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

Comparison of Methods for Model Uncertainty Quantification

Jannie S. Nielsen
Associate Professor, Dept. of the Built Environment, Aalborg University, Aalborg, Denmark

ABSTRACT: Model uncertainties are included in reliability assessments to account for the uncertainty introduced by the use of computational models to obtain e.g. resistances and load effects. Often, the model uncertainty is modeled as a lognormal variable and is quantified using a dataset consisting of measured outcomes from experiments and corresponding outcomes of the computational model. Unless the dataset is very large, statistical uncertainty also needs to be included. Approaches for model uncertainty quantification include the Bayesian analytical approach, Markov Chain Monte Carlo (MCMC) sampling, the maximum likelihood (ML) method, and the method in EN1990:2002 Annex D. In this paper, the methods are compared from a theoretical point of view, and a quantitative comparison is made based on simulated data. It is concluded that MCMC leads to the same result as the Bayesian analytical approach, when non-informative conjugate priors are used. The ML method underestimates the model uncertainty for small sample sizes. The method in EN1990:2002 has an inconsistency in the modelling and therefore leads to larger variation in the estimated bias and 5% quantile than the Bayesian method. This study motivates a change in EN1990 Annex D to the Bayesian approach, as it is equally simple, and the change has been included in the current draft PrEN1990:2022.

1. INTRODUCTION

Model uncertainties need consideration, as virtually no engineering models are perfect. Even models with a theoretical foundation are based on idealizations of reality. Often, engineering models are established on a combination of theoretical knowledge of the phenomena and data from experience or experiments. A model builder is often faced with the tradeoff between simplicity and accuracy. It is preferred to have a model that includes the most important variables, but it should not include variables which do not have a significant influence on the output of the model. There may be parameters which do have an influence but cannot be included as they cannot be measured. In order to fit an empirical model, data is needed. Ideally a large number of experiments are performed in a controlled environment, where all variables included in the model are measured accurately. However, the costs of testing limit the number of tests, especially when large scale tests are needed. In summary, model uncertainties are present due to 1) limited data, 2) omission of parameters, and 3) inaccurate functional relations in mathematical/computational models.

Model uncertainties may be estimated as part of the process of establishing a mathematical model based on data. Then it is usually desirable to obtain a model with no bias, and with a small model error. In other cases, it is preferred to use established models (e.g. from standards), and use data to quantify the model uncertainty and bias.

For many years, Bayesian analysis has been recognized as the most coherent approach for the modelling and quantification of uncertainties (Benjamin and Cornell, 1970; Der Kiureghian, 2022), and the approach is recommended in the JCSS Probabilistic model code (JCSS, 2001). Closed form solutions are available for simple cases where conjugate distributions exist, for example if accurate measurements of calculated and measured outcomes are available, and the model uncertainty is an additive normal distributed variable, or a multiplicative lognormal variable. In the normal case, the standard deviation of the error is considered constant regardless of the size of the theoretical resistance, and for the lognormal case, the coefficient of variation (COV) is constant, i.e. the standard deviation of the error increase linearly with the

theoretical resistance. The difference is illustrated in Figure 1.

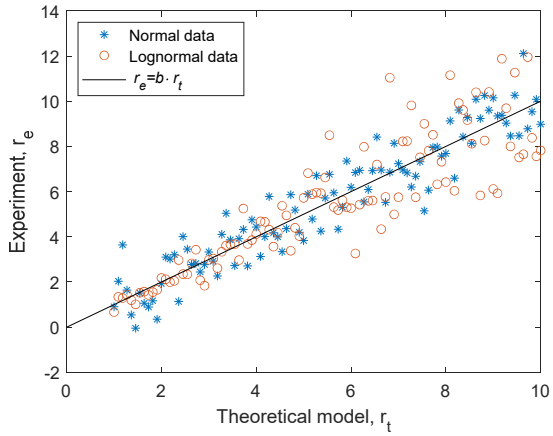

Figure 1:Experimental resistance r_e as a function of resistance from the theoretical model r_t for two samples of size 100 with respectively normal distributed additive model uncertainty and lognormal distributed multiplicative model uncertainty.

