

Uncertainty analysis of constant amplitude fatigue test data employing the six parameters random fatigue limit model

Davide Leonetti^{1,*}, Johan Maljaars^{1,2}, and H.H. (Bert) Snijder¹

¹Eindhoven University of Technology, Department of the Built Environment, Eindhoven, the Netherlands

²TNO, Delft, the Netherlands

Abstract. Estimating and reducing uncertainty in fatigue test data analysis is a relevant task in order to assess the reliability of a structural connection with respect to fatigue. Several statistical models have been proposed in the literature with the aim of representing the stress range vs. endurance trend of fatigue test data under constant amplitude loading and the scatter in the finite and infinite life regions. In order to estimate the safety level of the connection also the uncertainty related to the amount of information available need to be estimated using the methods provided by the theory of statistic. The Bayesian analysis is employed to reduce the uncertainty due to the often small amount of test data by introducing prior information related to the parameters of the statistical model. In this work, the inference of fatigue test data belonging to cover plated steel beams is presented. The uncertainty is estimated by making use of Bayesian and frequentist methods. The 5% quantile of the fatigue life is estimated by taking into account the uncertainty related to the sample size for both a dataset containing few samples and one containing more data. The S-N curves resulting from the application of the employed methods are compared and the effect of the reduction of uncertainty in the infinite life region is quantified.

1 Introduction

Performing constant amplitude (CA) fatigue tests on structural details, plain material or assemblies is a necessary step to either perform or validate the design of these components against fatigue failure. When undamaged specimens are tested, each fatigue test results in two types of data: failure and right censored (runout), which consists of a test terminated without failure. According to common practice [1] the fatigue test data related to the finite life region and those related to fatigue limit, the stress range at which no failure is assumed to occur, should be analyzed separately employing two different statistical techniques: the linear regression and the staircase method. Recently, several models have been proposed in order to infer CA fatigue test data irrespectively on the type of data. In this case, the inferential process results in the estimation of the parameters related to the location and scatter in the finite life region and in the fatigue limit. The estimation of the uncertainty underlying the analysis of fatigue test data is of primary importance in order to assess the safety of structures subjected to fatigue loads by employing probabilistic methods. According to the categorization made in [2], reference is made to the aleatory uncertainty to represent the inherent variability of a

*Corresponding author: d.leonetti@tue.nl

Table 1. Description of the dataset related to cover-plated steel beam with welded and unwelded ends, from [4] determining the “large sample” data.

Reference	Steel grade	Failures	Right Censored
[4], Table F1	A36	34	0
[4], Table F2	A441	33	0
[4], Table F9	A36	30	0
[4], Table F7	A36	29	0
[4], Table F8	A36	30	0
[5], Table 2	A36	12	14
[4], Table F10	A36	30	0

stochastic quantity and to the epistemic uncertainty for the lack of knowledge [3]. In practice, the aleatory uncertainty is quantified through the evaluation of the scatter produced by the experimental data. Instead, the epistemic uncertainty is quantified by evaluating the degree of belief attributed to the estimators of the model parameters, which increases with increasing the number of experimental observations. Estimating both the aleatory and the epistemic uncertainty allows to correctly estimate the safety margin to be applied in order to reach a predetermined safety level.

The present paper is concerned with the estimation and reduction of aleatory and epistemic uncertainties of constant amplitude fatigue test data of the cover-plated steel beam detail, which are summarized in Table 1. A “large sample” dataset containing all the test data, reported in Table 1, and a “small sample” containing a random subset of the data, see Table 2, are inferred using the linear regression model (LRM) for comparison with common practice and to explore the affinity of the employed methods in estimating the epistemic uncertainty. Then, a regression model for fatigue is employed to infer the “large sample” data, to estimate, to reduce the epistemic uncertainty, and to evaluate their effect on the predicted fatigue life distribution.

