup and measurement of the external load factor

The external load factor is determined by simultaneously measuring the bolt force and the applied external load, as shown in Figure 2.

To determine the external load factor, the bolted joint (3) is initially tightened to a preload V_0 to prevent separation and ensure stable contact between the clamped components. The preload is measured using a load cell sensor (4). Subsequently, an external load F is applied through a hydraulic cylinder (2), while the corresponding bolt force V is continuously monitored by the same sensor. As the external load increases, the bolt force rises from the initial preload V_0 to a new value V. Based on these measurements, the external load factor is calculated using Equation 1.

The experimental setup used to implement this measurement is shown in Figure 3. The system consists of a hydraulic loading unit, a rigid

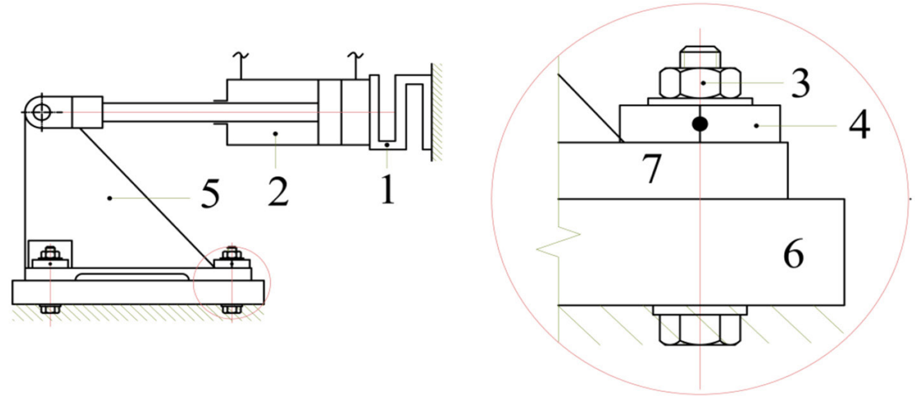

Figure 2. Schematic diagram of the system: 1 – load cell measuring the cylinder force; 2 – hydraulic cylinder generating force; 3 – bolt joint; 4 – load cell measuring bolt tightening force; 5 – frame structure; 6, 7 – clamping plates

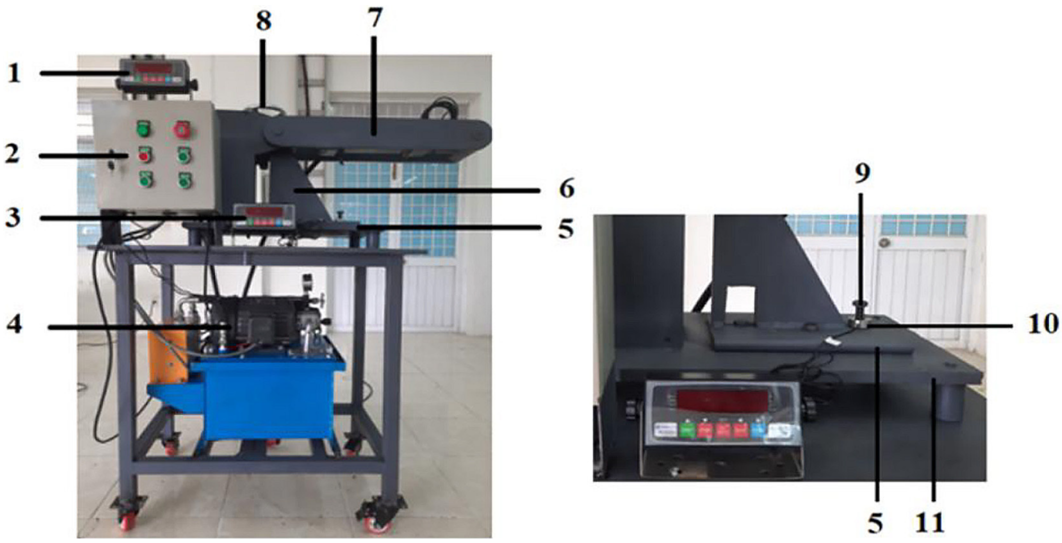

Figure 3. Experimental setup for external load factor measurement, previously reported in [3]: 1 – external force display meter; 2 – electrical control cabinet; 3 – bolt tightening force display meter; 4 – hydraulic system; 5,11 – clamping plates; 6 – support frame; 7 – force applicator arm; 8 – tilt angle adjuster (α); 9 – bolt; 10 – load cell measuring bolt tightening force

frame, and two independent load cells: a through-hole load cell for measuring the bolt force and an S-type load cell for measuring the applied external load. All sensors are factory-calibrated with full-scale accuracy within ±0.05% and repeatability better than ±0.02%.

The test specimen consists of two steel plates connected by an M16 bolt subjected to axial loading. Each experiment is repeated three times to ensure repeatability, and the average values are used for analysis. The overall measurement uncertainty is estimated to be within ±3%.

The experimental dataset serves as a reliable basis for model calibration, validation, and subsequent uncertainty analysis.

During the assembly of the bolted joint, an initial preload of 15 kN was applied. This preload was selected to ensure that no joint separation occurred throughout the loading process. The initial preload was measured directly using the coaxial donut-type load cell (Forsentek FBDA, 0–20 kN), rather than being estimated from the tightening torque. During the experiment, an external tensile load of 6 kN was subsequently applied to determine the external load factor.

Uncertainty modeling of input parameters

In practical applications, the parameters affecting the external load factor are inherently uncertain due to manufacturing variability, measurement errors, and operating conditions. To account for these effects, an uncertainty modeling framework is established.

The key input variables considered in this study are:

$$d = [\mu, Ra, HB]^T \tag{2}$$

where: μ is the friction coefficient, Ra is the surface roughness, and HB is the material hardness.

These variables are selected due to their significant influence on stiffness characteristics and load transfer behavior. The variables μ, Ra, and HB are assumed independent. Preliminary sensitivity analysis indicated that moderate correlations ($\rho \leq 0.3$) have a negligible influence on the predicted response, supporting the validity of this assumption.

Friction coefficient

In this study, the composite friction coefficient (μ) is determined from the measured tightening torque and bolt tension force instead of using the empirical nut factor (K). Torque (T) is measured using an electronic torque wrench (±1% accuracy), while the bolt tension force is obtained from a coaxial donut-type load cell (Forsentek FBDA, 0–20 kN) placed between the clamped surfaces. This approach better captures the effects of lubrication, surface roughness, and material properties on joint load transfer. According to technical references and ISO 16047, the composite friction coefficient in bolted joints typically ranges from 0.1 to 0.3 depending on the surface condition and lubrication state. The friction coefficient is calculated using:

$$\mu = \frac{T}{FD} \tag{3}$$

where: D is the nominal diameter of the bolt.

