Structural Reliability Analysis Using Machine Learning Techniques

Appalanaidu Yalla[1], Dr. Sharmila Sujatha[2]
[1]*M.Tech Scholar, Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Visakhapatnam, Andhra Pradesh, India*
[2]*Assistant Professor, Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Visakhapatnam, Andhra Pradesh, India*

Abstract— **Structural reliability analysis estimates the probability of failure P_f of a component whose material, geometric and loading parameters are uncertain. Monte Carlo simulation (MCS), the first- and second-order reliability methods (FORM and SORM) and importance sampling (IS) become costly or inaccurate when limit-state evaluations are expensive or strongly nonlinear. This paper develops a MATLAB framework in which machine-learning surrogates, namely the artificial neural network (ANN), support vector machine (SVM), support vector regression (SVR), Gaussian process regression (GPR/Kriging), active-learning Kriging (AK) and random forest (RF), replace the limit-state function in MCS. It is applied to two mechanical components: a beam in bending with ten uncertain variables (a 10,000-sample dataset with 1.46% failures) and a shaft under combined bending and torsion governed by the von Mises criterion (100,000 samples). For the beam, surrogates trained on 7,000 samples evaluated 100,000 Monte Carlo samples; the ANN, SVR and GPR estimates of P_f (0.01388, 0.01449 and 0.01448) lie within the sampling uncertainty of the dataset value 0.0146, and GPR fits best (coefficient of determination $R^2 = 0.99998$), whereas RF underestimates P_f (0.00517) despite $R^2 = 0.92252$. For the shaft, MCS gives $P_f = 0.33362$, and ANN-MCS, Kriging-MCS and AK-MCS lie within 0.4% of it, while FORM and SORM differ by 4.6% and 2.7%. Because both limit states are analytical, direct MCS was faster than every surrogate; surrogates are therefore justified mainly where each evaluation needs finite-element or other costly analysis.**

Index Terms— **Active learning, Artificial neural network, Gaussian process regression, Kriging, Monte Carlo simulation, Reliability index, Structural reliability, Support vector regression, Surrogate modelling**

I. INTRODUCTION

Engineering components work under uncertainty in material properties, dimensions, loads, defects and operating environment. Deterministic design absorbs this uncertainty in safety factors, which say little about how likely failure actually is. Structural reliability analysis instead treats the uncertain quantities as random variables and computes the probability that a performance requirement is violated [1].

Monte Carlo simulation (MCS) is robust and handles nonlinear limit states, but it needs very many samples when the failure probability is small, and its cost becomes prohibitive when each sample requires a finite-element or other nonlinear analysis. The first- and second-order reliability methods (FORM and SORM) replace sampling by a search for the most probable failure point, but their accuracy suffers when the limit-state surface is strongly nonlinear. Importance sampling (IS) concentrates samples near the failure region, yet a suitable sampling density is hard to choose for highly nonlinear or multiple failure regions; metamodel-assisted variants have therefore been proposed [2].

Machine-learning surrogates offer a way out. A model trained on a limited number of limit-state evaluations can afterwards be evaluated for millions of samples at negligible cost. Artificial neural networks (ANN) approximate nonlinear structural responses [3]–[5], support vector machines (SVM) and support vector regression (SVR) provide classification and regression surrogates [6]–[8], and Gaussian process regression (GPR, known in engineering as Kriging) also returns a prediction uncertainty that can steer the choice of

new training points [9], [10]. Reviews of this field are available in [11], [12].

Despite this progress, there remains a need for a unified comparison of several surrogate families with the conventional methods on common benchmark problems, because that comparison is what an engineer needs in order to choose a method. This paper addresses that need for two mechanical components. Its contributions are: (i) a modular MATLAB framework that couples ANN, SVM, SVR, GPR/Kriging, active-learning Kriging (AK) and random forest (RF) surrogates with MCS; (ii) its application to a ten-variable beam in bending and to a shaft under combined bending and torsion; and (iii) a comparison in terms of failure probability, reliability index, prediction error and computing time, which shows that a surrogate with a high coefficient of determination can still give a biased failure probability and that surrogates only pay off when the original model is costly.

Section II reviews related work, Section III states the reliability problem and the conventional methods, Section IV describes the surrogates and the evaluation protocol, Section V defines the datasets and the two examples, Section VI presents and discusses the results, and Section VII concludes.

II. RELATED WORK

A. Surrogates for structural reliability

Chojaczyk et al. [3] reviewed and applied ANN models to the reliability analysis of steel structures, an early demonstration of neural networks as response approximators. Zhang et al. [13] and Lieu et al. [4] used deep neural network surrogates; the latter added adaptive sampling to represent nonlinear limit states better. Roy and Chakraborty [6] reviewed SVM in structural reliability and stressed kernel selection, training and the accurate prediction of the failure boundary, while Bourinet [7] used adaptive SVR surrogates for rare-event probabilities. Afshari et al. [11] reviewed machine-learning methods more broadly and emphasised training data, model accuracy and computational efficiency. Hybrid schemes that combine machine learning with Monte Carlo sampling are also reported [14]–[17].