Table 2. “Small sample” dataset, R=0

ΔS [MPa]	N [kcycle]	Type
165.00	47,5	Failure
165.00	77,4	Failure
138.00	231,4	Failure
138.00	186,3	Failure
110.00	208,9	Failure
110.00	344,1	Failure
83.00	907,5	Failure
83.00	1113,3	Failure
41.40	100000,0	Runout
34.50	8451,0	Failure
34.50	89314,0	Runout
32.40	37714,0	Failure
32.40	34930,0	Failure
31.70	100000,0	Runout

2 Models and Methods

In order to model the trend of the fatigue test data, the Basquin relation [6] is often employed in its logarithmic form:

$$\log_{10}(N) = \beta_0 + \beta_1 \log_{10}(\Delta S) \quad (1)$$

Where ΔS is the stress range and β_0 and β_1 are parameters that respectively control the location and the slope of the curve. The aleatory uncertainty are often [1] quantified assuming that the logarithm of the fatigue life at a given stress range is normally distributed and the data are assumed to be heteroscedastic, i.e. the scatter is assumed to be independent on the stress range. On the other hand, the epistemic uncertainty are considered by determining the sampling distribution of the model parameters using ancillary functions. Indeed, this method is often employed to estimate prediction intervals when the Basquin relation and a normal log-fatigue life distribution are used to infer the fatigue test data, reference is made to this model as the LRM. Despite this, it is well established that the fatigue life shows an increasing variability at lower stress ranges, i.e. approaching the fatigue limit, and that the most suitable distribution function to model the fatigue life is actually unknown [7]. In order to better model the

trend of the test data, their variability and the fatigue limit, several models have been formulated in the last decades. Pascual and Meeker [8] proposed the Random Fatigue Limit model in order to mimic the curvature at low stress ranges and to better describe the variability of the fatigue life at stress ranges approaching the fatigue limit, ΔS_0 , which is defined as a random variable. To model the aleatory uncertainty, they formulated the distribution of the fatigue life (f_w) based on the distributions of the fatigue life given the fatigue limit ($f_{w|V}$) and the fatigue limit (f_V). Moreover, by using the maximum likelihood method (MLM), runouts and failure data can be analysed together. D’Angelo et alii [9-10] proposed the Bilinear Random Fatigue Limit model, which is based on the formulation of the Random Fatigue Limit model, in order to obtain linear S-N curves, similar those employed in the standard and to define a strategy to consider the epistemic uncertainty without assuming a predetermined confidence level. Moreover, their model implies a very weak correlation between the estimators of location the fatigue limit and the fatigue life. In [11] the Six-Parameter Random Fatigue Limit model (6PRFLM) has been introduced in order to explicitly model the curvature from the finite life region towards the fatigue limit and keep the added values of the two previous model: the increasing variability of the log-fatigue life at lower stress ranges, the description of the logarithm fatigue limit as a random variable f_V having location parameter μ_V and scale parameter σ_V , and the low correlation between model parameters related to the location $f_{w|V}$ and f_V . The equation proposed is:

$$\log_{10}(N) = \beta_0 + \beta_1 \log_{10}(\Delta S) - p \log_{10} \left(1 - \frac{\Delta S_0}{\Delta S} \right) \quad (2)$$

where p is the model parameter which controls the curvature from the finite life region to the fatigue limit. For the cover-plated steel beam datasets inferred in [11], the equation employed in the 6PRFLM revealed to better mimic the trend of the test data with respect to the Random Fatigue Limit model and the Bilinear Random Fatigue Limit model. The MLM has been employed to evaluate the estimator of the model parameters, that is the maximum likelihood estimator (MLE), $\hat{\theta}_{MLE}$. Pascual and Meeker alternatively employed the lognormal distribution and the smallest extreme value distribution to describe the variability of the fatigue limit and the fatigue life. Thus, the assumption made about the shape of these two random variables affects the description of the uncertainty, more details can be found in [5, 7, 8]. However, in the literature, less attention is paid to the estimation of the epistemic uncertainty. In [9-10] the authors proposed a Monte Carlo based approach to evaluate confidence bounds on quantiles by employing a Nested sampling of the epistemic and the aleatory uncertainty. As described later, the epistemic uncertainty was assumed to follow a normal distribution and was estimated using the Likelihood function.

