A Regression-Tree-Based Method for Structural Reliability Assessment

S. Russotto[1)], M. Di Paola[1)], and A. Pirrotta[1)]
[1)] *Engineering Department, Università degli Studi di Palermo, Palermo 90128, Italy,*
salvatore.russotto01@unipa.it mario.dipaola@unipa.it antonina.pirrotta@unipa.it

Abstract: Structural reliability plays a pivotal role in structural engineering, providing a quantitative basis for assessing safety, serviceability, and risk. Traditionally, the evaluation of a survival probability is carried out either through Monte Carlo simulations, which can be extremely computationally demanding, or through path integral approaches that account for absorbing barriers. However, for damped oscillators, path integral techniques typically require the probability density function (PDF) of the state variables, i.e., bi-dimensional PDF, which significantly increases both the mathematical complexity of the problem and the associated computational cost.
In this paper, a regression tree is specifically trained on a large dataset to estimate, within a fraction of a second, the survival probability of a damped oscillator at a given time instant when subjected to white noise excitation. The proposed method aims to retain the accuracy of reliability computations while drastically reducing runtime. The accuracy of the proposed method is validated by comparing its predictions against a Monte Carlo simulation, showing strong agreement and confirming the proposed approach as a fast and effective tool for structural reliability assessment.

Keywords: Structural Reliability, Regression Trees, Computational Efficiency.

1. Introduction

Structural reliability has long represented one of the fundamental pillars of modern structural engineering, because it provides a rational probabilistic framework for quantifying the safety, serviceability, and risk of structures in the presence of unavoidable uncertainties in loads, material properties, geometry, boundary conditions, deterioration mechanisms, and environmental actions (Cornell, 1969; Hasofer and Lind, 1974; Rackwitz, 2001; Ditlevsen and Madsen, 1996; Melchers and Beck, 2018; Pirrotta, 2007; Biondini and Frangopol, 2016; Chen and Li, 2007). In contrast with purely deterministic safety checks, reliability analysis explicitly recognizes that both resistance and demand are random quantities and that structural performance must therefore be evaluated in terms of the probability of violating a prescribed limit state. This viewpoint has profoundly influenced the development of probability-based design formats, reliability indices, and life-cycle performance assessment strategies for civil and mechanical systems (Cornell, 1969; Hasofer and Lind, 1974; Rackwitz, 2001; Ditlevsen and Madsen, 1996; Melchers and Beck, 2018; Biondini and Frangopol, 2016). In this sense, structural reliability is not merely a verification tool, but also a decision-support instrument for design, maintenance planning, optimization, and risk-informed management of engineering systems throughout their service life (Rackwitz, 2001; Ditlevsen and Madsen, 1996; Melchers and Beck, 2018; Biondini and Frangopol, 2016).

Despite its conceptual importance, structural reliability assessment remains computationally challenging, especially when the structural response is nonlinear. Over the years, a broad spectrum of methods has been developed, ranging from first-order and second-order reliability approximations to direct simulation methods and advanced rare-event strategies (Hasofer and Lind, 1974; Rackwitz, 2001; Ditlevsen and Madsen, 1996; Melchers and Beck, 2018; Doliński, 1983; Breitung, 1984; Der Kiureghian and De Stefano, 1991; Au and Beck, 2001; Spanos and Kougioumtzoglou, 2013; Bucher et al., 2016; Song and Kawai, 2023; Afshari et al., 2022). Classical first-order and second-order reliability approximations methodologies are attractive because of their analytical elegance and reduced computational cost, but their accuracy may deteriorate when the limit-state surface is strongly nonlinear or when multiple failure regions exist (Hasofer and Lind, 1974; Doliński, 1983; Breitung, 1984; Der Kiureghian and De Stefano, 1991). On the other hand, Monte Carlo simulation retains a nearly universal applicability and can be regarded as a benchmark method, yet it is notoriously computationally expensive when small failure probabilities must be estimated with adequate accuracy (Au and Beck, 2001; Song and Kawai, 2023). This issue becomes even more severe in stochastic dynamics and first-passage problems, where repeated time-domain simulations or repeated evaluations of transition probabilities are required (Au and Beck, 2001; Spanos and Kougioumtzoglou, 2013; Bucher et al., 2016; Song and Kawai, 2023). Recent reviews have in fact confirmed that Monte Carlo and related sampling procedures remain accurate but often computationally demanding, thereby motivating the search for faster surrogate-based alternatives (Song and Kawai, 2023; Afshari et al., 2022).