B. Kriging and active learning

Wei et al. [18] developed a multiple-response Gaussian process model for system reliability and reliability-based importance analysis. Moustapha et al. [9] surveyed active learning for structural reliability and proposed a general framework and benchmark; Jiang et al. [10] proposed failure-pursuing sampling, and Marelli and Sudret [19] combined sparse polynomial chaos expansions with bootstrap. Kriging-based active learning has since been made more efficient and versatile [20], [21], extended with quantified stopping criteria [22], parallel sampling [23] and noisy responses [24]. The most recent work targets small failure probabilities [25], multi-fidelity models [26], epistemic uncertainty [27], stochastic Kriging [28], pseudo-Kriging strategies [29] and the selection of the learning function [30].

C. Mechanical components and fatigue

Hu et al. [31] studied reliability-based design optimisation of wind turbine blades under fatigue and dynamic wind-load uncertainty, and Dong et al. [32] used adaptive surrogates for the time-variant fatigue reliability of welded joints with surface cracks. These studies show that surrogate modelling carries over to mechanical components, and they suggest natural extensions of the static models used here.

D. Research gap

The reviewed work shows that ANN, SVR and Kriging/GPR are well established as surrogates, with most effort going into sampling strategies, accuracy and cost. A side-by-side study on a beam and on a shaft under combined bending and torsion is still useful: the shaft stresses depend on the inverse cube of the diameter, so the limit state is nonlinear, and failure is governed by uncertainty in strength, bending moment, torque, diameter and stress-concentration factors. The present work supplies such a comparison, with attention to the accuracy of failure-boundary prediction and to computing cost.

III. RELIABILITY FORMULATION

A. Limit state and reliability index

Let X denote the vector of random variables and g(X) the limit-state function, defined so that $g \leq 0$ means failure and $g > 0$ means safety. The probability of failure is

$$P_f = P[g(X) \le 0] \quad (1)$$

With Φ the standard normal cumulative distribution function, the reliability index is

$$\beta = -\Phi^{-1}(P_f) \quad (2)$$

This index is used throughout to compare methods; it equals the distance from the origin to the limit-state surface in standard normal space when that surface is linear.

B. Monte Carlo simulation

For N samples X_i drawn from the joint distribution, with $I[\cdot]$ the failure indicator,

$$P_f \approx \frac{1}{N}\sum_{i=1}^{N} I[g(X_i) \le 0] \quad (3)$$

MCS needs no linearity or differentiability of g, but its cost grows quickly as P_f decreases or as each evaluation of g becomes expensive. The estimate itself is uncertain; its standard error is

$$SE_P = \sqrt{\frac{P_f(1-P_f)}{N}} \quad (4)$$

so a failure fraction obtained from a finite dataset must be read together with this sampling uncertainty.

C. FORM and SORM

FORM transforms the variables to independent standard normal variables u and finds the design point u^* that minimises $\|u\|$ subject to $g(u) = 0$. Then

$$\beta = \|u^*\|, \quad P_f \approx \Phi(-\beta) \quad (5)$$

SORM adds the curvature of the limit-state surface at the design point. With principal curvatures κ_i, the Breitung approximation is

$$P_f \approx \Phi(-\beta)\prod_{i=1}^{n-1}(1+\beta\kappa_i)^{-1/2} \quad (6)$$

D. Importance sampling

IS draws samples from an alternative density h(X) placed near the failure region and corrects them with the likelihood ratio f/h, where f is the original density:

$$P_f \approx \frac{1}{N}\sum_{i=1}^{N} I[g(X_i) \le 0]\frac{f(X_i)}{h(X_i)} \quad (7)$$

IV. MACHINE-LEARNING SURROGATES AND EVALUATION

A. Surrogate concept

A surrogate $\hat{g}(X) = f_{ML}(X)$ is trained on pairs of inputs and limit-state values obtained from the original model. After training, it replaces the original model during sampling. Its quality depends on how well the training data cover the input space; errors near $g = 0$ matter most, because they flip the safe/failure classification of a sample and distort P_f.

B. Artificial neural network

The ANN is a feed-forward network with a hidden layer and an output layer that maps the input variables to g. It was trained with MATLAB's train function using scaled conjugate gradient back-propagation and the mean squared error as performance measure. The same network can be trained on the binary failure label for classification.

C. SVM and SVR

The SVM classifies samples as safe or failed through the decision function sign $[\Sigma\, \alpha_i\, y_i\, K(X_i, X) + b]$, where K is a kernel [33]. SVR extends the idea to continuous outputs and predicts g directly; a nonlinear kernel such as the radial basis function lets both models follow a curved safe/failure boundary.

D. GPR/Kriging

GPR models the limit state as a Gaussian process with mean function m(X) and covariance function k(X, X′). At any input the prediction is Gaussian,

$$g(X) \sim N\left(\mu_g(X), \sigma_g^2(X)\right) (8)$$

so the model returns both a prediction μ_g and an uncertainty σ_g.

E. Active-learning Kriging

AK starts from a small training set and a Kriging model. Candidate samples are scored by the learning function

$$U(X) = \frac{|\mu_g(X)|}{\sigma_g(X)+\varepsilon} \quad (9)$$

where ε is a small positive constant. A small U marks a point that is close to the predicted limit state and uncertain. The candidate with the smallest U is evaluated with the original model, added to the training set, and the model is retrained; the loop stops when the reliability estimate meets a convergence criterion.

F. Random Forest

RF averages the predictions of many decision trees grown on bootstrap samples. It captures nonlinear interactions and ranks the importance of the input variables, but its predictions are piecewise constant.

