Árpád RÓZSÁS[1], Miroslav SÝKORA[2]

NEGLECT OF PARAMETER ESTIMATION UNCERTAINTY CAN SIGNIFICANTLY
OVERESTIMATE STRUCTURAL RELIABILITY

Abstract

Parameter estimation uncertainty is often neglected in reliability studies, i.e. point estimates of distribution parameters are used for representative fractiles, and in probabilistic models. A numerical example examines the effect of this uncertainty on structural reliability using Bayesian statistics. The study reveals that the neglect of parameter estimation uncertainty might lead to an order of magnitude underestimation of failure probability.

Keywords

Structural reliability, Bayesian inference, parameter estimation uncertainty, posterior predictive distribution.

1 INTRODUCTION

1.1 Problem statement

In probabilistic engineering analysis the physical properties are typically modelled as random variables. Even if selected variables are derived from advanced models, the basic variables are commonly represented as random variables to express their uncertainty in the chosen modelling space. The parameters of the related probabilistic models are estimated from observations and experimental data. The scarcity of this information inevitably leads to uncertainty regarding the parameter estimates. This uncertainty is often neglected in probabilistic analyses, e.g. in reliability case studies [10; 18], and in deriving representative fractiles [4; 21]. A Venezuelan rainfall study [4] highlights how this neglect can cost thousands of lives and cause extensive damage as a component of the "naïve use of extreme value techniques". Additionally, this uncertainty appears to be neglected or inadequately addressed in some standards and standardization processes as well. Therefore, the aim of this paper is to analyse the effect of parameter estimation uncertainty on the failure probability of structures. For simplicity this study focuses solely on this effect, other uncertainties are not considered:

- Model uncertainty, e.g. how accurately a resistance formulae describes experiments;
- Probability model uncertainty, e.g. selection of an appropriate distribution type;
- Measurement uncertainty, e.g. resolution of the measuring device;
- Human error, e.g. negligence, error in calculations.

---

[1] Árpád Rózsás, M.Sc., Department of Structural Engineering, Faculty of Civil Engineering, Budapest University of Technology and Economics, 3-9. Műegyetem rkp., Kmf. 85, H-1111 Budapest, Hungary, e-mail: rozsas.arpad@epito.bme.hu.

[2] Doc. Ing. Miroslav Sýkora, Ph.D., Department of Structural Reliability, Klokner Institute, Czech Technical University in Prague, Solinova 7, 166 08, Prague-Dejvice, Czech Republic, phone: (+420) 224 353 850, e-mail: miroslav.sykora@cvut.cz.

1.2 Parameter uncertainty in standards

EN 1990 for basis of structural design contains an informative annex related to parameter estimation uncertainty; however, it is restricted to normally or lognormally distributed resistance variables, and general principles are not included. It focuses on estimating characteristic and design values, and does not mention other than resistance random variables nor how this uncertainty should be taken into account in reliability calculations. The following background documents for Eurocodes on variable actions suggest that parameter estimation uncertainty is neglected in drafting the code:

- The report on the joint European research that produced harmonized ground snow map for member countries uses method of moments point estimates, thus disregards parameter estimation uncertainty [20].
- The same approach is adopted for thermal actions [19], the method is also included in the annex of EN 1991-1-5 for thermal actions on structures.

In Eurocodes there is no reference to studies regarding the recommended partial factors; therefore, it is unclear whether or not they are accounting for parameter uncertainty. However, it seems unlikely considering the treatment of snow and thermal actions, and lack of regulations in EN 1990.

A background document [7] of American ASCE 7-10 [2] standard states that the partial factors (load and resistance factors) take into account the parameter estimation uncertainty by increasing the variance of random variables. The applied approximate technique seems to be based on the posterior predictive distribution of a normally distributed random variable with known variance and vague priors, but again no general rationale is provided, hence it is of limited applicability.

To our knowledge the effect of neglecting parameter estimation uncertainty has not yet been studied and the performance of the mentioned approximate techniques has not yet been investigated.