2.1 Epistemic uncertainty in Frequentist Statistic

2.1.1 The linear regression theory

The linear regression model is based on the assumption that the observations are normally distributed. The estimators are evaluated by minimizing the sum of the residuals squared and assuming the stress range to be the independent variable. Thus, by employing a linear relation, the estimators of the slope and the intercept are a linear combination of the observation. Then, it is possible to construct in a closed form their sampling distribution. Because of the normality of the residuals, $\varepsilon \sim N(0, \sigma)$, given an n -dimensional dataset, the Student-T distribution is employed as ancillary function and the following relations apply:

$$t_{n-2} \sim \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \quad (3)$$

$$t_{n-2} \sim \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\log_{10} \Delta S_i}{S_{xx}} \right)}} \quad (4)$$

where n is the number of tests, t_{n-2} is the Student-T distribution with $n-2$ degrees of freedom, S_{xx} is the corrected sum of squares of the $\log_{10}(\Delta S_i)$, the hat denotes the estimator of the parameters of Equation (1) and the upper-score denotes the mean value. Equations (3) and (4) denote the sampling distributions for the model parameters

In the same way, the sampling distribution for the standard deviation, σ , is denoted by the following:

$$\chi_{n-2} \sim \frac{(n-2)\hat{\sigma}^2}{\sigma} \quad (5)$$

where χ_{n-2} is the chi-square distribution (ancillary function) with $n-2$ degrees of freedom. By combining these formulations it is possible to determine confidence bounds on any quantile of the model.

2.1.1 The Maximum Likelihood Method

When the distribution of the fatigue life is not normal, the Student-T distribution is no longer suitable to describe the epistemic uncertainty. Thus, the distribution of the model parameter has to be known. The likelihood function is a function of the model parameter θ proportional to the probability of the data that were observed. The maximum likelihood method (MLM) involves the maximization of the likelihood function in order to estimate the model parameter. Moreover, the MLM can be employed to evaluate the uncertainty affecting the estimation of the model parameter. For this purpose, the Wald statistic can be employed. It assumes that the (long-run) distribution of the model parameters is a multivariate normal distribution. For each model parameter θ_i , the following relation apply:

$$N \sim \frac{\hat{\theta}_{MLE} - \theta_i}{\sqrt{[I(\hat{\theta}_{MLE})^{-1}]_{ii}}} \quad (6)$$

Where $I(\hat{\theta}_{MLE})$ is the observed Fisher information matrix and \mathcal{N} is the standard normal distribution. The multivariate normal distribution has mean equal to the MLE and the covariance matrix can be obtained from the Fisher information matrix, which is calculated from the second order derivatives of the log-likelihood function at the MLE. More details can be found in [7, 8] where the epistemic uncertainty have been estimated by employing the Wald statistic. The main drawback of this approach consist in the fact that the resulting distribution might be non-meaningful in some cases [12], especially when a parameter is close to the boundary of its domain. This has been found for the 6PRFLM fitting cover plate data [11]. More details about these methods can be found in textbooks dedicated to the MLM, such as [12], among others.

2.1.1 The Bootstrap Method

Another way to estimate the epistemic uncertainty is the Bootstrap [14]. It is a resampling method, which is able to provide non-parametric statistical inferences, without any assumption about the distribution describing the quantity under consideration. In order to estimate a certain statistic and its uncertainty, the method involves the use of the sampling distribution, which is a probability distribution that shows how the statistic of interest would vary if the sample was collected several times from the population. The bootstrap distribution, an approximation of the sampling distribution, is obtained by sampling with replacement a certain number of bootstrap samples from the original dataset and estimate the statistic of interest for each one of them. It is necessary that the dimension of each bootstrap sample is the same as the dataset. Thus, a certain number of estimators of the statistic of interest are obtained, determining the bootstrap distribution, which can be intended as a frequentist estimation of the epistemic uncertainty. The main drawback of the bootstrap is the intense computational effort with respect to the other methods, which increases with increasing the dimension of the original dataset. Indeed, since the resampling is performed with replacement, the number of possible bootstrap samples would be equal to $(2n - 1)!/[n!(n - 1)!]$, which makes the analysis time consuming if the original dataset consists of more than 25-30 data. For this reason, the Monte Carlo method is employed to randomly draw bootstrap samples in order to approximate the bootstrap distribution.

