



Simulation-based sensitivity analysis: Methods and software tools

Drahomír Novák^{1,*}, David Lehký¹ and Ondřej Slowik¹

¹Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology, Veverří 95, Brno, 60200, Czech Republic

*Corresponding author. Email address: novak.d@fce.vutbr.cz

Abstract

The topic of the paper is simulation-based sensitivity analysis with emphasize on the use of the small-sample Latin Hypercube Sampling simulation method. Three approaches are described in the paper: Spearman's rank-order correlation, covariance-based sensitivity analysis and input perturbation-based sensitivity analysis. Software tools are briefly described, especially SEAN software as an effective sensitivity analysis environment developed to simplify sensitivity analysis of a user-defined numerical model. An example application is presented.

Keywords: Sensitivity analysis; Statistical analysis; Latin Hypercube simulation.

1. Introduction

Sensitivity analysis (SA) is a crucial part of computational modelling and simulation-based structural reliability assessment. Therefore, it has received much attention in the literature in the past decade. SA is important to reduce a space of random variables for stochastic calculation, building the response surface, training neural networks, etc. Several interrogations are possible, and several SA methods have been developed, giving rise to a vast and growing literature. Overview of available methods is given in review papers, e.g. Novák et al. (1993), Kleijnen (2010), Borgonovo and Plischke (2016), Antucheviciene et al. (2015).

Versatile sensitivity indices for nonlinear non-monotonic problems are known to be based on Sobol' decomposition of functions (Sobol 1993, Sobol 2001). Sudret (2008) introduced generalized polynomial chaos expansions to build surrogate models that allow one to compute the Sobol' indices analytically.

Important topics are influence of statistical correlation in SA and artificial neural network-based sensitivity analysis (Pan et al., 2021) or usage of Cramér-von Mises distance (Novák and Novák, 2019). Kala (2021) proposed a new sensitivity measure based on the difference between superquantile and subquantile.

There are generally two types of sensitivity analyses. Local sensitivity analysis focuses on behavior of function around a point of interest (e.g. one-at-a-time and screening). Global sensitivity analysis investigates the whole design domain considering probability distribution of input random variables, e.g. Kala (2016), Kala and Valeš (2017). Global SA represents powerful tool for uncertainty quantification of mathematical model e.g. regression-based methods and analysis of variance (ANOVA). Local and global sensitivity analysis has different purpose, and the interpretation of their results is frequently inaccurate or even erroneous. This is because the user often uses just one "available" method and states global conclusions without a deeper



knowledge of the SA problem.

The common definition of an engineering problem involving uncertainty and randomness, which is to be numerically analysed is as follows. A random variable Z represents a random response of the studied engineering system (e.g. a structure). In statistical analyses, Z may represent the random response of a system. Random variable Z is a function of random variables $\mathbf{X} = X_1, X_2, \dots, X_{N_{var}}$:

$$Z = g(\mathbf{X}) \quad (1)$$

where the function $g(\mathbf{X})$, a computational model, is a function of a random vector \mathbf{X} . Random vector \mathbf{X} follows a joint probability distribution function (PDF) $f_{\mathbf{X}}(\mathbf{X})$. In general, its marginal variables can exhibit a statistical correlation too.

With respect to the treatment of uncertainties in stochastic computational mechanics, three main task categories can be generally distinguished:

- Statistical analysis: Approaches focused on the estimation of the statistical moments of Z such as means, variances, etc.;
- Sensitivity analysis: Approaches aimed at the quantification of the sensitivity of outputs (response Z) due to the variation of input basic variables \mathbf{X} ;
- Reliability analysis: Approaches aimed at the calculation of theoretical failure probability.

The aim of this paper is to present quite simple and easy-to-use techniques for sensitivity analysis based on a small sample Monte Carlo type simulation. Because they are easy to understand and use, they can compete with other advanced SA methods due to their simplicity. The methods presented here were implemented into SEAN software for straightforward and fast utilization together with a statistical simulation environment. Therefore, basic information on a small-sample statistical simulation and a related software tool is provided first.