G. ML-assisted Monte Carlo simulation

The trained surrogate is evaluated for a large set of independent Monte Carlo samples, and the number of predicted failures gives

$$P_f^{ML} \approx \frac{1}{N}\sum_{i=1}^{N} I[\hat{g}(X_i) \le 0] \quad (10)$$

The reliability index follows from (2). The overall workflow is shown in Fig. 1. Evaluating the surrogate rather than the original model is what allows 100,000 samples to be used even when only a few thousand original evaluations are available for training.

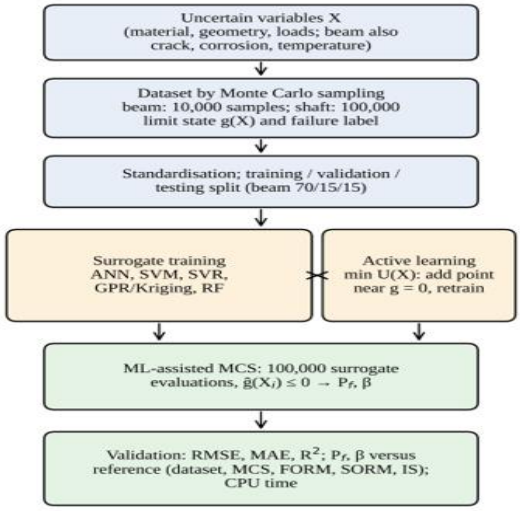


Fig. 1. Workflow of the machine-learning-assisted reliability framework.

H. Evaluation protocol

Each surrogate is evaluated in two steps. First, its regression accuracy is measured on held-out test samples by the root-mean-square error, mean absolute error and coefficient of determination, where g_i is the original and $\hat{g}_i$ the predicted limit state and $\bar{g}$ is the mean of g_i:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(g_i - \hat{g}_i)^2} \quad (11)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|g_i - \hat{g}_i| \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(g_i - \hat{g}_i)^2}{\sum_{i=1}^{N}(g_i - \bar{g})^2} \quad (13)$$

Second, the surrogate is used in (10) and the resulting P_f and β are compared with a reference, using the relative error

$$\varepsilon_{Pf} = \frac{\left|P_f^{ML} - P_f^{ref}\right|}{P_f^{ref}} \times 100\% \quad (14)$$

and the same definition for β. A high R^2 alone does not guarantee an accurate P_f, since the error may be concentrated near $g = 0$; the two steps are therefore reported side by side.

V. DATASETS AND NUMERICAL EXAMPLES

A. Uncertainty modelling and data generation

Mechanical components are affected by uncertainty in material properties (yield strength, Young's modulus), geometry (length, width, height, diameter), loading (forces, moments, torque), manufacturing defects and environment. This study mainly models material, geometric and loading uncertainty; for the beam, crack length, corrosion rate and temperature are included as well. Random samples of the variables are drawn by MCS, the analytical limit state is evaluated for each sample, and a failure label (1 if $g \le 0$, else 0) is stored together with the inputs. The framework was implemented in MATLAB R2022b in a modular form, so data generation, surrogates and reliability methods can be run separately or from one main program. Inputs are standardised to zero mean and unit variance. The beam dataset has 10,000 samples and contains only 146 failures (1.46%). It is split into 7,000 training, 1,500 validation and 1,500 testing samples, with the validation set used to select model settings; because failures are rare, the split is stratified by the failure label. The shaft dataset has 100,000 samples. ANN, SVR, GPR and RF predict g, and SVM predicts the failure label. Surrogate hyperparameters were:

B. Example 1: beam in bending

The first problem is a beam of rectangular section $b \times h$ and length L loaded by a point load P at distance $a = L(a/L)$ from a support. Ten variables are uncertain: Young's modulus E, yield strength f_y, length L, width b, height h, load P, normalised load position a/L, crack length c, corrosion rate r and temperature T. Table I gives their sample means and maxima. The section modulus is $Z = b\,h^2/6$, the resisting moment is $M_R = f_y Z$ and the demand is $M_D = P\,a$, so the limit state is

$$g(X) = \frac{M_R}{M_D} - 1 = \frac{f_y b\,h^2}{6P\,a} - 1 \quad (15)$$

Crack length, corrosion rate and temperature enter through modification factors on the resistance or the demand. The limit state is dimensionless: $g \le 0$ means that the demand reaches the resistance.

Table I
Uncertain Variables of the Beam Problem

Symbol	Description	Unit	Mean	Max.
E	Young's modulus	MPa	199,897.86	240,000
f_y	Yield strength	MPa	250.51	330
L	Beam length	m	2.0012	2.2946
b	Section width	m	0.04997	0.065
h	Section height	m	0.09988	0.120
P	Applied load	N	10,190.10	18,000
a/L	Load position	–	0.5005	0.7902
c	Crack length	m	0.001099	0.004
r	Corrosion rate	–	0.009996	0.020
T	Temperature	°C	24.80	50

C. Example 2: shaft under combined bending and torsion

The second problem is a circular shaft of diameter d carrying a bending moment M and a torque T, as found in gearboxes, pumps and turbines. The yield strength S_y, M, T, d and the stress-concentration factors K_b (bending) and K_t (torsion) are uncertain and were sampled from prescribed probability distributions. The normal and shear stresses are

$$\sigma_b = \frac{K_b \cdot 32M}{\pi d^3}, \quad \tau = \frac{K_t \cdot 16T}{\pi d^3} \quad (16)$$

and the von Mises criterion gives the limit state

$$g(X) = S_y - \sqrt{\sigma_b^2 + 3\tau^2} \quad (17)$$

Failure occurs when the equivalent stress reaches the yield strength. The dataset contains roughly 30% failure samples by design (70,000 safe and 30,000 failed), which helps classifier training but must not be read as the failure probability; P_f was estimated by MCS on samples drawn from the input distributions.

