14th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP14
Dublin, Ireland, July 9-13, 2023

Applicability of meta-model assisted reliability assessment for dynamic problems: a comparison between regression-based methods

Reza Allahvirdizadeh
PhD Student, Division of Structural Engineering & Bridges, KTH Royal Institute of Technology, Stockholm, Sweden

Andreas Andersson
Researcher, Division of Structural Engineering & Bridges, KTH Royal Institute of Technology, Stockholm, Sweden

Raid Karoumi
Professor, Division of Structural Engineering & Bridges, KTH Royal Institute of Technology, Stockholm, Sweden

ABSTRACT: There is a growing intent among engineers, stakeholders, and decision makers to use probabilistic methods for infrastructure assessment or design objectives. However, the corresponding limit state for such problems usually requires the construction of complex computational models, usually using commercial software without parallelization capability. Such a requirement makes performing reliability analysis computationally prohibitive, which is even more challenging for dynamic problems, since a very short time step is required to obtain sufficiently accurate predictions. This concern has led to several methods being proposed to surrogate the limit state function with a generally black box called a meta-model. A variety of them, such as Kriging, Polynomial Chaos Expansion (PCE), Artificial Neural Networks (ANN), and response surfaces (e.g., polynomial, spline, or radial-base functions), have been adopted for this purpose. These meta-models are typically trained on a limited data set collected by computing the true responses of carefully selected input variables. Their applicability for assessing the probability of failure has been studied individually in the literature for both benchmark and practical problems. However, as far as the authors are aware, no comparison has been made between them for dynamic problems. This comparison needs to be made from the point of view of both accuracy and performance (number of calls to the limit state function). In this context, this paper takes a systematic approach to evaluate their performance under identical conditions, i.e., with similar training datasets. For this purpose, the dynamic response of railway bridges with different spans excited by the passage of trains with a wide range of speeds is used as a reference problem.

1. INTRODUCTION

The built environment faces a variety of uncertainties during its lifetime. These uncertainties include, for example, induced impacts, experienced damage, or changes in properties. Therefore, the evaluation of their influence on the desired performance of the infrastructure seems essential. This statement is further emphasized when considering that the infrastructures currently in operation are quite old, the future environmental conditions are very uncertain, and the expectations of modern society are much higher than before.

Traditionally, partial safety factors are used to implicitly account for the above uncertainties. However, their applicability is not guaranteed if the design scenario is far from those considered in the calibration process. Moreover, they cannot help in assessing the level of risk. In addition, the design process is inherently an optimization process, usually aimed at minimizing total cost or maximizing utilities. Other researchers have discussed that conventional optimization approaches do not necessarily result in safe designs, even when all regulations (based on partial safety factor approaches) are followed. This has led to reliability-based design optimization (RBDO) approaches gaining significant attention in recent decades. In this approach, the satisfaction of the target reliability for each design situation (limit state) is implemented as an additional constraint in the optimization process. The interested reader can find more about RBDO in Enevoldsen and Sørensen (1994).

Therefore, depending on the approach that aims to address safety-related concerns of the infrastructure, the estimation of the probability of failure is unavoidable. Such estimation requires the calculation of a multiple integral that reads as Eq. (1).

$$p_f = P\big[g(\mathbf{X}) \leq 0\big] = \int_{\mathscr{D}_f} f_{\mathbf{X}}(\boldsymbol{x})d\boldsymbol{x} \qquad (1)$$

where $f_{\mathbf{X}}(\boldsymbol{x})$ is the joint probability distribution of the random variables $\mathbf{X} \in \mathbb{R}^{N_X}$ (N_X is problem dimensionality), $g(\bullet)$ is the limit state function defining the boundary between safe and failure domains and $\mathscr{D}_f = \{x|g(x) \leq 0\} \subset \Omega$ is the failure domain.