Although the Bayesian approach has many advantages, other methods are also used. Annex D in Eurocode 0 (EN 1990, 2002) describes a method to estimate characteristic values of the resistance based on data accounting for the model uncertainty. Here, the COV of the model uncertainty is estimated using a Bayesian approach, but the bias (referred to as mean value correction factor) is found in a different way.

In other cases, the lack of closed form solutions necessitates the use of other approaches. Bayesian regression can be performed for model fitting and uncertainty quantification using discretized distributions or using Markov Chain Monte Carlo (MCMC) methods (Schubert et al., 2020). The maximum likelihood (ML) method is also often highlighted as a preferred method for estimating model parameters, due to the property that the estimated parameters are asymptotically normal distributed (Lindley, 1965), and thereby the method includes statistical uncertainty through the parameter uncertainty. The ML method was used for estimating model parameters for timber structures (Sørensen et al., 2003), for finding parameters in the distribution for cable tensile strength (Faber et al., 2003), and for the shear capacity for concrete (Melhem et al., 2020).

In is unclear, to what degree the choice of method for uncertainty quantification affects the result, and this motivates the current study where four methods are considered: Bayesian analytical, MCMC, EN1990:2002, and ML. The aim is to compare the obtained predictive distributions numerically, and explain observed differences based on theory.

2. METHODS

The notation from EN1990:2002 Annex D is adapted here, where the model uncertainty is divided into a variable for the bias b, and variable with mean one for the model uncertainty δ, such that the probabilistic model for the resistance, r, is given as:

$$r = b\delta r_t \tag{1}$$

where r_t is the resistance found using the theoretical model. It is assumed that the population for the model uncertainty $b\delta$ follows a lognormal distribution with mean b and coefficient of variation V_δ. The task is to estimate the predictive distribution of $b\delta$, using a dataset of n corresponding values of the theoretical resistance r_t and resistance found using an experiment r_e. In the basis case, it is assumed that there is no measurement uncertainty on neither the input to the theoretical model nor on the experimental resistance. Also, there is no prior information on the model uncertainty.

2.1. Bayesian analytical

The basic case has a simple analytical closed form solution, thus the Bayesian approach provides the most accurate and consistent solution. A realization of the model uncertainty can be calculated for each pair of theoretical and experimental resistance:

$$(b\delta)_i = \frac{r_{ei}}{r_{ti}} \tag{2}$$

Because $b\delta$ follows a lognormal distribution, $\ln(b\delta)$ follows a normal distribution with parameters $m_{\ln(b\delta)}$ and $s_{\ln(b\delta)}$, where the parameters are estimated as the sample mean and

sample standard deviation of $\ln(b\delta)$. The predictive distribution for $\ln(b\delta)$ is then given by:

$$\ln(b\delta) = m_{\ln(b\delta)} + T_{n-1} \cdot s_{\ln(b\delta)} \sqrt{1 + \frac{1}{n}} \quad (3)$$

where T_{n-1} follows a Student's t distribution with n-1 degrees of freedom. Consequently, the predictive distribution for $b\delta$ is:

$$b\delta = \exp\left(m_{\ln(b\delta)} + T_{n-1} \cdot s_{\ln(b\delta)} \sqrt{1 + \frac{1}{n}} \right) \quad (4)$$

The estimates of b and V_δ are calculated through the well-known relations between the moments and parameters in a lognormal distribution.

2.2. Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a sampling-based method for Bayesian inference. The strength of the method is that it can be used, where closed form solutions do not exist. For the basic case, the inference model can be represented graphically as in Figure 2.