1.3 Adopted conceptual framework to explore parameter uncertainty

A simplified numerical example is selected and the probabilistic models, i.e. distribution function, for resistance and effects are assumed to be known. From each of these a random sample is generated and probabilistic models are inferred with the only assumptions that the distribution type is known, and the parameters have uniform prior on wide support. This latter condition is adopted for the fair comparison between the Bayesian and maximum likelihood based approaches. In practical applications stronger prior information is almost always available and its incorporation is strongly recommended. The obtained probabilistic models with and without parameter estimation uncertainty are then applied to calculate and critically compare the failure probabilities. Bayesian approach is promoted here as it can naturally incorporate parameter estimation uncertainties using the posterior predictive distribution. The posterior mean is selected as a point estimate that neglects parameter uncertainty.

Additionally, a more widely applied maximum likelihood based approach is used as well where the parameter estimation uncertainty is approximately taken into account by using 75% level confidence intervals. The effect of sample size and sampling variability is then analysed in more details.

2 STATISTICAL INFERENCE

2.1 Bayesian inference

Bayesian statistics views parameters of a probabilistic model as random variables and uses Bayes' rule for statistical inference. For continuous random variables it is given as:

$$f(\mathbf{\theta}|\mathbf{x}) = \frac{f(\mathbf{\theta}) \cdot f(\mathbf{x}|\mathbf{\theta})}{\int_{\Theta} f(\mathbf{\theta}) \cdot f(\mathbf{x}|\mathbf{\theta}) \cdot d\mathbf{\theta}} \tag{1}$$

where:

$f(.)$ – probability density function;

$\mathbf{x}$ – vector of observations;

$\boldsymbol{\theta}$ – vector of model parameters;

$\boldsymbol{\Theta}$ – space of model parameters.

The mean of the parameter's posterior distribution is selected as point estimate, the corresponding probabilistic model does not take into account parameter uncertainty.

The posterior predictive distribution is obtained by integrating over the posterior distribution of the parameters [1]:

$$f(\tilde{x}|\mathbf{x}) = \int_{\Theta} f(\boldsymbol{\theta} \mid \mathbf{x}) \cdot f(\tilde{x}|\boldsymbol{\theta}) \cdot d\boldsymbol{\theta} \qquad (2)$$

The parameter estimation uncertainty is thus incorporated through this averaging. Vague priors are applied to have a more or less sound base for comparison with the maximum likelihood based approach. This means uniform distribution on sufficiently wide interval.

2.2 Maximum likelihood method

For comparison, maximum likelihood parameter estimation is applied as well. The distribution function corresponding to maximum likelihood parameter estimates is given as:

$$F_{\mathrm{ML}}(x) = F(x, \hat{\boldsymbol{\theta}}) \qquad (3)$$

where:

$F(.)$ – cumulative probability distribution function;

$\hat{\boldsymbol{\theta}}$ – vector of parameters' maximum likelihood estimates.

The parameter uncertainty is taken into account using confidence intervals. This approach is included in the EN 1990, adopted for S-N curves in EN 1993-1-9, and used for representative wind loads in South Africa as well [13]. It is an approximate technique that is used for particular problems, but no general method or rationale is provided in the referred documents. Herein, it is extended to cumulative distribution functions using the delta method [3]. The variance of a derived parameter can be estimated from the variance-covariance matrix of the parameters:

$$\mathrm{Var}\left(F_{\mathrm{ML}}(x)\right) = \nabla F_{\mathrm{ML}}(x) \cdot \mathbf{I}_{\mathrm{o}}^{-1} \cdot \nabla F_{\mathrm{ML}}(x)^{\mathrm{T}} \qquad (4)$$

where:

$\mathbf{I}_{\mathrm{o}}$ – observed Fisher information matrix.

Utilizing the asymptotic normality property of the maximum likelihood estimator the 75% level one-sided confidence value can be calculated as follows:

$$F_{75\mathrm{CI}}(x) = F_{\mathrm{ML}}(x) \pm \Phi^{-1}(0.75) \cdot \sqrt{\mathrm{Var}\left(F_{\mathrm{ML}}(x)\right)} \qquad (5)$$

$\Phi(.)$ standard normal cumulative distribution function.