A variety of methods have been developed to calculate this integral, of which the first/second order reliability method (FORM /SORM) and direct Monte Carlo simulation (MCS) are probably the best known. The former expresses the shortest distance from the point with largest likelihood to the most probable point (MPP) located on the limit state; this is approximated by its first (or second) order Taylor series expansion around the MPP. Such an approximation may affect the accuracy of FORM /SORM results for strongly nonlinear limit state functions (Melchers and Beck, 2018). Moreover, the MPP is an unknown point that should be determined iteratively, and the method requires the calculation of gradients of the limit state function, which can lead to high computational costs, especially for problems with complex computational models (e.g., nonlinear finite element models).

On the other hand, MCS estimates the expected value of an indicator function fed by pointwise estimates of the limit state function. The method is very stable, but its accuracy is inversely proportional to the square root of the pointwise estimates (number of samples), which makes it computationally impractical for problems with low failure probability and computationally intensive models.

The efficiency drawback of direct MCS is attempted to be addressed by proposing methods such as important sampling (Melchers, 1989) and subset simulation (Au and Beck, 2001). Despite the significant reduction in computational cost, these methods still require significant calls to the computational model. Therefore, an alternative approach is taken by adapting statistical learning concepts to surrogate the limit state function with a function that is very cheap to evaluate, known as metamodeling (Hurtado, 2004).

Most often, the trained meta-models are used in conjunction with MCS; however, previous studies have also used this concept for variance reduction and subset simulation approaches. As mentioned earlier, only the sign of the pointwise estimate of the limit state function is required, making both regression-based and binary classification-based surrogate models applicable to reliability assessment objectives. The former represents a continuous function that interpolates the result of the limit state function, while the latter finds a boundary separating the state space into safe and failure domains. It should be emphasized here that the current study focuses only on regression-based models.

A variety of regression-based surrogate models have been used for reliability assessment problems, such as Gaussian Process Regression (also known as Kriging), Polynomial Chaos Expansion (PCE), Artificial Neural Network (ANN), and Polynomial Response Surface (RS). The following sections provides a brief overview of their concepts.

All these methods have shown acceptable performance on benchmark problems (mostly with explicit limit state functions), especially in terms of computational cost. However, the previous studies do not seem to answer in detail how and which methods should be selected for practical problems. Assuming that the training phase of these models is much less computationally intensive than the calculation of the true value of the limit state function at an unknown point, cross-validation techniques can be used to select the most appropriate meta-model. Despite the applicability of this approach, some practical problems may arise.

Most often, meta-models are trained according to an active learning scheme that starts with a small experimental design (DOE). Then, at each iteration, a learning function is used to find the most informative points, its true limit state value is computed, the DOE is enriched, and then the meta-model is retrained until a stopping criterion is satisfied (Moustapha et al., 2022). In this approach, there is no guarantee that the best model selected by cross-validation in the first step is still the best one at the end of the active learning phase. Moreover, DOE is usually a very imbalanced dataset. Therefore, it is possible that the stratified training/validation datasets contain points that all belong to one class (most likely points within safe domain), leading to misleading conclusions during cross-validation.

Given these concerns, a systematic approach is taken here to compare the performance of regression-based meta-models. It is clear that their performance also depends on the nature of the problem. Therefore, the current study focuses only on dynamic problems, which often have a nonlinear limit state with computationally expensive models. The latter is mainly due to the very short time steps required to achieve acceptable accuracy in the numerical integration of the equations of motion. In this context, the dynamic behavior of high-speed railway bridges is considered.

The remainder of this article is organized to provide a brief overview of the considered meta-models, describe the problem of dynamic behavior of high-speed railway bridges, compare the performance of the meta-models considered for these problems, and then conclude the article with a discussion of the highlights of the study.

2. REGRESSION-BASED META-MODELS

This section is devoted to briefly introduce the concept of some common regression-based meta-models used for structural reliability assessment.