VI. RESULTS AND DISCUSSION

A. Beam

The 7,000 training samples were used to build the surrogates, which then evaluated 100,000 independent Monte Carlo samples. The reference is the failure fraction of the dataset, $P_f = 146/10{,}000 = 0.0146$ ($\beta \approx 2.18$), whose standard error from (4) is 0.0012, giving a 95% interval of about 0.0122 to 0.0170. Table II and Fig. 2 show that the ANN, SVR and GPR estimates (0.01388, 0.01449 and 0.01448) all fall inside this interval, with errors of −4.9%, −0.8% and −0.8%. RF gives 0.00517, which is 65% below the reference and far outside the interval, and consequently overestimates β (2.5642 against 2.18).

Table II
Failure Probability and Reliability Index of The Beam

Method	P_f	Reliab. (%)	β	Error in P_f (%)
Reference (dataset)	0.01460	98.540	2.1816	–
ANN	0.01388	98.612	2.2007	−4.93
SVR	0.01449	98.551	2.1838	−0.75
GPR/Kriging	0.01448	98.552	2.1840	−0.82
Random forest	0.00517	99.483	2.5642	−64.59

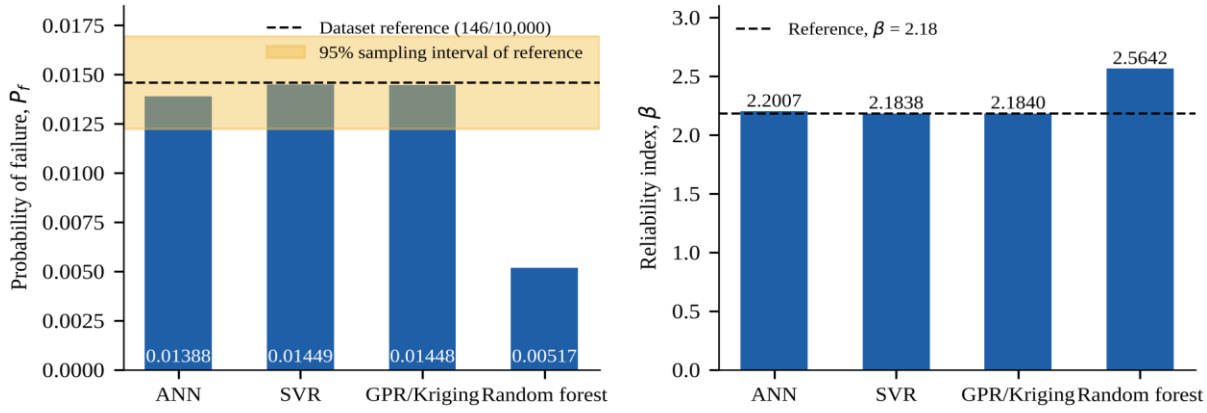


Fig. 2. Beam: (a) failure probability and (b) reliability index from the surrogates, with the dataset reference (dashed line) and its 95% sampling interval (shaded).

Table III gives the regression accuracy on the test data. GPR is best ($R^2 = 0.99998$, RMSE = 0.0031), followed by ANN (0.99946) and SVR (0.99614), while RF is clearly the weakest ($R^2 = 0.92252$, RMSE = 0.194). Fig. 3 shows the same ordering: GPR and ANN points lie on the ideal line, SVR departs slightly at large limit-state values, and RF scatters visibly, as expected of piecewise-constant tree predictions. The accuracy ranking carries over to P_f only partly: ANN has a higher R^2 than SVR but a larger P_f error, so R^2 cannot replace a direct check of P_f. RF's bias is probably due to large errors near g = 0, where fewer than about a hundred failure samples are available for training.

Table III
Prediction Accuracy of The Beam Surrogates

Method	RMSE	MAE	R^2
ANN	0.016161	0.0096333	0.99946
SVR	0.043322	0.017254	0.99614
GPR/Kriging	0.0031466	0.001716	0.99998
Random forest	0.19412	0.13695	0.92252