2. Statistical simulation

2.1. Small-sample simulation

An effective method for statistical analysis of computationally demanding structural systems is small sample simulation technique of the Monte Carlo type. A special implementation of this method, called Latin Hypercube Sampling (LHS), seems to be the most effective. The technique can efficiently cover a multidimensional space of random variables with a small number of simulations (McKay et al., 1979; Stein, 1987).

The main feature of the LHS method is that the range of univariate random variables is divided into N_{sim} intervals (N_{sim} is number of simulations). The values from the intervals (random selection, the median or the

mean value) are then used in the simulation process. The selection of the intervals is performed in such a way that the range of the probability distribution function of each random variable is divided into intervals of equal probability $1/N_{sim}$. The samples are chosen directly from the distribution function based on an inverse transformation of the univariate distribution function (Figure 1). The representative parameters of variables are selected randomly, being based on random permutations of integers $k=1, 2, \dots, N_{sim}$. Every interval of each variable must be used only once during the simulation. A preferable LHS strategy is the approach suggested by Huntington and Lyrintzis (1998), where the representative value of each interval is the mean value (Figure 1):

$$x_{i,k} = \frac{\int_{y_{i,k-1}}^{y_{i,k}} x f_i(x) dx}{\int_{y_{i,k-1}}^{y_{i,k}} f_i(x) dx} = N_{sim} \int_{y_{i,k-1}}^{y_{i,k}} x f_i(x) dx \quad (2)$$

Here, f_i is the probability density function of variable X_i , and the integration limits are:

$$y_{i,k} = F_i^{-1}\left(\frac{k}{N_{sim}}\right), \quad k = 1, \dots, N_{sim} \quad (3)$$

The sample averages exactly equal the mean values of the variables, and the variances of the sample sets are much closer to the target values than in other selection schemes. A robust technique to impose statistical correlation based on the stochastic method of optimization called simulated annealing has been proposed by Vořechovský and Novák (2009).

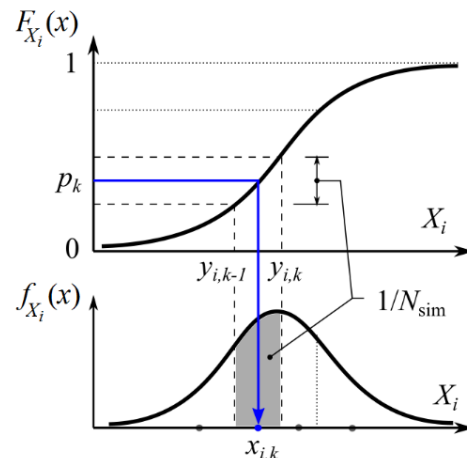


Figure 1. Samples as probabilistic means of intervals.

2.2. Software FReET

LHS simulation technique, described above, is essential methods implemented in software FReET (Novák et al., 2014). It is designed as a user-friendly tool for simulation of random variables according to their

probability distribution functions (vector \mathbf{X}) and to process numerically user-defined response/limit state function $g(\mathbf{X})$. Uncertainties propagation concept is illustrated in Figure 2 by “input and output” software windows.