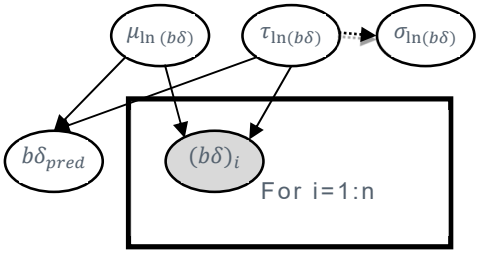

Figure 2: Graphical representation of the MCMC inference model. Full lines represent stochastic relationships, dotted lines represent logic relationship, and shaded nodes are observed nodes.

The node $(b\delta)_i$ has *n* instantiations; one for each set of values in the sample, and it follows a lognormal distribution with parameters $\mu_{\ln(b\delta)}$ and $\tau_{\ln(b\delta)}$, which are the mean and precision of the associated normal distribution for $\ln(b\delta)$. The precision is often used in MCMC as a second parameter instead of the standard deviation $\sigma_{\ln(b\delta)}$, and they are related through $\tau_{\ln(b\delta)} = 1/\sigma^2_{\ln(b\delta)}$. The node $b\delta_{pred}$ follow the same distribution as $(b\delta)_i$, and is included to directly obtain outcomes of the predictive distribution for $b\delta$.

Because no prior information is available on the parameters, they are given non-informative priors. For the mean value, the conjugate distribution is a normal distribution, and for the precision, a gamma distribution. Appropriate non-informative priors are therefore the conjugate distributions with parameters resulting in a rather flat distribution within the probable range of the parameter. The values in Table 1 are initially used in the comparison study. In the comparison, 10^4 samples were used for burn-in, and 10^5 samples were used to estimate the distributions.

Table 1: Stochastic model.

Variable	*Distri-bution*	*Parameter 1*	*Parameter 2*
$\mu_{ln\,R_0}$	*Normal*	*Mean:* $\mu = 0$	*Precision:* $\tau = 10^{-6}$
$\tau_{ln\,R_0}$	*Gamma*	*Shape:* $\alpha = 0.001$	*Rate:* $\beta = 0.001$
$R_{0,i}$ $R_{0,pred}$	*Log-normal*	*Mean of* $ln\,R_0$*:* $\mu_{ln\,R_0}$	*Precision of* $ln\,R_0$*:* $\tau_{ln\,R_0}$

2.3. EN1990:2002 Annex D

In EN1990:2002 Annex D (EN 1990, 2002), an approach is presented for estimating the characteristic value of the resistance based on data, such that it accounts for the model uncertainty, statistical uncertainty, and physical uncertainty. Although the Bayesian approach is used when accounting for statistical uncertainty, an alternative method is used for the estimation of the bias *b*.

First, the experiment values r_e are plotted as function of theoretical values r_t, in the same way as in Figure 1. The mean value correction factor b is then found as the slope of the zero-intercept "Least squares" best-fit:

$$b = \frac{\sum_{i=1}^{n}(r_e r_t)}{\sum_{i=1}^{n}(r_t^2)} \quad (5)$$

This formula is obtained by minimization of $F = \sum_{i=1}^{n}(b \cdot r_{ti} - r_{ei})^2$ w.r.t. b. The error term for each data point δ_i is now calculated from:

$$\delta_i = \frac{r_{ei}}{b \cdot r_{ti}} \tag{6}$$

Subsequently, Δ is defined by:

$$\Delta = \ln \delta \tag{7}$$

and s_Δ is calculated as the sample standard deviation of Δ. Next, the coefficient of variation of δ is calculated as:

$$V_\delta = \sqrt{\exp(s_\Delta^2) - 1} \tag{8}$$

As this is the well-known procedure for calculating the coefficient of variation of a lognormal variable, it is evident from this procedure that δ is implicitly assumed to follow a lognormal distribution. However, the sample mean of Δ is not used; instead it is implicitly assumed that the mean value of δ is equal to one.

In the comparison, the predictive distribution for this method is established using Eq. 4 with lognormal parameters calculated using the bias b and COV V_δ.