The sign of the second term is determined by the sign of the sensitivity factor of the random variable to get a conservative estimate. If the approximated value of the distribution function falls outside the [0,1] it is constrained to the closer bound.

3 NUMERICAL EXAMPLE

3.1 Mechanical and probabilistic models

A simple limit state function is selected to demonstrate the effect of parameter estimation uncertainty on reliability. It can be viewed as the cross-section level limit state function of a structure subjected to 50-year maxima of a variable action:

$$g = R - (\mu \cdot Q_{50} + G) \qquad (6)$$

where:

g – limit state function;

R – resistance;

μ – conversion factor, e.g. shape factor, pressure coefficient;

Q_{50} – 50-year maxima of a variable action;

G – permanent action.

For simplicity it is assumed that these actions are directly convertible to effects, e.g. internal forces, by a factor of 1.0, thus actions and their effects are interchangeable here. The connection between 1 year and 50-year maxima distributions is established by assuming statistical independence of annual maxima:

$$F(q_{50}) = F^{50}(q_1) \qquad (7)$$

The probabilistic models are summarized in Tab. 1. With the exception of variable action the sample sizes are selected to reflect information from other sources as well since these are not site-specific, these numbers express our experience, engineering judgement. Generalised extreme value distribution is adopted for the variable action instead of the in Europe commonly applied Gumbel since it was shown to be non-conservative and to have deceptively narrow confidence interval for modelling extremes in some cases [4; 16; 17; 21]. For annual maxima of variable action coefficient of skewness of 1.6 is considered. This value is obtained from meteorological data of snow load on the ground for Hungarian lowlands and is deemed representative for lowlands in Central Europe. The two-parameter lognormal and normal distributions are parametrized in a standard way, the generalised extreme value distribution is defined by shape, scale and location parameters [3].

The load-ratio χ is defined as:

$$\chi = \frac{\mu \cdot Q_{50}}{G + \mu \cdot Q_{50}} \qquad (8)$$

For the selected problem the load-ratio is 0.73 using mean values, and 0.84 using characteristic values, i.e. mean for permanent action and 0.98 fractile for variable action. These ratios correspond to lightweight steel structures.

Tab. 1: Properties of the generating probabilistic models, and selected sample sizes

	Distribu-tion	Mean	Coeff. of var.	Sample size	Reference
Resistance, R	LN	200	0.15	20	[11]
Conversion factor, μ	LN	0.80	0.17	20	[8]
Variable action (annual), Q_1	GEV	20	0.70	50	[20]
Permanent action, G	N	20	0.10	50	[12]

N – normal; LN – two-parameter lognormal; GEV – generalised extreme value.

3.1 Numerical analysis, implementation

The Bayesian inference is completed using numerical integration. Uniform, mutually independent priors are applied for all parameters, the support of a prior function is determined using the parameter's maximum likelihood estimate and its approximate standard deviation (based on the observed Fisher information matrix). The numerical integration is validated using Markov Chain Monte Carlo simulation where besides the uniform priors unbounded normal distributions with very large standard deviation are applied as well. If a parameter is bounded by definition then truncated

normal distribution is applied. The calculations confirmed the validity of the numerical integration that is applied in further analysis because even for three parameters it is more efficient than Markov Chain Monte Carlo simulation, especially because the tails of the distributions are needed in the reliability analysis.

First order reliability method (FORM) [14] is used to calculate the failure probability, the results are verified by importance sampling simulations. All presented results are corresponding to FORM analyses. When non-parametric distribution functions are applied multiple discretizations are tested to verify the convergence of calculations.

4 ANALYSIS RESULTS

Initially a single random sample is generated from each random variable using sample sizes given in Tab. 1. These are hereinafter referred to as "selected realizations".

