

Article

A Unified Surrogate Framework for Data-Driven Reliability Analysis of Mechanical Systems from Low to Multi-DOF

Lun Shao ¹ , Alexandre Saidi ², Abdel-Malek Zine ³  and Mohamed Ichchou ^{1,*}¹ Laboratory of Tribology and Dynamics of Systems, Ecole Centrale Lyon, 69130 Ecully, France; lun.shao@ec-lyon.fr² Computer Science Laboratory for Image Processing and Information Systems, Ecole Centrale Lyon, 69134 Ecully, France; alexandre.saidi@ec-lyon.fr³ Institut Camille Jordan, Ecole Centrale de Lyon, 69134 Ecully, France; abdel-malek.zine@ec-lyon.fr

* Correspondence: mohamed.ichchou@ec-lyon.fr

Abstract: This paper proposes a unified reliability analysis framework for mechanical and structural systems equipped with Tuned Mass Dampers (TMDs), encompassing single-degree-of-freedom (1-DOF), two-degrees-of-freedom (2-DOF), and ten-degrees-of-freedom (10-DOF) configurations. The methodology integrates four main components: (i) probabilistic uncertainty modeling for mass, damping, and stiffness, (ii) Latin Hypercube Sampling (LHS) to efficiently explore parameter variations, (iii) Monte Carlo simulation (MCS) for estimating failure probabilities under stochastic excitations, and (iv) machine learning models, including Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), and Neural Networks (NNs), to predict structural responses and failure probabilities. The results demonstrate that ensemble methods, such as RF and XGBoost, provide high accuracy and can effectively identify important features. Neural Networks perform well for capturing nonlinear behavior, although careful tuning is required to prevent overfitting. The framework is further extended to a 10-DOF structure, and the simulation results confirm that machine learning-based models are highly effective for large-scale reliability analysis. These findings highlight the synergy between simulation methods and data-driven models in enhancing the reliability of TMD systems under uncertain inputs.



Academic Editor: Aleksandar Pavic

Received: 16 January 2025

Revised: 10 February 2025

Accepted: 18 February 2025

Published: 20 February 2025

Citation: Shao, L.; Saidi, A.; Zine, A.-M.; Ichchou, M. A Unified Surrogate Framework for Data-Driven Reliability Analysis of Mechanical Systems from Low to Multi-DOF. *Vibration* **2025**, *8*, 7. <https://doi.org/10.3390/vibration8010007>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: structural reliability; tuned mass damper; machine learning; multi-degree-of-freedom; surrogate modeling

1. Introduction

Modern engineering structures, ranging from tall buildings and long-span bridges to industrial machinery and aerospace components, are becoming increasingly complex and are subjected to diverse and often stochastic loading conditions [1,2]. Ensuring that these systems maintain their structural integrity and reliability over their intended service life is of paramount importance. Traditionally, Structural Reliability Analysis (SRA) has employed deterministic or semi-probabilistic methods under simplified assumptions [3,4]. While these foundational approaches offer valuable insights, they often fail to capture the true complexity of systems due to multiple sources of uncertainty, such as material properties, geometric variations, and environmental forces.

As industrial demands push for more sophisticated structures and integrate performance-based design principles, the need for more accurate and computationally efficient reliability

assessment tools grows markedly [5,6]. Monte Carlo simulations, polynomial chaos expansions, and other uncertainty quantification methods have expanded the scope of SRA [7]; however, they frequently become prohibitively expensive when the number of uncertain parameters and the complexity of the structural model increase. This challenge is especially pronounced in multi-degree-of-freedom (MDOF) systems, where interactions among numerous components must be accurately accounted for. For example, many large-scale infrastructures and mechanical systems employ Tuned Mass Dampers to mitigate vibrational responses [8–10]. Although TMDs can significantly reduce the likelihood of catastrophic failures under dynamic loading, their performance depends on a range of uncertain factors, including mass, damping, stiffness, and external excitation characteristics.

Tuned Mass Dampers [11–14] are widely used to reduce the harmful effects of dynamic forces from wind, earthquakes, and other external loads. The literature [15,16] highlights the use of TMDs to mitigate vibrations. Liu et al. investigated the application of TMDs in high-rise buildings [17]. Miguel et al. discussed the role of TMDs in bridge structures [18,19]. Many studies focus on simple models like 1-DOF or 2-DOF systems to understand basic vibration control. However, real structures often have many degrees of freedom, making modeling and reliability analysis more complex and computationally demanding.

In the presence of uncertainties in parameters and external loads, Monte Carlo simulation has established itself as a robust technique for assessing structural reliability, particularly for random vibration problems [20,21]. However, the high computational cost of exhaustive simulations can become impractical as the number of degrees of freedom and thus the complexity increases. Recent research [22–27] has explored hybrid strategies that combine simulation-based approaches with machine learning to address this computational burden. In these approaches, ML models serve as surrogate predictors for system responses [28–31], thereby reducing the overall number of high-fidelity simulations required.

Despite these advancements, several limitations persist in existing reliability analysis methods for TMD-equipped systems: Firstly, traditional methods like MCS are computationally expensive, especially for MDOF systems. Secondly, many studies focus on low-DOF systems, which do not adequately represent the complexity of real-world structures. Thirdly, existing methods often fail to comprehensively account for the uncertainties in mass, damping, and stiffness parameters.

This article addresses these limitations by unifying the concepts of uncertainty quantification, Monte Carlo simulation, and machine learning within a single framework. There are three specific objectives of this work:

- To demonstrate how uncertainty in mass, damping, and stiffness affects the vibration response and failure probability of TMD systems.
- To compare different machine learning models as efficient tools for predicting structural responses and failure probabilities.
- To demonstrate the scalability of the method by applying it to systems with both low and high degrees of freedom.

The practical implications of this study are significant. The proposed framework can be directly applied to real-world engineering problems, such as the design and optimization of TMDs for high-rise buildings, long-span bridges, and industrial machinery. By providing a computationally efficient tool for reliability analysis, this work enables engineers to assess the performance of TMD-equipped systems under uncertain loading conditions, thereby improving safety and reducing the risk of catastrophic failures. Additionally, the scalability of the method to high-DOF systems makes it suitable for complex structures that are difficult to analyze using traditional approaches.

In the subsequent sections, the methodologies ranging from uncertainty modeling and Monte Carlo simulation to machine learning are elaborated. The results for each system size are then presented, concluding with an analysis of model performance, feature importance, and the role of TMD parameters in mitigating vibrations.

2. Proposed Methodology

Figure 1 outlines the procedural framework. The process begins with the initiation of the reliability analysis framework for the mechanical system in question. This serves as the entry point for the subsequent steps that will systematically model uncertainties, conduct simulations, and apply machine learning techniques.