2.4. Maximum likelihood

In the ML method, the distribution parameters are estimated such that the probability of obtaining the sample is maximized. Therefore, the distribution parameters $\boldsymbol{\theta}$ are estimated to maximize the likelihood function, which is the product over the probability density function $f_X(x;\boldsymbol{\theta})$ evaluated in the n sample points:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_X(x_i;\boldsymbol{\theta}) \tag{9}$$

For some distribution types, analytical closed form solutions exist for the maximum likelihood estimate of the parameters. For other distributions, optimization algorithms can be used to estimate the parameters, based on a sample. In both cases it is convenient to use the loglikelihood function as it eases the computations for analytical solutions and avoids numerical underflow for numerical calculations:

$$LL(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln\big(f_X(x_i;\boldsymbol{\theta})\big) \tag{10}$$

The maximum likelihood estimates of the parameters are found by maximizing the loglikelihood function with respect to the parameters. The statistical uncertainty (standard deviations and correlation coefficients) related to the estimated parameters can be estimated based on the hessian matrix (consisting of the second derivatives of the loglikelihood function in the point of the maximum likelihood estimate).

The hessian matrix for two parameters θ_1, θ_2 is given by:

$$\boldsymbol{H} = \begin{bmatrix} \dfrac{\partial^2 LL(\boldsymbol{\theta})}{\partial \theta_1^2} & \dfrac{\partial^2 LL(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} \\ \dfrac{\partial^2 LL(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} & \dfrac{\partial^2 LL(\boldsymbol{\theta})}{\partial \theta_2^2} \end{bmatrix} \tag{11}$$

The covariance matrix is then given by:

$$\boldsymbol{C} = [-\boldsymbol{H}]^{-1} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \tag{12}$$

It can be shown that the parameters are asymptotically normal distributed, i.e. for sample sizes approaching infinity (Lindley, 1965).

In the comparison, an optimization algorithm is used to estimate the parameters and the hessian matrix. The predictive distribution is established numerically by assuming a multivariate normal distribution for the parameters. Then samples from the predictive distribution can be drawn by first drawing from the multivariate distribution for $\mu_{\ln(b\delta)}, \sigma_{\ln(b\delta)}$, and then from $(b\delta) \sim logn\big(\mu_{\ln(b\delta)}, \sigma_{\ln(b\delta)}\big)$.

3. COMPARISON OF METHODS

In this section, the four methods will be compared through numerical studies, and the results will be discussed. First the methods are compared using a single sample, and next the methods are compared based on simulations.

3.1. Comparison for one sample

Initially, a sample of size $n = 10$ was generated with a coefficient of variation $V_\delta = 0.15$ and $b = 1$. Figure 3 shows the (r_t, r_e)-plot for the random sample, and the bias found using each method is shown with a line. Figure 4 shows the cumulative distribution function for the predictive distribution found using each method, and Figure 5 shows the lower part of the figure, to clarify the differences. For this random sample, the EN1990:2002 method results in a value of b which is smaller than for the other methods, whereas the other methods leads to similar values of b (b is the slope in Figure 3 and the mean/median in Figure 4). Looking also at the COV of the distribution function, it is seen that the Bayesian method and MCMC lead to very similar results, whereas the ML method has a smaller COV. The effect on the 5% quantile can be seen in Figure 5, where EN1990:2002 gives a 4.9% smaller characteristic value, and ML gives a 3.2% larger value. Nothing in general can be concluded by looking at just one sample, except that the ML and EN1990:2002 clearly can lead to different results than the Bayesian methods.

To examine the effect of the prior distributions used in MCMC, Figure 6 shows the found predictive distributions, when flat priors are used instead of the conjugate priors. It is seen that here the distribution for the Markov chain does not converge to the Bayesian analytical predictive distribution. It has been verified with additional random samples that the MCMC method converges towards the Bayesian analytical solution, when the conjugate priors are used, whereas an error is seen when inappropriate diffuse priors are used. The reason is that the priors also are used as sampling distributions. The MCMC method will not be included in the simulation-based evaluation, because it converges to the same distribution as the Bayesian analytical solution, and it is time demanding to run each analysis.

