Comparative Investigation of Uncertainty Analysis with Different Methodologies on Fatigue Data of Rebars

Sima Rastayesh*[1], Amol Mankar[1], John Dalsgaard Sørensen[1]

[1]Department of Civil Engineering, Aalborg University, Aalborg, Denmark.
*sir@civil.aau.dk

Abstract

Stochastic modeling of uncertainties for fatigue strength parameters is vital step as basis for reliability analyses. In this paper, the Maximum Likelihood Method (MLM) is used for fitting the statistical parameters in a regression model for the fatigue strength. Furthermore, application of the Bootstrap method is investigated. The results indicate that the latter methodology does not work well in the considered case study because of runout tests within the test data. Moreover, use of Bayesian inference with Markov Chain Monto Carlo implementation is studied. These results indicate reduction in the uncertainty and thus better parameter estimations.

Key words: Uncertainty; Maximum Likelihood Method; Bootstrap; Bayesian inference; Reinforcement Concrete

1. Introduction

This paper presents statistical analyses performed on fatigue data obtained from literature [1], where laboratory fatigue tests were performed on reinforcement bars.

The statistical analyses are the first step for stochastic modeling of the material fatigue uncertainties, which next can be used as basis for a probabilistic modelling and reliability analysis of structures with reinforced concrete components.

The Maximum Likelihood Method (MLM) is used for fitting the statistical parameters [2] in a regression model for the fatigue strength. Typically, the statistical analyses are based on a limited number of data, for which MLM can provide estimates of the uncertainties associated with each of these parameters and correlation between the parameters [3].

This paper also presents use of the Bootstrapping method, which is applied to generate synthetic data with the available measurements from the lab. The results indicate that this methodology does not work well in this case because of runout tests in the test database. Finally, the parameters are re-estimated along with their uncertainties using the synthetic data with results showing that the uncertainty is increased using this method.

This paper also presents application of Bayesian inference with a Markov Chain Monto Carlo implementation. These results show reduction in the uncertainty and thus better parameter estimation.

2. Materials and methods

2.1. Case study

In this paper, test data on fatigue strength test for steel reinforcement from Hansen and Heshe [1] is utilized for statistical analysis to determine typical fatigue strength uncertainties. The steel reinforcing bars with 16 mm of diameter is chosen and the yield strength is around 570 MPa.

The uncertainty modelling is important in order to determine the characteristic fatigue design curves as well as to perform reliability analyses and risk analysis.

2.2. Statistical analysis of fatigue data of steel reinforcing bars

For steel reinforcement bars used in concrete S-N curves are recommended by various international codes (Model code 2010, Model code 1990, DNV OS C 502 and EN 1992-1 etc.) [4] [5] [6] [7] written as:

$$N_i = K \times \Delta s_i^{-m} \quad (1)$$

or

$$\log(N_i) = \log(K) - m\log(\Delta s_i) \quad (2)$$

where N_i is the number of cycles to failure with stress range Δs_i in test number i, K and m are fatigue parameters to be fitted by e.g. MLM using the tests results.

To account for uncertainties in the fatigue life, Equation (2) can be rewritten [8]:

$$\log(N_i) = \log(K) - m\log(\Delta s_i) + \varepsilon \quad (3)$$

where ε represents the uncertainty of the fatigue life model and is modelled by a stochastic variable with mean value equal to zero and standard deviation σ_ε. ε is often assumed Normal distributed.

The Likelihood function to be used to estimate the optimal values of the parameters K, m and σ_ε from test data becomes the following Equation(4):

$$L(K, m, \sigma_\varepsilon) = \prod_{i=1}^{n_F} P[\log(K) - m\log(\Delta s_i) + \varepsilon = \log(n_i)] \times \prod_{i=n_F+1}^{n_F+n_R} P[\log(K) - m\log(\Delta s_i) + \varepsilon > \log(n_i)] \quad (4)$$

Here n_i is the number of stress cycles to failure or to run-out with stress range Δs_i in test number i. n_F is the number of tests where failure occurs, and n_R is the number of tests where failure did not occur after n_i stress cycles (run-outs). The total number of tests is $n = n_F + n_R$. K, m and σ_ε are obtained from the optimization problem $max_{K,m,\sigma_\varepsilon} L(K, m, \sigma_\varepsilon)$, which can be solved using a standard non-linear optimizer.