2.1. *Gaussian process regression (Kriging)*

Gaussian process regression is a nonparametric model that, unlike conventional regression analyzes, does not assume independent errors between predictions. It considers a spatial correlation between predictions implemented using an error function that follows a Gaussian process. Thus, the meta-model can be formulated as Eq. (2).

$$g(\mathbf{X}) \approx \hat{\mathscr{M}}_{\mathrm{GPR}}(\mathbf{X}) = \mathscr{H}(\mathbf{X}, \mathbf{w}) + z(\mathbf{X}; \sigma_n^2, \boldsymbol{\theta}) \quad (2)$$

where $\mathscr{H}(\mathbf{X}, \mathbf{w})$ is the matrix of basis functions with coefficients of $\mathbf{w}$ representing the general trend of the model and $z(\mathbf{X}; \sigma_n^2, \boldsymbol{\theta}) \sim \mathscr{GP}(0, \Sigma)$ is the error function following a Gaussian process with zero mean and covariance of Σ. The latter is accounted for by a variety of kernel functions (e.g., Gaussian and Matern families) with process variance of σ_n^2 and hyperparameters of $\boldsymbol{\theta}$.

As mentioned earlier, the model is trained by defining the coefficients of the basis functions, the variance of the process, and the hyperparameters of the autocorrelation function. It is worth noting that the first two parameters can be estimated knowing the hyperparameters, which are determined by the maximum likelihood method. The interested reader can find details of this method in Rasmussen and Williams (2006) and its application for structural reliability assessment in Gaspar et al. (2014); Allahvirdizadeh et al. (2022).

2.2. *Polynomial chaos expansion (PCE)*

Polynomial chaos expansion approximates the limit state function by a set of orthogonal basis functions whose probability density functions of the random variables are their inner product weighting functions. It can thus be formulated as Eq.(3).

$$\hat{\mathscr{M}}_{\mathrm{PCE}}(\mathbf{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^{N_{\mathbf{X}}}} w_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{u}) \approx \sum_{|\boldsymbol{\alpha}| \leq q} w_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{u}) \quad (3)$$

where $\Psi_{\alpha}(\cdot)$ is a multivariate polynomial basis function resulting from the tensor product of the univariate polynomial basis functions, $\mathbf{u}$ is the vector of transformed random variables, and $w_{\boldsymbol{\alpha}}$ are the coefficients of the basis functions. The model is trained by defining the type of the basis function and calculating the deterministic coefficients. The former depend on the type of the random variable and the latter are computed either by minimizing the least square losses or by Galerkin projection methods.

Note that the number of coefficients increases exponentially with the dimensionality of the problem, so PCE suffers from the curse of dimensionality. Therefore, the highest order of polynomials in practical problems is restricted to q (see Eq. (3)). This truncation scheme is further refined by the rank-based sparse representation or hyperbolic truncation (Blatman and Sudret, 2011). The interested reader can find details of this method in Xiu and Karniadakis (2002) and some applications of that for structural reliability assessment in Sudret (2012); Allahvirdizadeh et al. (2021).

2.3. *Artificial neural network (ANN)*

An artificial neural network consists of multiple layers with the middle ones known as hidden layers, each containing multiple parallel processing nodes called neurons that are interconnected and map the input space to the output space. The network where all neurons in subsequent layers are interconnected is called a multilayer perceptron network, which has already been used for structural reliability assessment. Determining the number of hidden layers and also the number of neurons per layer has a significant impact on the performance of ANN, commonly known as network architecture; which can be determined using cross-validation techniques. It is worth noting that ANN models with few hidden layers are known as shallow neural networks, while models with many hidden layers are known as deep learning. Clearly, the runtime complexity of ANN increases dramatically with the number of layers; therefore, only the application of shallow neural networks is of interest in this study.