2.2 Posterior estimation using Bayesian Analysis. The Random-Walk Metropolis-Hastings algorithm.

The Bayesian statistic explicitly treats the model parameters as probability density functions. For this reason, the output of an inference performed using Bayesian statistics already includes the information about the uncertainty underlying the model parameters. The Bayesian statistics involves the use of the Bayes theorem and gives a method for updating the probabilities of unobserved events, given that another related event has occurred. The Bayes theorem allows the evaluation of the conditional probability of observing the event A given that the event B has occurred:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (7)$$

where P(A), the probability of observing A, is the prior information about the event A; P(B) is the probability of observing B and P(B|A) is the conditional probability of observing B given that A has occurred. P(A|B) in Equation (7) is usually referred to as the “posterior”. Recently, it has been applied to infer fatigue test data in [8, 11]. It is usually referred as non-informative (NBI) when no prior information is employed. Instead, the informative (IBI) involves the use of prior information on the model parameters. The random walk Metropolis-Hastings algorithm [12, 13] has been employed to perform the analysis. It is a Markov Chain Monte Carlo (MCMC) method giving as output a Markov Chain which allows estimating the posterior distribution. The algorithm is based on the following steps:

- 1) Start from an initial arbitrary set of values for the model parameters;
- 2) Using a symmetric candidate distribution (e.g. normal), whose mean is the current state of the Markov Chain and with standard deviation to be tuned, sample a new candidate set for the model parameters;
- 3) Based on the criterion described in [16–18], accept or reject the candidate value based on the likelihood function and the prior knowledge. In case of acceptance, the candidate value becomes the state of the Markov Chain at the current time, otherwise, the previous state is kept at the current time.
- 4) Repeat the steps 2 and 3 until the Markov Chain converges to its limiting distribution.

In order to reach convergence, the candidate distribution should not be too narrow or too wide. The limiting distribution obtained through his procedure is a non-parametric estimation of the posterior. More details about the random walk Metropolis-Hasting algorithm and the tuning process can be found in [12–16], among others. As result of this procedure, the posterior distribution is approximated and its mean value identifies the Bayesian mean estimator of the parameter.

3. Results and Discussion

In order to show the degree of affinity of the proposed methods, the two fatigue datasets are inferred by employing the LRM, which is based on the Basquin relation, Equation (1), assuming that the base 10 logarithm of the fatigue life follows a normal distribution. The analysis involves the estimation of the parameters and both the aleatory and epistemic uncertainty according to Table 3. The effect of the epistemic uncertainty on the quantiles of the fatigue life is estimated using the approach proposed in [6, 7]. Then, the 6PRFLM is employed in order to infer the “large sample” data. The uncertainty is estimated and reduced following the scheme in Table 3. The different methods are discussed and the effect of the reduction of the uncertainty is also quantified.

Table 3. Simulation scenario

Fit and epistemic uncertainty		LRM	6PRFLM
Least Square Method	Student T-distribution	Y	N
Maximum Likelihood Method	Wald Statistic	Y	Y
	Bootstrap	Y	Y
Non-Informative Bayesian Inference	MCMC	Y	Y
Informative Bayesian Inference	MCMC	N	Y

3.1 The Linear Regression Model

The methods summarized in Table 3 can be employed to infer the data, estimate the model parameters β_0, β_1, σ and the associated uncertainty. In order to perform the fit using the Bootstrap and the Bayesian inference, the number of samples generated is equal to one million.

3.1.1 The large sample dataset

For the failure data considered in the large sample

Table 4. Results of the LRM fitting the “large sample” data. The estimators of the model parameters and the COV are reported for each method employed.

	Model Parameters		
	β_0	β_1	σ
	Estimators		
LSM	1.22E+01	-3.29E+00	1.47E-01
MLM	1.22E+01	-3.28E+00	1.46E-01
NBI	1.22E+01	-3.29E+00	1.47E-01
	Coefficient of Variation [%]		
LSM	0.99	1.83	5.05
Wald	0.98	1.81	4.98
Bootstrap	1.04	2.53	6.11
NBI	0.97	1.80	5.09

dataset, the estimators of the LRM parameters and the coefficient of variation are reported in Table 4. The value of the log-likelihood function at the MLE is 102.18. However, such value is related to the failure data only. By inferring the whole dataset using the LRM and

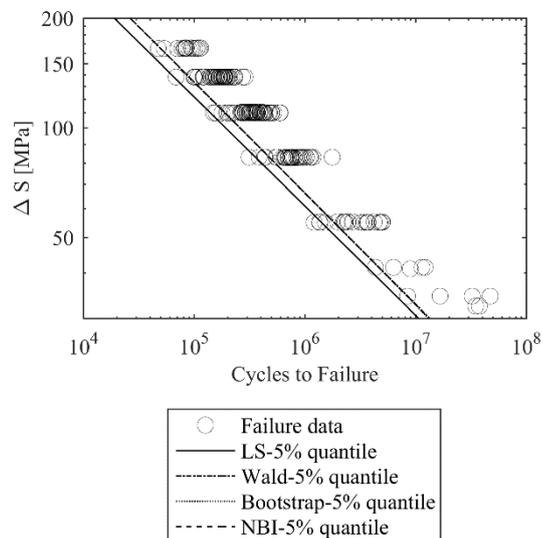


Figure 1. Fatigue test data belonging to the large sample dataset and 5% quantile of the fatigue life obtained for the linear regression model considering the epistemic uncertainty.