The two terms in the previous Equation can be obtained from Equation (5):

$$P[\log(K) - m\log(\Delta s_i) + \varepsilon = \log(n_i)] = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\exp\left(-\frac{1}{2}\left(\frac{\log(K) - m\log(\Delta s_i) - \log(n_i)}{\sigma_\varepsilon}\right)^2\right)$$

$$P[\log(K) - m\log(\Delta s_i) + \varepsilon > \log(n_i)] = \Phi\left(\frac{\log(K) - m\log(\Delta s_i) - \log(n_i)}{\sigma_\varepsilon}\right) \quad (5)$$

K, m and σ_ε are parameters determined using a limited number of data; consequently, they are subject to statistical uncertainty. Since the parameters are estimated by the Maximum-Likelihood Method, they become asymptotically (number of data should be > 25 – 30) Normal distributed stochastic variables with expected values equal to the maximum-likelihood estimator and a covariance matrix equal to:

$$C_{K,m,\sigma_\varepsilon} = \left[-H_{K,m,\sigma_\varepsilon}\right]^{-1} = \begin{bmatrix} \sigma_K^2 & \rho_{K,m}\sigma_K\sigma_m & \rho_{K,\sigma_\varepsilon}\sigma_K\sigma_{\sigma_\varepsilon} \\ \rho_{K,m}\sigma_K\sigma_m & \sigma_m^2 & \rho_{m,\sigma_\varepsilon}\sigma_m\sigma_{\sigma_\varepsilon} \\ \rho_{K,\sigma_\varepsilon}\sigma_K\sigma_{\sigma_\varepsilon} & \rho_{m,\sigma_\varepsilon}\sigma_m\sigma_{\sigma_\varepsilon} & \sigma_{\sigma_\varepsilon}^2 \end{bmatrix} \quad (6)$$

$H_{K,m,\sigma_\varepsilon}$ is the Hessian matrix with second-order derivatives of the log-likelihood function. σ_K, σ_m, and $\sigma_{\sigma_\varepsilon}$ denote the standard deviations of K, m and σ_ε , respectively, and e.g. $\rho_{K,m}$ is the correlation coefficient between K and m.

2.3. Bootstrap Methodology

One of the key assumptions for using the MLM method as well as its simplified version, the nonlinear least squares method, is that the underlying distribution of errors is assumed to follow a normal (Gaussian) distribution.

In many practical applications, however, this condition is rarely satisfied. Hence, theoretically the MLM method for parameter estimation cannot be applied without compromising its assumptions, which may lead to over or underestimation of the parameter estimation errors and their covariance structure.

An alternative to this approach is the bootstrap method developed by Efron [9], which removes the assumption that the residuals follow a normal distribution. Instead, the bootstrap method works with the actual distribution of the measurement errors, which are then propagated to the parameter estimation errors by using an appropriate Monte Carlo scheme.

Fatigue tests take very long time as it would take millions of cycle for failure of one specimen and changing frequency of load application could have different results. The Bootstrap method is used to generate more data to estimate parameters, and next, to estimate the parameter uncertainties.

Residuals are estimated by subtracting calculated number of cycles to failure from observed number of cycles (Residuals = N_Exp – N_Cal). These residuals are plotted in Figure 1 considering the case when runouts are ignored in the calculation. This histogram shows that the assumption of residuals as white noise is true and it's well distributed with mean zero. In this case, the bootstrap methodology can be used, but in reality in future calculations it is mandatory to consider runouts as well which the bootstrap technique cannot fulfill as seen in Figure 2.

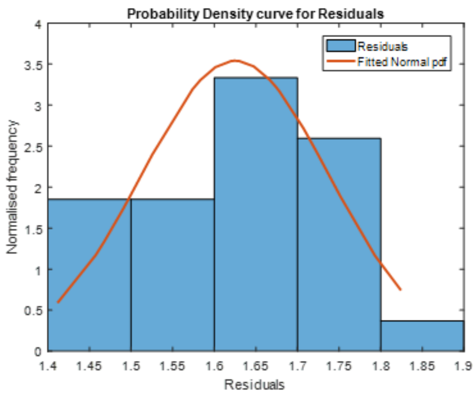

Figure 1: Histogram for Residuals without runouts

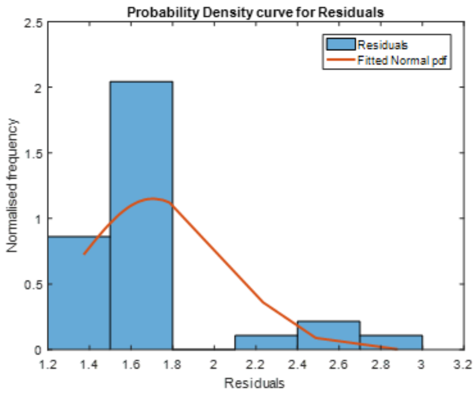

Figure 2: Histogram for Residuals with runouts

If we plot the residuals along its own index, they are random without considering runouts, this is basic requirement for using Bootstrap methodology that the residuals should be random. Random in this context mean that residuals should not follow a pattern. By looking figure 3 which is without consideration of run-outs, residuals do not follow any pattern.