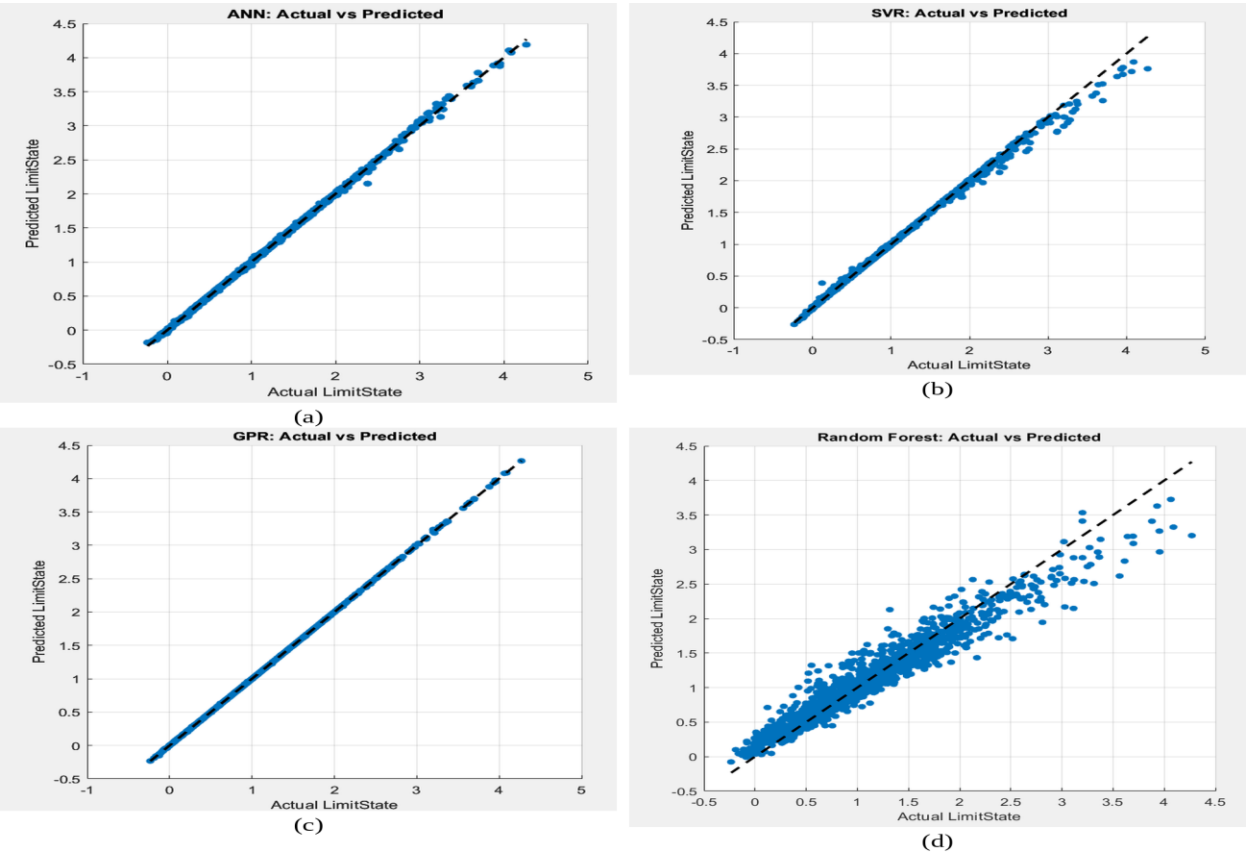


Fig. 3. Beam: predicted versus actual limit state for (a) ANN, (b) SVR, (c) GPR/Kriging and (d) random forest. The dashed line is the ideal prediction.

B. Shaft: failure probability and reliability index

Table IV gives the results for the shaft. The failure probabilities lie between 0.31585 and 0.33362 and β between 0.42993 and 0.47932 (Fig. 4). Taking MCS (0.33362) as reference, IS, AK-MCS, Kriging-MCS and ANN-MCS are within 0.4%. FORM and SORM underestimate the failure probability by 4.6% and 2.7%, consistent with a nonlinear limit state. SVR-MCS differs by 2.5%, and SVM-MCS has the largest deviation (5.3%), probably because a classifier misassigns samples near the boundary.

Table IV
Reliability Of the Shaft by Different Methods

Method	P_f	Reliability	β	Error in P_f (%)
MCS	0.33362	0.66638	0.42993	–
FORM	0.31832	0.68168	0.47239	4.59
SORM	0.32470	0.67530	0.45459	2.67

Method	P_f	Reliability	β	Error in P_f (%)
Importance sampling	0.33318	0.66682	0.43116	0.13
ANN + MCS	0.33231	0.66769	0.43354	0.39
SVM + MCS	0.31585	0.68415	0.47932	5.33
SVR + MCS	0.32531	0.67469	0.45290	2.49
Kriging + MCS	0.33277	0.66723	0.43228	0.25
AK + MCS	0.33309	0.66691	0.43138	0.16

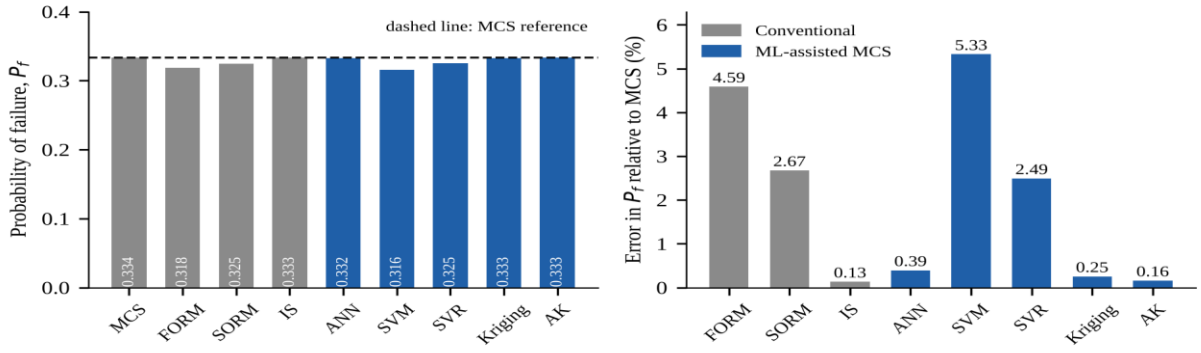


Fig. 4. Shaft: (a) failure probability and (b) error in P_f relative to MCS for the conventional and surrogate-based methods.