4.1 Statistical analysis

The results of the statistical inferences are illustrated in Fig. 1. The observations and fitted models for each random variable are plotted in a transformed space, e.g. normal space refers to a space where the normally distributed random variables form a straight line. This presentation has the advantage that (i) the models are easier to visually compare; (ii) deviation from the particular distribution type is clear; (iii) the crucial tail regions are enlarged by the logarithmic like scale of the horizontal axis. In the Gumbel space convex curves correspond to Fréchet while concave to Weibull distribution.
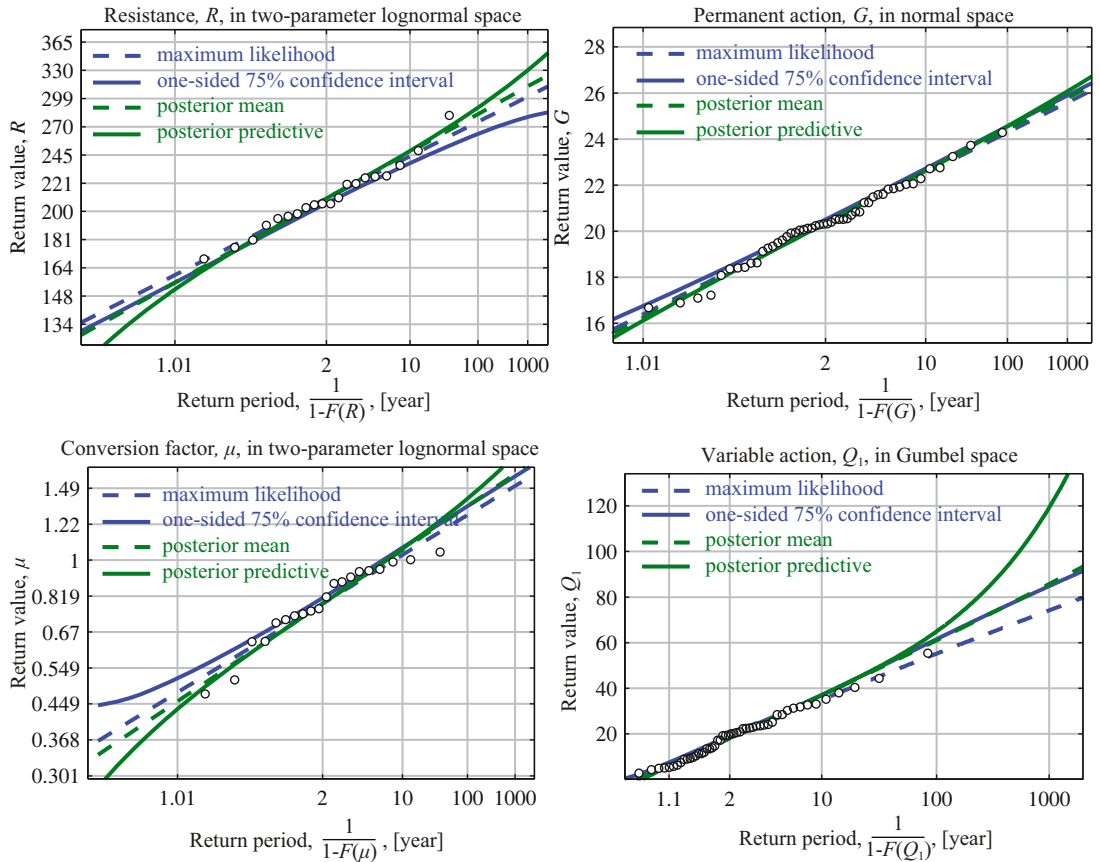

Fig. 1: Observations and fitted distributions in matching space. Maximum likelihood (dashed blue), one-sided 75% confidence interval (solid blue), posterior mean (dashed green) and posterior predictive (solid green) distributions

For all random variables the maximum likelihood model with 75% confidence interval is close to the Bayesian posterior mean, while the maximum likelihood point estimate model runs consistently under the Bayesian point estimate. The distance between the Bayesian point estimate and posterior predictive models is increasing as moving away from the mean. The difference between the maximum likelihood model with 75% confidence interval and Bayesian posterior predictive is salient in these regions especially for the variable action.