the MLM thus, including the contribution of runouts, the value of the log-likelihood function drops to 3.43. The estimators obtained using the different methods are equal. It is known that the mean Bayesian posterior estimator tends to be equal to the MLE with increasing the sample size. The small difference observed is addressed to the use of the numerical procedure employed to determine the MLM, the Bootstrap and the NBI based estimators.

The normality of the distributions describing the epistemic uncertainty has been verified in each case by using the Kolmogorov-Smirnoff test, leading in any case

to “p-values” close to 1. In this case, the LSM, the MLM and the NBI are equivalent and would lead to a close estimation of each quantile of the fatigue life,

Table 5. Results of the LRM fitting the “small sample” data. The estimators of the model parameters and the COV are reported for each method employed.

	Model Parameters		
	β_0	β_1	σ
	Estimators		
LSM	1.28E+01	-3.56E+00	2.05E-01
MLM	1.28E+01	-3.56E+00	1.84E-01
	Coefficient of Variation [%]		
LSM	3.77	6.96	29.74
Wald	3.01	5.56	21.32
Bootstrap	4.32	7.64	26.75

when epistemic uncertainty is not considered. In order to consider both the epistemic and the aleatory

uncertainty, a nested sampling has been performed in a Monte Carlo framework as described in [9-10].

The 5% quantile of the linear regression model has been estimated considering the effect of the epistemic uncertainty. From Figure 1 it results that, despite the estimators of the model parameters and the coefficient of variation of the marginal distribution are comparable for each model parameter, the 5% quantile predicted by the least square method predicts a lower fatigue life (-21.93% at $\Delta\sigma=100$ MPa) than the other methods, that at maximum differ by 3%, in absolute value. This large difference obtained by employing the linear regression theory is mostly associated with the fact that in the linear regression theory the correlation between the model parameters is not taken into account. Instead, the correlation is estimated when the Wald statistic and the Bootstrap method are employed. However, in the former case the correlation coefficient is estimated at the mode of the Likelihood function through the Fisher Information Matrix, and, in the latter case, the correlation is evaluated through the resampling procedure.

3.1.1 The small sample dataset

For the ‘small sample’ analysis, the estimators of the model parameters are evaluated only using the LSM and the MLM because the NBI using the Random Walk Metropolis-Hastings algorithm results in a Markov-Chain that hardly converges to its limiting distribution. In other words, the amount of test data is so small that the resulting Markov-Chain was very sensitive to the values of the tuning parameter, the standard deviation of the candidate distribution. For this reason, it was not possible to obtain a converging Markov Chain by using a “trial and fail” procedure.

As shown in Table 5, the estimators of the parameters β_0 and β_1 are in good agreement. Because of the small amount of test data, the LS estimator of the parameter σ is 1.1 times higher than that one obtained by the MLM. Moreover, as it can be seen in Figure 2, also the estimation of the epistemic uncertainty led to different results. In general, the coefficient of variation is in any case higher than for the “large sample” dataset, meaning that due to the smaller amount of data the epistemic uncertainty increased. When the Wald statistic is employed, the epistemic uncertainty are normally distributed. Instead, the distribution resulting by employing the Bootstrap is of a different type. For example, in Figure 2 it results that the bootstrap distribution associated to the parameter σ is bi-modal; this can be seen because the trend of the cumulate distribution function shows two inflection points, the

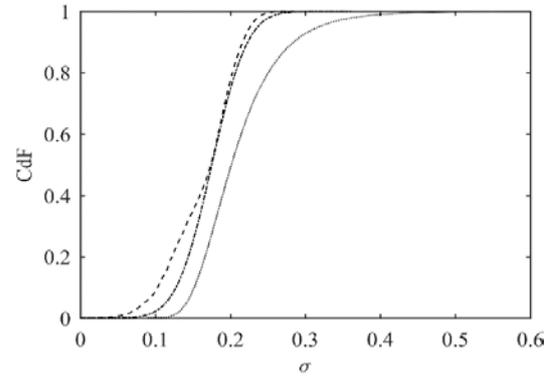


Figure 2. Epistemic uncertainty related to the parameter σ of the LRM fitting the small sample data.