The input of each neuron results from the weighting of the outputs of the neurons in the previous layer, which is fed by a function known as the activation function, which reads as Eq.(4).

$$\hat{x}_{ij} = \mathscr{A}\left[\sum_{n=1}^{N_{i-1}} w_{nj}\hat{x}_{(i-1)n} + b\right] \tag{4}$$

where $\hat{x}_{ij}$ is the input of the jth neuron in the ith layer, w_{nj} is the weight of the output of the nth neuron in the previous layer fed to the jth neuron in the next layer, N_{i-1} is the number of neurons in the previous layer, b is the bias, and $\mathscr{A}(\bullet)$ is the activation function. It should be noted that a variety of activation functions have been proposed, of which sigmoid, hyperbolic tangent, rectified linear unit (ReLU), and radial basis functions are the most commonly used. The meta-model based on ANN can then be formulated as Eq.(5).

$$\begin{gathered}\hat{\mathscr{M}}_{\text{ANN}}(\mathbf{X}) = \boldsymbol{A}' \circ \boldsymbol{A}_{N_L} \circ \ldots \boldsymbol{A}_2 \circ \boldsymbol{A}_1(\mathbf{X}) \\ \boldsymbol{A}_i : \mathbb{R}^{N_{i-1}} \to \mathbb{R}^{N_i} \quad i = 1, \ldots, N_L\end{gathered} \tag{5}$$

where $\boldsymbol{A}'(\bullet)$ is the set of output functions that map the output of the last hidden layer to the desired output format, $\boldsymbol{A}_i(\bullet)$ is the set of activation functions for the ith hidden layer, and N_L is the number of hidden layers. Note that the weights are computed by minimizing the loss function (e.g., mean square error or cross entropy loss), which is often achieved by the backpropagation method. For brevity, the details of the method are not discussed here; the interested reader is therefore referred to Papadrakakis et al. (1996); Hurtado and Alvarez (2001).

2.4. *Polynomial response surface*

The response surface method often uses low-order polynomial functions (usually up to a maximum order of 2) to surrogate the limit state function. Therefore, in the case of considering interactions between terms, the meta-model can be formulated as Eq.(6).

$$\hat{\mathscr{M}}_{\text{RS}}(\mathbf{X}) = b + \sum_{i=1}^{N_X} w_i x_i + \sum_{i=1}^{N_X}\sum_{j=i}^{N_X} w'_{ij} x_i x_j \tag{6}$$

where b is a bias or intercept term and $\mathbf{w}$ is a set of deterministic weights determined by either least squares or maximum likelihood methods.

Note that when considering a full quadratic model, $(N_X+1)(N_X+2)/2$ terms must be determined, which consequently require at least the same number of calls to the computational model, causing this model to suffer from the curse of dimensionality. This problem can be solved by neglecting the interaction terms, which reduces the number of undetermined terms to $2N_X + 1$. The interested reader can find details of this method and its application to structural reliability assessment in Bucher and Bourgund (1990); Gaspar et al. (2014).

3. BENCHMARK PROBLEM

As mentioned earlier, the dynamic benchmark problem used here to compare the performance of the regression-based meta-models focuses on high-speed railway bridges; a schematic representation of it can be seen in Figure 1. In this context, the limit state considered is devoted to the phenomenon of ballast instability (also known as running safety). Previous experimental studies by Zacher and Baeßler (2009) have shown that the phenomenon occurs when the vertical acceleration of the bridge deck exceeds certain limits. Therefore, the limit state function is formulated based on a limiting vertical acceleration and the maximum vertical acceleration of the bridge deck. A detailed explanation of the phenomenon and a preliminary reliability assessment using FORM can be found in Allahvirdizadeh et al. (2020).

It should be noted that only single-span, simply supported reinforced concrete bridges with spans of 10 m, 20 m, and 30 m are considered. The structural behavior of the bridges is modeled using 2D Euler-Bernoulli beams. Moreover, the stiffness of track and rails is neglected and their contribution is considered as additional mass. Furthermore, passing trains are simplified by neglecting the structure of the coaches and representing them as a series of equidistant moving loads.

The dynamic response of the bridge to the passing trains is then determined by the solution of the equation of motion, which can be read in Eq.(7). At this point, it should be emphasized that in EN 1991-2 (2003) it is recommended to filter out the contribution of higher frequencies from the obtained responses, which leads to the consideration of a low-pass filter with a cutoff frequency of $\max(30\,\mathrm{Hz}, 1.5f_1, f_3)$, where f_1 and f_3 are the frequencies of the first and third vibration modes, respectively.