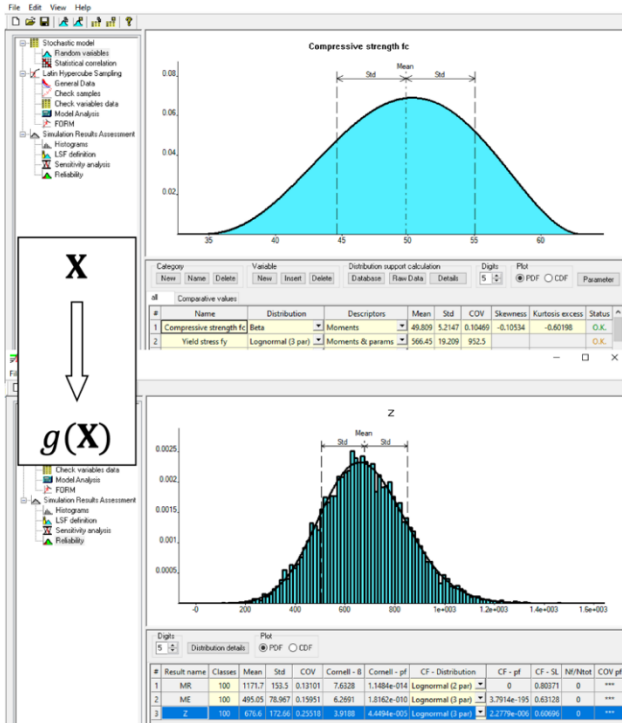


Figure 2. Uncertainties propagation concept in FReET: “Random variables” window - INPUT (above); “Reliability” window with empirical histogram - OUTPUT, (below)

The concept of program FReET is characterized by the aim of providing a very practical software tool that can be used and handled very easily – in a straightforward feasible way. State-of-the-art probabilistic algorithms are implemented in FReET to compute the probabilistic response and reliability. FReET is a modular computer system for performing probabilistic analysis developed mainly for computationally intensive deterministic modeling and the running of user-defined tasks/subroutines.

3. Sensitivity analysis

3.1. Non-parametric rank-order statistical correlation

The relative effect of each basic random variable on structural response can be measured using the partial correlation coefficient between each basic input variable and the response variable. With respect to the small-sample simulation techniques of the Monte Carlo type utilized for the reliability assessment of time-consuming nonlinear problems, the most straightforward and simplest approach uses non-

parametric rank-order statistical correlation (Iman and Conover, 1980). This method is based on assumption that the random variable which influences the response variable most considerably (either in a positive or negative sense) will have a higher correlation coefficient than the other variables. For a detailed discussion of rank-order statistical correlation see Vořechovský (2012). Non-parametric correlation is more robust than linear correlation and more resistant to defects in data. It is also independent of probability distribution. Because the model for the structural response is generally nonlinear, a non-parametric rank-order correlation is used by means of the Spearman correlation coefficient:

$$r_{s,i} = \frac{6 \sum_{j=1}^N (q_{ji} - p_j)^2}{N^3 - N}, \quad r_{s,i} \in (-1,1) \quad (4)$$

where q_{ji} is the rank of a representative value of the random variable X_i in an ordered sample of N simulated values used in the j th simulation and p_j is the rank of the response variable obtained in the same simulation.

Although the crude Monte Carlo simulation method can be used to prepare random samples, it is recommended to use an appropriate sampling scheme, such as the stratified Latin hypercube sampling described above. This method utilizes random permutations of the number of layers of the distribution function of the basic random variables to obtain representative values for the simulation. When using this method, the ranks q_{ji} in Equation (4) are directly equivalent to the permutations used in sampling.

Non-parametric rank-order correlation can be depicted using parallel coordinates (Wegman, 1990). Instead of an orthogonal representation of (q_{ji}, p_j) , they are drawn in parallel and pairs of points are joined by lines. A strong positive influence (high correlation coefficient) results in parallel lines between the input variable and the response variable, while a strong negative influence results in a bundle of intersecting lines.

3.2. Covariance-based sensitivity analysis

SA in terms of coefficient of variation (Novák et al., 1993) is another simulation-based sensitivity method widely utilized for the optimum selection of dominant random variables. In this approach, the ratio between the partial coefficient of variation of resistance and the coefficient of variation of a selected basic variable is calculated for a case in which the selected random variable is the only one treated as random in the simulation process.