The measured data exhibit a symmetric distribution, and the Shapiro–Wilk test yields a p-value of 0.784 (> 0.05), indicating that the normality assumption cannot be rejected. Accordingly, μ is modeled as a normally distributed random variable For lubricated surfaces, the friction coefficient typically varies within a tolerance of ±0.05. Based on this range, the standard deviation is estimated as ($\sigma_\mu = 0.0167$), which is adopted to represent the inherent variability arising from surface interactions and lubrication conditions.

Surface roughness

In this study, two surface roughness levels were considered: Ra = 1.6 μm (fine machining with lower friction) and Ra = 6.3 μm (rough machining with higher friction). Roughness was measured using a contact profilometer (Mitutoyo SJ-210) following ISO standards, with three measurements per specimen and averaged results. The data show an approximately normal distribution (Shapiro–Wilk p = 0.184 > 0.05), so Ra is modeled as a normally distributed random variable Ra ~ $N(\overline{R}a, \sigma_\mu)$.

For surfaces with nominal roughness in the range Ra ∈ [1.6, 6.3] μm, a tolerance of ±0.5 μm is adopted to reflect the typical process capability of conventional turning and fine milling operations. This tolerance is consistent with practical machining experience and the recommended control levels specified in ISO 4288.

The allowable variation range is used to define the probabilistic model of surface roughness in the reliability analysis. Assuming a normal distribution, the standard deviation is estimated as: $\sigma_{Ra} = 0.167$.

Material hardness

Two hardness levels were considered: HB ≈ 120 (aluminum alloys or mild steels) and HB ≈ 250 (hardened steels), representing typical engineering materials. Brinell hardness (HB) was measured according to ISO 6506-1 using a 10 mm steel ball, 1000 kgf load, and 15-s dwell time, with three measurements per specimen and averaged results. The Shapiro–Wilk test yields a p-value of 0.908 (> 0.05), indicating no evidence against the normality assumption. Therefore, hardness is modeled as a normally distributed random variable $HB \sim N(\overline{HB}, \sigma_{HB})$.

To account for manufacturing variability, an absolute tolerance of ±10 HB is adopted, reflecting typical quality control capability and recommendations from ISO 6506. Accordingly, the allowable variation range is defined as [HB−10, HB+10]. Assuming a normal distribution, the corresponding standard deviation is estimated as: $\sigma_{HB} = 3.33$.

In summary, all three input variables are modeled as normally distributed, which is consistent with the random and unbiased nature of variability in mechanical systems. This assumption facilitates the application of reliability analysis methods, such as the first-order reliability method (FORM) and Monte Carlo simulation, in subsequent analyses. The corresponding statistical parameters are summarized in Table 1.

The three input parameters, friction coefficient (μ), surface roughness (Ra), and material hardness (HB), are selected due to their significant influence on load distribution in threaded joints. The friction coefficient governs the conversion of tightening torque into clamping force (μ ∈ [0.1, 0.3]), while surface roughness and material hardness affect contact stiffness and load-sharing behavior. Accordingly, representative ranges of Ra ∈ [1.6, 6.3] μm and HB ∈ [120, 250] are considered to evaluate their effects on the external load factor χ. In summary, the three selected input parameters are as follows:

$$0.1 \leq \mu \leq 0.3;\ 1.6 \leq Ra \leq 6.3;\ 120 \leq HB \leq 250 \quad (4)$$

The external load factor is expressed as a nonlinear function of the uncertain input parameters:

$$\chi = g(d) \quad (5)$$

where: $g(d)$ represents the stiffness-governed mechanical relationship.

Due to the presence of nonlinear interactions among friction coefficient, surface roughness, and material hardness, particularly through stiffness formulations involving power-law and exponential terms, the response function exhibits strong nonlinearity and potential non-Gaussian behavior.

To propagate uncertainty, Monte Carlo Simulation (MCS) is employed. Unlike linearization-based methods such as FORM or Taylor expansion, which may introduce significant errors in highly nonlinear systems, MCS provides a non-intrusive and assumption-free approach by directly sampling the input space. This enables accurate estimation of the full probability distribution of χ, including tail behavior critical for reliability analysis. From a reliability perspective, the performance is defined by:

$$0.2 < \chi < 0.3 \quad (6)$$

This requirement is expressed probabilistically as:

$$P(0.2 < \chi < 0.3) \geq 0.999 \quad (7)$$

Failure occurs when the response falls outside the admissible range. The corresponding probability is estimated using MCS. To improve computational efficiency, a surrogate model is developed and presented in the following section.

Reliability-based robust design optimization

Building upon the uncertainty modeling framework in Section "Experimental setup and measurement of the external load factor", the design of bolted joints is formulated as a reliability-based robust design optimization problem. Unlike conventional RBDO, RBRDO accounts for both reliability and robustness by considering the mean response, variability, and tail risk. The design vector is defined as:

$$d = [\mu, Ra, HB]^T \quad (8)$$

where: μ, Ra, and HB denote the nominal values of the friction coefficient, surface roughness, and material hardness, respectively.

Due to manufacturing variability and operational conditions, the actual system inputs are modeled as random variables:

$$X = d + \varepsilon \quad (9)$$

where: ε represents random deviations characterized by known probability distributions.

This formulation allows explicit representation of uncertainty around the design point. The objective of the RBRDO problem is to identify a design that minimizes not only the expected value

Table 1. Manufacturing tolerances and standard deviations

Design Variables	Tolerances	Standard deviation (σ)	Probability distribution
Combined friction coefficient (μ)	±0.05	0.0167	Normal $\left(\overline{\mu}, \sigma_{\mu}\right)$
Surface roughness (Ra), μm	±0.5	0.167	Normal $\left(\overline{Ra}, \sigma_{Ra}\right)$
Material hardness (HB), HB	±10	3.33	Normal $\left(\overline{HB}, \sigma_{HB}\right)$

of χ but also its variability and extreme behavior. Accordingly, the objective function is defined as:

$$\min_{X} \quad E[\chi]+\lambda_1 Var[\chi]+\lambda_2 CVaR_\alpha(\chi) \quad (10)$$

where: $E[\chi]$ is expected value (performance metric); $Var[\chi]$ is variance (robustness measure); $CVaR_\alpha(\chi)$ is conditional value-at-risk at confidence level α (tail risk measure) and λ_1, λ_2 are weighting coefficients controlling the trade-off.