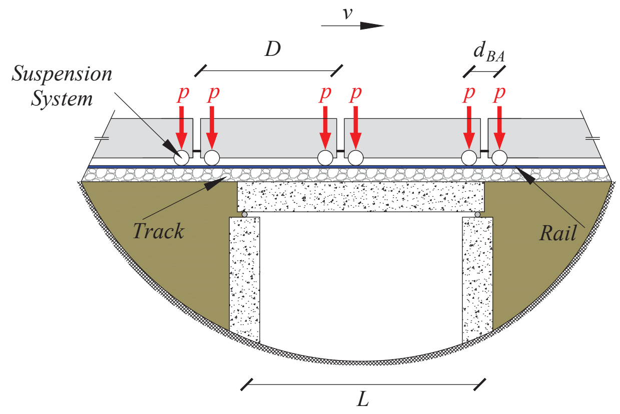


Figure 1: Schematic view of the benchmark problem.

$$\mathbf{M}\ddot{\mathbf{y}}(x,t) + \mathbf{C}\dot{\mathbf{y}}(x,t) + \mathbf{K}\mathbf{y}(x,t)$$
$$= \sum_{j=1}^{N_t} p\left[\Delta_j(t,v,L) + \Delta_j(t-(D-d_{BA})/v,v,L)\right]$$
$$\Delta_j(t,v,L) = \delta\left[x - v(t-t_j)\right]\left[h(t-t_j) - h(t-t_j-\frac{L}{v})\right] \quad (7)$$

where $\mathbf{M}$,$\mathbf{C}$ and $\mathbf{K}$ are mass, damping and stiffness matrices. N_t is the number of coaches, p is the axle load of the train, v is the train speed, D is the length of the coaches, $\delta(\bullet)$ is the Dirac delta function, $h(\bullet)$ is the Heaviside function, and $t_j = (j-1)D/v$. It should be mentioned here that a closed form solution to this problem was developed in Frỳba (2001). This method is adopted here; however, its details are not explained for the sake of brevity.

As mentioned earlier, the structure of passing coaches, rails, and tracks is neglected in the constructed computational models. This approach makes it impossible for the computational model to take into account the effects of train-track-bridge- interaction (TTBI), in particular the additional damping due to the train's suspension system, the amplification of responses due to unevenness on the rail (rail irregularities), and the reduction of responses due to the distribution of axle loads on the track structure. These effects are implicitly implemented in the computational models by Eqs. (8)-(10) (Yau et al., 2019; EN 1991-2, 2003; ERRI D 214/RP 9, 1999). At this point, it should be emphasized that the additional damping should be added to the damping recommended by the regulations in Eq. (11).

$$\Delta\xi = \mu r\big|(r+2i\xi_v)/[(1-r^2)-2i\xi_v r]\big| \quad (8)$$

$$\gamma_r = 1+\alpha'\big[56e^{-(\frac{L}{10})^2}+50(\frac{Lf_1}{80}-1)e^{-(\frac{L}{20})^2}\big] \quad (9)$$

$$\gamma_t = -1.1\lambda'^4+5.3\lambda'^3-8.3\lambda'^2+5.7\lambda'-0.54 \quad (10)$$

$$\xi_b(L) = \max\left[1.5+0.07(20-L),1.5\right] \quad (11)$$

where μ is the ratio of the modal mass of the coach to the bridge, r is its fundamental frequency ratio in percent, ξ_v is the modal damping of the coach, $\alpha' = \alpha/200 = \min(v/22,1)/200$, $\lambda' = 0.1\lambda$, $\lambda = v/f_1$ is the wavelength, and $0.14 \leq \gamma_t(\lambda') \leq 1.0$.