C. Shaft: surrogate accuracy and computing cost

Table V lists the error measures and total CPU time. All surrogates reach R^2 between 0.9899 and 0.99566. The ANN shows the smallest RMSE and MAE, followed by Kriging and AK, while SVM and SVR show errors about two orders of magnitude larger. The ANN training record (Fig. 5) ends at the maximum of 500 epochs with a mean squared error of about 51, and its regression coefficient is 0.99975. The analytical-versus-predicted plots (Fig. 6) show the same picture as for the beam: ANN and GPR follow the ideal line closely, SVR deviates slightly and RF scatters most.

Table V
Accuracy And Computing Time for The Shaft

Method	RMSE	MAE	R^2	Time (s)
MCS	–	–	–	0.0649
FORM	–	–	–	0.0086
SORM	–	–	–	0.0113
Importance sampling	–	–	–	0.0203
ANN + MCS	11,926	4,952	0.99566	17.4
SVM + MCS	3.13×10^6	2.26×10^6	0.99566	2.54
SVR + MCS	3.28×10^6	2.36×10^6	0.9899	0.773
Kriging + MCS	1.55×10^5	82,981	0.99233	4.40
AK + MCS	7.92×10^5	3.86×10^5	0.99144	1.33

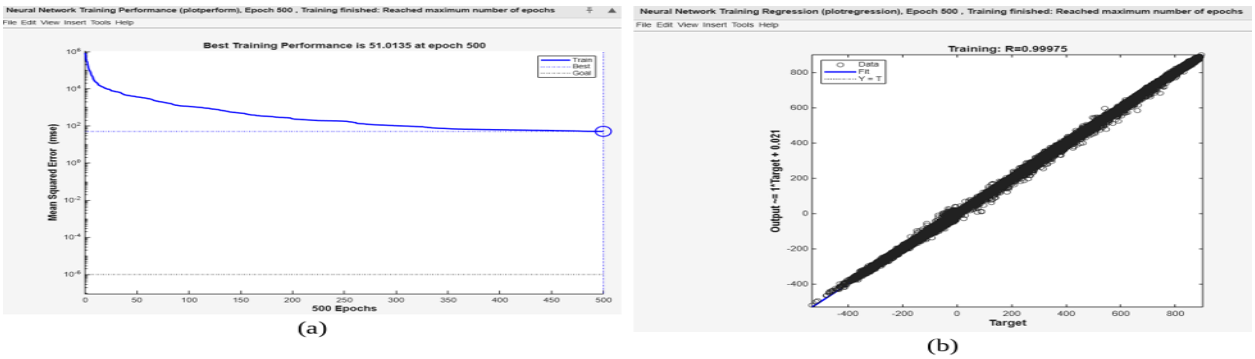


Fig. 5. ANN for the shaft: (a) best training performance (mean squared error versus epoch) and (b) regression of network output on target.