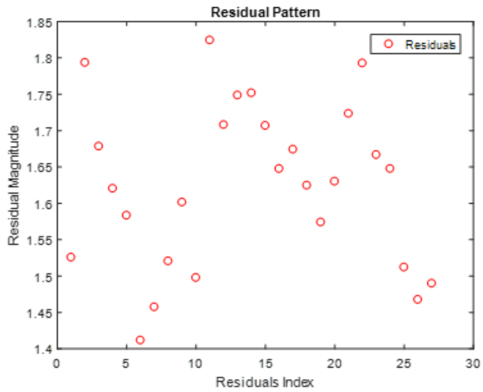

Figure 3: Residual Pattern without run-outs

Whereas, in Figure 4 with runouts, residuals are following a pattern, so the first requirement to use Bootstrapping will fail.

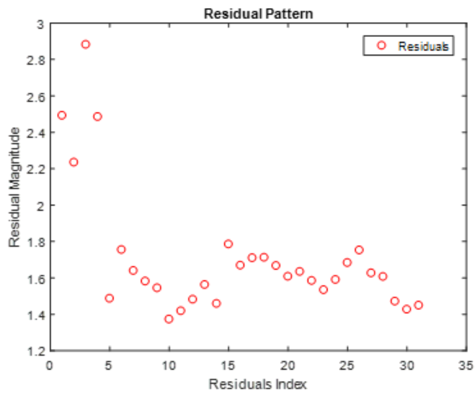

Figure 4: Residuals Pattern with run-outs

As explained above generating synthetic data is possible when the residuals do not follow a pattern, so in this case study Bootstrap method can not be used.

2.4. Bayesian inference with Markov Chain Monto Carlo implementation

2.4.1 Bayesian updating MCMC theory Bays' rule

Bays' rule provides the mathematical basis to update beliefs (prior information) about a variable, θ, given observations, y. Mathematically Bays' rule calculates the posterior probability of θ given observations, $p(\theta|y)$ as follows

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (7)$$

Bayesian inference for predictive distributions

$$p(y^*|y) = \int p(y^*|\theta)p(\theta|y)d\theta \quad (8)$$

Future predictions are modelled using the updated probability $p(\theta|y)$ similar to making a prediction for y* using a single value of θ in the traditional sense.

This integral is calculated using Monto Carlo simulations strategies such as Markov-Chain Monto-Carlo algorithms.

Markov chain simulations in brief

By definition Markov chain simulation is a sequence of random variables θ^1, θ^2, θ^3, ..., for which for any *k*, the distribution of θ^k depends only on the most recent one θ^{k-1}.

In practice, several independent sequences of Markov chain simulations are created.

Metropolis algorithm

The Metropolis algorithm is used for the transition distribution. It is an adaption of a random walk that uses an acceptance / rejection rule to converge to the specified target distribution. The step by step procedure is as follows [4]:

1. Select initial parameter vector
2. Iterate as follows for k=1, 2, 3…..
    a. Create a new trial position $\theta^* = \theta^{k-1} + \Delta\theta$, where Δθ is randomly sampled from jumping distribution q(Δθ).
    b. Create metropolis ratio

$$r = \frac{\pi\left((\theta^*|y)\right)}{\pi\left((\theta^{k-1}|y)\right)} \quad (9)$$

3. Accept new sample if:

$$\theta^k = \begin{cases} \theta^* & \text{with probability } \min(r, 1) \\ \theta^{k-1} & \text{otherwise} \end{cases} \quad (10)$$

*Note that this requires that the jumping distribution is symmetric: $q(\theta^*, \theta^{k-1}) = q(\theta^{k-1}, \theta^*)$; if the jumping distribution is not symmetric then the Metropolis-hasting algorithm can be used where both side jumping distributions are part of the ratio.

Implementing Bayesian updating using Markov Chain Monte Carlo:

Since the posterior distribution can be calculated by,

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (11)$$

where p(y) is a normalizing constant it also follows that the posterior density can be written:

$$p(\theta|y) \propto p(\theta)p(y|\theta) \quad (12)$$

i.e. the posterior distribution is proportional to product of the prior and the likelihood functions.