4.2 Reliability analysis

Reliability analyses are conducted using all four probabilistic model sets determined in Section 4.1. These correspond to the selected realizations, the results are summarized in Tab. 2.

Tab. 2: Summary of reliability indices and failure probabilities using maximum likelihood based and Bayesian probabilistic models

	Maximum likelihood	Maximum likelihood with 75% confidence interval	Bayesian posterior mean	Bayesian posterior predictive
Reliability index, β	4.56	3.94	3.80	2.43
Failure prob., P_f	$2.2 \cdot 10^{-6}$	$4.1 \cdot 10^{-5}$	$7.1 \cdot 10^{-5}$	$7.6 \cdot 10^{-3}$

Comparison of the Bayesian models shows that the failure probability is increased by two order of magnitude by incorporating the parameter estimation uncertainty. The dominant component is the variable action (Fig. 2). Moreover, using Bayesian posterior predictive distributions for all but the variable action the reliability index is 3.70, this shows that the main source of difference is the parameter uncertainty in the variable action model. Fig. 2 also illustrates that the incorporation of parameter estimation uncertainty can substantially change the sensitivity factors.

The ratio between failure probabilities of maximum likelihood model with and without 75% confidence interval is about 20. The Bayesian point estimate yields to about 30 times larger failure probability than the maximum likelihood estimate. This highlights the significance of the selected statistical approach.
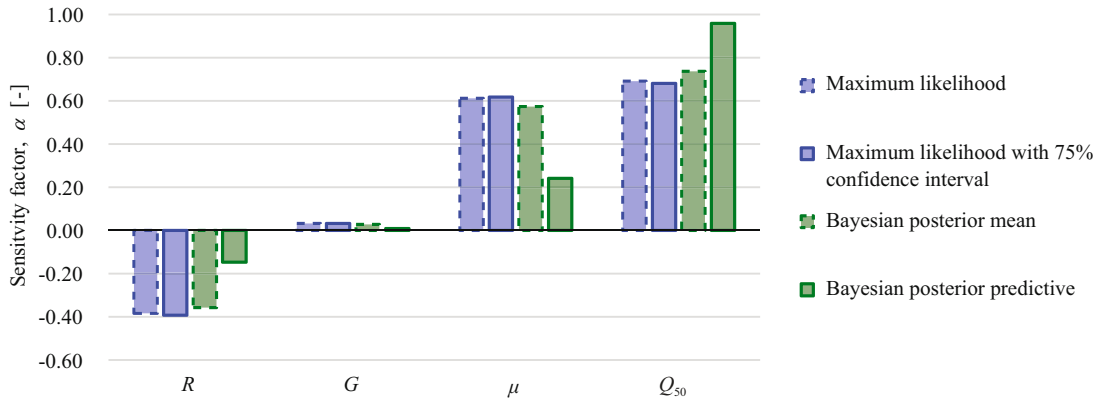

Fig. 2: Sensitivity factors for probabilistic models inferred using maximum likelihood (blue) and Bayesian techniques (green)

4.3 Effect of sample size and sampling variability

The preceding analyses are related to a particular set of observations from the underlying distributions given in Tab. 1. In this section this is extended in two directions: first, the sample size of the most important variable (Q) is varied between 25 and 150, this covers the practically available annual maxima for most of the actions. Second, to assess the effect of sampling variability the model fittings and reliability analyses are repeated for multiple samples from the underlying distributions.