Given the formulated problem, the contributing basic random variables are categorized as given in Table 1. It should be emphasized that all of them are assumed to be independent, with the exception of mass per length and moment of inertia. These two random variables are both dependent on the cross section dimensions and therefore cannot be considered independent. Therefore, a Student's t-copula function with parameters ($\rho = 0.93, \nu = 21$) is used to model this dependence.

Table 1: Considered random variables (parameters of ξ_b are in physical space and χ_M is model uncertainty).

Variable	Dist.	Param.		Trunc.
I(m^4)	$\mathcal{LN}$	-1.12	0.46	0.083
		0.097	0.34	0.580
		0.72	0.40	1.40
m(kg/m)	$\mathcal{LN}$	9.29	0.32	8900
		9.72	0.26	12900
		9.92	0.23	16900
E (GPa)	$\mathcal{N}$	29.7	3.56	-
ξ_b (%)	$\mathcal{LN}$	Eq.(11)	0.30	-
p (kN)	Gumbel	196	20.4	120
d_{BA} (m)	$\mathcal{U}$	2.0	3.50	-
D (m)	$\mathcal{U}$	17.0	28.0	-
f_v (Hz)	Weibull	1.04	3.07	-
m_v (kg)	Gamma	2.2e3	8.60	-
ξ_v (%)	Weibull	16.4	1.59	-
χ_M (-)	$\mathcal{LN}$	0.0	0.10	-
$a_{\lim}$(m/s^2)	$\mathcal{N}$	0.8g	0.1g	[0.6-1.0]g

4. METHODOLOGY

The performance of surrogate models depends on the DOE used for their training. In this context, active learning methods have been developed to improve the performance of the meta-model through gradual DOE enrichment. However, these methods are developed exclusively for each type of surrogate model. Therefore, it was decided to adopt a general approach applicable to all meta-models. In this context, an identical initial DOE is used for all meta-models. The latter is created by generating $(N_X+1)(N_X+2)/2$ uniformly distributed sample points using Latin hypercube sampling technique. Then the meta-models are trained and a randomly generated and identical sample pool is passed to the trained meta-models. The points within 5% quantiles of the absolute calculated values are selected. These points are the candidates closest to the approximated limit state at this stage. Then, the point with the farthest distance from the DOE of the previous iteration is selected to enrich the DOE. This approach is continued until DOE reaches a size of 200. It should be emphasized here that the final size of DOE was chosen based on the authors' experience and no stopping criteria are considered here. Therefore, it is not the absolute accuracy of the meta-models that is of interest here, but their performance relative to each other. Then, the failure probability of the considered bridges is calculated for train speeds up to 200 km/h, 250 km/h and 300 km/h. This procedure is repeated 10 times to investigate the stability of the results and to estimate the confidence interval of the predictions of each meta-model. In addition, a subset simulation is performed to evaluate the actual failure probability of the system, which allows a comparison of the performance of the meta-models. In Allahvirdizadeh et al. (2021), the subset simulation was shown to perform acceptably with a fixed conditional probability of 0.1 for intermediate limit states (p_0) and a number of samples equal to 4000 per iteration (N_s) for the benchmark problems considered. Therefore, the sensitivity of the results with respect to these parameters is not investigated here.

It should be highlighted here that the structure of ANN is designed using the cross-validation technique only in the first stage of the described procedure for each train speed range and remains similar in the remaining steps. In this context, the number of hidden layers is changed between one and three. Moreover, the number of neurons in each hidden layer is varied between $[N_X - 1.5N_X]$.