It is assumed that the prior distribution is the multivariate Normal distribution, then the likelihood function is defined by:

$$p(y|\theta, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2\sigma^2} SS(\theta)\right) \quad (13)$$

where,

$$SS(\theta) = \sum_i^n \left(y - f(S, \theta)\right)^2 \quad (14)$$

The Metropolis ratio becomes:

$$r = \frac{p(\theta^*|y, \sigma^2)}{p(\theta^{k-1}|y, \sigma^2)} = exp\left(-\frac{1}{2\sigma^2}\left(SS(\theta^*) - SS(\theta^{k-1})\right)\right) \quad (15)$$

Conversion statistics by potential scale reduction factor, R

Reference is made to [10] for theory and more detailed descriptions, the scale reduction factor *R* indicate the potential scale reduction factor for the current distribution if the sampling were to continue to infinity. The sampling is said to converge if *R* is close to one.

Therefore the number of simulations should be chosen such that *R* becomes as close to one as possible and thereby the Monte Carlo error close to zero.

The parameters fitted are *K* and *m*. The correlation between them is illustrated in Figure 5. In this context the Monte Carlo simulation strategy like the Markov Chain Monte Carlo algorithm is used. Furthermore, Metropolis algorithm is applied for obtaining the transition distribution. Based on the Reference [10], the scale reduction factor *R* is also calculated.

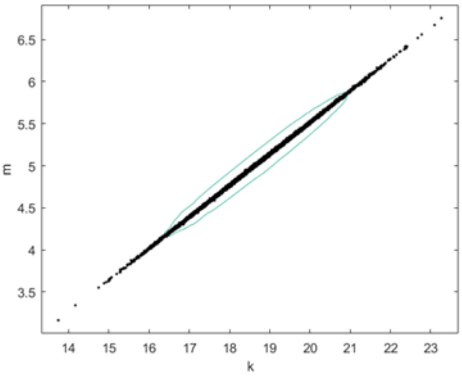

Figure 5: correlation between k and m

3. Results and discussions

The results provide estimates of the statistical parameters by MLM accounting for runouts.

Table 1 shows a comparison between values based on the Eurocode and parameters estimated by MLM assuming that the slope of Wohler curve is fixed to 5, see Figure 5. If a 5% quantile is estimated using the MLM results then log k = 18.1 is obtained; which is on the safe side compared to the Eurocode value.

The Bootstrap method is generally beneficial especially in cases where measurements are costly or time consuming [11]. The investigations in this paper indicate that the Bootstrap method applied for the fatigue test can not be used because of runouts implying that the residuals are not random.

Markov Chain Monte Carlo simulation results show that log k is mostly in the interval 18-19, and m around 5, which is the same as the results obtained from MLM.

Posterior marginal density function is also depicted in Figure 7.
Uncertainties associated with all these statistical parameters are reduced by using Bayesian inference. For instance log(k) is calculated 18.72. Results is shown in Table 1 in comparison with MLM method.
Using results obtained from MLM, reliability in a bridge as a case study is studied in [12].

Table 1: Results

Parameter	Value based on Eurocode	Mean by MLM	Mean by Bayesian	Standard Deviation by MLM	Standard Deviation by Bayesian	Distribution	Remark
ε	0	0	0			Normal	Error term
σ_ε		0.39	0.21	0.06	0.09	Normal	Standard deviation of error term
$\log k$	17.054	18.77	18.72	0.07	0.05	Normal	Location parameter in Wohler curve
m	5	Fixed = 5	5			Fixed / Deterministic	Slope of Wohler curve
$\rho_{\log k,\sigma_\varepsilon}$		0.06	0.04			Deterministic	Correlation coefficient between location and standard deviation of error