600 simulations are used for each sample size, and only the Bayesian models with and without parameter estimation uncertainty are considered. The results of simulation study are illustrated in Fig. 3. The transparent bands encompass 90% of the simulations based on equal tailed fractiles. The dotted lines represent the selected realizations extended to multiple sample sizes. The extension is done by keeping the old realizations, e.g. in moving from 50 to 75 only 25 new realizations are added. The dashed line shows the reliability index obtained using models without parameter estimation uncertainty.
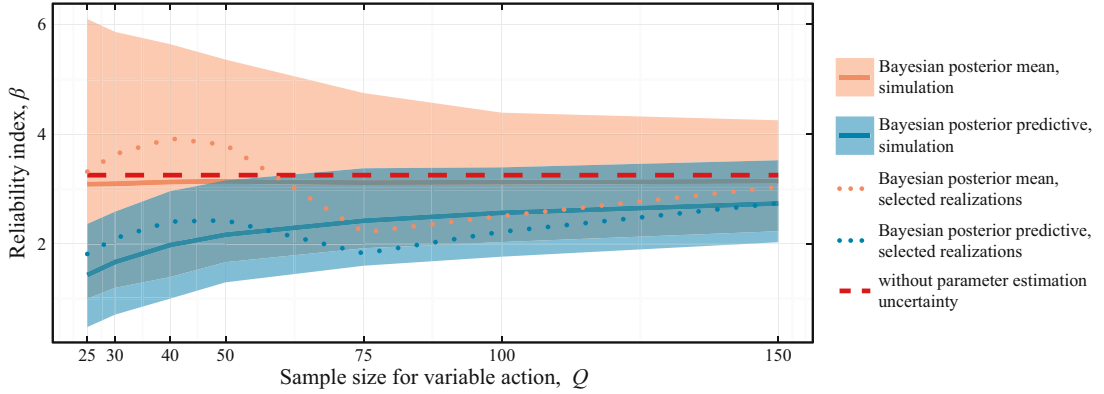

Fig. 3: Simulation based comparison of posterior mean (red) and posterior predictive (blue) probabilistic models in respect of reliability index and sample size of the variable action (Q)

As expected, the posterior mean results are fluctuating symmetrically around the reliability index without parameter uncertainty while the posterior predictive model is predicting consistently lower reliability indices. All the models are converging to the reliability index without parameter uncertainty with increasing sample size. The posterior mean estimate based analysis yields to a slightly biased estimate of the reliability index without parameter estimation uncertainty. The dotted lines show the considerable variability of reliability index as new observations become available.

The ratio of the posterior predictive and posterior mean failure probabilities for sample size of 50 for the variable action are calculated for all 600 simulations, and its empirical cumulative distribution function is given in Fig. 4. Greenwood's formula [9] is used to approximate its variance and to construct 90% level confidence intervals. The plot shows that for about 50% of the cases the ratio is larger than 10 and for 20% it is larger than 100.
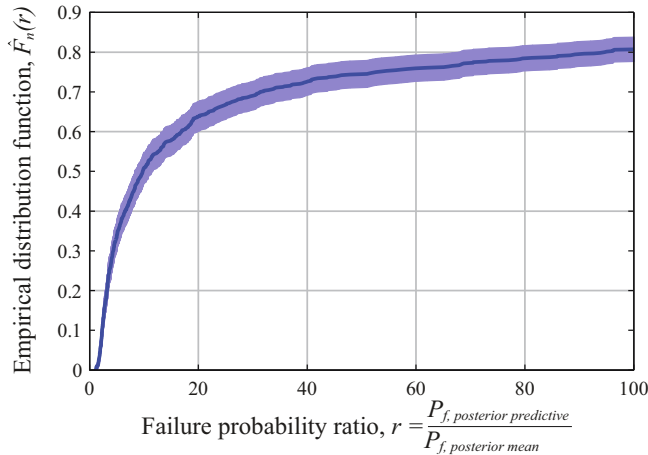

Fig. 4: Empirical cumulative distribution function (solid blue) and 90% confidence interval (blue band) of the failure probability ratio for sample size 50 (Q)

5 DISCUSSION

The numerical results imply that the parameter estimation uncertainty has significant effect on the failure probability as it can increase it by several orders of magnitude. In some sense it is the matter of consensus whether this uncertainty is taken into account in reliability calculations since the calculated failure probability is just a decision analysis tool and cannot be directly compared to actual failure frequencies. This question is similar to the problem of the selection of distribution type. The limited number of available observations does not make it possible to unambiguously select the type which typically has larger effect on failure probability than that of the parameter uncertainty [16]. The pragmatic consensus among reliability experts seems to be that the distribution type should be fixed for each type of random variables. This approach could be adopted in regard of parameter estimation uncertainty as well, e.g. it could be agreed that point estimates should be applied. However the following arguments are supporting its incorporation into failure probability:

- It is often misleading to use point estimates to identify probabilistic models and to apply these in reliability analysis. Although the point estimates provide the best-fit to the observations in some sense, e.g. they maximize the likelihood that the data were generated by the distribution or the Bayesian posterior mean minimizes the expected quadratic loss, these are related to the explanatory power of the models while in reliability calculations the predictive power should be emphasized. For example, the design of a new structure is in large extent the prediction of future extreme actions, and it should be based on a prediction appreciating the uncertainties stemming from scarcity of data, not just on a model which fits best to the historical observations. As Fig. 3 shows the point estimate models very often underestimate the failure probability.
- If point estimates are used the models derived from 5 realizations would convey the same confidence as those based on 1000 data. In contrast, the Bayesian posterior predictive distribution automatically penalizes the small sample size based predictions.
- The completeness requirement of the reliability index [5], i.e. all sources of uncertainties should be taken into account supports the incorporation of parameter uncertainty.
- The Bayesian approach provides a natural way to rationally incorporate parameter estimation uncertainty. There is no such ambiguity involved as in probabilistic model uncertainty, hence independent analysts would calculate the same reliability index; besides the prior distributions there is no room for subjective treatment.

Fig. 1 and Tab. 2 show that the maximum likelihood model with 75% confidence interval is a poor approximation to incorporate parameter uncertainty compared to the posterior predictive that has rational basis. The former can considerably underestimate the failure probability (Tab. 2).

The numerical results would likely change if different distribution types were adopted, especially in case of the dominant variable action. It is expected that the effect of parameter estimation uncertainty would be smaller for real structures since prior information is frequently available. Further research is needed to generalise the findings and to assess the sensitivity of the numerical results to these assumptions.

6 CONCLUSIONS

The effect of parameter estimation uncertainty on failure probability seems to be neglected in reliability studies and in structural standards. Using Bayesian approach it is shown that this neglect can lead to several orders of magnitude underestimation of failure probability. The maximum likelihood model with 75% confidence interval, promoted in Eurocodes, underestimates this effect and consequently may lead to overestimation of a reliability level. Bayesian statistics proves to be a suitable tool for treating parameter estimation uncertainty, and Bayesian posterior predictive distribution is recommended for its incorporation in reliability studies. However, further research and consideration of practical examples are needed to generalise the conclusions. The incorporation of this type of uncertainty could be especially important for critical facilities such as nuclear power plants where site-specific data are used to construct the probabilistic models.

REFERENCES

[1]   AITCHISON, J., and DUNSMORE, I. R. 1980. Statistical Prediction Analysis. New York, USA: Cambridge University Press.

[2]   ASCE. 2010. Minimum Design Loads for Buildings and Other Structures. ASCE/SEI 7-10. Reston, Virginia: American Society of Civil Engineers.

[3]   COLES, S. 2001. An Introduction to Statistical Modeling of Extreme Values. London: Springer-Verlag.

[4]   COLES, S., and PERICCHI, L. 2003. Anticipating catastrophes through extreme value modelling. Journal of the Royal Statistical Society. Series C (Applied Statistics), 52(4), 405-416.

[5]   DER KIUREGHIAN, A. 1989. Measures of Structural Safety Under Imperfect States of Knowledge. Journal of Structural Engineering, 115(5), 1119-1140. doi: 10.1061/(ASCE)0733-9445(1989)115:5(1119)

[6]   DER KIUREGHIAN, A., HAUKAAS, T., and FUJIMURA, K. 2006. Structural reliability software at the University of California, Berkeley. Structural Safety, 28(1–2), 44-67. doi: http://dx.doi.org/10.1016/j.strusafe.2005.03.002

[7]   ELLINGWOOD, B., GALAMBOS, T. V., MACGREGOR, J. G., and CORNELL, C. A. 1980. Development of a Probability-Based Load Criterion for American National Standard A58. Washington, DC, USA: National Bureau of Standards.