5. RESULTS AND CONCLUSION

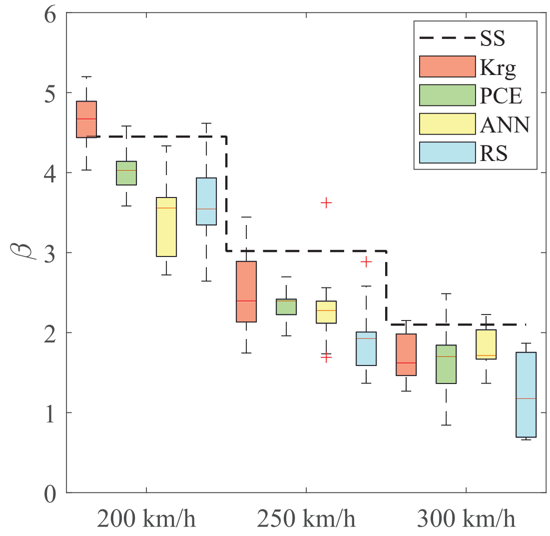
The results obtained for each bridge span, the train speed range, and type of the surrogate model are summarized as box plots in Figure 2. As mentioned earlier, no stopping criteria are considered when training the surrogate models; therefore, their relative performance should be assessed using these results without paying attention for possibility of over/under-estimation.

It appears that Kriging and PCE perform better than other meta-models in terms of both accuracy and stability. Similarly, polynomial response surfaces appear to have fewer advantages in this regard. Moreover, in some cases, a large discrepancy is observed in the estimated failure probabilities of ANN. This shortcoming could be due to the dependence of ANN on its architecture, which has not been optimized in all iterations of the training. As a general conclusion, it seems possible for Kriging, PCE, and ANN to accurately surrogate the computational model in dynamic problems, but when no further information is available, Kriging may be the first choice.

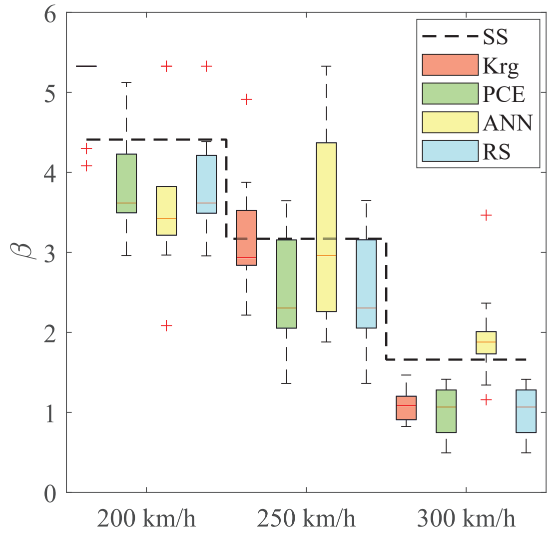
It should be noted that for bridges with shorter spans, there is usually no resonance phenomenon at the considered operating train speed ranges. Therefore, the considered train speed is likely to result in maximum responses. Similarly, for longer span bridges, the resonance phenomenon almost always occurs. This is not the case for medium span bridges, which makes it difficult for the adopted surrogate models to mimic the limit state function. Therefore, a much larger deviation in the estimated failure probabilities is observed for bridges with 20 m span length.

ACKNOWLEDGMENT

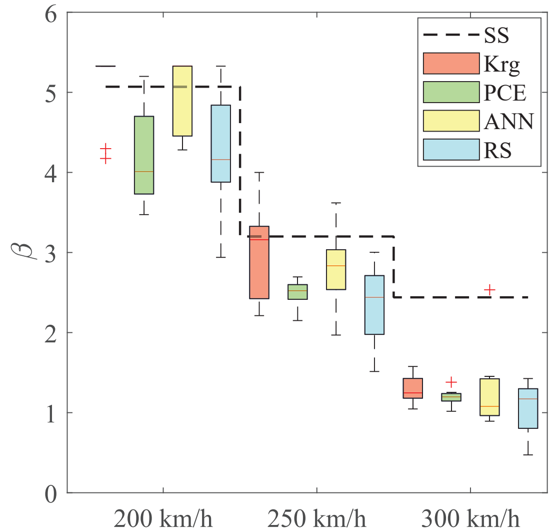
This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101012456.



(a) L=10m



(b) L=20m



(c) L=30m

Figure 2: Comparison between performance of the meta-models (SS stands for subset simulation).