The above difficulty is particularly evident for nonlinear oscillatory systems under stochastic excitation. In such cases, reliability is often associated with a first-passage or survival probability, namely the probability that the response does not cross an assigned threshold up to a given time instant (Spanos and Kougioumtzoglou, 2013; Bucher et al., 2016). Path-integral formulations have proved to be powerful tools for this class of problems because they allow one to propagate the response probability density in time and, with suitable absorbing barriers, to evaluate first-passage quantities in a rigorous probabilistic setting (Spanos and Kougioumtzoglou, 2013; Bucher et al., 2016; Di Paola and Santoro, 2008; Pirrotta and Santoro, 2011; Kougioumtzoglou and Spanos, 2014; Kougioumtzoglou et al., 2015; Di Matteo et al., 2016; Kougioumtzoglou, 2017; Zhang et al., 2023). However, for dynamical systems excited by stochastic forces, these methods generally require the joint probability density of the state variables, typically displacement and velocity, which increases the dimensionality of the probabilistic problem and consequently the computational burden associated with discretization, numerical integration, and repeated evaluation in time (Spanos and Kougioumtzoglou, 2013; Bucher et al., 2016). Therefore, although path-integral approaches are mathematically sound and highly valuable, their practical implementation may still be onerous when many reliability evaluations are needed, for example in parametric analyses, sensitivity studies, optimization loops, or real-time decision environments.

For these reasons, the last decade has witnessed a growing interest in surrogate and machine-learning strategies for structural reliability analysis (Song and Kawai, 2023; Afshari et al., 2022). The central idea is to replace an expensive physical or numerical model, or even the reliability mapping itself, with an inexpensive predictor trained on representative data. In structural reliability, machine-learning methods have been employed to approximate limit-state functions, classify safe and failure domains, guide adaptive sampling, and accelerate Monte Carlo-based estimators (Song and Kawai, 2023; Afshari et al., 2022). Broadly speaking, the motivation is twofold: first, to preserve an acceptable level of accuracy; second, to reduce runtime to a level compatible with large-scale uncertainty propagation or near-real-time applications. In parallel, the broader structural-engineering literature has shown that machine learning is becoming increasingly relevant for nonlinear structural analysis, health monitoring, damage detection, resistance prediction, and performance evaluation, especially when large datasets are available and repeated evaluations are required (Afshari et al., 2022; Thai, 2022).

Within the wide family of machine-learning methods, decision trees are especially appealing because they combine conceptual simplicity, interpretability, and low computational cost at inference time (Morgan and Sonquist, 1963; Friedman, 1977; Breiman et al., 1984; Quinlan, 1986; Kass, 1980; Loh and Shih, 1997; Quinlan, 1993; Frank et al., 1998; Breiman, 2001; Song and Ying, 2015). A decision tree organizes the prediction process through a hierarchical structure composed of a root node, internal decision nodes, branches, and terminal leaves. At each split, the dataset is partitioned according to a decision rule based on one predictor, with the objective of increasing the homogeneity of the resulting subsets. The final prediction is then associated with the terminal leaf reached by a given input pattern (Friedman, 1977; Breiman et al., 1984; Quinlan, 1986; Song and Ying, 2015). In decision trees, the terminal output is a class label or class probability, whereas in regression trees it is a continuous value, commonly represented by the mean response within the leaf or by a local predictive model (Breiman et al., 1984; Quinlan, 1986; Frank et al., 1998; Song and Ying, 2015). Owing to their nonparametric nature, trees can capture highly nonlinear relationships without imposing a global functional form, while still preserving a transparent rule-based structure that is generally easier to interpret than many black-box alternatives (Breiman et al., 1984; Quinlan, 1986; Song and Ying, 2015).

The historical development of tree-based methods is rich and diversified. Early roots can be traced back to Automatic Interaction Detection (AID), introduced by Morgan and Sonquist (Morgan and Sonquist, 1963), followed by recursive partitioning approaches for classification (Friedman, 1977). A decisive milestone was the Classification and Regression Trees (CART) proposed by Breiman, Friedman, Olshen, and Stone (Breiman et al., 1984), which established the modern formulation of classification and regression trees based on binary recursive splitting and pruning. Quinlan's Iterative Dichotomiser 3 (ID3) and later C4.5 algorithms extended tree induction through information-theoretic criteria and became particularly influential in the machine-learning community (Quinlan, 1986; Quinlan, 1993). Alternative families were subsequently proposed to address specific needs: Chi-squared Automatic Interaction Detection (CHAID) relies on chi-squared tests and naturally allows multiway splits (Kass, 1980); Quick Unbiased Efficient Statistical Tree (QUEST) was designed to reduce variable-selection bias and improve computational efficiency (Loh and Shih, 1997); model trees place local linear models at the leaves instead of constant predictions (Frank et al., 1998); and random forests extend single-tree predictors into an ensemble framework with improved generalization performance (Breiman, 2001). Accordingly, the literature now distinguishes, at least, between decision trees, regression trees, model trees, statistically driven multiway trees such as CHAID, unbiased split-selection trees such as QUEST, and ensemble tree methods (Breiman et al., 1984; Quinlan, 1986; Kass, 1980; Loh and Shih, 1997; Quinlan, 1993; Frank et al., 1998; Breiman, 2001; Song and Ying, 2015).