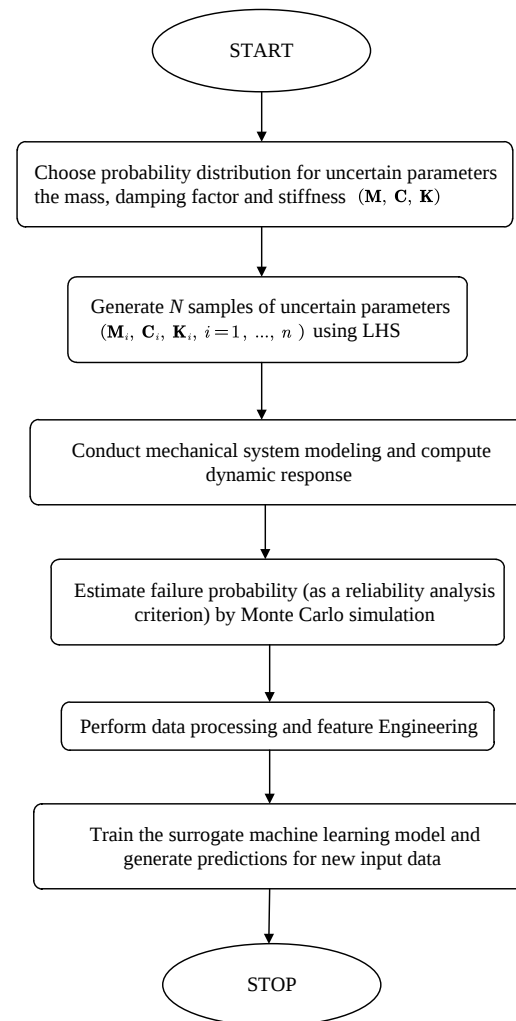


Figure 1. Flowchart of the proposed reliability analysis procedure, combining Monte Carlo simulation, TMD-based system modeling, and ML surrogates.

Figure 1 illustrates the theoretical framework for reliability analysis of a mechanical system. The figure provides a step-by-step explanation of the process, including the following:

- Identification of uncertain parameters and their probability distributions.
- Use of Latin Hypercube Sampling for generating input samples.
- Dynamic response analysis and failure probability calculation using Monte Carlo simulation.
- Training and validation of the machine learning surrogate model.

The process begins by identifying the uncertain parameters, such as mass (M), damping coefficient (C), and stiffness (K), and assigning suitable probability distributions to

represent their variability. These distributions help capture practical uncertainties that can arise due to variations in materials or operating conditions.

Next, based on the selected distributions, Latin Hypercube Sampling is used to create a wide range of samples. It could provide a better coverage of the input than random sampling. By using these sampled parameters, the mechanical system is analyzed to evaluate its dynamic response.

Once the dynamic responses are obtained, Monte Carlo simulation is performed to estimate the failure probability. By running multiple simulations across the entire set of sampled parameters, a statistical measure of the system's likelihood of failing to meet specified performance criteria is calculated. This probability helps explain the risk linked to different parameter setups.

Next, the response data are cleaned, normalized, and processed to prepare them for machine learning. These steps turn the raw simulation results into a dataset that works for training the model. Finally, a machine learning model is created to predict how the system behaves. It runs much faster than the original simulations. Once trained, it can quickly predict the system's response for new input parameters, enabling efficient and practical reliability assessments in complex mechanical systems.

2.1. Uncertainty Modeling and Latin Hypercube Sampling

In a deterministic scenario, where the structural properties are fully known, the likelihood of failure can be measured directly by calculating the failure probability. However, this becomes more complex when uncertainties in the material's properties are introduced. To quantify these uncertainties, multiple samples are drawn from the uncertainty space. Reliability analysis is then performed using first-passage probability theory and Monte Carlo simulations.

The present analysis assumes that uncertainties in mass (m), damping (c), and stiffness (k) are each represented by normal distributions, $m \sim \mathcal{N}(m_0, \sigma_m^2)$, etc., following [32]. These distributions allow for a realistic portrayal of the inherent variability in these physical properties. LHS generates N samples of (m, c, k) , ensuring a more thorough exploration of input space than simple random sampling. Each sample is then used to simulate the structural response under random excitation.

2.2. Failure Mechanism

A failure mechanism describes the physical process through which stresses cause damage to the components of a system, ultimately leading to its failure. According to [33,34], overstress failures occur when the applied stress exceeds the material's maximum allowable strength, causing significant elastic deformation, yielding, or fracturing. When the applied stress remains below this threshold, the material usually recovers without any lasting damage.

In the context of random vibration, similar failure mechanisms have been identified, as discussed in [35]. This study focuses on the first-passage failure mechanism. Specifically, it analyzes the overstress failure mechanism when a structure is subjected to static loads. For dynamic loads, the first-passage problem becomes the primary focus. Failure is defined as the random response exceeding a critical threshold for the first time.

We assume the first-passage failure: the structure fails if the maximum displacement $y_{\max}$ surpasses a threshold y_{limit} at any time within $t \in [0, T]$ [35]. Mathematically,

$$P_f = \Pr(\max |Y(t)| > y_{\text{limit}}, t \in [0, T]) \quad (1)$$

The threshold often reflects design codes or operational limits. If $y_{\max} \leq y_{\text{limit}}$, the system remains safe.

2.3. Monte Carlo Simulation for Reliability

Monte Carlo simulation obtains an empirical estimate of P_f by repeating the dynamic analysis for N_{mc} random excitations (Gaussian white noise or similar). Each trial yields a pass/fail outcome. The fraction of trials exceeding y_{limit} approximates the failure probability:

$$\hat{P}_f = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} I_F(x^{(i)}), \quad (2)$$

where $I_F = 1$ if failure occurs and 0 otherwise. As N_{mc} increases, $\hat{P}_f$ becomes closer to the true probability of failure [36].

This method calculates the failure probability by adding the indicator function values and dividing by the number of samples, N_{mc} . The Monte Carlo method works well for complex systems. It is helpful when solving the problem directly is not possible. It gives reliable results by following the law of large numbers. As the sample size increases, the estimate becomes more accurate. In the Monte Carlo simulation, 10,000 samples were generated for each random variable to ensure statistical stability of the reliability estimates while maintaining computational efficiency.

2.4. State-Space Formulation and TMD Modeling

The TMD system, shown in Figure 2, includes two main parts: the primary structure, modeled as an N -degree-of-freedom (N-DOF) system, and the TMD unit, represented as an n -degree-of-freedom (n-DOF) system.

The TMD is attached to the primary structure in parallel. This allows for dynamic interaction between the two systems. The TMD could absorb vibrational energy from the main structure so that it can reduce the amplitude of vibrations and improve the stability.

This mathematical framework serves as a foundation for analyzing and optimizing the dynamic response of structures equipped with TMDs.

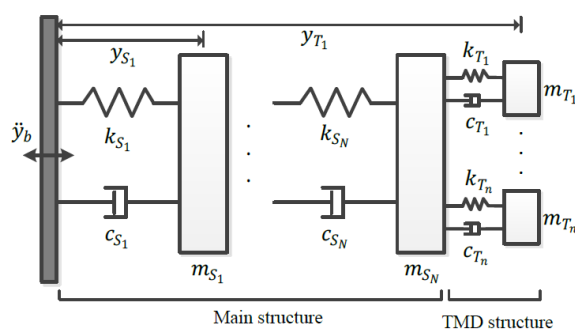


Figure 2. Multi-degree-of-freedom structural system equipped with TMDs.