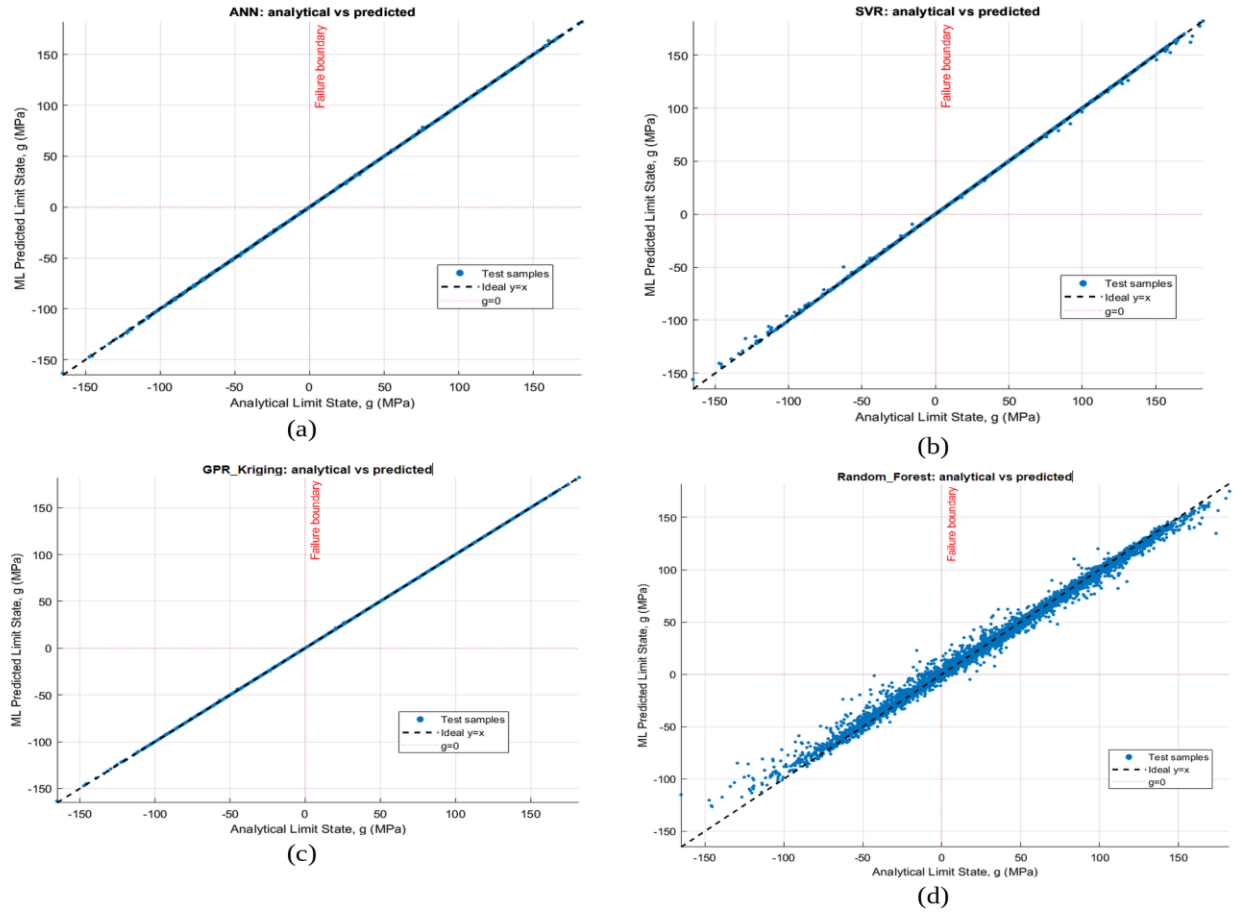


Fig. 6. Shaft: analytical versus predicted limit state for (a) ANN, (b) SVR, (c) GPR/Kriging and (d) random forest. The vertical line marks the failure boundary g = 0.

D. Discussion and limitations

For both problems the best surrogates reproduce the reference failure probability well: GPR, SVR and ANN on the beam, and Kriging-type models and ANN on the shaft. The computing times, however, show the limit of the approach. The limit states used here are closed-form expressions, so direct MCS needed 0.065 s, whereas the surrogate schemes needed between 0.77 s (SVR) and 17.4 s (ANN) (Fig. 7). These times are totals that include training and prediction, so a fairer comparison would fix the sample size, hardware and timing procedure. A surrogate save time only when one evaluation of g is much more expensive than training and predicting, as in finite-element or nonlinear analysis, which is the setting that motivates the method [11].

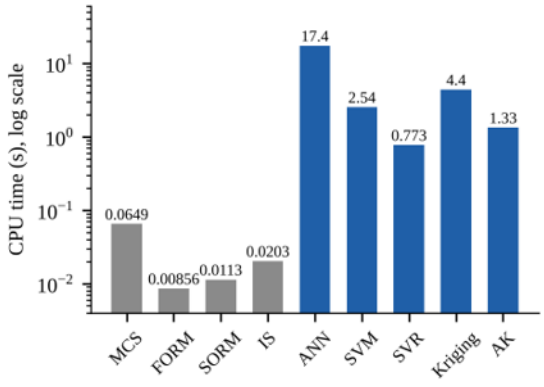


Fig. 7. Computing time of the reliability methods for the shaft (logarithmic scale).

Four limitations should be stated. First, the beam reference is itself a finite-sample estimate with a standard error of about 8% of its value, so differences of a few percent between surrogates are not

significant. Second, the failure probabilities (0.0146 for the beam and 0.33 for the shaft) are far larger than those usually targeted in design, so the study does not show performance for rare events, for which adaptive and active-learning strategies were developed [7], [25]. Third, the results depend on the assumed input distributions and failure criteria. Fourth, the benchmark models are analytical; validation on finite-element models or measured data is still required.

VII. CONCLUSION

This paper coupled ANN, SVM, SVR, GPR/Kriging, active-learning Kriging and random forest surrogates with Monte Carlo simulation and compared them with conventional reliability methods for a beam in bending and a shaft under combined bending and torsion. On the beam, ANN, SVR and GPR reproduced the dataset failure probability (0.0146) within its sampling uncertainty, GPR gave the best fit ($R^2 = 0.99998$), and RF underestimated P_f by 65% despite $R^2 = 0.92252$. On the shaft, ANN-MCS, Kriging-MCS and AK-MCS were within 0.4% of MCS, whereas FORM and SORM were off by 4.6% and 2.7% and SVM-MCS by 5.3%. Because the limit states were cheap to evaluate, direct MCS was faster than all surrogates, so the benefit of surrogates is expected for costly models rather than for these benchmarks.

Future work should use finite-element or experimental data, add fatigue, crack growth and corrosion to the shaft model, study rare-event probabilities with active learning, integrate the surrogates in reliability-based design optimisation, and consider Bayesian updating with monitoring data [34].

ACKNOWLEDGMENT

The authors thank Dr. Sharmila Sujatha and the Department of Computer Science and Systems Engineering, Andhra University College of Engineering for guidance and support.

REFERENCES

[1] Haldar and S. Mahadevan, Probability, Reliability, and Statistical Methods in Engineering Design. New York, NY, USA: Wiley, 2000.

[2] V. Dubourg et al., “Metamodel-based importance sampling for structural reliability analysis,” Probabilistic Engineering Mechanics, 2013.