Beyond their methodological appeal, tree-based methods have already demonstrated their usefulness in several engineering domains. Model trees have been employed for hydrologic forecasting problems, where they showed the ability to provide accurate predictions together with an interpretable piecewise structure (Solomatine and Xue, 2004). In structural and civil engineering, decision-tree-based approaches have been used for seismic damage prediction of reinforced concrete buildings, preliminary vulnerability assessment, and vibration-based damage localization in structural health monitoring (Karbassi et al., 2014; Mariniello et al., 2021; Su and He, 2019). More generally, recent reviews on machine learning for structural engineering have emphasized that tree-based models and their ensemble variants now constitute a relevant class of predictors across structural analysis, monitoring, damage assessment, and material-property estimation tasks (Mariniello et al., 2021; Su and He, 2019; Thai, 2022). These applications are particularly significant because they highlight two properties that are highly desirable also in structural reliability: fast evaluation and transparent partitioning of complex nonlinear input-output relationships (Mariniello et al., 2021; Su and He, 2019; Thai, 2022).

In the specific field of structural reliability, tree-based surrogates have already begun to attract attention. Keshtegar and Kisi (Keshtegar and Kisi, 2017) proposed an M5 model tree combined with Monte Carlo simulation for efficient structural reliability analysis, showing that a tree-based metamodel can significantly improve computational efficiency while maintaining satisfactory accuracy. The same authors later introduced the RM5Tree approach, in which the training strategy was refined to enhance the predictive capability of the model tree in reliability applications (Keshtegar and Kisi, 2018). These contributions clearly indicate that tree-based surrogates are not only conceptually attractive, but also practically competitive in reliability problems governed by complex and implicit performance functions (Keshtegar and Kisi, 2017; Keshtegar and Kisi, 2018). Nevertheless, much remains to be explored in the use of regression trees as standalone and accurate surrogate models for the estimation of time-dependent survival probabilities associated with dynamical systems excited by stochastic inputs.

Starting from this background, the present study proposes a regression-tree-based method for the rapid estimation of the survival probability of a damped oscillator subjected to white noise excitation. Rather than repeatedly solving the full reliability problem at each query point, the proposed approach trains a regression tree on a sufficiently large dataset so that, once learned, the mapping between the relevant input parameters and the corresponding survival probability can be evaluated in a fraction of a second. In this way, the method aims to retain the reliability information delivered by classical probabilistic analyses while drastically reducing computational time. The predictive capability of the proposed regression-tree-based method is then assessed against Monte Carlo simulations, which are taken as the reference solution for validation. Such a framework is particularly attractive for engineering scenarios in which reliability must be estimated many times under varying conditions, since it combines data-driven speed, limited inference cost, and a naturally interpretable model structure. From this perspective, the proposed method can be viewed as a bridge between rigorous stochastic reliability analysis and fast decision-oriented computational tools suitable for modern structural engineering applications.

2. Classical Structural Reliability Assessment via Monte Carlo Simulation

In this section, some preliminary concepts and definitions are introduced, including the classical method for the structural reliability assessment based on the Monte Carlo simulation.

Consider a damped oscillator, quiescent at $t=0$, excited by a zero-mean Gaussian white noise process. The governing equation of motion is

$$\begin{cases} \ddot{X}(t)+2\zeta_0\omega_0\dot{X}(t)+\omega_0^2X(t)=W(t) \\ X(0)=0 \qquad \text{w.p.1} \\ \dot{X}(0)=0 \qquad \text{w.p.1} \end{cases}, \tag{1}$$

in which $X(t)$, $\dot{X}(t)$ and $\ddot{X}(t)$ represent the stochastic response process in terms of displacement, velocity and acceleration, respectively, ω_0 is the natural frequency of the oscillator, ζ_0 is the damping ratio, and

$W(t)$ indicates the structural excitation: a zero-mean Gaussian white noise process. The latter is a delta-correlated stochastic process, fully described in probabilistic setting by the correlation function or, equivalently, by the power spectral density (PSD)

$$S_W(\omega) = S_0, \tag{2}$$

being $S_0 > 0$.

The first step of the Monte Carlo simulation consists in the generation of n samples of the input process, performed starting from the PSD by the Shinozuka's formula. Therefore, the *r*-th sample of the structural excitation, labeled as $W^{(r)}(t)$, is expressed in the form

$$W^{(r)}(t) = \sum_{k=1}^{m_c} \sqrt{4S_W(\omega_k)\Delta\omega} \sin\left(\omega_k t + \theta_k^{(r)}\right) \tag{3}$$

in which $\Delta\omega$ is the frequency sampling step, $\omega_k = k\Delta\omega$, $\theta_k^{(r)}$ is the *r*-th realization of a random variable uniformly distributed in $(0, 2\pi)$, and $m_c = \omega_c / \Delta\omega$, being ω_c the sampling frequency.