TMDs are placed to mitigate resonant frequencies. For an N -DOF structure plus TMD, mass ($\mathbf{M}$), damping ($\mathbf{C}$), and stiffness ($\mathbf{K}$) matrices can be written in block form to include the TMD degrees of freedom [27]. The equation of motion is

$$\mathbf{M} \ddot{\mathbf{y}}(t) + \mathbf{C} \dot{\mathbf{y}}(t) + \mathbf{K} \mathbf{y}(t) = -\mathbf{M} \mathbf{r} \ddot{y}_b(t), \quad (3)$$

where $\ddot{y}_b(t)$ is the stochastic base acceleration, and $\mathbf{r}$ is a vector distributing base excitation. The mass matrix $\mathbf{M}$ is block-diagonal and can be written as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_S & 0 \\ 0 & \mathbf{M}_T \end{bmatrix} \quad (4)$$

where $\mathbf{M}_S$ is the mass matrix for the structure and $\mathbf{M}_T$ is the mass matrix for the TMD system. Specifically,

$$\mathbf{M} = \text{diag}(m_{s_1}, m_{s_2}, \dots, m_{s_n}, m_{T_1}, \dots, m_{T_m}) \quad (5)$$

Similarly, the damping matrix $\mathbf{C}$ and stiffness matrix $\mathbf{K}$ can be expressed in block-diagonal form:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_S & 0 \\ 0 & \mathbf{C}_T \end{bmatrix} \quad (6)$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_S & 0 \\ 0 & \mathbf{K}_T \end{bmatrix} \quad (7)$$

Next, we derive the state-space representation of the system by defining the state vector $\mathbf{X}$, which contains both displacements and velocities:

$$\mathbf{X} = [y_S, y_T, y_f, \dot{y}_S, \dot{y}_T, \dot{y}_f]^T \quad (8)$$

The dynamic equation in state-space form becomes

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{f} \quad (9)$$

where $\mathbf{f}$ is the input vector containing zeros except for the last element, $-\omega(t)$, which represents Gaussian white noise. The matrix $\mathbf{A}$ is the state system matrix, and the output equation is

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}\mathbf{X} + \mathbf{f} \\ \mathbf{Y} &= \mathbf{C}\mathbf{X} \end{aligned} \quad (10)$$

Here, $\mathbf{C}$ is a diagonal matrix that selects the degrees of freedom of interest. The state system matrix $\mathbf{A}$ is structured as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{H}_k & \mathbf{H}_c \end{bmatrix} \quad (11)$$

where

$$\mathbf{H}_k = \begin{bmatrix} -(\mathbf{M}^{-1}\mathbf{K}) & \omega_f^2 \\ 0 & -\omega_f^2 \end{bmatrix} \quad (12)$$

$$\mathbf{H}_c = \begin{bmatrix} -(\mathbf{M}^{-1}\mathbf{C}) & 2\varepsilon_f\omega_f \\ 0 & -2\varepsilon_f\omega_f \end{bmatrix} \quad (13)$$

Here,

- ε_f is the damping ratio of the seismic model;
- ω_f is the natural frequency of the seismic model;
- $\mathbf{H}_k$ describes the stiffness dynamics;
- $\mathbf{H}_c$ captures the damping dynamics.

The state-space formulation is implemented by MATLAB R2024b and supported by numerical libraries. The displacement of the structure is determined by solving the dynamic equations of motion for the mechanical system. These equations are derived based on the system's mass ($\mathbf{M}$), damping ($\mathbf{C}$), and stiffness ($\mathbf{K}$) matrices as well as the applied external forces.

In the seismic model, Gaussian white noise serves as the excitation source. After deriving the dynamic equations and forming the state system matrix, the state-space

model of the system can be constructed. Using this model, Monte Carlo simulation can be employed to calculate the structure's displacement and failure probability.

2.5. Surrogate Modeling via Machine Learning

To minimize the need for repeated full-scale simulations, a surrogate model is trained using Monte Carlo simulation data. Models such as Random Forest, Gradient Boosting, XGBoost, and Neural Networks are employed. Each training sample consists of uncertain parameters and, optionally, time-domain signal features (e.g., RMS, kurtosis) mapped to observed outcomes, such as failure states or peak displacement. Once trained, the surrogate model enables rapid reliability predictions for new realizations of the input parameters, which were sampled from their defined statistical distributions. The surrogate model is used to offset the computational difficulty in predicting the structure's displacement by approximating the input–output relationship, thereby avoiding repeated expensive simulations during the MCS process.

The machine learning model acts as a surrogate to predict the failure probability of the structure for different input values, replacing the need for expensive simulations during reliability analysis.

Comparison of Machine Learning Algorithms

- Random Forest constructs an ensemble of decision trees using bootstrap sampling and randomized feature splits. It provides reliable performance and feature importance insights, although it may show slightly higher errors in highly nonlinear scenarios.
- Gradient Boosting sequentially builds trees to reduce the residual errors of the ensemble. It often achieves high accuracy but requires longer training times.
- Extreme Gradient Boosting (XGBoost) is a faster and more efficient boosting algorithm. It uses regularization and parallelization to maintain a balance between accuracy and efficiency.
- Neural Networks are good at modeling complex, nonlinear patterns. Their performance depends on proper tuning of hyperparameters and scaling of input data. When tuned well, Neural Networks often achieve the best accuracy.

The selection of machine learning models, such as Random Forest, Gradient Boosting, XGBoost, and Neural Networks, was driven by our focus on ensemble methods, which are particularly well suited for reliability analysis. RF, as a bagging-based ensemble method, provides robustness to noise and overfitting while offering interpretability through feature importance scores. GB and XGBoost, as boosting-based ensemble methods, deliver high predictive accuracy by sequentially correcting errors from previous models. XGBoost, in particular, is optimized for computational efficiency and scalability, making it ideal for large-scale reliability analysis. NNs were included as a complementary approach to capture highly complex, nonlinear patterns in the data. This combination of ensemble methods and NN ensures a balance between accuracy, interpretability, and computational efficiency for reliability analysis of TMD-equipped systems.

The models are compared using metrics like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Relative Error.

3. Numerical Study and Benchmarks

This section shows the simulation results for 1-DOF and 2-DOF structural systems. The input is random, and probabilities are used to model the structural properties. This method helps study how the system responds. Machine learning models are trained on data from simulations to predict the system's behavior. This approach is faster and less expensive than running many simulations.

3.1. Part A: Low-DOF Systems

3.1.1. Single Degree-of-Freedom System

We assume that the single-degree-of-freedom (1-DOF) system is subjected to Gaussian white noise. The nominal values for the system are summarized in Table 1. Uncertainties in the mass (m), damping (c), and stiffness (k) parameters are modeled as normal distributions, which are denoted by $\mathcal{N}(\cdot)$. The standard deviations for the parameters are

$$\sigma_m = 0.1, \quad \sigma_c = 0.001, \quad \sigma_k = 5$$

These values show the uncertainty in each physical property. The parameters of the base structure and the TMD are selected to reflect a practical vibration control setup. To include variability, Latin Hypercube Sampling is used to create a wide range of parameter combinations. Each combination is tested with random base acceleration to analyze how the system behaves under uncertain conditions.

Table 1. Sample nominal parameters for a 1-DOF TMD system.

Component	Mass (kg)	Damping (N·s/m)	Stiffness (N/m)
Base structure	1.0	0.03	696.4
TMD	0.02	0.0695	12.725

When using machine learning algorithms with random elements, like Random Forest, it is important to validate their performance reliably. The randomness in these algorithms can cause results to vary between runs. To handle this, repeated k -fold cross-validation is used. This method calculates performance metrics by averaging results over several runs. It reduces the effect of randomness and gives a more consistent measure of how well the model works.