The inclusion of variance reduces sensitivity to input uncertainty, while CVaR penalizes extreme high-load scenarios that may lead to failure. To ensure safe operation, a probabilistic constraint is imposed such that the external load factor remains within a prescribed interval with high confidence:

$$P(0.2<\chi<0.3)\geq 0.999$$

The design variables are constrained within feasible engineering bounds: μ ∈ [0.1, 0.3], Ra ∈ [1.6, 6.3], HB ∈ [120, 250]

The complete RBRDO formulation is given by: to summarize, reliability-based robust design optimization problem for the bolted joint can be expressed as follows:

Design variables:

μ ∈ [0.1, 0.3], Ra ∈ [1.6, 6.3], HB ∈ [120, 250]

Objective function:

$$\min_{d} \quad E[\chi]+\lambda_1 Var[\chi]+\lambda_2 CVaR_\alpha(\chi) \quad (11)$$

Subject to:

$$\begin{cases} 0.1\leq \mu \leq 0.3 \\ 1.6\leq Ra\leq 6.3 \\ 120\leq HB\leq 250 \\ \mathrm{P}(0.2<\chi<0.3)\geq 0.999 \end{cases}$$

Direct MCS is computationally expensive for repeated evaluation of the objective and reliability constraints. Therefore, a physics-informed surrogate model is developed to approximate g(X), providing an efficient computational framework for RBRDO.

PHYSICS-INFORMED SURROGATE MODELING

This section presents the development of the proposed physics-informed surrogate model for predicting the external load factor in bolted joints. It describes the mechanics-guided surrogate formulation, the physics-informed loss function, the model training and validation procedures, and its integration into the RBRDO framework

Mechanics-guided surrogate formulation

To improve physical consistency and extrapolation capability, a mechanics-embedded physics-informed surrogate model is developed for predicting the external load factor in bolted joints. The model directly incorporates the stiffness-based load transfer mechanism into its structure, resulting in a compact and data-efficient formulation.

Problem definition

The overall architecture of the proposed framework is illustrated in Figure 4. It consists of two complementary components: (i) a data-driven surrogate model that learns complex nonlinear relationships from data, and (ii) a physics-based model that encodes the underlying stiffness-driven load transfer mechanism. These two components are coupled through a hybrid loss formulation to ensure both data fidelity and physical consistency.

Physics-based reference model

The load transfer behavior in bolted joints is fundamentally governed by the relative stiffness of the bolt and the clamped members. Based on classical stiffness-based theory, the external load factor can be expressed as:

$$\chi = g_{phys}(X)=\frac{k_b}{k_b+k_c(X)} \quad (12)$$

where: k_b is the bolt stiffness and k_c is the stiffness of the clamped members.

In this study, the bolt stiffness is assumed to be constant, with $k_b = 1.2\times 10^7$ N/m, as its variation is relatively small compared to that of the clamped members.

To account for the influence of contact conditions and material properties, the clamped member stiffness k_c is modeled using a semi-empirical formulation:

$$k_c(\mu, Ra, HB)=k_0\left(\frac{HB}{HB_0}\right)^{\alpha}\left(\frac{e^{-\gamma\mu}}{1+\beta Ra}\right) \quad (13)$$

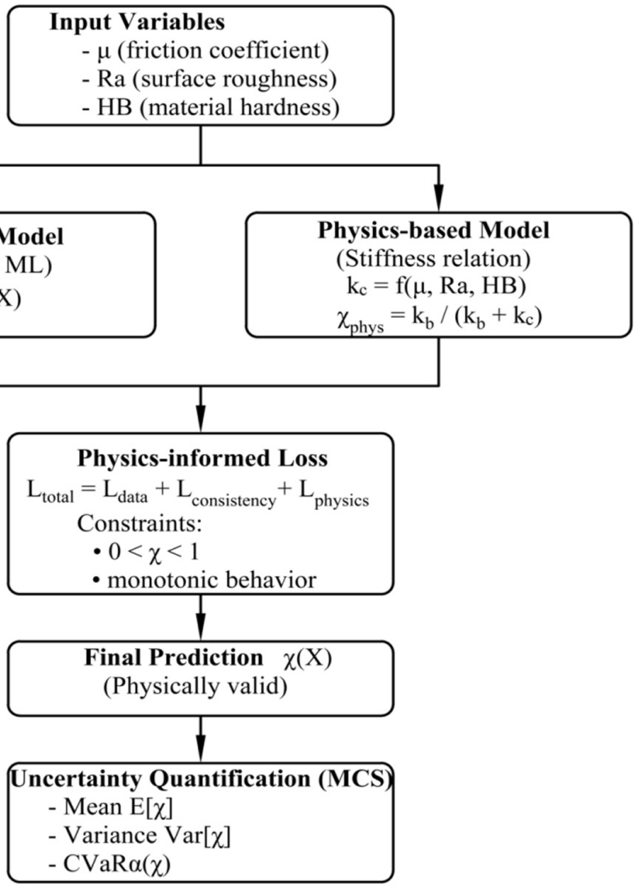

Figure 4. Architecture of the proposed physics-informed surrogate modeling framework

where: μ, Ra, and HB denote the nominal values of the friction coefficient, surface roughness, and material hardness, respectively and α,β,γ are model parameters. The reference values are $k_0 = 4.0 \times 10^7$ N/m and $HB_0 = 180$. The calibrated parameters based on representative datasets are $\alpha = 0.85$, $\beta = 0.05$, $\gamma = 0.4$.

Equation 13 is derived from rough-surface contact mechanics, where contact stiffness depends on asperity deformation. The term $(HB/HB_0)^\alpha$ captures the nonlinear influence of material hardness, with higher hardness resulting in greater contact stiffness. A decoupled structure is adopted to represent interfacial effects, where friction and surface roughness are modeled separately through $e^{-\gamma\mu}$ and $(1+\beta Ra)^{-1}$, respectively. This formulation reduces parameter coupling, improves interpretability, and ensures stable behavior under extrapolation.

Overall, the model reflects a hierarchical sensitivity structure: material hardness governs the global stiffness scaling, while friction and surface roughness introduce nonlinear attenuation effects. This behavior is consistent with the experimentally observed variability and uncertainty propagation of the external load factor.

Substituting Equation 13 into Equation 12 defines a mechanics-consistent mapping:

$$g_{phys}: R^3 \rightarrow (0,1) \tag{14}$$

where: $g_{phys}(\mu,Ra,HB)$ represents the physics-based prediction of the external load factor.

This mapping inherently guarantees physically admissible predictions, i.e., $0 < \chi < 1$.

It is emphasized that the objective is not to derive an exact first-principles solution, but to construct a physically consistent and computationally efficient reference model that captures the dominant stiffness-governed load transfer behavior. This model serves as a physics-based backbone within the hybrid surrogate framework, enabling data-driven refinement while preserving interpretability and robustness under uncertainty.

Hybrid surrogate structure

The surrogate is constructed as a physics-guided correction model:

$$f_\theta(X) = g_{phys}(X) + \delta_\theta(X) \tag{15}$$

where: δ_θ captures residual nonlinearities.

The proposed physics-informed hybrid surrogate directly incorporates the stiffness-driven load transfer mechanism, ensuring physically consistent and interpretable predictions. The compact formulation captures nonlinear interactions among the governing parameters and enables efficient repeated evaluations in RBRDO.

Physics-informed loss formulation

To integrate physical knowledge into the learning process, a composite loss function is constructed as:

$$L_{Total} = L_{data} + \lambda_1 L_{consistency} + \lambda_2 L_{physics} \tag{16}$$

where: λ_1 and λ_2 are weighting coefficients.