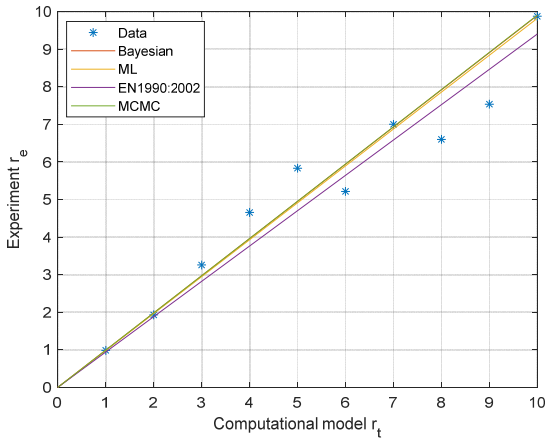

Figure 3: Sample with n=10 and lines indicating the bias estimated using each method.

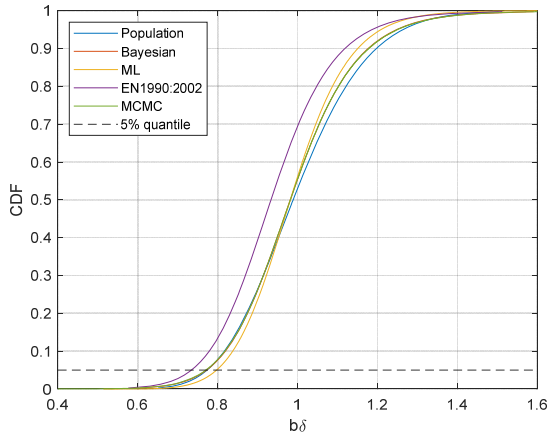

Figure 4: Cumulative distribution function for the model uncertainty.

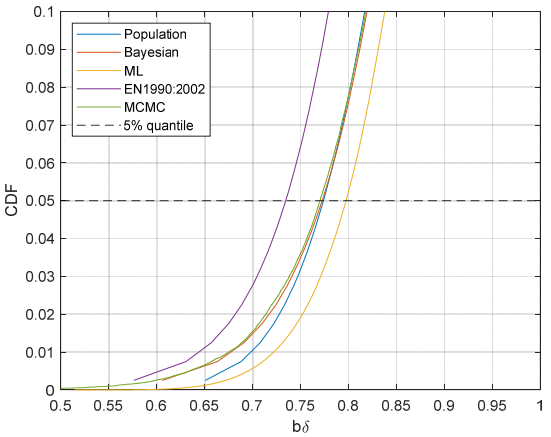

Figure 5: Lower part of the cumulative distribution function for the model uncertainty.

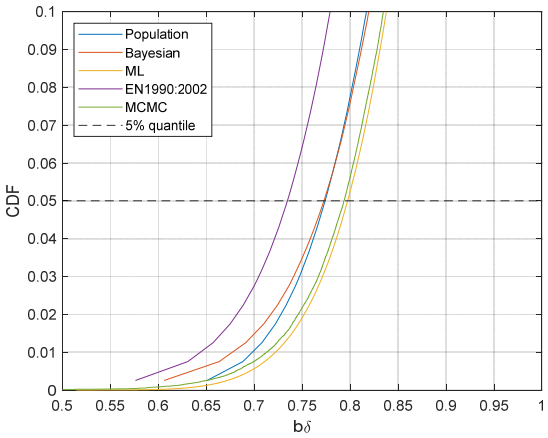

Figure 6. Cumulative distribution function for the model uncertainty, when flat priors are used in MCMC.