Figure 3 shows the performance of the Random Forest model with different datasets. In Figure 3a, the correlation coefficient measures the linear relationship between predicted and actual values. We can see that as the dataset size increases from 500 to 2000 samples, the correlation coefficient improves and remains stable around 0.9990 when it comes to 2000 samples. After that, the risk of overfitting should be considered.

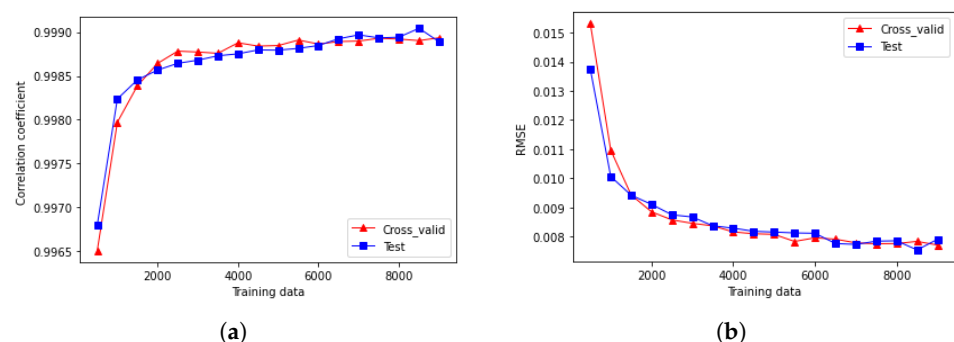


Figure 3. Performance metrics of the Random Forest model: (a) correlation coefficient between predicted and observed values. (b) RMSE representing the average prediction error magnitude.

Figure 3b depicts the RMSE, which measures prediction accuracy. RMSE decreases significantly with larger datasets, stabilizing around 0.008 when the sample size exceeds 2000. The close agreement between cross-validation and test dataset values further demonstrates the model's robustness and generalization capability.

Hyperparameter optimization was performed for each machine learning model to ensure optimal performance. For Random Forest, Gradient Boosting, and XGBoost, a grid

search cross-validation approach was employed with a 5-fold cross-validation strategy to evaluate model performance. Key hyperparameters such as the number of trees, maximum depth, and learning rate were tuned. For Neural Networks, Bayesian Optimization was used to optimize the number of layers, number of neurons per layer, activation functions, learning rate, and dropout rate. The optimization process was implemented using Python libraries such as Scikit-learn and Optuna with RMSE as the performance metric for regression tasks. This rigorous tuning process ensures that each model is well suited for predicting structural responses and failure probabilities.

Additionally, we compare the performance of three machine learning models—Random Forest, Gradient Boosting (GB), and XGBoost—in predicting system responses across different sample sizes. The comparison is based on the Mean Absolute Error (MAE), where lower MAE values represent higher prediction accuracy.

Table 2 shows the Mean Absolute Error values for three machine learning models with different sample sizes. The results show that as the sample size increases, the MAE of the models becomes smaller. XGBoost performs the best with the lowest MAE. Gradient Boosting also does well, especially with larger datasets. Random Forest works well for smaller datasets, but its performance shifts as the dataset grows.

Table 2. Comparison of MAE for different ML models across varying sample sizes.

Sample Size	Random Forest	Gradient Boosting	XGBoost
2000	0.004830	0.004479	0.004447
5000	0.004204	0.004013	0.003911
8000	0.003755	0.003840	0.003819
10,000	0.003763	0.003708	0.003794

These findings show that boosting models, especially XGBoost, are better for larger datasets and complex patterns. This is because they have advanced ways to reduce errors. On the other hand, Random Forest may be a better choice for scenarios with smaller datasets where computational efficiency is a key requirement despite its relatively lower accuracy for larger datasets.

3.1.2. Two-Degree-of-Freedom TMD System

This simulation analyzes a 2-DOF primary structure combined with a 1-DOF TMD. In this setup, the two primary masses are subjected with dynamic forces, while the TMD absorbs vibrational energy from both masses. This system resembles real-world structures, where multiple connected components interact dynamically and require coordinated vibration control methods.

Table 3 provides the nominal values of structural parameters, including mass, damping coefficient, and stiffness. The system consists of two primary structure and a TMD.

Table 3. Nominal values of the 2-DOF structure.

Parameter	Mass (kg)	Damping (N · s/m)	Stiffness (N/m)
1st DOF	6	62	6500
2nd DOF	6	62	6500
TMD	1.38	38.997	1.8327

Parameter Tuning for Feature Selection

The structural features m_{s1} , m_{s2} , k_{s1} , k_{s2} , c_{s1} , and c_{s2} represent the mass, stiffness, and damping coefficients in the system. These parameters are very important to describe the dynamic behavior and response when it is subjected to external forces.

Figure 4 presents the RMSE values for different machine learning models as the number of features increases. We found the following:

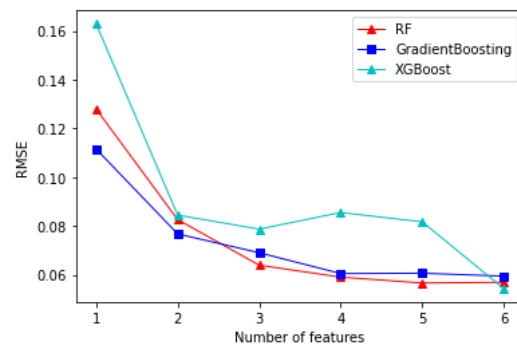


Figure 4. Performance of various ML algorithms based on the number of features considered.

- All models show a reduction in RMSE as the feature count increases from 1 to 6; therefore, we can achieve a better performance of prediction by adding features.
- XGBoost has the highest RMSE when only one feature is used, but then its performance improves a lot after adding a second feature. However, it maintains a slightly higher RMSE compared to the other models as more features are added.
- Gradient Boosting and Random Forest exhibit similar performance, achieving consistently lower RMSE values throughout, and converge to nearly identical levels when all six features are included.
- The convergence of model performance with more features suggests diminishing returns beyond a certain point where additional features have a minimal impact on reducing the RMSE.

The results show that all models improve with more features. However, Gradient Boosting and Random Forest are more robust and achieve better accuracy with lower RMSEs than XGBoost. The diminishing returns suggest that selecting the right number of features is key to balancing accuracy and efficiency. While increasing the number of features improves model accuracy, it also necessitates a larger training dataset to mitigate the curse of dimensionality. In this study, we addressed this challenge through feature selection, data augmentation, and regularization, ensuring robust model performance even in high-dimensional spaces.

Comparison of Machine Learning Models

Table 4 presents the performance of various machine learning models in predicting the response of the 2-DOF system. The Neural Network performs the best with an RMSE of 0.003703 and an MAE of 0.003264. This shows its strong ability to predict accurately in handling complex and nonlinear patterns.

Table 4. Comparison of RMSE and MAE for different ML models.

Model	RMSE	MAE
Random Forest	0.011663	0.007297
Gradient Boosting	0.009969	0.006335
Xtrem Gradient Boosting	0.009211	0.006150
Neural Network	0.003703	0.003264

XGBoost achieves the second-best performance, with an RMSE of 0.009211 and an MAE of 0.006150, which was followed by Gradient Boosting and Random Forest. Although these models provide accurate predictions, they do not perform as well as the Neural

Network, especially in capturing the nonlinear dynamics and complex interactions of the 2-DOF system.