[8]   ELLINGWOOD, B., and O'ROURKE, M. 1985. Probabilistic models of snow loads on structures. Structural Safety, 2(4), 291-299. doi: http://dx.doi.org/10.1016/0167-4730(85)90015-3

[9]   GREENWOOD, M. 1926. A report on the natural duration of cancer Reports on Public Health and Medical Subjects. His Majesty's Stationery Office, London: Ministry of Health.

[10]   HOLICKY, M., and SYKORA, M. 2009. Failures of Roofs under Snow Load: Causes and Reliability Analysis Forensic Engineering 2009 (pp. 444-453): American Society of Civil Engineers.

[11]   JCSS. 2000. JCSS Probabilistic Model Code Part III - Resistance Models: Joint Committee on Structural Safety.

[12]   JCSS. 2001. JCSS Probabilistic Model Code Part II - Load Models: Joint Committee on Structural Safety.

[13]   KRUGER, A. C. 2010. Wind climatology of South Africa relevant to the design of the built environment. (Ph.D.), Stellenbosch University, South Africa. Retrieved from http://hdl.handle.net/10019.1/6847

[14]   LEMAIRE, M., CHATEAUNEUF, A., and MITTEAU, J.-C. 2010. Structural Reliability: Wiley.

[15]   MATLAB. 2014. The Language of Technical Computing (Version R2014a (8.3.0.532)). Natick, Massachusetts: MathWorks Inc. Retrieved from http://www.mathworks.com/products/matlab/

[16]   RÓZSÁS, Á., and SÝKORA, M. 2015. Model comparison and quantification of statistical uncertainties for annual maxima of ground snow loads. In L. Podofillini, B. Sudret, B. Stojadinovic, E. Zio and W. Kröger (eds.), Safety and Reliability of Complex Engineered Systems - Proceedings of the European Safety and Reliability Conference ESREL 2015, Zurich, 7-10 September 2015. Leiden: Balkema/CRC Press, p. 2667-2674, 2015.

[17]   RÓZSÁS, Á., and SÝKORA, M. 2015. Effect of Statistical Uncertainties in Ground Snow Load on Structural Reliability. In Proceedings of IABSE Geneva Conference 2015, 23-25 September 2015, Geneva. Zurich: IABSE.

[18]   SADOVSKÝ, Z., and PÁLEŠ, D. 2008. Probabilistic optimization of partial safety factors for the design of industrial buildings. International Journal of Reliability, Quality and Safety Engineering, 15(05), 411-424. doi: doi:10.1142/S0218539308003131

[19]   SANPAOLESI, L., and COLOMBINI, S. 2005. Implementation of Eurocodes. Handbook 3. Action effects for buildings. Chapter IV - Thermal actions on buildings. Aachen, Germany.

[20]   SANPAOLESI, L., CURRIE, D., SIMS, P., SACRÉ, C., STIEFEL, U., LOZZA, S., EISELT, B., PECKHAM, R., SOLOMOS, G., HOLAND, I., SANDVIK, R., GRÄNZER, M., KÖNIG, G., SUKHOV, D., DEL CORSO, R., and FORMICHI, P. 1998. Scientific Support Activity in the Field of Structural Stability of Civil Engineering Works: Snow Loads. Final Report Phase I. Brussels: Commission of the European Communities. DGIII-D3.

[21]   SISSON, S. A., PERICCHI, L. R., and COLES, S. G. 2006. A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean. Stochastic Environmental Research and Risk Assessment, 20(4), 296-306. doi: 10.1007/s00477-005-0246-4

Reviewers:

Doc. Ing. Pustka David, Ph.D., Department of Building Structures, Faculty of Civil Engineering, VŠB – Technical University of Ostrava, Czech Republic.

Dr. Ing. Rostislav Rypl, Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology, Czech Republic.