When using a Monte Carlo type simulation, a simulated set of realizations of structural response variable R_j ($j = 1, 2, \dots, N_{sim}$), where N_{sim} is the number of simulations, is statistically evaluated and its coefficient

of variation COV_R can be determined. The number of variables N_{var} , representing e.g. material properties or load, can be defined as random in the simulation process. Let us designate the partial coefficient of variation ($i = 1, 2, \dots, N_{var}$) for a case in which the variable X_i is the only one treated as random and is defined using its mean value and coefficient of variation. The other basic variables are kept as deterministic constants at their mean values. The partial sensitivity factor α_i^{COV} for the basic random variable X_i is then defined as:

$$\alpha_i^{COV} = \frac{COV_{R_i}}{COV_{X_i}} \quad (5)$$

This procedure requires additional computational effort, since the set of values of structural response variable R_i and its coefficient of variation need to be evaluated with additional N_{var} sets of simulations. Therefore, in cases when the evaluation of an original model is time-consuming, e.g. a nonlinear finite element method (FEM) model is employed, a suitable type of surrogate model is needed in order to reduce the computational time to an acceptable level.

Sensitivity factor α_i^{COV} in Equation (5) expresses the relative influence of individual input random variables on the variability of structural response. If N_{var} basic variables are considered as random ones, the coefficient of variation of the response variable COV_R can be calculated using an approximate formula in the form:

$$COV_R \approx \sqrt{\sum_{i=1}^{N_{var}} (\alpha_i^{COV} COV_{X_i})^2} \quad (6)$$

It can be seen from Equation (6) that the actual influence (not the relative one) of random variable X_i is represented by the $COV_{R_i}^2$ value. Such sensitivity may be easily depicted using a pie chart.

3.3. Input perturbation-based sensitivity analysis

The input perturbation algorithm (Scardi and Harding, 1999), is the simplest way to interrogate a model. It produces sensitivity analysis results based on the assessment of the effect of input perturbation in each input on the neural network output (Gevrey et al., 2003). The proper adjustment of the values of each explicative variable while keeping all the others unchanged allows the effect of the output variables corresponding to each perturbation in the input variable to be recorded. The result of sensitivity analysis is yielded by ranking the effect on output induced by the same manner of perturbation in every input variable. The input variable, whose perturbation influences the output most, possesses the highest sensitivity or importance.

In principle, the output increases or decreases as the selected input variable increases. The changes to the input variable take the form of $X_i^* = X_i + \delta$, where X_i is the selected input variable, and δ is the perturbation value. The input variables can be ranked according to the increasing magnitude of the output due to each input variable change. In other words, the result is a sensitivity analysis outcome.

It is common to choose a perturbation value δ as an increment or decrement in the percentage of the input variable. However, this approach fails to take into account the specific variance of random variables. Therefore, an alternative approach can be recommended where δ is represented by a standard deviation σ , e.g. $\pm 3\sigma$. This alternative approach directly reflects the variability of every input random variable (Downing et al., 1985).

4. Software SEAN

Above-described sensitivity analysis methods were implemented into the SEAN software. SEAN is a software environment that has been developed to simplify sensitivity analyses of user-defined numerical models. It uses FReET software as a simulation processor, thanks to which it significantly extends possibilities of model definition and subsequent sensitivity analysis in a user-friendly way. The software architecture is built on visual scripting language (Liu et al., 2007), which allows the user to easily create user-defined scripts. So, users are not required to have any programming experience. Visual scripting languages might be understood as programming languages that help to build program operating multiple other programs encapsulated within separated nodes. The environment then handles mutual communication and data transfer among separate applications. The structure of the environment allows adding and development of the single nodes (programs) without impact on the inner architecture of the environment. It is, therefore, suitable for gradual development, and it allows us to implement even existing applications relatively quickly into the environment. SEAN was developed based on brother visual scripting library dedicated for reliability-based optimization of the general tasks (Slowik and Novák, 2019). The further versions of SEAN will likely become part of this brother software solution. The first version, however, works as the stand-alone software meant to be utilized for sensitivity analysis of general models.

GUI of SEAN is based on standard QMainWindow class in order to ensure basic dialog functionality necessary for possible further implementation of SEAN within another software solution of dialog window. The example of appearance of the script created in SEAN environment is captured in the Figure 3, documenting node architecture of solution.