Evaluation of Predictive Performance

The predictive performance of machine learning models for failure probability is evaluated, including Neural Networks, Random Forest, Gradient Boosting, and XGBoost. Figure 5 shows scatter plots comparing the tested and predicted values for each model, providing a clear visual of their accuracy and capabilities. In the evaluation of predictive performance, the true values are the failure probabilities calculated using Monte Carlo simulation. These values are used to check the accuracy of the machine learning model's predictions.

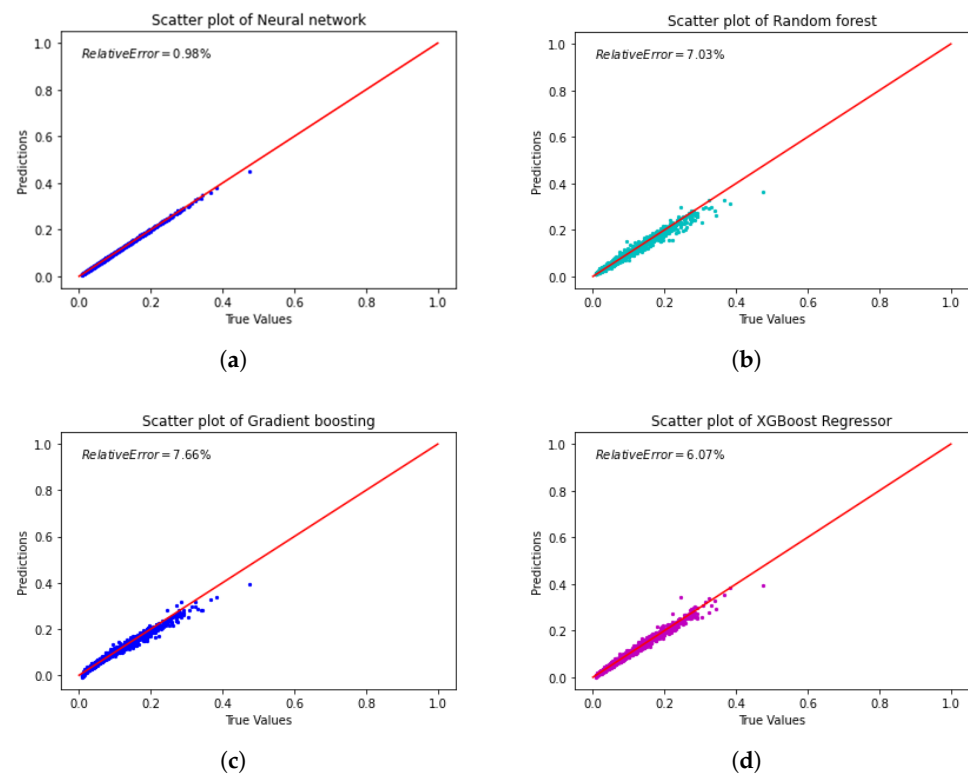


Figure 5. Scatter plots showing the relationship between tested and predicted values for failure probability across different ML models: (a) Neural Network. (b) Random Forest. (c) Gradient Boosting. (d) Extreme Gradient Boosting.

From the scatter plots, we can see the following: The Neural Network has the highest accuracy. The relative error is the lowest at 0.98%, which indicates its strong ability to handle nonlinear relationships. However, there are signs of overfitting, as it may have learned some noise from the training data, which could affect its performance on new data. Generally, the methods to reduce overfitting include the following:

- **Regularization:** Adding a penalty to the loss function to keep the model weights small. This helps prevent the model from becoming too complex.
- **Dropout:** Randomly turning off some neurons during training. This makes the model less dependent on specific neurons and improves its ability to generalize.
- **Cross-Validation:** Splitting the data into multiple parts and training the model on different subsets. This helps check if the model works well on new data.
- **Early Stopping:** Stopping the training process when the model's performance on validation data stops improving. This prevents the model from learning noise in the training data.

- **Data Augmentation:** Adding more data or creating new data by making small changes to the existing data. This helps the model learn better patterns.
- **Simplifying the Model:** Using fewer layers or neurons in the neural network. A simpler model is less likely to overfit.

As for Random Forest, Gradient Boosting, and XGBoost, they also perform well but have slightly higher prediction errors compared to the Neural Network. The scatter plots show bigger deviations from the ideal prediction line, which means lower accuracy. The relative errors are 7.03% for Random Forest, 7.66% for Gradient Boosting, and 6.07% for XGBoost. XGBoost is the best at capturing complex patterns, while Random Forest generalizes better because of its ensemble approach.

Some outliers are present in the scatter plots. These could be due to the data complexity, limited training samples, measurement noise, or model limitations. For instance, simpler models like Random Forest may have difficulty capturing highly nonlinear relationships, leading to inaccuracies in predictions.

3.2. Part B: Ten-Degree-of-Freedom Structure with TMD

This section shows the simulation results for a 10-degree-of-freedom structural model combined with a TMD. The primary objective is to evaluate the predictive performance of various machine learning models for these systems under stochastic excitations. The analysis is conducted under the following conditions:

- A 10-DOF structural model coupled with a single TMD unit.
- Base excitations modeled as Gaussian white noise to simulate random external forces.

The normal distribution is commonly used to model random variables in engineering and natural systems. It is a reasonable assumption when the true distribution is unknown, as it often approximates real-world behaviors well. Many random processes tend to follow a normal distribution when influenced by multiple independent factors, making it a suitable choice for complex systems like the ten-degree-of-freedom system analyzed here.

3.2.1. Impact of Parameter Uncertainty on Structural Dynamics

To evaluate the impact of parameter uncertainties, the mass (m), damping coefficient (c), and stiffness (k) of the primary structure are modeled as independent normal distributions. Table 5 provides the nominal values of the structural parameters along with their mean values and standard deviations (SD). Adding these uncertainties makes the model closer to real-world conditions and helps improve the accuracy and reliability of failure probability predictions.

Table 5. Nominal values of the TMD-based structure.

Variables	Mean	SD
$m_{s_1}, \dots, m_{s_{10}}$	360×10^4 kg	36×10^3 kg
$c_{s_1}, \dots, c_{s_{10}}$	620×10^4 Ns/m	62×10^4 Ns/m
$k_{s_1}, \dots, k_{s_{10}}$	650×10^6 N/m	65×10^6 N/m

Each mass parameter (m_{s_1} to $m_{s_{10}}$) is modeled with a mean value of 360×10^4 kg and a standard deviation of 36×10^3 kg, reflecting a consistent distribution across the structural elements. The damping coefficients (c_{s_1} to $c_{s_{10}}$) have a mean of 620×10^4 Ns/m and a standard deviation of 62×10^4 Ns/m, representing the system's energy dissipation capabilities and the TMD's role in minimizing oscillations. Similarly, the stiffness parameters (k_{s_1} to $k_{s_{10}}$) are characterized by a mean value of 650×10^6 N/m and a standard deviation of 65×10^6 N/m, contributing to the overall rigidity and load-bearing capacity.

of the structure. To account for these uncertainties, N Latin Hypercube Sampling samples (e.g., 8000–10,000) are generated, covering all floors' mass, damping, and stiffness parameters. LHS was selected for its ability to provide better coverage of the input space compared to random sampling. A pass/fail outcome is recorded for each sample based on whether the top-floor displacement ($y_{\max}$) exceeds a predefined limit (y_{limit} , e.g., 0.28 m). While the training dataset consisted of 8000–10,000 samples, the surrogate model was validated against high-fidelity MCS results to ensure its accuracy. This allowed us to use the surrogate model for large-scale MCS without significant loss of precision.