The data loss ensures agreement between predictions and observed/simulated data:

$$L_{data} = \frac{1}{N}\sum_{i=1}^{N}\left(f_\theta(X_i) - \chi_i\right)^2 \tag{17}$$

To enforce consistency with the stiffness-based load-sharing mechanism:

$$L_{consistency} = E\left[\left(f_\theta(X) - \frac{k_b}{k_b + k_c(X)}\right)^2\right] \tag{18}$$

This term ensures that the learned model remains close to the physically derived relationship across the input space. Fundamental physical constraints are imposed as:

$$L_{physics} = L_{bound} + L_{monotonic} \tag{19}$$

Bound constraint:

$$L_{bound} = E\left[\max(0, -\hat{\chi})^2 + \max(0, \hat{\chi} - 1)^2\right] \tag{20}$$

Monotonicity constraint**:**

$$L_{monotonic} = E\left[\max\left(0, \frac{\partial\hat{\chi}}{\partial H}\right)^2 + \max\left(0, -\frac{\partial\hat{\chi}}{\partial R}\right)^2\right] \tag{21}$$

These constraints ensure that the surrogate model respects known physical behavior, even in regions with limited training data.

The framework combines: data-driven learning and physics-based regularization leading to improved robustness and generalization.

Selection of weighting parameters

The weighting parameters λ_1 and λ_2 control the balance between data fidelity and physical consistency. Improper selection may lead to overfitting or excessive bias toward the physics-based model. To ensure a balanced contribution of each loss term, all components are first normalized:

$$\tilde{L}_i = \frac{L_i}{L_i^{(0)}} \tag{22}$$

where: $L_i^{(0)}$ is the initial loss value.

A grid search is then conducted over: $\lambda_1, \lambda_2 \in [0.01, 10]$. The optimal values are selected based on a trade-off between predictive accuracy (RMSE, R^2) and physical consistency (constraint satisfaction and monotonicity preservation). Based on this procedure, the weighting parameters are set to $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$. These values were selected to achieve a balanced trade-off between the data-fitting objective and the physics-based consistency. A slightly higher weight is assigned to λ_1 to prioritize prediction accuracy while ensuring sufficient enforcement of the physics-based consistency term. The model exhibits stable performance within a moderate range of λ_1 and λ_2, indicating robustness of the proposed framework. A detailed sensitivity analysis is presented in Section Results and Discussion.

Training, validation, and RBRDO integration

The dataset is generated using DoE and LHS. The model is trained with cross-validation and early stopping, and evaluated using RMSE, (R^2), maximum absolute error, physical consistency, and comparisons with RSM and Kriging.

Once trained, the surrogate replaces the original model:

$$g(X) \rightarrow f_\theta(X)$$

enabling efficient evaluation of:

$$E[\chi], \mathrm{Var}[\chi], \mathrm{CVaR}\alpha(\chi)$$

This significantly reduces computational cost in Monte Carlo simulation and enables efficient RBRDO.

The overall workflow of the proposed framework is illustrated in Figure 5. This integrated framework provides an efficient and physically consistent solution for the design optimization of bolted joints under uncertainty.

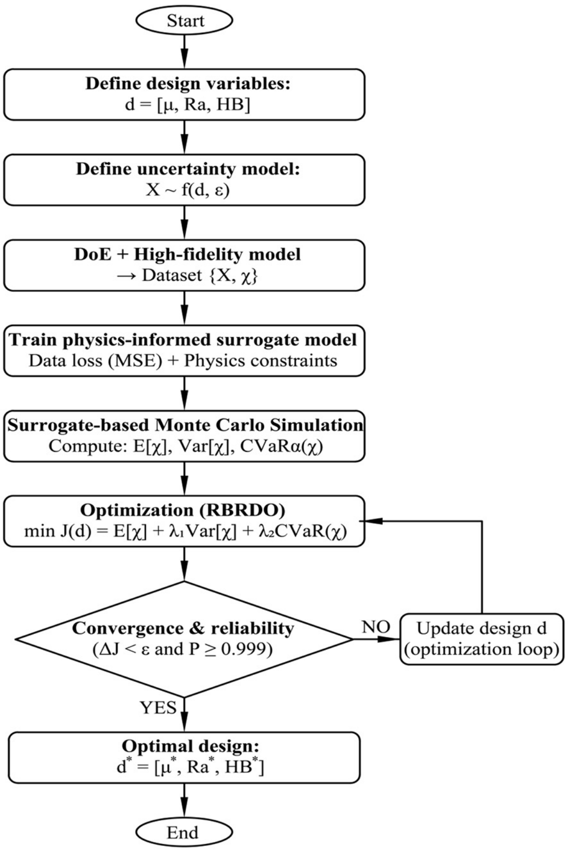

Figure 5. Workflow of the proposed physics-informed surrogate-based RBRDO framework

RESULTS AND DISCUSSION

This section presents the experimental validation and performance evaluation of the proposed physics-informed surrogate model. The results include comparisons with experimental measurements and conventional surrogate models, analyses of physical consistency and sensitivity, and the application of the proposed framework to reliability-based robust design optimization.

Experimental data, physical insights, and model validation

The dataset is generated using a CCD to capture the nonlinear interactions among the μ, Ra, and HB. The experimental data and the corresponding predictions of the physics-based and surrogate models are summarized in Table 2.

The experimental values (χ_{exp}) are obtained from direct measurements, while the physics-based predictions (χ_{phys}) are calculated using Equations 12–13, and the surrogate predictions (χ_{surr}) are obtained from the data-driven model. The absolute prediction errors are defined as $|Error| = |\chi_{pred} - \chi_{exp}|$. Statistical indicators (mean and maximum errors) are provided to quantify the overall predictive performance.

Each experimental condition was repeated three times, and the reported values represent the average, with a standard deviation within ±3%, confirming good repeatability and measurement reliability. Unlike many previous studies relying on simulated data, the present work utilizes directly measured values of χ, thereby enhancing the physical credibility of the analysis. The experimental results indicate clear trends in the influence of input parameters on χ.

To evaluate model performance, a comprehensive comparison between experimental measurements, the original physics-based model, the refined physics-based model, and the surrogate model is conducted in Table 2. The original physics-based model shows limited predictive capability due to its inability to capture the full variation range of χ.

A physics-based model is introduced by reformulating the clamped member stiffness and adding a scaling factor to correct underestimation. This modification improves predictive accuracy, reducing the mean absolute error to 0.058 (from ~0.12). The refined model also captures the correct trends with respect to all input variables, indicating improved physical consistency.

The surrogate model demonstrates the highest predictive accuracy, with a mean absolute error of 0.016. However, unlike the physics-based model, it lacks direct physical interpretability. Therefore, the refined physics-based model serves as a physically consistent and interpretable approximation, while the surrogate model provides superior predictive performance.