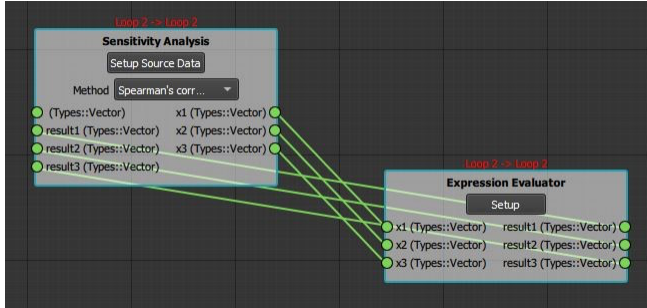


Figure 3. The script created in SEAN environment.

5. Example of application

We present an application of sensitivity analysis for shear ultimate capacity of LDE7 TT roof girders, that were produced by Franz Oberndorfer GmbH & Co KG, details for deterministic modelling can be found in Strauss et al. (2017) and stochastic modelling has been performed recently by Slowik et al. (2020). Computational modelling was extremely time consuming as regular hexagonal finite element (FEM) mesh consisted of 61784 finite elements (Figure 4) and one simulation took app. 12 hours. Utilization of small-sample simulation technique and related simulation-based sensitivity analysis was therefore necessary.

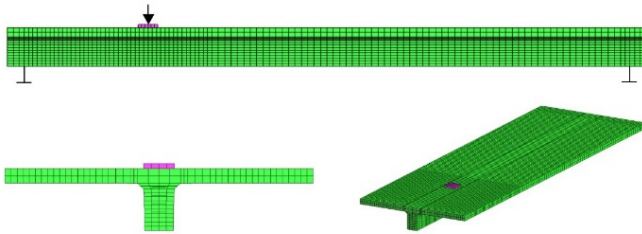


Figure 4. Computational mode – FEM mesh.

The utilized complete stochastic model is summarized in Table 1, which displays random quantities and their stochastic parameters. The employed variables are E – Young's modulus (E – concrete, E_s – steel reinforcement, E_t – tendons), f_t – tensile strength, f_c – compressive strength, G_f – fracture energy, ρ – density of the concrete mixture, f_{ys} – yield strength of steel reinforcement, f_{yt} – yield strength of prestress tendons, $I.L.$ – model uncertainty for immediate losses of prestress, $L.T.L.$ – model uncertainty for long term losses of prestress and P – initial prestressing force. Note that COV stands for coefficient of variation and PDF is probability density function. Table 2 displays utilized statistical correlation matrix for concrete.

Table 1. Utilized stochastic model.

Parameter	Mean	COV in %	PDF
E (GPa)	34.8	10.6	Lognormal (2 par)
f_t (MPa)	3.9	10.6	Lognormal (2 par)
f_c (MPa)	77.0	6.4	Lognormal (2 par)
G_f (J/m ²)	219.8	12.8	Lognormal (2 par)
ρ (ktons/m ³)	0.0023	4.0	Normal
E_s (GPa)	200.0	2.0	Normal
f_{ys} (MPa)	610.0	4.0	Normal
E_t (GPa)	195.0	2.0	Normal
f_{yt} (MPa)	1387.9	2.5	Normal
P (MN)	0.0835	6.0	Normal
$I.L.$ (-)	1	10.0	Lognormal (2 par)
$L.T.L.$ (-)	1	10.0	Lognormal (2 par)

Table 2. Utilized correlation matrix for concrete.

	E	f_t	f_c	G_f	ρ
E	1	0.5	0.8	0.5	0
f_t	0.5	1	0.7	0.8	0
f_c	0.8	0.7	1	0.6	0
G_f	0.5	0.8	0.6	1	0
ρ	0	0	0	0	1

Let us emphasize here that sensitivity analysis may be profoundly affected by the correlation among input random variables. The correlation among variables might be understood as a stochastic description of the complex natural relations which are not directly involved within a numerical model. Sensitivity analysis of the correlated model captures the cumulative effects of interdependent parameters. Such analysis corresponds to real-life behavior of the modelled entity, but it does not provide objective information of the influence of numerical model parameters to the observed output. Thus, it is necessary to analyze uncorrelated space as well to identify the actual role of input random variables.