6. REFERENCES

Allahvirdizadeh, R., Andersson, A., and Karoumi, R. (2020). "Reliability assessment of the dynamic behavior of high-speed railway bridges using first order reliability method." *Eurodyn 2020, 11th International Conference on Structural DynamicsAthens, 23-26 November, 2020.*, Vol. 2, 3438–3450.

Allahvirdizadeh, R., Andersson, A., and Karoumi, R. (2022). "Improved dynamic design method of high-speed railway bridges using surrogate-assisted reliability-based design optimization of dependent variables." *Unpublished manuscript*.

Allahvirdizadeh, R., Karoumi, R., and Andersson, A. (2021). "Surrogate-assisted versus subset simulation-based stochastic comparison between running safety and passenger comfort design criteria of high-speed railway bridges." *31st European Safety and Reliability Conference (ESREL 2021)*.

Au, S.-K. and Beck, J. L. (2001). "Estimation of small failure probabilities in high dimensions by subset simulation." *Probabilistic engineering mechanics*, 16(4), 263–277.

Blatman, G. and Sudret, B. (2011). "Adaptive sparse polynomial chaos expansion based on least angle regression." *Journal of computational Physics*, 230(6), 2345–2367.

Bucher, C. G. and Bourgund, U. (1990). "A fast and efficient response surface approach for structural reliability problems." *Structural safety*, 7(1), 57–66.

EN 1991-2 (2003). "Eurocode 1: Actions on structures - part 2: Traffic loads on bridges." *European Committee for Standardization*.

Enevoldsen, I. and Sørensen, J. D. (1994). "Reliability-based optimization in structural engineering." *Structural safety*, 15(3), 169–196.

ERRI D 214/RP 9 (1999). "Rail bridges for speeds > 200 km/h." *European Rail Research Institute*.

Frỳba, L. (2001). "A rough assessment of railway bridges for high speed trains." *Engineering Structures*, 23(5), 548–556.

Gaspar, B., Teixeira, A. P., and Soares, C. G. (2014). "Assessment of the efficiency of kriging surrogate models for structural reliability analysis." *Probabilistic Engineering Mechanics*, 37, 24–34.

Hurtado, J. E. (2004). *Structural reliability: statistical learning perspectives*, Vol. 17. Springer Science & Business Media.

Hurtado, J. E. and Alvarez, D. A. (2001). "Neural-network-based reliability analysis: a comparative study." *Computer methods in applied mechanics and engineering*, 191(1-2), 113–132.

Melchers, R. (1989). "Importance sampling in structural systems." *Structural safety*, 6(1), 3–10.

Melchers, R. E. and Beck, A. T. (2018). *Structural reliability analysis and prediction*. John wiley & sons.

Moustapha, M., Marelli, S., and Sudret, B. (2022). "Active learning for structural reliability: Survey, general framework and benchmark." *Structural Safety*, 96, 102174.

Papadrakakis, M., Papadopoulos, V., and Lagaros, N. D. (1996). "Structural reliability analyis of elastic-plastic structures using neural networks and monte carlo simulation." *Computer methods in applied mechanics and engineering*, 136(1-2), 145–163.

Rasmussen, C. E. and Williams, C. K. (2006). *Gaussian processes for machine learning*. MIT press Cambridge, MA.

Sudret, B. (2012). "Meta-models for structural reliability and uncertainty quantification." *arXiv preprint arXiv:1203.2062*.

Xiu, D. and Karniadakis, G. E. (2002). "The wiener–askey polynomial chaos for stochastic differential equations." *SIAM journal on scientific computing*, 24(2), 619–644.

Yau, J., Martínez-Rodrigo, M. D., and Doménech, A. (2019). "An equivalent additional damping approach to assess vehicle-bridge interaction for train-induced vibration of short-span railway bridges." *Engineering Structures*, 188, 469–479.

Zacher, M. and Baeßler, M. (2009). *Dynamic Behaviour of Ballast on Railway Bridges. In: Delgado, R., Calçada, R., Goicolea, J.M. and Gabaldón, F. (Ed.), Dynamics of High-Speed Railway Bridges*. London: Taylor & Francis.