Once the *r*-th sample of the structural excitation has been generated, the correspondent sample of the response process, labeled as $X^{(r)}(t)$, can be obtained by solving the following differential equation

$$\begin{cases} \ddot{X}^{(r)}(t) + 2\zeta_0\omega_0\dot{X}^{(r)}(t) + \omega_0^2 X^{(r)}(t) = W^{(r)}(t) \\ X^{(r)}(0) = 0 \\ \dot{X}^{(r)}(0) = 0 \end{cases}. \tag{4}$$

To solve Eq.(4), different numerical techniques can be used. One of the most used is the explicit Runge-Kutta technique: a step-by-step integration that, by discretizing the time axis with a time sampling step Δt, allows to find the solution of Eq.(4) in terms of the state space variable $\mathbf{Y}^{(r)}(t) = \left[X^{(r)}(t) \quad \dot{X}^{(r)}(t)\right]^T$.

When all the n samples of the response process $X(t)$ have been calculated, the structural reliability assessment can be performed. To this end, first of all a value for the absorbing barrier, labeled as $x_{\lim}$, must be selected. Subsequently, for each time instant $t_j = j\Delta t$, examined sequentially from the earliest ($j = 0$) to the latest ($j = n_f$, being $t_{n_f} = n_f \Delta t$ the final time instant of the process), the samples that satisfy the condition

$$\left|X^{(r)}(t_j)\right| \geq x_{\lim} \tag{5}$$

are first identified and counted, and their number is denoted by n_j^{fail}. Once counted, all samples satisfying the condition in Eq.(5) are removed from the process before moving to the next time instant. Finally, the value of the survival probability in t_j, labeled as SP_j, is calculated, for each $j = 1, \ldots, n_f$, as

$$SP_j = 1 - \frac{1}{n}\sum_{u=0}^{j} n_u^{fail} \ . \tag{6}$$

Eq.(6) represents the survival probability in the time instant t_j, i.e. the probability that $\left|X\left(t\right)\right|$ does not exceed the barrier $x_{\lim}$ in $\left(0,t_j\right)$.

Although for high values of n the structural reliability assessment performed with the aid of the Monte Carlo simulation allows to obtain accurate results in terms of survival probability, it requires a very high computational burden. Therefore, in the next section, a novel and computationally efficient method for structural reliability assessment, based on regression tree models, is proposed.

3. Development of the Proposed Regression-Tree-Based Method for Structural Reliability Assessment

In order to drastically reduce the computational time required to perform a structural reliability assessment, a regression tree has been properly trained on a very large dataset.

A regression tree is a non-parametric supervised learning model for predicting a continuous response variable. The algorithm operates by recursively splitting the input feature space into smaller, non-overlapping regions. Each split is chosen by selecting the input feature and a threshold that best reduce the prediction error, while each terminal leaf is then associated with a constant prediction (output feature).

The proposed regression-tree-based method allows us to obtain in a fraction of a second, by providing ω_0, ζ_0, S_0, $x_{\lim}$, and t_j as input features, the output feature SP_j.

To construct the dataset required for training the regression tree, a large number (N_{TR}) of different combinations of the input features must be considered, and the corresponding values of the output feature must be generated. To this end, the combinations of the input features have been generated as reported in Figure 1.

To obtain the structural response process necessary for the estimation of the survival probability, which is the output feature needed for training of the regression tree, the relationship

$$S_X\left(\omega\right) = \left|H\left(\omega\right)\right|^2 S_W\left(\omega\right) \tag{7}$$

has been initially used, in which $S_X\left(\omega\right)$ represents the PSD of the structural response process, while $H\left(\omega\right)$ is the frequency response function (FRF) of the structural system which can be evaluated in the form

$$H\left(\omega\right) = \frac{1}{\omega_0^2 - \omega^2 + 2i\zeta_0\omega_0\omega} \ . \tag{8}$$

Once the PSD of the structural response process has been obtained, the formula of Shinozuka can be used to directly generate samples of the structural response as

$$X^{(r)}\left(t_j\right) = \sum_{k=1}^{m_c}\sqrt{4S_X\left(\omega_k\right)\Delta\omega}\,\sin\left(\omega_k t_j + \theta_k^{(r)}\right). \tag{9}$$