3.2.2. Feature Importance

Random Forest regression is employed to identify which floors significantly influence the structural behavior. The mass parameters of the upper floors (e.g., m_9, m_{10}) are usually the most important features. This is because the TMD is located there, and higher levels experience larger displacements. Figure 6 shows that the mass of the upper floors ranks highest in importance. This suggests that design efforts should prioritize optimizing the properties of the top floors to improve how the structure performs.

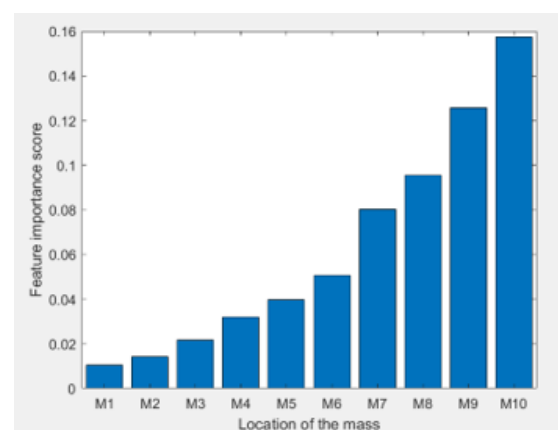


Figure 6. Feature importance of different mass units.

Below is an analysis of the mass units (m_1 to m_{10}) based on their location within the structure:

- Mass Units m_1 to m_5 : These lower-level mass units exhibit lower importance scores due to their positioning at the base of the structure. They are less influenced by significant lateral displacements that typically occur at higher levels during dynamic loading.
- Mass Units m_6 to m_8 : Positioned in the mid-levels of the structure, these mass units show moderate importance scores. They act as intermediaries, facilitating the transmission of vibrational energy from the lower levels to the upper floors.
- Mass Units m_9 and m_{10} : The top-level mass units have the highest importance scores. Due to their location, m_9 and m_{10} are significantly affected by large lateral displacements caused by external dynamic forces such as seismic or wind loads. The TMD, strategically placed at the top, interacts directly with these mass units to mitigate vibrations.

3.2.3. Time Domain Feature Analysis

In the previous section, we examined key physical properties like mass, stiffness, and damping, which determine the dynamic behavior of the structure. Here, we focus on time-domain features from vibration signals to explore their impact on model predictions. Each feature provides unique insights into the system's vibrations. For example, kurtosis highlights sharp peaks in the signal. Skewness checks if the signal is uneven and identifies

any bias in vibration. The Root Mean Square (RMS) measures the energy of the signal. The crest factor compares the signal's peak value to its RMS.

Figure 7 shows a plot matrix that illustrates how time-domain features (kurtosis, RMS, skewness, crest factor) relate to failure probability (P_f). The RMS has the strongest connection to P_f , which is shown by a clear linear trend. This makes the RMS a dependable way of predicting structural health. Higher RMS values, which indicate more vibrational energy, often signal potential structural issues or a greater chance of failure.

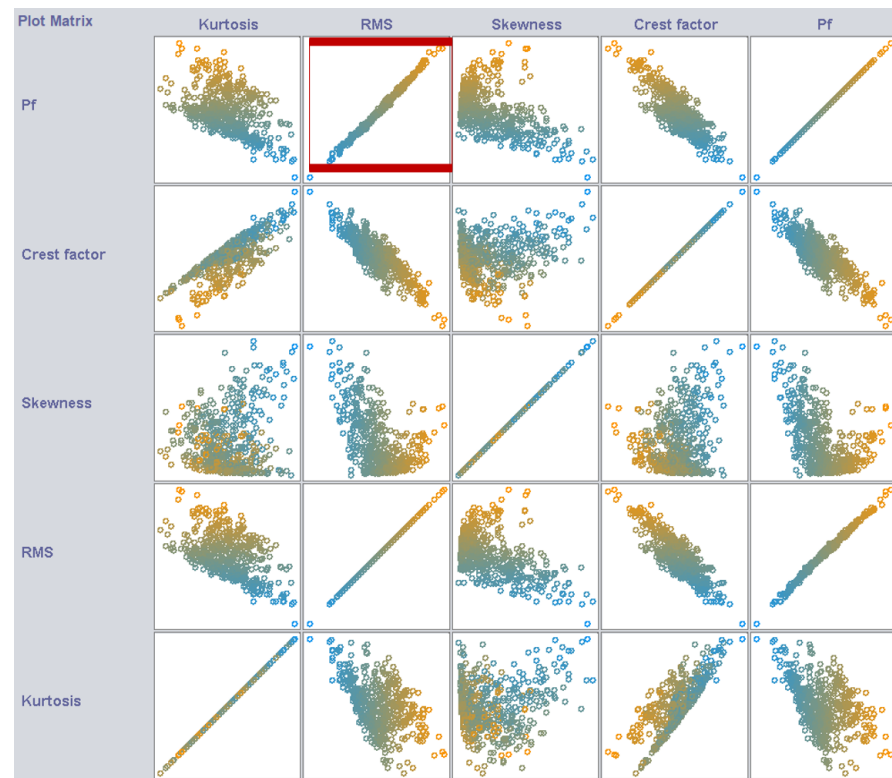


Figure 7. Plot matrix of inputs and output, showing the relationship between different parameters.

The crest factor also demonstrates a notable influence on failure probability, particularly during transient events. Its sensitivity to sudden peaks makes it a valuable feature for identifying potential structural issues, although its impact is less consistent compared to the RMS. In contrast, skewness and kurtosis provide complementary information about the distribution shape of the vibration signal but exhibit weaker direct correlations with P_f . In summary, the RMS and crest factor emerge as the most critical predictors of failure probability, while skewness and kurtosis offer additional insights that enhance the model performance overall.

3.2.4. Comparative Results of ML Models

Figure 8 compares the tested and predicted values across various machine learning models.

Random Forest (Figure 8a): The scatter plot illustrates a close alignment of predicted values with actual values with points clustering tightly around the diagonal red line representing perfect predictions. A relative error of 0.75% demonstrates reasonable predictive accuracy. However, minor deviations suggest limitations in capturing more complex nonlinear interactions within the dataset.

XGBoost (Figure 8b): The scatter plot for XGBoost similarly shows strong alignment between predicted and actual values, achieving the same relative error of 0.75% as Random

Forest. The tight clustering around the diagonal line underscores XGBoost’s reliability and its comparable performance to Random Forest in managing the dataset’s characteristics.

Neural Network (Figure 8c): The Neural Network outperforms other models, as evidenced by the tight clustering of predicted values along the diagonal line. The Neural Network achieves the lowest relative error of 0.23%, showing excellent predictive accuracy. Its ability to handle nonlinear relationships explains why it outperforms Random Forest and XGBoost.