A quantitative comparison between the experimental data, the physics-based model, and the surrogate model is presented in Figure 6. As shown in Figure 6a, all three exhibit similar probability density distributions of χ, with the surrogate closely matching the experimental results, while the physics-based model shows a slightly narrower distribution. The parity plot in Figure 6b further demonstrates that the surrogate achieves excellent agreement with the experimental data, whereas the physics-based model preserves the overall physical trend. These results confirm that the proposed surrogate provides accurate predictions while maintaining physical consistency.

Table 2. Comparison between experimental data, physics-based model, and surrogate model predictions of the external load factor χ

No.	μ	Ra (μm)	HB	χ_{exp}	χ_{phys}	χ_{surr}	$\lvert Error_{phys} \rvert$	$\lvert Error_{surr} \rvert$
1	0.100	1.60	120.00	0.368	0.322	0.374	0.046	0.006
2	0.300	1.60	120.00	0.348	0.340	0.345	0.008	0.003
3	0.100	6.30	120.00	0.337	0.367	0.345	0.030	0.008
4	0.300	6.30	120.00	0.308	0.386	0.317	0.078	0.009
5	0.100	1.60	250.00	0.160	0.203	0.169	0.043	0.009
6	0.300	1.60	250.00	0.132	0.216	0.140	0.084	0.008
7	0.100	6.30	250.00	0.143	0.237	0.140	0.094	0.003
8	0.300	6.30	250.00	0.106	0.252	0.112	0.146	0.006
9	0.032	3.95	185.00	0.241	0.262	0.267	0.021	0.026
10	0.368	3.95	185.00	0.193	0.289	0.219	0.096	0.026
11	0.200	0.10	185.00	0.255	0.242	0.266	0.013	0.011
12	0.200	7.90	185.00	0.207	0.307	0.219	0.100	0.012
13	0.200	3.95	75.67	0.449	0.448	0.415	0.001	0.034
14	0.200	3.95	294.33	0.104	0.204	0.070	0.100	0.034
15	0.200	3.95	185.00	0.289	0.275	0.243	0.014	0.046
Mean	—	—	—	—	—	—	0.058	0.016

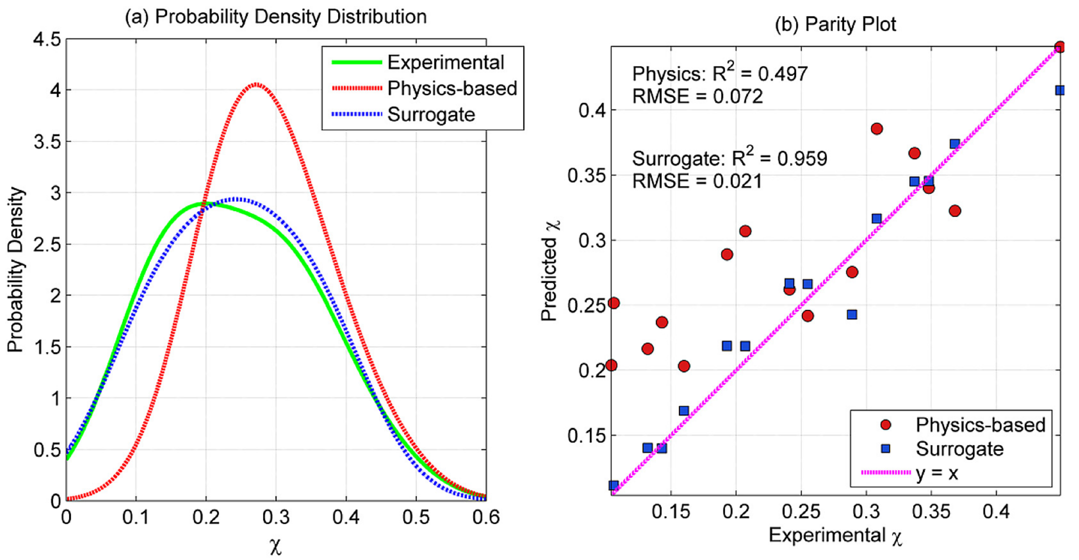

Figure 6. Comparison of experimental measurements, physics-based model, and surrogate model for the external load factor χ: a) probability density distribution and b) parity plot with respect to experimental data

Model performance and comparison

The predictive performance of the proposed surrogate model is evaluated against conventional approaches, including response surface methodology (RSM), Kriging, and a baseline data-driven neural network. The results are summarized in Table 3.

As shown in Table 3, the proposed model achieves a root mean square error (RMSE) of 0.020 and a coefficient of determination (R^2) of 0.959, which are comparable to those of the Kriging model and superior to RSM and the baseline model. Although the improvement in global accuracy metrics is moderate, the proposed model provides additional advantages in terms of physical consistency and extrapolation capability.

A more detailed evaluation of the prediction quality is presented in Figure 7, including parity, residual, and probability density analyses.

As shown in Figure 7a, the predicted values closely follow the ideal agreement line (y = x),

Table 3. Predictive performance of surrogate models evaluated against experimentally measured external load factor

Model	RMSE	R²	Extrapolation Error	Physics
RSM	0.021	0.931	High	No
Kriging	0.015	0.958	Moderate	No
Baseline NN	0.0185	0.942	High	No
Proposed model	0.020	0.959	Low	Yes

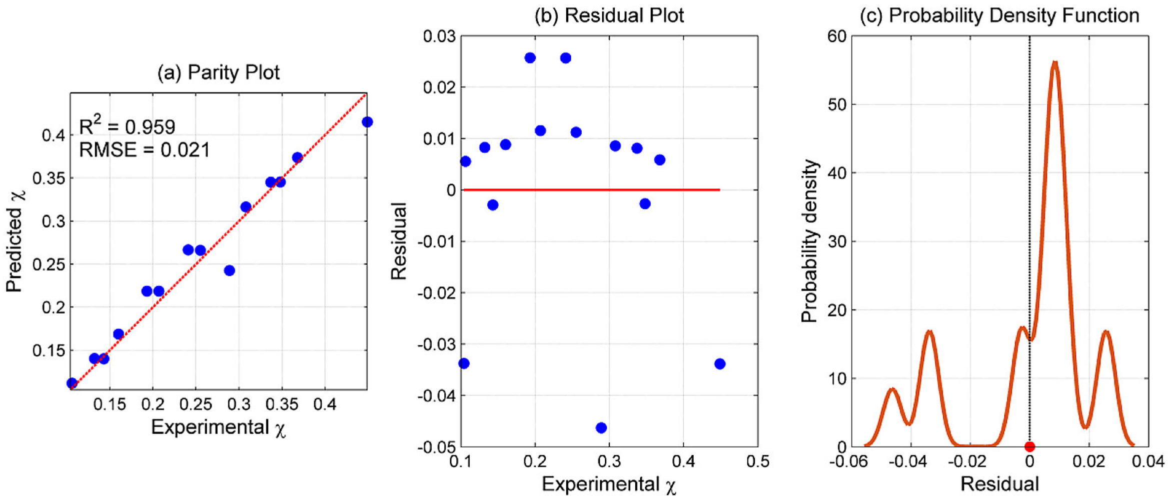

Figure 7. Prediction diagnostics of the proposed model: a) parity plot, b) residual plot, and (c) probability density of residuals

confirming a high level of accuracy. The residual plot in Figure 7b indicates that the prediction errors are generally small and distributed around zero, with no evident systematic bias. The probability density function of residuals in Figure 7c further shows a concentrated distribution with a narrow spread, suggesting stable and reliable predictions.