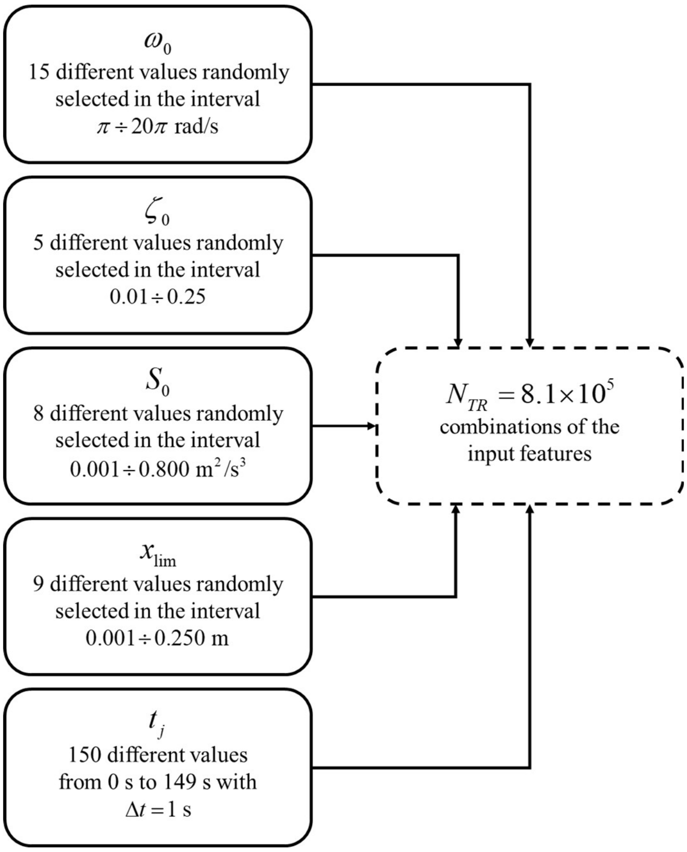

Figure 1. Selection of the input features.

For each value of ω_0, ζ_0, and S_0, a total of 10000 samples of the structural response have been generated with Eq.(9). Then, the survival probability has been evaluated, for each value of $x_{\lim}$ and for each value of t_j, as reported in Eq.(6). In this way, the 8.1×10^5 output features necessary to train the regression tree have been generated.

Before the training of the regression tree, the input features were organized into the matrix

$$\mathbf{F}_{IN} = \begin{bmatrix} \omega_0^{(1)} & \zeta_0^{(1)} & S_0^{(1)} & x_{\lim}^{(1)} & t_j^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \omega_0^{(N_{TR})} & \zeta_0^{(N_{TR})} & S_0^{(N_{TR})} & x_{\lim}^{(N_{TR})} & t_j^{(N_{TR})} \end{bmatrix}, \tag{10}$$

where the $\overline{q}$-th row represents the $\overline{q}$-th combination of the input features, being $\overline{q} = 1,\ldots,N_{TR}$. Similarly to the input features, the corresponding values of the output feature were organized into the vector

$$\mathbf{F}_{OUT} = \left[SP_j^{(1)} \quad \cdots \quad SP_j^{(N_{TR})} \right]^T, \tag{11}$$

in which the $\overline{q}$ -th value represents the survival probability associated with the $\overline{q}$ -th combination of the input features.

For the training of the regression tree, the regression learning tool of MATLAB has been used. Specifically, a “fine tree” was selected, with a minimum leaf size equal to four. During the training, both the principal component analysis (PCA) and the surrogate decision splits have been disabled.

After the training, to avoid values of the survival probability major than one or minor than zero, the following restriction has been added

$$SP_j = \begin{cases} 0 & \text{for} \quad SP_j < 0 \\ 1 & \text{for} \quad SP_j > 1 \end{cases}. \tag{12}$$

In the next section, the validation of the proposed regression-tree-based method for structural reliability assessment is reported.

4. Validation of the Proposed Regression-Tree-Based Method and results

In order to validate the proposed regression-tree-based method, $N_{VAL} = 50000$ different combinations of input features were randomly generated by considering the same intervals adopted for the training stage. For each validation combination $q = 1, \ldots, N_{VAL}$, the value of the survival probability was first predicted by the proposed model, thus obtaining $SP_{j(PM)}^{(q)}$. Then, the corresponding value $SP_{j(MC)}^{(q)}$ was evaluated by means of the classical Monte Carlo-based procedure described in Section 2. In this way, two sets of values for the survival probability were obtained and directly compared, over the whole validation set, in Figure 2.

In Figure 2, each point $P^{(q)} = \left(SP_{j(PM)}^{(q)}, SP_{j(MC)}^{(q)} \right)$ represents, simultaneously, the survival provability associated with the q-th validation combination of the input features obtained with the proposed method and that obtained from the Monte Carlo simulation. As for the continuous line in Figure 2, it is the least-squares regression line fitted to the points $P^{(q)}$. The slope of the latter represents the Pearson correlation coefficient, which is calculated as

$$\rho = \frac{\sum_{q=1}^{N_{VAL}} \left(SP_{j(PM)}^{(q)} - \mu_{SP_{j(PM)}} \right) \left(SP_{j(MC)}^{(q)} - \mu_{SP_{j(MC)}} \right)}{\sqrt{\sum_{q=1}^{N_{VAL}} \left(SP_{j(PM)}^{(q)} - \mu_{SP_{j(PM)}} \right)^2} \sqrt{\sum_{q=1}^{N_{VAL}} \left(SP_{j(MC)}^{(q)} - \mu_{SP_{j(MC)}} \right)^2}}, \tag{13}$$

being $\mu_{SP_{j(PM)}}$ and $\mu_{SP_{j(MC)}}$ the mean values over the whole validation set of $SP_{j(PM)}^{(q)}$ and $SP_{j(MC)}^{(q)}$, respectively.