[3] Chojaczyk et al., “Review and application of artificial neural networks models in reliability analysis of steel structures,” Structural Safety, vol. 52, pt. A, pp. 78–89, 2015.

[4] Q. X. Lieu et al., “An adaptive surrogate model to structural reliability analysis using deep neural network,” Expert Systems with Applications, vol. 189, 2022.

[5] J. B. Cardoso et al., “Structural reliability using Monte Carlo and neural networks,” Advances in Engineering Software, 2008.

[6] Roy and S. Chakraborty, “Support vector machine in structural reliability analysis: A review,” Reliability Engineering & System Safety, vol. 233, 2023.

[7] J.-M. Bourinet, “Rare-event probability estimation with adaptive support vector regression surrogates,” Reliability Engineering & System Safety, vol. 150, pp. 210–221, 2016.

[8] H. Li et al., “Reliability analysis using SVM and Monte Carlo simulation,” Structural Safety, 2016.

[9] M. Moustapha, S. Marelli, and B. Sudret, “Active learning for structural reliability: Survey, general framework and benchmark,” Structural Safety, vol. 96, 2022.

[10] Jiang et al., “A general failure-pursuing sampling framework for surrogate-based reliability analysis,” Reliability Engineering & System Safety, vol. 183, pp. 47–59, 2019.

[11] S. Saraygord Afshari et al., “Machine learning-based methods in structural reliability analysis: A review,” Reliability Engineering & System Safety, vol. 219, 2022.

[12] J. Joseph and C. K. Swalih, “Implementation of machine learning in structural reliability analysis,” Journal of Civil Engineering and Materials Application, 2023.

[13] Y. Zhang et al., “Deep neural network surrogate models for structural reliability analysis,” Computer Methods in Applied Mechanics and Engineering, 2019.

[14] Luo et al., “Hybrid enhanced Monte Carlo with machine learning,” Computer Methods in Applied Mechanics and Engineering, 2022.

[15] Q. Wang et al., "Machine learning-aided static structural reliability analysis," Applied Mathematical Modelling, 2020.

[16] W. You, L. Shao, and F. Qi, "A hybrid method combining importance sampling and ensemble learning for structural reliability analysis," Quality and Reliability Engineering International, vol. 42, no. 5, pp. 1908–1922, 2026.

[17] N. Zemed, "Enhanced active learning for structural reliability analysis: An ensemble SVR metamodel–Monte Carlo approach," Quality and Reliability Engineering International, 2025.

[18] P. Wei, F. Liu, and C. Tang, "Reliability and reliability-based importance analysis of structural systems using multiple response Gaussian process model," Reliability Engineering & System Safety, vol. 175, pp. 183–195, 2018.

[19] S. Marelli and B. Sudret, "An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis," Structural Safety, vol. 75, pp. 67–74, 2018.

[20] J. Wang et al., "An efficient and versatile Kriging-based active learning method for structural reliability analysis," Reliability Engineering & System Safety, vol. 241, 2024.

[21] Z. Xiang et al., "Active learning method for structural reliability analysis," Mechanical Systems and Signal Processing, 2020.

[22] Prentzas and M. Fragiadakis, "Quantified active learning Kriging model for structural reliability analysis," Probabilistic Engineering Mechanics, vol. 78, 2024.

[23] Y. Meng et al., "An active learning Kriging model with approximating parallel strategy for structural reliability analysis," Reliability Engineering & System Safety, vol. 247, 2024.

[24] Chun, "Active learning-based Kriging model with noise responses and its application to reliability analysis of structures," Applied Sciences, vol. 14, no. 2, Art. no. 882, 2024.

[25] Y. Pang et al., "An active learning Kriging method based on uniform sampling and local refinement for efficient reliability analysis with small failure probability," Computer Methods in Applied Mechanics and Engineering, vol. 444, 2025.

[26] S. Feng et al., "An active-learning method based on hierarchical Kriging model for multi-fidelity reliability analysis," Reliability Engineering & System Safety, vol. 266, pt. B, 2026.

[27] H. Zhu et al., "A dynamic-switching active-learning Kriging considering epistemic uncertainty for structural reliability analysis," Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 240, no. 4, 2026.

[28] Wan et al., "A new multiple stochastic Kriging model for active learning surrogate-assisted reliability analysis," Reliability Engineering & System Safety, vol. 260, 2025.

[29] S. Yu, W. Fan, and Z. Li, "A new active learning Kriging method with pseudo Kriging strategy for structural reliability analysis," Structural and Multidisciplinary Optimization, vol. 68, no. 6, 2025.

[30] Y. Ma, C. Hai, and L. Mei, "Adaptive learning function selection for Kriging-assisted structural reliability analysis," Applied Mathematical Modelling, vol. 152, 2026.

[31] W. Hu, K. K. Choi, and H. Cho, "Reliability-based design optimization of wind turbine blades for fatigue life under dynamic wind load uncertainty," Structural and Multidisciplinary Optimization, vol. 54, pp. 953–970, 2016.

[32] Y. Dong et al., "Application of adaptive surrogate models in time-variant fatigue reliability assessment of welded joints with surface cracks," Reliability Engineering & System Safety, vol. 195, 2020.

[33] P. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press, 2012.

[34] Straub and I. Papaioannou, "Bayesian updating in structural reliability," Structural Safety, 2015.