The obtained results for correlated space are depicted in Figure 5 and for uncorrelated space in Figure 6 (for the most dominating variables only). In both cases, the essential material characteristics of concrete are apparent. It is possible to see the significant difference between correlated and uncorrelated space. Generally, in correlated space, there is a high correlation among concrete material characteristics, and thus their influence is together dominant in comparison to other variables. For correct interpretation of such results, correlated variables must be assumed as a group of variables. Note that, information about sensitivity in correlated space is valid only for this one stochastic model, including the given dependency structure. The mutual influence of concrete material parameters observed in model with

correlation was verified by analysis of model without correlation showing that introduced uncertainty has a significant influence on immediate losses of prestressing. According to expectation, the compressive strength of concrete does not influence the model's performance (for given limit state). The significant influence of Young's modulus of concrete might be explained by the fact that it is involved within the utilized model for evaluation of prestressing losses. The assumption of uncorrelated material characteristics is not realistic, but such information may be crucial for the reduction of the stochastic model.

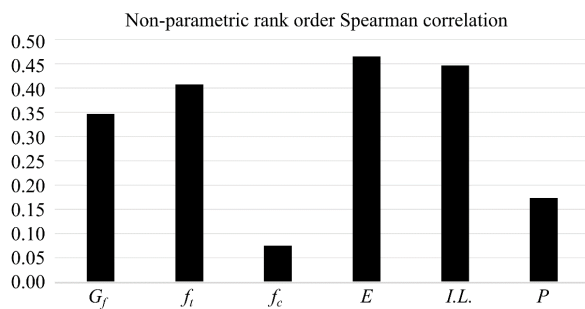


Figure 5. Spearman rank-order correlation between input random variables and the ultimate shear strength of precast prestressed concrete roof girders for correlated space.

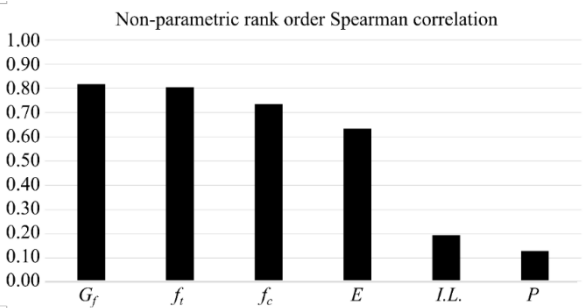


Figure 6. Spearman rank-order correlation between input random variables and the ultimate shear strength of precast prestressed roof girders for uncorrelated space.

6. Conclusions

The paper presents three simulation-based sensitivity analysis methods which can be easily used to real-world structural engineering problems. These are often computational demanding taking into account the nonlinearity of the materials. Efficiency is emphasized using the stratified LHS simulation method. Information is provided on the developed software tools that will allow to analyze the influence of input parameters on the response of the analyzed structure. The presented techniques were successfully applied to several engineering tasks, one of them is briefly shown in the paper. The presented software tools may be routinely applied to any problem in the advanced design/assessment of engineering tasks.

Funding

The authors acknowledge Czech Science Foundation project “Probability oriented global sensitivity measures of structural reliability”, No. 20-01734S, principal investigator prof. Z. Kala. The software development was supported by the Chinese (Jiangsu)–Czech Bilateral Co-funding R&D Project DELTA “Advanced system for monitoring, diagnosis and assessment of large-scale concrete infrastructures” (No. SBZ2018000220 – China side, No. TF06000016 – Czech side), principal investigator prof. D. Novák.