In the present case, ρ is equal to 0.9992, thus indicating an almost perfect positive linear correlation between the survival probability estimated with the proposed method and those obtained from the Monte Carlo simulation, Moreover, it can be observed that the points are strongly concentrated around the least-squares regression line: a clear visual evidence of the high predictive accuracy of the proposed regression tree model.

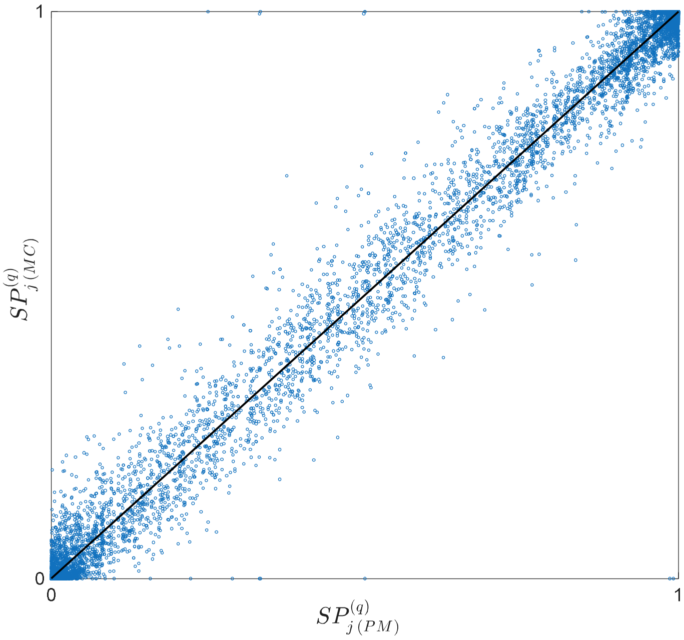

Figure 2. Graphical representation of the points $P^{(q)}$ (circular marker) and least-squares regression line fitted to the points $P^{(q)}$ (continuous line).

To quantify the accuracy, for each point $P^{(q)}$, the absolute difference between the value predicted by the proposed model and the corresponding benchmark value was computed as

$$D^{(q)} = \left| SP^{(q)}_{j(PM)} - SP^{(q)}_{j(MC)} \right|. \tag{14}$$

Then, the percentage of validation combinations for which $D^{(q)}$ is smaller than a selected threshold ε, denoted by P_ε, was evaluated as

$$P_\varepsilon = \frac{100}{N} \sum_{q=1}^{N_{VAL}} \mathbf{1}\left(D^{(q)} < \varepsilon\right), \tag{15}$$

in which

$$\mathbf{1}\left(D^{(q)} < \varepsilon\right) = \begin{cases} 1, & \text{if } D^{(q)} < \varepsilon \\ 0, & \text{if } D^{(q)} \geq \varepsilon \end{cases}. \tag{16}$$

The results obtained are reported in Figure 3.

From Figure 3, it can be observed that 89.89% of the validation combinations exhibit a value of $D^{(q)}$ smaller than 0.01, the 97.07% exhibit a value of $D^{(q)}$ lesser than 0.05, smaller than 0.05, while the 99.39% and the 99.91% exhibit a value of $D^{(q)}$ smaller than 0.1 and 0.2, respectively. These results indicate that the discrepancy between the proposed regression tree model and the benchmark Monte Carlo solution remains very limited for the overwhelming majority of the examined cases.

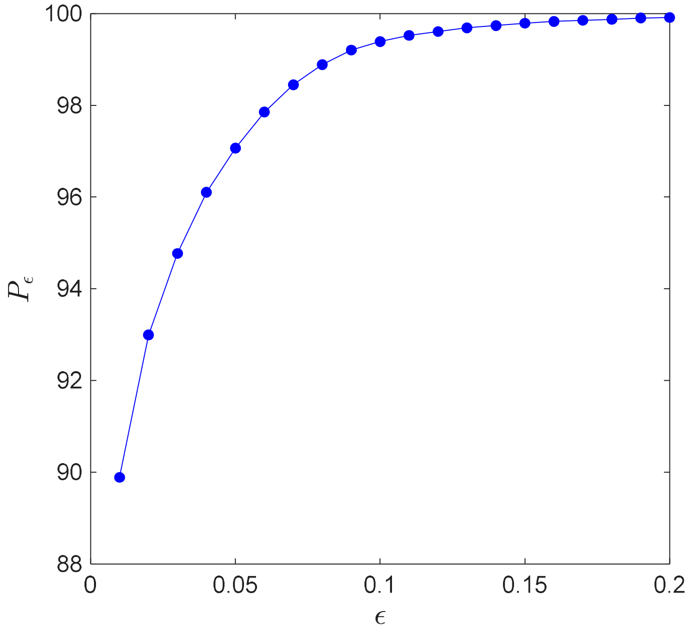

Figure 3. Evaluation of P_ε for different values of ε .