3.2. Comparison using simulations

For comparison of the approaches, a range of outcomes of the theoretical model r_t is defined on a uniform grid of n values from 1 to 10. Realizations of the model uncertainty are generated from a lognormal population with mean $b = 1$ and coefficient of variation V_δ equal to 10%, 15%, and 20%. Experiment outcomes r_e are generated by multiplying each value from the theoretical model by a random outcome of the model uncertainty drawn from the population. For each random sample, the distribution and 5% quantile of the model uncertainty $b\delta$ was estimated using the Bayesian approach, the ML method, and the method in EN1990:2002.

To investigate for systematic deviations between the methods, 1000 random samples was simulated from the same assumed population, and the bias b, the uncertainty V_δ, and the 5% quantile in the predictive distribution for the model uncertainty was found for each sample using each method. The mean and coefficient of variation of the estimated b, V_δ, and 5% quantile values were calculated and are shown in Figure 7 to Figure 12 as function of the sample size n.

Figure 7 and Figure 8 show that all methods lead to estimates of b and V_δ that are consistent in the sense that the mean estimates converge to the population values for increasing n. However, V_δ is systematically underestimated for the ML method for sample sizes below 25-30. Figure 9 and Figure 10 show that the COV of the estimates decrease with increasing sample size, and for V_δ all methods lead to same uncertainty on the estimate. For the bias b, the ML method has the same precision as the Bayesian method, whereas the EN1990:2002 method has a lower precision, i.e. the COV of the estimate of b is systematically larger.

Figure 11 and Figure 12 show the effect on the estimated 5% quantile of the model uncertainty. It is seen that the Bayesian method and EN1990:2002 lead to same mean of the estimated 5% quantile, whereas the ML method systematically leads to larger values, but approaches the correct values for sample sizes larger than 25-30. The COV of the estimated 5% quantile is systematically larger for the EN1990:2002 method than for the Bayesian method, and slightly smaller for the ML method.

3.3. Discussion

Generally, the Bayesian analytical approach is considered the most consistent and accurate approach, thus the other methods are compared to this.

The ML method underestimated V_δ for sample sizes less than 25-30. The reason is that the ML method finds the most likely values of the parameters, i.e. the mode values. The correct posterior distribution for the standard deviation of the model uncertainty is not symmetric, as the inverse of the squared standard deviation follows a gamma distribution, and consequently the squared standard deviation follows an inverse-gamma distribution. Here, the mode value is smaller than mean value, thus the ML method underestimates V_δ. However, for a sample size approaching infinity, the gamma distribution approaches a normal distribution, and the mode approaches the mean. Therefore, the estimate of V_δ approaches the correct value, as the sample size approaches infinity.

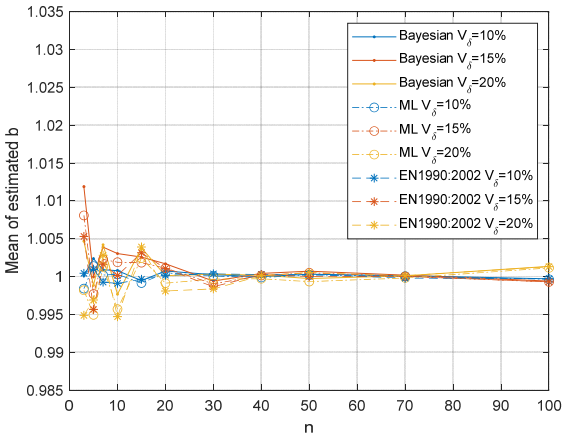

Figure 7: Mean of estimated bias b.