References

- Antucheviciene, J., Kala, Z., Marzouk, M. and Vaidogas, E.R. (2015). Solving civil engineering problems by means of fuzzy and stochastic MCDM methods: Current state and future research. *Mathematical Problems in Engineering*, 2015 (Article ID 362579), 1–16.
- Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, 248:869–887.
- Downing, D. J., Gardner, R. H. and Hoffman, F. O. (1985). An examination of response-surface methodologies for uncertainty analysis in assessment models. *Technometrics*, 27(2):151–163.
- Gevrey, M., Dimopoulos, I. and Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological modelling*, 160(3):249–264.
- Huntington, D.E. and Lyrintzis C.S. (1998). Improvements to and limitations of Latin Hypercube Sampling”, *Probabilistic Engineering Mechanics*; 13 (4):245–253.
- Iman, R.L. and Conover, W.J. (1980). Small sample sensitivity analysis techniques for computer models with an application to risk assessment. *Communications in Statistics – Theory and Methods*, 9(17):1749–1842.
- Kala, Z. (2016). Higher-order approximations methods for global sensitivity analysis of nonlinear model outputs, *International Journal of Mathematics and Computers in Simulation*, 10:260–264.
- Kala, Z. (2021). Global Sensitivity Analysis of Quantiles: New Importance Measure Based on Superquantiles and Subquantiles. *Symmetry* 13(2):263.
- Kala, Z., Valeš, J. (2017). Global sensitivity analysis of lateral-torsional buckling resistance based on finite element simulations, *Engineering Structures*, 134:37–47.
- Kleijnen, J.P.C (2010). Sensitivity Analysis of Simulation Models: An Overview. *Procedia Social and Behavioral Sciences*, 2:7585–7586.
- Liu, N., Hosking, J. and Grundy, J. (2007). A visual

- language and environment for specifying user interface event handling in design tools. *Conferences in Research and Practice in Information Technology*.
- McKay, M.D., Conover, W.J. and Beckman, R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code”, *Technometrics*; 21:239–245.
- Novák, L. and Novák, D. (2019). On the possibility of utilizing Wiener-Hermite polynomial chaos expansion for global sensitivity analysis based on Cramér-von Mises Distance. In 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE), 1–9.
- Novák, D., Teplý, B. and Shiraiishi, N. (1993). Sensitivity analysis of structures: A review. In *Proc. Of Int. Conference CIVIL COMP'93*: Edinburgh, Scotland, 201–207.
- Novák, D., Vořechovský, M., and Teplý, B. (2014). FReET: Software for the statistical and reliability analysis of engineering problems and FReET-D: Degradation module, *Advances in Engineering Software*, 72:179–192.
- Pan, L., Novák, L., Novák, D., Lehký, D. and Cao, M. (2021). Neural network ensemble-based sensitivity analysis in structural engineering: Comparison of selected methods and the influence of statistical correlation. *Computers & Structures*, 242(1):1–19.
- Scardi, M. and Harding Jr, L. W. (1999). Developing an empirical model of phytoplankton primary production: a neural network case study. *Ecological modelling*, 120(2–3):213–223.
- Slowik, O. and Novák, D. (2019). Node based software concept for general purpose reliability-based optimization tool. In *Proc. of 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*.
- Slowik, O., Novák, D, Novák, L and Strauss, A. (2020). Stochastic modelling and assessment of long-span precast prestressed concrete elements failing in shear. *Engineering Structures*, 2020, 111500.
- Sobol, I.M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiment*, 1, pp. 407–414. Translated from Russian: I. M. Sobol'. 1990. Sensitivity estimates for nonlinear mathematical models, *Matematicheskoe Modelirovanie* 2:112–118.
- Sobol, I.M (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55:271–280.
- Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, 29(2):143–151.
- Strauss, A., Krug, B., Slowik, O. and Novák D. (2017). Combined shear and flexure performance of prestressing concrete T-shaped beams: Experiment and deterministic modeling. *Structural Concrete*, 1–20.
- Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93:964–979.
- Vořechovský, M. and Novák, D. (2009). Correlation control in small sample Monte Carlo type simulations I: A Simulated Annealing approach, *Probabilistic Engineering Mechanics*, 24(3):452–462.
- Vořechovský, M. (2012). Correlation control in small sample Monte Carlo type simulations II: Analysis of estimation formulas, random correlation and perfect uncorrelatedness. *Probabilistic Engineering Mechanics*, 29:105–120.
- Wegman, E.J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85:664–675.