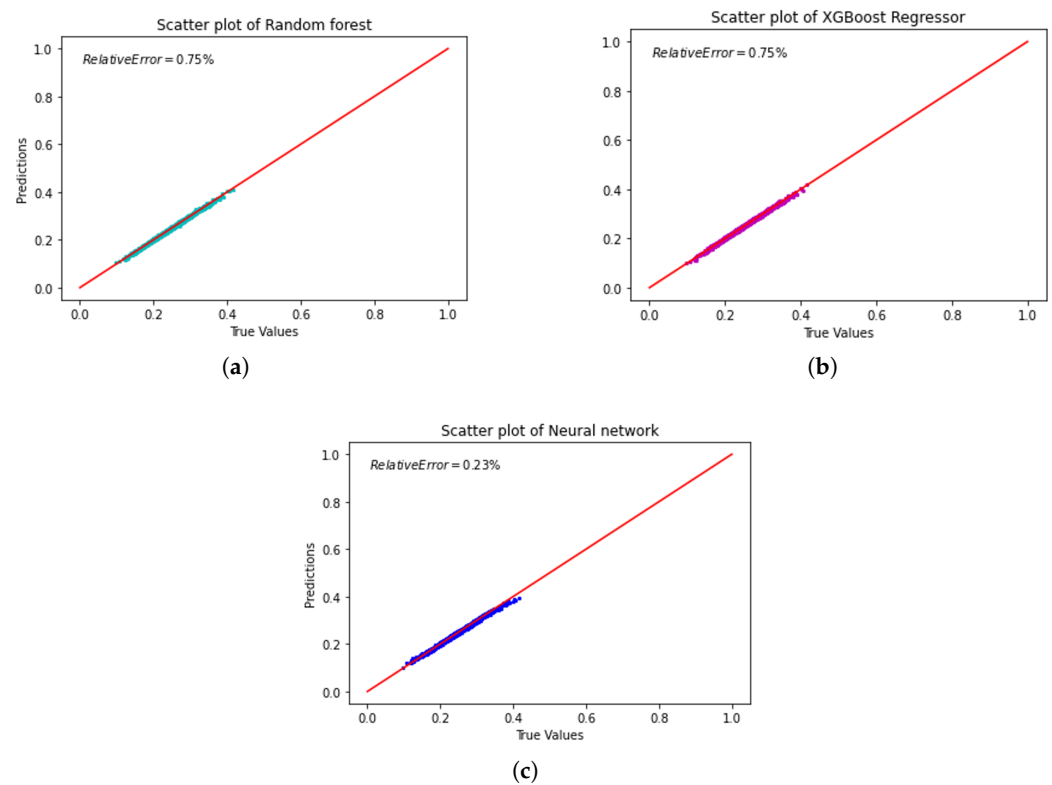


Figure 8. Scatter plots comparing tested vs. predicted values for failure probability using different ML models. (a) Test vs. predicted values (Random Forest). (b) Test vs. predicted values (XGBoost). (c) Test vs. predicted values (NN).

The performance of different machine learning models for predicting the system response is summarized in Table 6. The metrics used for evaluation include RMSE, MAE, and relative error. These metrics provide insights into the accuracy of each model with lower values indicating better performance.

Table 6. Comparison of RMSE, MAE, and relative error for different ML models structure.

Model	RMSE	MAE	Relative Error
Random Forest	0.002204	0.001703	0.0075
XGBoost	0.002160	0.001678	0.0075
Neural Network	0.001651	0.001302	0.0023

These results highlight the trade-offs between machine learning models. The Neural Network stands out for capturing complex nonlinear patterns, making it the best choice for tasks needing high accuracy, as shown by its low relative error. Ensemble methods such as Random Forest and XGBoost perform well with the low RMSE (0.002204 and 0.002160), making them suitable for scenarios where reducing significant prediction errors is critical.

ML-based models are useful for large-scale reliability analysis, which is substantiated by the following evidence. First, the computational efficiency of the framework was

evaluated by comparing traditional MCS without surrogates to MCS with ML surrogates. For the 10-DOF system, the use of ML surrogates reduced the computational time by over 90% from approximately 28 h to less than 2 h. Second, the training times for ML models were recorded with XGBoost requiring 1.5 h for training and less than 0.01 s for each prediction. Third, the accuracy of ML surrogates was validated against traditional MCS, achieving RMSE values below 0.005. Finally, the framework demonstrates excellent scalability with prediction times remaining nearly constant as the number of DOFs increases.

4. Conclusions

This study has established a robust reliability analysis framework for TMD-equipped structures, ranging from simple 1-DOF and 2-DOF systems to a complex 10-DOF building-like model. By using probabilistic parameter modeling, the proposed method achieves both accurate and efficient estimates of structural failure probabilities. Models like Random Forest and XGBoost performed well. They gave accurate predictions. They also identified the most important features. Neural Networks handle complex nonlinear behaviors well. However, they required careful tuning to avoid overfitting. These results show the importance of including uncertainty in TMD design. Real-world conditions often differ from ideal assumptions. Surrogate modeling makes it easy to quickly evaluate different scenarios. It removes the need for repeated, time-consuming simulations. The proposed methodology reduced the computational effort by over 90%, from 28 h to 1.1 h, by leveraging the surrogate model for prediction.

In the future, this method can be used for larger and more complex structures. It can also consider changing properties over time and material wear. Also, we plan to investigate the effects of different distributions, such as uniform, log normal, and Weibull, on the system's reliability analysis. This will help us better understand the robustness of our framework under various conditions. Real-time monitoring data can be added to create digital twin models for adaptive TMD adjustments. This approach paves the way for more reliable, data-driven solutions to structural safety and vibration control challenges in dynamic and uncertain environments.

Author Contributions: Conceptualization, methodology, data curation, visualization, investigation, validation, writing original draft, L.S.; supervision, resources, A.S., A.-M.Z. and M.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funds from the China Scholarship Council.

Data Availability Statement: The data are not publicly available as they are also part of ongoing research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, L.; Caracoglia, L. Layered Stochastic Approximation Monte-Carlo method for tall building and tower fragility in mixed wind load climates. *Eng. Struct.* **2021**, *239*, 112159. [[CrossRef](#)]
2. Hacıfendioglu, K. Stochastic dynamic response of short-span highway bridges to spatial variation of blasting ground vibration. *Appl. Math. Comput.* **2017**, *292*, 194–209. [[CrossRef](#)]
3. Šykora, M.; Holický, M.; Marková, J. Verification of existing reinforced concrete bridges using the semi-probabilistic approach. *Eng. Struct.* **2013**, *56*, 1419–1426. [[CrossRef](#)]
4. Strauss, A.; Hauser, M.; Täubling, B.; Ivanković, A.M.; Skokandić, D.; Matos, J.; Galvão, N.; Benko, V.; Dobrý, J.; Wan-Wendner, R.; et al. Probabilistic and semi-probabilistic analysis of slender columns frequently used in structural engineering. *Appl. Sci.* **2021**, *11*, 8009. [[CrossRef](#)]
5. Nguyen, T.T.; Dang, V.H.; Nguyen, H.X. Efficient framework for structural reliability analysis based on adaptive ensemble learning paired with subset simulation. *Structures* **2022**, *45*, 1738–1750. [[CrossRef](#)]