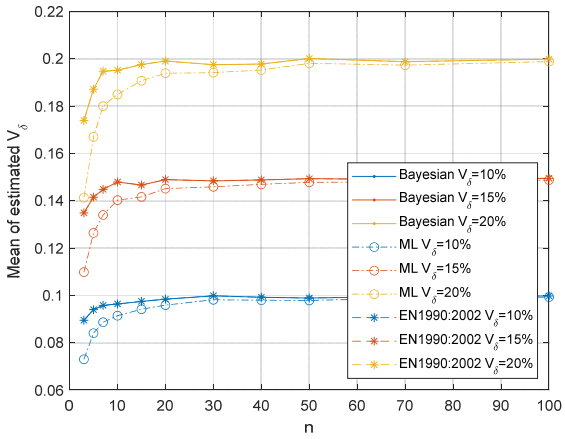

Figure 8: Mean of estimated COV of the model uncertainty V_δ.

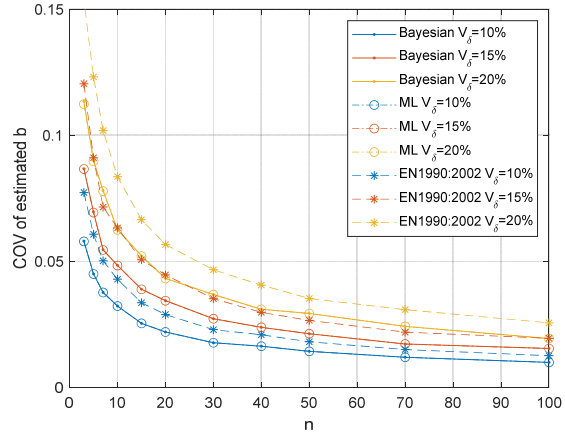

Figure 9: COV of estimated bias b.

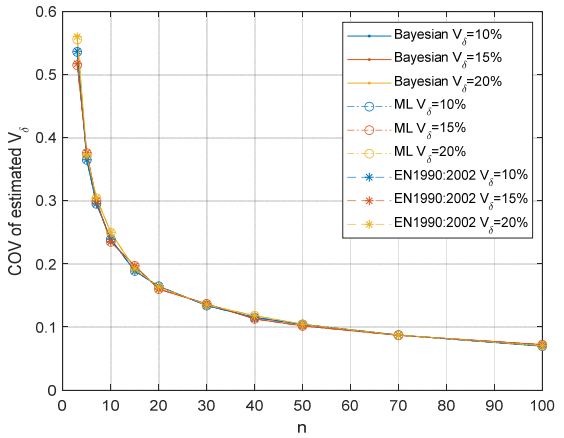

Figure 10: COV of estimated COV of the model uncertainty V_δ.

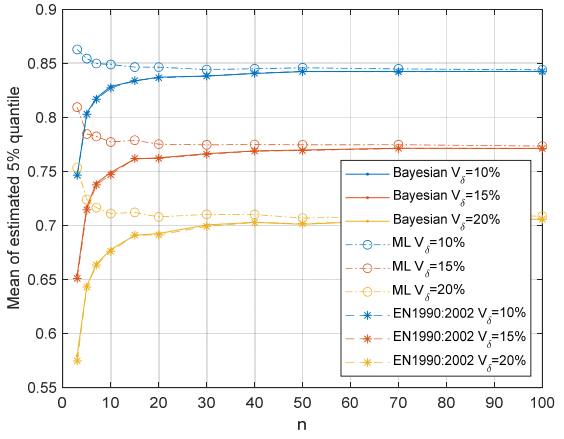

Figure 11: Mean of the estimated 5% quantile.

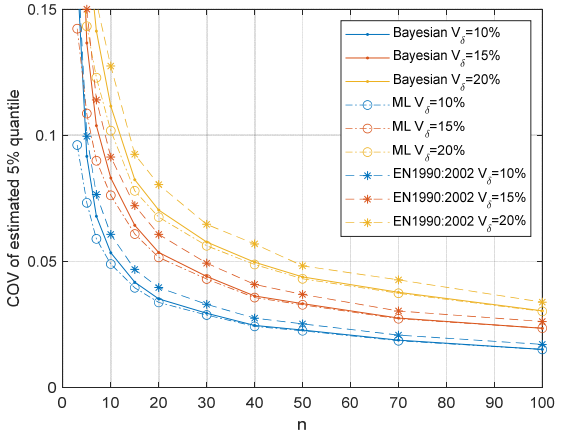

Figure 12: COV of the estimated 5% quantile.