Overall, the results show that the proposed surrogate achieves a balance between predictive accuracy, interpretability, and generalization, making it suitable for reliability analysis and robust design optimization.

Physical consistency, sensitivity, and interpretability

Beyond predictive accuracy, the physical consistency of the proposed model is examined to ensure its suitability for engineering applications. The model satisfies the physical constraint ($0 < \chi < 1$) throughout the design space and reproduces trends that are consistent with the underlying load transfer mechanism.

As shown in Figure 8, increasing material hardness significantly decreases the external load factor (χ) because higher stiffness reduces the load transmitted to the bolt. In contrast, increasing surface roughness increases (χ) by reducing effective contact stiffness. The friction coefficient (μ) has a comparatively weaker influence, mainly affecting interfacial load redistribution.

The relative importance of the input variables is quantified through sensitivity analysis, as shown in Table 4. Material hardness is identified as the dominant factor, with a sensitivity index of 0.47, followed by surface roughness (0.33) and friction coefficient (0.20). These results are consistent with the trends observed in Figure 8 and confirm the primary role of stiffness in governing the external load factor. The input parameters exhibit nonlinear interactions that cannot be captured by simple linear models. By combining a physics-based model with a data-driven correction term, the proposed hybrid surrogate preserves physical consistency and improves extrapolation capability, enabling physically reliable RBRDO solutions.

To further assess robustness beyond the training domain, the extrapolation performance is evaluated in Figure 9, focusing on regions not covered by the experimental data.

As shown in Figure 9a, the baseline data-driven model exhibits unstable predictions under

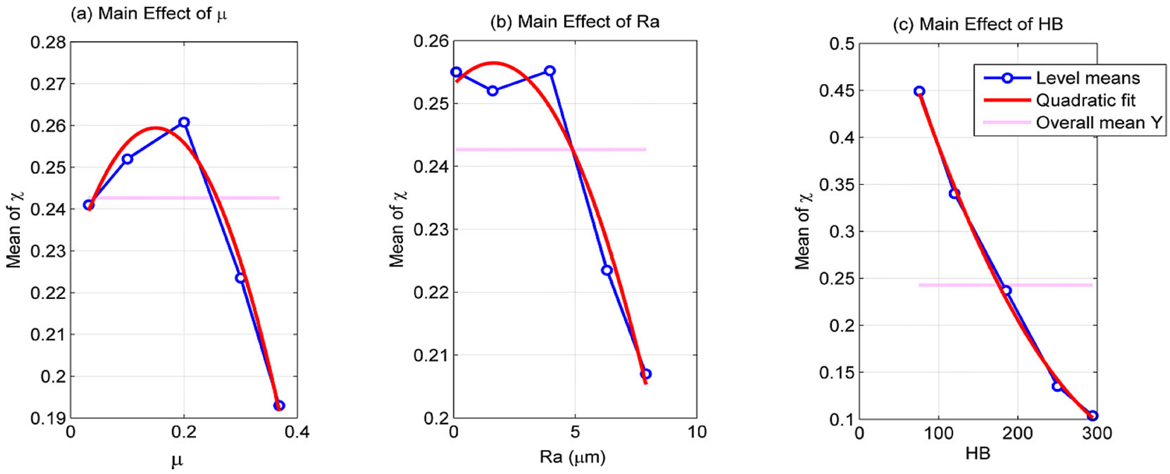

Figure 8. Main effects of input parameters on the external load factor χ: a) friction coefficient (μ), b) surface roughness (Ra), and c) material hardness (HB)

Table 4. Sensitivity indices of input parameters

Parameter	Sensitivity index
HB	0.47
Ra	0.33
μ	0.20

extrapolation, indicating poor generalization. In contrast, Figure 9b shows that the proposed physics-informed model maintains smooth, physically consistent predictions by incorporating physical knowledge into the model structure. These results demonstrate superior extrapolation capability and physical reliability, which are essential for obtaining robust and trustworthy RBRDO solutions.

Reliability analysis and RBRDO results

Monte Carlo-based uncertainty propagation indicates a high probability of satisfying the target range ($0.2 < \chi < 0.3$). To enhance robustness, the RBRDO framework incorporates variance and CVaR alongside the expected value. As shown in Table 5, the optimized design exhibits increased friction coefficient, reduced surface roughness, and higher material hardness, resulting in improved load distribution and mechanical stability.

Quantitatively, the optimized design achieves a reduction in the expected external load factor to approximately $E[\chi] \approx 0.24$, accompanied by a substantial decrease in variance and tail risk. In particular, the $CVaR_{0.95}$ is significantly reduced,

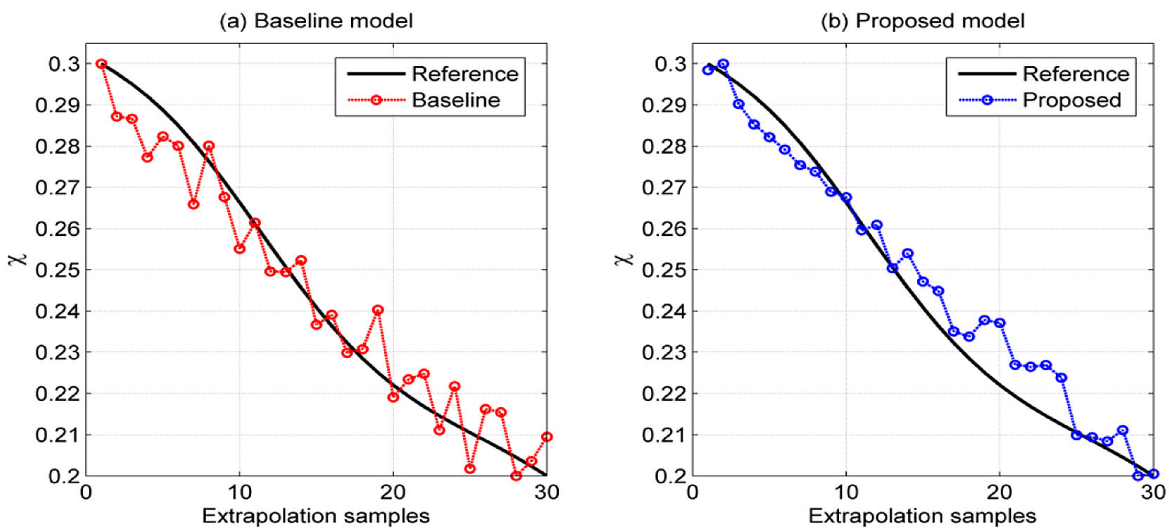

Figure 9. Comparison of extrapolation performance between: a) baseline data-driven model, and b) proposed physics-informed model in regions outside the training domain