6. Zhou, T.; Peng, Y. Structural reliability analysis via dimension reduction, adaptive sampling, and Monte Carlo simulation. *Struct. Multidiscip. Optim.* **2020**, *62*, 2629–2651. [\[CrossRef\]](#)
7. Xu, J.; Wang, D. Structural reliability analysis based on polynomial chaos, Voronoi cells and dimension reduction technique. *Reliab. Eng. Syst. Saf.* **2019**, *185*, 329–340. [\[CrossRef\]](#)
8. Chen, Z.; Han, Z.; Zhai, W.; Yang, J. TMD design for seismic vibration control of high-pier bridges in Sichuan–Tibet Railway and its influence on running trains. *Veh. Syst. Dyn.* **2019**, *57*, 207–225. [\[CrossRef\]](#)
9. Zuo, L.; Tang, X. Large-scale vibration energy harvesting. *J. Intell. Mater. Syst. Struct.* **2013**, *24*, 1405–1430. [\[CrossRef\]](#)
10. Stanikzai, M.H.; Elias, S.; Chae, Y. Recent advances in hybrid vibration-control systems. *Pract. Period. Struct. Des. Constr.* **2022**, *27*, 03122003. [\[CrossRef\]](#)
11. Peng, Y.; Sun, P. Reliability-based design optimization of tuned mass-damper-inerter for mitigating structural vibration. *J. Sound Vib.* **2024**, *572*, 118166. [\[CrossRef\]](#)
12. de Salles, H.B.; Miguel, L.F.F.; Lenzi, M.S.; Lopez, R.H. Reduced-order model for RBDO of multiple TMDs on eccentric L-shaped buildings subjected to seismic excitations. *Mech. Syst. Signal Process.* **2024**, *206*, 110906. [\[CrossRef\]](#)
13. Miguel, L.F.F.; Elias, S.; Beck, A.T. Reliability-based optimization of supported pendulum TMDs' nonlinear track shape using Padé approximants. *Eng. Struct.* **2024**, *306*, 117861. [\[CrossRef\]](#)
14. Ma, W.; Yang, Y.; Yu, J.; Yang, G. Improving processing stability of thin-walled casing workpieces by sucker-mounted TMD: Optimization, design, and implementation. *Mech. Syst. Signal Process.* **2024**, *216*, 111512. [\[CrossRef\]](#)
15. Malliotakis, G.; Alevras, P.; Baniotopoulos, C. Recent advances in vibration control methods for wind turbine towers. *Energies* **2021**, *14*, 7536. [\[CrossRef\]](#)
16. Lei, Z.; Liu, G.; Yang, Q.; Law, S.; Li, Y.; Cui, T. Experimental investigation on vibration suppression of a new prestressed TMD for wind turbine towers. *Thin-Walled Struct.* **2024**, *202*, 112080. [\[CrossRef\]](#)
17. Liu, Y.; Wang, K.; Mercan, O.; Chen, H.; Tan, P. Experimental and numerical studies on the optimal design of tuned mass dampers for vibration control of high-rise structures. *Eng. Struct.* **2020**, *211*, 110486. [\[CrossRef\]](#)
18. Miguel, L.F.; Lopez, R.H.; Torii, A.J.; Miguel, L.F.; Beck, A.T. Robust design optimization of TMDs in vehicle–bridge coupled vibration problems. *Eng. Struct.* **2016**, *126*, 703–711. [\[CrossRef\]](#)
19. Zhang, S.Y.; Sheng, X.; Jiang, J.Z.; Zhou, H.; Ren, W.X.; Zhang, Z.H. Vibration suppression of bridges under moving loads using the structure-immittance approach. *Int. J. Mech. Sci.* **2021**, *211*, 106792. [\[CrossRef\]](#)
20. Yu, H.; Gillot, F.; Ichchou, M. Reliability based robust design optimization for tuned mass damper in passive vibration control of deterministic/uncertain structures. *J. Sound Vib.* **2013**, *332*, 2222–2238. [\[CrossRef\]](#)
21. Bansal, S.; Cheung, S.H. Stochastic simulation algorithm for robust reliability updating of structural dynamic systems based on incomplete modal data. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2017**, *3*, 04017008. [\[CrossRef\]](#)
22. Lieu, Q.X.; Nguyen, K.T.; Dang, K.D.; Lee, S.; Kang, J.; Lee, J. An adaptive surrogate model to structural reliability analysis using deep neural network. *Expert Syst. Appl.* **2022**, *189*, 116104. [\[CrossRef\]](#)
23. Li, X.Q.; Song, L.K.; Bai, G.C. Recent advances in reliability analysis of aeroengine rotor system: A review. *Int. J. Struct. Integr.* **2022**, *13*, 1–29. [\[CrossRef\]](#)
24. Peng, Y.; Zhou, T.; Li, J. Surrogate modeling immersed probability density evolution method for structural reliability analysis in high dimensions. *Mech. Syst. Signal Process.* **2021**, *152*, 107366. [\[CrossRef\]](#)
25. Li, Z.A.; Dong, X.W.; Zhu, C.Y.; Chen, C.H.; Zhang, H. Vectorial surrogate modeling method based on moving Kriging model for system reliability analysis. *Comput. Methods Appl. Mech. Eng.* **2024**, *432*, 117409. [\[CrossRef\]](#)
26. Fan, H.; Wang, C.; Li, S. Novel method for reliability optimization design based on rough set theory and hybrid surrogate model. *Comput. Methods Appl. Mech. Eng.* **2024**, *429*, 117170. [\[CrossRef\]](#)
27. You, W. Reliability Assessment of TMD-Based Control Structures: A Statistical Learning Perspective. Ph.D. Thesis, Université de Lyon, Lyon, France, 2020.
28. Samadian, D.; Muhit, I.B.; Dawood, N. Application of data-driven surrogate models in structural engineering: A literature review. *Arch. Comput. Methods Eng.* **2024**, 1–50. [\[CrossRef\]](#)
29. Mottaghi, H.; Masoodi, A.R.; Gandomi, A.H. Multiscale analysis of carbon nanotube-reinforced curved beams: A finite element approach coupled with multilayer perceptron neural network. *Results Eng.* **2024**, *23*, 102585. [\[CrossRef\]](#)
30. Samadian, D.; Muhit, I.B.; Occhipinti, A.; Dawood, N. Meta databases of steel frame buildings for surrogate modelling and machine learning-based feature importance analysis. *Resilient Cities Struct.* **2024**, *3*, 20–43. [\[CrossRef\]](#)
31. TQ, A.D.; Masoodi, A.R.; Gandomi, A.H. Unveiling the potential of an evolutionary approach for accurate compressive strength prediction of engineered cementitious composites. *Case Stud. Constr. Mater.* **2023**, *19*, e02172.
32. Venanzi, I. Robust optimal design of tuned mass dampers for tall buildings with uncertain parameters. *Struct. Multidiscip. Optim.* **2015**, *51*, 239–250. [\[CrossRef\]](#)
33. Dasgupta, A.; Pecht, M. Material failure mechanisms and damage models. *IEEE Trans. Reliab.* **1991**, *40*, 531–536. [\[CrossRef\]](#)
34. Blischke, W.R.; Murthy, D.P. *Reliability: Modeling, Prediction, and Optimization*; John Wiley & Sons: Hoboken, NJ, USA, 2011.

35. Crandall, S.H.; Mark, W.D. *Random Vibration in Mechanical Systems*; Academic Press: Cambridge, MA, USA, 2014.
36. Au, S.; Beck, J.L. First excursion probabilities for linear systems by very efficient importance sampling. *Probabilistic Eng. Mech.* **2001**, *16*, 193–207. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.