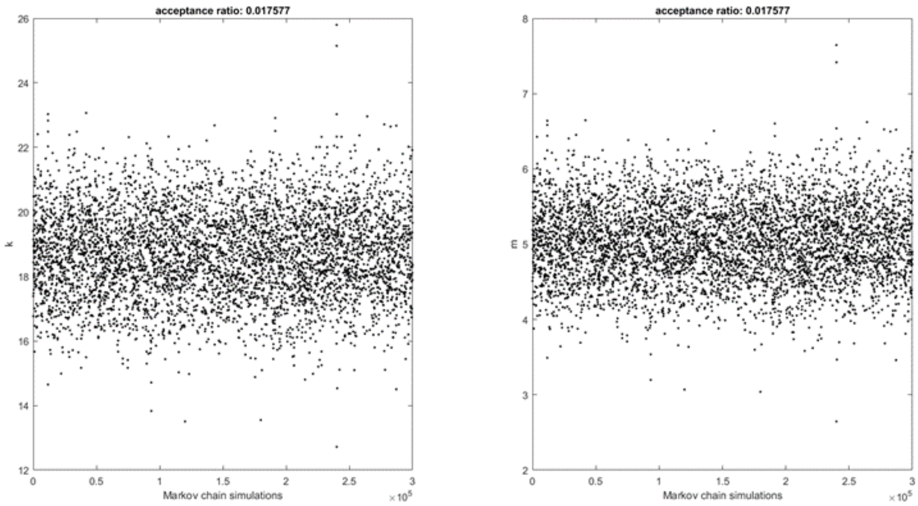

Figure 6: Markov Chain Simulation

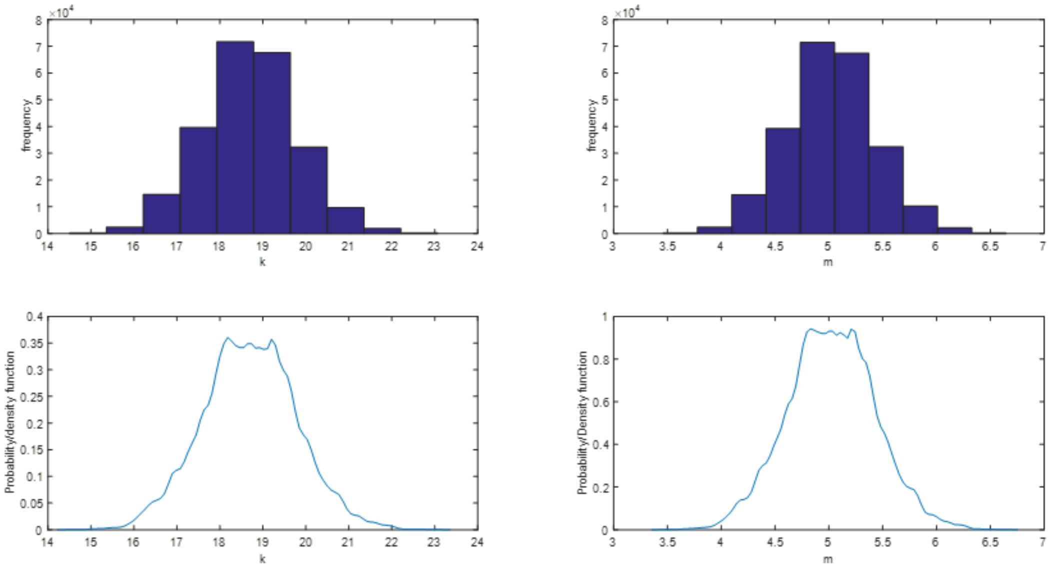

Figure 7: Posterior marginal density functions

Acknowledgements

The project INFRASTAR (infrastar.eu) has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676139. The grant is gratefully acknowledged.

References

[1] Hansen, L. P., & Heshe, G., "Static, Fire and Fatigue Tests of Ultra High-Strength Fibre Reinforced Concrete and Ribbed Bars," *Journal of Nordic Concrete Research,* pp. 17-37, 2001.

[2] J. D. Sørensen, Notes in Structural Reliability Theory and Risk Analysis, Aalborg University, 2011.

[3] S. G. a. W. C., Non-linear regression, New York: Wiley, 1989.

[4] Metropolis, N., Ulam, S., "The Monte Carlo method," *Journal of the American*

Statistical Association, vol. 44, no. 247, pp. 335-341, 1949.

[5] MC1990, FIB Model Code for concrete structures 1990, Berlin, 1993.

[6] MC2010, FIB Model Code for concrete structures 2010, Berlin: Ernst & Sohn, 2013.

[7] S. Høvik, DNV OS C 502, DNV OS C 502, Offshore concrete structures, DNVGL, 2012.

[8] S. M´arquez-Dom´ınguez, Reliability-Based Design and Planning of Inspection and Monitoring of OffshoreWind Turbines, Aalborg: Aalborg University, 2013.

[9] B. Efron, "Bootstrap methods: another look at the jackknife," *The annals of Statistics,* 1979.

[10] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D., Bayesian Data Analysis, New York: A CHAMPMAN & HALL BOOK, 2003.

[11] Sin, G., & Gernaey, K. V., " Data Handling and Parameter Estimation," *Experimental Methods in Wastewater Treatment,* vol. 281780404745, 2016.

[12] Mankar A., Rastayesh S. &. Sørensen, J.D, "Fatigue Reliability analysis of Cret De l'Anneau Viaduct: a case study," in *IALCCE 2018*, Ghent, 2018.