Table 5. Optimal design parameters and performance

Parameter	Initial design	Optimal design (RBRDO)	Change
μ	0.18–0.22	0.238	↑
Ra (μm)	3.0–4.0	2.75	↓
HB	170–190	205	↑
E[χ]	0.250–0.280	0.240	↓
Var[χ]	0.0015–0.0020	0.00062	↓↓↓
$CVaR_{0.95}(χ)$	0.290–0.310	0.262	↓↓↓
P(0.2 < χ < 0.3)	0.94–0.97	0.9993	↑

indicating improved control over extreme loading scenarios. More importantly, the reliability constraint is satisfied with a high confidence level, with $P(0.2<\chi<0.3) \approx 0.9993$

Unlike standard RBDO, which often yields boundary solutions, the RBRDO solution lies within the interior of the design space. This reflects the effect of robustness terms in penalizing extreme parameter values, resulting in more realistic designs with reduced sensitivity to surrogate modeling errors. The feasible design region under the reliability constraint is illustrated in Figure 10. It can be observed that the feasible region remains concentrated in the domain of relatively high material hardness and low-to-moderate surface roughness, which is consistent with the stiffness-controlled load transfer mechanism in bolted joints. However, compared to pure RBDO, the RBRDO solution avoids extreme parameter combinations, leading to a more balanced and stable design.

Overall, these results demonstrate that the proposed RBRDO framework effectively integrates reliability and robustness considerations, providing a more comprehensive and realistic basis for engineering design under uncertainty.

Robustness, uncertainty, and design trade-off

The robustness of the optimized design is evaluated by varying the CVaR confidence level (\alpha). As summarized in Table 6, increasing (α) from 0.90 to 0.99 leads to more conservative solutions, resulting in slightly higher expected values of (χ) but improved reliability. The probability $P(0.2 < \chi < 0.3)$ increases from 0.997 to 0.9995, indicating enhanced robustness against extreme loading conditions.

As expected, higher confidence levels yield larger ($CVaR\alpha$) values because more severe tail events are considered. These results demonstrate that the proposed RBRDO framework effectively balances mean performance and tail-risk reduction.

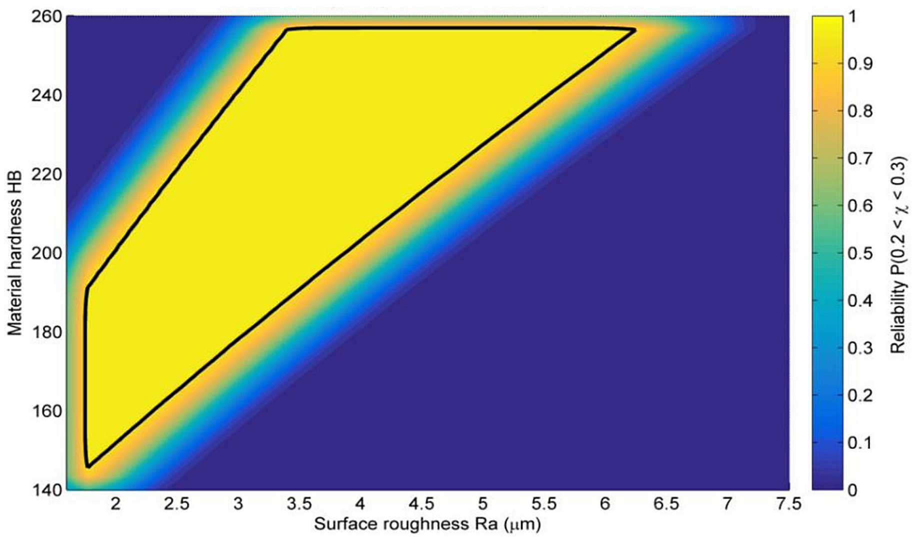

Figure 10. Feasible design region in the Ra–HB space under the reliability constraint

Table 6. Effect of the CVaR confidence level on performance and reliability in the RBRDO framework

CVaR confidence level (α)	Expected value E[χ]	ΔE(%)	CVaRα(χ)	Reliability P(0.2<χ<0.3)
0.90	0.236	–	0.258	0.997
0.95	0.240	+1.7%	0.262	0.999
0.99	0.245	+3.8%	0.274	0.9995

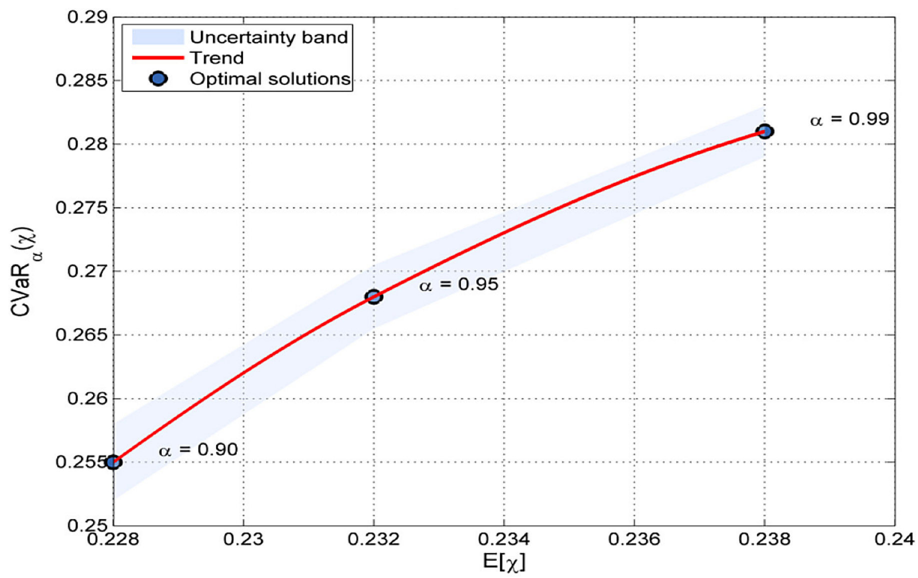

Figure 11. Trade-off between the expected external load factor E[χ] and the corresponding tail risk CVaRα(χ) for different confidence levels α

The trade-off between expected performance and robustness is illustrated in Figure 11. A Pareto-like relationship is observed, where lower expected values of (χ) are associated with increased sensitivity to extreme conditions, whereas more robust solutions achieve lower risk at the expense of slightly reduced performance.