The joint evidence provided by the point concentration around the least-squares regression line in Figure 2, the Pearson correlation coefficient ρ equal to 0.9992, and the distribution of the absolute difference $D^{(q)}$ demonstrates that the proposed regression-tree-based method is capable of predicting the survival probability with high accuracy over the whole considered range of input features. Therefore, the proposed approach can be regarded as a reliable and computationally efficient surrogate tool for the structural reliability assessment of damped oscillators subjected to white noise excitation.

5. Concluding Remarks and Possible Future Developments

In this paper, a novel structural reliability assessment method based on a regression tree has been proposed for the rapid estimation of the survival probability of a damped oscillator subjected to white noise excitation. The main objective of the study was to develop a computationally efficient alternative to classical reliability assessment procedures, which are often associated with a significant computational burden, especially when repeated evaluations are required over a large number of input configurations. In this respect, the proposed approach has shown that, once properly trained, a regression tree can provide the value of the survival probability in a fraction of a second, thus making the estimation process extremely fast and suitable for applications in which computational speed is of primary importance.

The predictive capability of the proposed model has been assessed through a systematic comparison with results obtained by means of Monte Carlo simulation. The comparison has shown a very strong agreement between the predicted values and the reference ones, thereby confirming that the proposed method is able to retain a high level of accuracy while drastically reducing the required computational effort. These results demonstrate that regression-tree-based models can represent a valid surrogate tool for structural reliability assessment and may constitute a promising bridge between rigorous stochastic analysis and fast engineering-oriented prediction.

Possible future developments may include an extension of the proposed methodology to more general classes of dynamical systems. In particular, a natural continuation of the present research consists in adapting the approach to the prediction of the survival probability of linear and nonlinear multi-degree-of-freedom structural systems subjected to seismic excitation.

References

Afshari, S. S., F. Enayatollahi, X. Xu, and X. Liang. Machine Learning-Based Methods in Structural Reliability Analysis: A Review. Reliability Engineering and System Safety, 219:108223, 2022.

Au, S. K. and J. L. Beck. Estimation of Small Failure Probabilities in High Dimensions by Subset Simulation. Probabilistic Engineering Mechanics, 16(4):263-277, 2001.

Biondini, F. and D. M. Frangopol. Life-Cycle Performance of Deteriorating Structural Systems Under Uncertainty: Review. Journal of Structural Engineering, 142(9):04016068, 2016.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and Regression Trees. Wadsworth, Belmont, CA, 1984.

Breiman, L. Random Forests. Machine Learning, 45(1):5-32, 2001.

Breitung, K. Asymptotic Approximations for Multinormal Integrals. Journal of Engineering Mechanics, 110(3):357-366, 1984.

Bucher, C., A. Di Matteo, M. Di Paola, and A. Pirrotta. First-Passage Problem for Nonlinear Systems Under Lévy White Noise Through Path Integral Method. Nonlinear Dynamics, 85(3):1445-1456, 2016.

Chen, J. B. and J. Li. Development-Process-of-Nonlinearity-Based Reliability Evaluation of Structures. Probabilistic Engineering Mechanics, 22(3):267-275, 2007.

Cornell, C. A. A Probability-Based Structural Code. ACI Journal, 66:974-985, 1969.

Der Kiureghian, A. and M. De Stefano. Efficient Algorithm for Second-Order Reliability Analysis. Journal of Engineering Mechanics, 117(12):2904-2923, 1991.

Di Matteo, A., M. Di Paola, and A. Pirrotta. Path Integral Solution for Nonlinear Systems Under Parametric Poissonian White Noise Input. Probabilistic Engineering Mechanics, 44:89-98, 2016.

Di Paola, M. and R. Santoro. Non-Linear Systems Under Poisson White Noise Handled by Path Integral Solution. Journal of Vibration and Control, 14(1-2):35-49, 2008.

Ditlevsen, O. and H. O. Madsen. Structural Reliability Methods. John Wiley and Sons, Chichester, 1996.

Doliński, K. First-Order Second-Moment Approximation in Reliability of Structural Systems: Critical Review and Alternative Approach. Structural Safety, 1(3):211-231, 1983.

Frank, E., Y. Wang, S. Inglis, G. Holmes, and I. H. Witten. Using Model Trees for Classification. Machine Learning, 32(1):63-76, 1998.