For the method in EN1990:2002 Annex D, the COV of the estimate of the bias b is larger than for the other methods. The explanation can be found in the way the b is estimated. The estimate of b is the least squares best-fit to the slope in the (r_t, r_e)-diagram. This would be appropriate for a normal distributed additive error, but it is inconsistent with the assumption of a lognormal distribution for the error term δ. With the current least squares approach, too much weight is put on the points with large theoretical resistance at the expense of points with lower theoretical resistance. The issue is that with the applied method for estimating the bias b, the error term δ will generally not have a mean value equal to one, but this is also implicitly assumed.

4. CONCLUSIONS

It is concluded that the Bayesian analytical approach is most correct and efficient and that the MCMC converges to the same results, but only if the appropriate conjugate priors are used. In the maximum likelihood method, the estimated parameters are assumed to follow normal distributions as commonly done, and the parameter uncertainties are estimated based on the Hessian matrix. Because the maximum likelihood method identifies the mode value of the parameters instead of the mean, and because the standard deviation does not follow a normal distribution, this approach consistently leads to slightly non-conservative results for small sample sizes.

The method in EN1990:2002 estimates the coefficient of variation of the model uncertainty in correspondence with a Bayesian approach implicitly assuming a lognormal distribution for the error term. However, the bias is found using the least squares best fit, which is inconsistent with the lognormal assumption. The quantitative analysis reveals that the current method in EN1990:2002 does not systematically lead to a higher or lower estimated 5% quantile compared to the Bayesian approach, but it leads to a higher variation of the estimate. Since the Bayesian analytical approach is equally simple and more consistent, this work has motivated a proposal for changing the method in EN1990 annex D, which is included in the current draft PrEN1990:2022.

5. ACKNOWLEDGEMENTS

The work on this paper has been supported by the Danish Energy Agency through the EUDP project ProbWind: Data-driven probabilistic design of wind turbines (Grant Number 64019-0587). The support is greatly appreciated.

6. REFERENCES

Benjamin, J. R., and Cornell, C. A. (1970). *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill.

Der Kiureghian, A. (2022). Structural and System Reliability. In *Structural and System Reliability*. Cambridge University Press. https://doi.org/10.1017/9781108991889

EN 1990. (2002). Eurocode - Basis of structural design. In *En*. https://doi.org/10.1680/cien.144.6.8.40609

Faber, M. H., Engelund, S., and Rackwitz, R. (2003). Aspects of parallel wire cable reliability. *Structural Safety*, *25*(2), 201–225. https://doi.org/10.1016/S0167-4730(02)00057-7

JCSS. (2001). *Probabilistic model code*.

Lindley, D. V. (1965). *Introduction to Probability and Statistics from a Bayesian Viewpoint*. Cambridge University Press.

Melhem, M. M., Caprani, C., and Stewart, M. G. (2020). Model Error for Australian Code Shear Capacity of Concrete Structures. In *Lecture Notes in Civil Engineering* (Vol. 37, pp. 327–336). Springer. https://doi.org/10.1007/978-981-13-7603-0_33

Schubert, M., Wu, Y., Tychsen, J., Dixen, M., Faber, M. H., Sørensen, J. D., and Jonathan, P. (2020). On the distribution of maximum crest and wave height at intermediate water depths. *Ocean Engineering*, *217*, 107485. https://doi.org/10.1016/j.oceaneng.2020.107485

Sørensen, J. D., Stang, B. F. D., and Svensson, S. (2003). Effect of Load Duration on Timber Structures in Denmark. In A. Der Kiureghian, S. Madanat, and J. M. Pestana (Eds.), *Proceedings of the 9th international conference on applications of statistics and probability in civil engineering* (pp. 1185–1192). Millpress.