The use of the surrogate model reduces the computational cost of uncertainty propagation by more than 90% compared with direct evaluation of the physics-based model, making repeated probabilistic analyses computationally feasible.

Despite these advantages, the current framework is limited by the relatively small dataset and the simplified representation of contact mechanics. Future work will focus on expanding the experimental database and incorporating more detailed interface behavior into the modeling framework.

CONCLUSIONS

This study presents a comprehensive framework for modeling, analysis, and optimization of the external load factor (χ) in bolted joints by integrating experimental data, physics-based understanding, and surrogate modeling within a reliability-based robust design optimization framework.

A structured experimental design was employed to investigate the nonlinear and coupled effects of μ, Ra, and HB. The results show that material hardness plays a dominant role in controlling χ, while surface roughness and friction coefficient influence contact stiffness and load redistribution.

A physics-informed surrogate model was developed to approximate the system response, demonstrating high predictive accuracy and strong generalization. Compared to conventional data-driven approaches, the incorporation of physical knowledge improves interpretability and extrapolation performance.

Based on this model, an RBRDO framework was formulated by incorporating both variance and Conditional value-at-risk into the objective function, enabling simultaneous control of mean performance, variability, and extreme responses.

The optimal design is characterized by moderately increased friction, reduced surface roughness, and increased material hardness, leading to improved stiffness and load distribution. As summarized in Table 5, the optimized solution reduces the expected external load factor and variability while satisfying the reliability constraint $P(0.2 < \chi < 0.3) \approx 0.9993$.

Sensitivity analysis of the CVaR confidence level reveals a clear trade-off between performance and robustness. Higher confidence levels lead to more conservative designs with improved reliability and reduced sensitivity to extreme conditions, at the cost of a slight increase in the expected response. An important outcome is that the optimal solution lies within the interior of the design space, indicating improved stability and robustness compared to boundary solutions.

From a computational perspective, the surrogate model significantly reduces the cost of uncertainty propagation, enabling efficient evaluation of reliability and robustness metrics. Despite these advantages, limitations include the relatively small dataset and simplified physical assumptions, which may affect accuracy in extreme conditions.

Acknowledgements

We acknowledge the support of time and facilities from Pham Van Dong University for this study.

REFERENCES

1. Shigley JE, Mitchell LD, Saunders H. Mechanical engineering design. 2015.
2. Loc NH. Fundamental of machine design: National University Publisher HCMC (in Vietnamese); 2016.
3. Tran VT, Nguyen HL. Experimental investigation of the influence of friction, surface roughness and material hardness on the external load factor in threaded joints. PLoS One. 2026;21(1):e0340073.
4. Liu Z, Yan X, Chen W, Niu N, Li M, Li Y. A prediction model of bolted joint loosening based on deep learning network. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2024;238(9):4240–9.
5. Alsardia T, Lovas L. Investigation of the effect of the surface treatment and lubrication during repeated tightening on the nut coefficient of a bolted joint using the Taguchi method. Jordan Journal of Mechanical & Industrial Engineering. 2024;18(1).
6. Wettstein A, Matthiesen S. Investigation of the thread coefficient of friction when impact tightening bolted joints. Forschung im Ingenieurwesen-Engineering Research. 2020;84(1):55-63.
7. Wang C, Sun Q, Han S. Effect of surface parameters on anti-loosening performance of bolts for engineering machinery. Matéria (Rio de Janeiro). 2023;28(3):e20230159.
8. Jalali H, Khodaparast HH, Friswell MI. The effect of preload and surface roughness quality on linear joint model parameters. J Sound Vib. 2019;447:186–204.
9. Stang N. Effect of surface roughness and external loading on embedment in steel and aluminum bolted joints: South Dakota State University; 2021.
10. Yılmaz M, Çorapsız MF. A robust MPPT method based on optimizable Gaussian process regression and high order sliding mode control for solar systems under partial shading conditions. Renewable Energy. 2025;239:122339.
11. Samadian D, Muhit IB, Dawood N. Application of data-driven surrogate models in structural engineering: a literature review. Archives of Computational Methods in Engineering. 2025;32(2):735–84.
12. Shirzadi N, Lau D, Stylianou M. Surrogate modeling for building design: Energy and cost prediction compared to simulation-based methods. Buildings. 2025;15(13):2361.
13. Zhang H, Li J, Yan M, Miao Y. Optimization of supersonic combustor configuration based on Gaussian process regression and genetic algorithm. Aerospace Science and Technology. 2025;159:109980.
14. Ma Z, Jiang G, Hu Y, Chen J. A review of physics-informed machine learning for building energy modeling. Applied Energy. 2025;381:125169.
15. Meng C, Griesemer S, Cao D, Seo S, Liu Y. When physics meets machine learning: A survey of physics-informed machine learning. Machine Learning for Computational Science and Engineering. 2025;1(1):20.
16. Zhang W, Suo W, Song J, Cao W. Physics-informed neural networks (PINNs) as intelligent computing technique for solving partial differential equations: Limitation and future prospects. Science China Physics, Mechanics & Astronomy. 2026;69(1):214602.
17. Chen Y, Wang H, Chen Z. Sensitivity analysis of physical regularization in physics-informed neural networks (PINNs) of building thermal modeling. Build. Environ. 2025;273:112693.
18. Chen J, Chen Z, Jiang W, Guo H, Chen L. A reliability-based design optimization strategy using quantile surrogates by improved PC-kriging. Reliability Engineering & System Safety. 2025;253:110491.
19. Gholami P, Yengejeh SE, Farsi MA. Reliability-based design optimization of composite pipes subjected to internal hydrostatic pressure fatigue. Composites Communications. 2025:102484.

20. Chen Y, Huang X, Ma M, Luo J, Ding B. Reliability-based design optimization of key components in a gantry machining center. Probabilistic Engineering Mechanics. 2025:103786.
21. Tran VT, Nguyen HL. Experimental investigation and reliability-based optimization of the nut factor in bolted joints considering friction coefficient, surface roughness and material hardness. PLoS One. 2026;21(6):e0351082.
22. Mulumudi DK, Manupriya P, Aminian G, Raj A. On the generalization and robustness in conditional Value-at-Risk. arXiv preprint arXiv:2602.18053. 2026.
23. Wang Q, Cao H, Hu J. Statistical inference in conditional Value-at-Risk optimization. Available at SSRN 5168942. 2025.
24. Khalid MA, Bansal S. A two-stage simulation-based framework for reliability-based robust design optimization of systems. Optimization and Engineering. 2025:1–36.
25. Hassani H, Khodaygan S. Reliability-based robust design considering performance degradation: a machine learning-driven approach for high-temperature helical springs. Research in Engineering Design. 2026;37(1):2.
26. Fang J, Xu Y, Li K, Zou S, Liu C, Wang T, et al. Vibration fatigue reliability-based robust design optimization of turbine blades in rocket engines using Kriging surrogate model with sequential sampling. Structural and Multidisciplinary Optimization. 2026;69(4):94.