Friedman, J. H. A Recursive Partitioning Decision Rule for Nonparametric Classification. IEEE Transactions on Computers, 26:404-408, 1977.

Hasofer, A. M. and N. C. Lind. Exact and Invariant Second-Moment Code Format. Journal of the Engineering Mechanics Division, ASCE, 100(1):111-121, 1974.

Karbassi, A., B. Mohebi, S. Rezaee, and P. Lestuzzi. Damage Prediction for Regular Reinforced Concrete Buildings Using the Decision Tree Algorithm. Computers and Structures, 130:46-56, 2014.

Kass, G. V. An Exploratory Technique for Investigating Large Quantities of Categorical Data. Applied Statistics, 29(2):119-127, 1980.

Keshtegar, B. and O. Kisi. M5 Model Tree and Monte Carlo Simulation for Efficient Structural Reliability Analysis. Applied Mathematical Modelling, 48:899-910, 2017.

Keshtegar, B. and O. Kisi. RM5Tree: Radial Basis M5 Model Tree for Accurate Structural Reliability Analysis. Reliability Engineering and System Safety, 180:49-61, 2018.

Kougioumtzoglou, I. A. A Wiener Path Integral Solution Treatment and Effective Material Properties of a Class of One-Dimensional Stochastic Mechanics Problems. Journal of Engineering Mechanics, 143:04017014, 2017.

Kougioumtzoglou, I. A. and P. D. Spanos. Nonstationary Stochastic Response Determination of Nonlinear Systems: A Wiener Path Integral Formalism. Journal of Engineering Mechanics, 140:04014064, 2014.

Kougioumtzoglou, I. A., A. Di Matteo, P. D. Spanos, A. Pirrotta, and M. Di Paola. An Efficient Wiener Path Integral Technique Formulation for Stochastic Response Determination of Nonlinear MDOF Systems. Journal of Applied Mechanics, 82:101005, 2015.

Loh, W. Y. and Y. S. Shih. Split Selection Methods for Classification Trees. Statistica Sinica, 7:815-840, 1997.

Mariniello, G., T. Pastore, C. Menna, P. Festa, and D. Asprone. Structural Damage Detection and Localization Using Decision Tree Ensemble and Vibration Data. Computer-Aided Civil and Infrastructure Engineering, 36(9):1129-1149, 2021.

Melchers, R. E. and A. T. Beck. Structural Reliability Analysis and Prediction. 3rd edition. John Wiley and Sons, Hoboken, NJ, 2018.

Morgan, J. N. and J. A. Sonquist. Problems in the Analysis of Survey Data, and a Proposal. Journal of the American Statistical Association, 58(302):415-434, 1963.

Pirrotta, A. Multiplicative cases from additive cases: Extension of Kolmogorov-Feller equation to parametric Poisson white noise processes. Probabilistic Engineering Mechanics, 22(2):127-135, 2007.

Pirrotta, A. and R. Santoro. Probabilistic Response of Nonlinear Systems Under Combined Normal and Poisson White Noise via Path Integral Method. Probabilistic Engineering Mechanics, 26(1):26-32, 2011.

Quinlan, J. R. Induction of Decision Trees. Machine Learning, 1:81-106, 1986.

Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, CA, 1993.

Rackwitz, R. Reliability Analysis: A Review and Some Perspectives. Structural Safety, 23(4):365-395, 2001.

Solomatine, D. P. and Y. Xue. M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China. Journal of Hydrologic Engineering, 9(6):491-501, 2004.

Song, C. and R. Kawai. Monte Carlo and Variance Reduction Methods for Structural Reliability Analysis: A Comprehensive Review. Probabilistic Engineering Mechanics, 73:103479, 2023.

Song, Y. Y. and L. Ying. Decision Tree Methods: Applications for Classification and Prediction. Shanghai Archives of Psychiatry, 27(2):130-135, 2015.

Spanos, P. D. and I. A. Kougioumtzoglou. Response and First-Passage Statistics of Nonlinear Oscillators via a Numerical Path Integral Approach. Journal of Engineering Mechanics, 139(9):1207-1217, 2013.

Su, L. and H. J. He. Decision Tree-Based Seismic Damage Prediction for Reinforcement Concrete Frame Buildings Considering Structural Micro-Characteristics. Advances in Structural Engineering, 22:2097-2109, 2019.

Thai, H. T. Machine Learning for Structural Engineering: A State-of-the-Art Review. Structures, 38:448-491, 2022.

Zhang, Y., I. A. Kougioumtzoglou, and F. Kong. A Wiener Path Integral Technique for Determining the Stochastic Response of Nonlinear Oscillators with Fractional Derivative Elements: A Constrained Variational Formulation with Free Boundaries. Probabilistic Engineering Mechanics, 71:103410, 2023