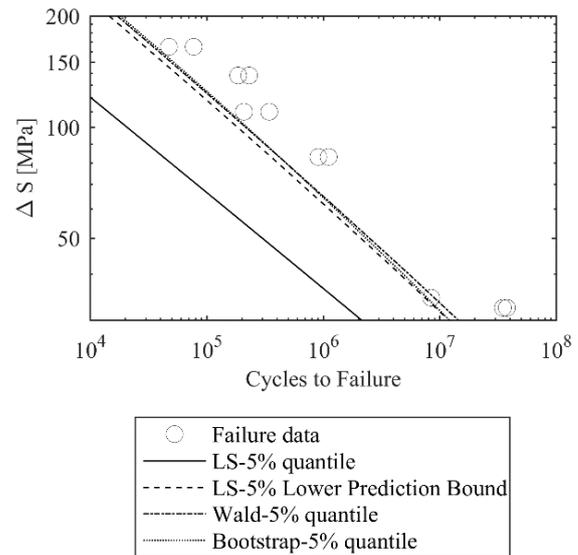


Figure 3. Fatigue test data belonging to the small sample dataset and 5% quantile of the fatigue life obtained for the linear regression model considering the epistemic uncertainty. For comparison, also the 5% lower prediction bound is plotted.

first at a cumulate probability value around 0.2 and a second one which is around 0.6.

The estimated quantiles of the fatigue life are shown in Figure 3. As for the “large sample” dataset, the results are obtained considering the epistemic uncertainty. LSM predicts a 5% quantile of the fatigue life that is 0.11 times the life predicted by the lower prediction bound as defined in the linear regression theory. The 5% quantile of the fatigue life estimated considering the evaluation of the uncertainty given by the Bootstrap and the Wald statistic are closer to each other and predict a fatigue life respectively 1.18 and 1.19 times higher than the 5% lower prediction bound as defined in the linear regression theory.

3.2 The Six-Parameters Random Fatigue Limit Model

The 6PRFLM has been employed to infer the large sample data according to the scheme reported in Table

Table 6. Prior distributions employed in the informative Bayesian inference, from [11].

6PRFLM Parameter	Estimators	
$-\beta_1$	$\mu_{\log(-\beta_1)}$	1.15E+00
	$\sigma_{\log(-\beta_1)}$	2.00E-01
μ_v	$m = \mu_{\mu_v}$	1.55E+00
	n'	20
σ_v	$s' = \mu_{\sigma_v}$	7.24E-02
	v'	20

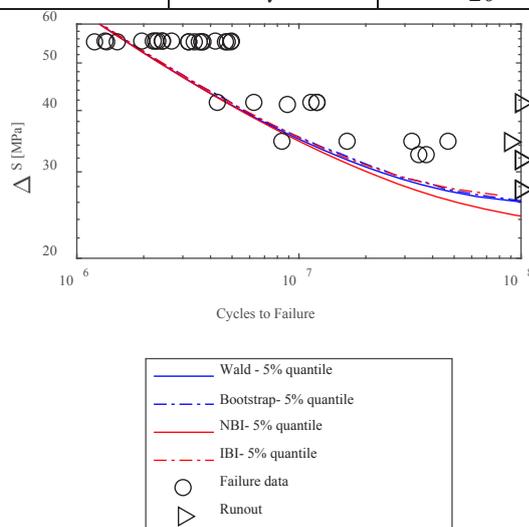


Figure 4. Large sample data inferred using the 6PRFLM. Only data at $\Delta S < 60$ MPa are shown. The median curve is obtained using the MLE. Moreover, the 5% quantile of the fatigue life is plot considering the epistemic uncertainty.

3. In the IBI, two prior distributions have been formulated for the parameters related to the slope of the finite life region and the parameters associated with the location and the scatter of the fatigue limit.

In the first case, the prior information is obtained from the fatigue test data of transverse loaded welded connections reported in [19], where the LRM has been employed in order to estimate the slope parameter for each sample. To formalize the prior distribution of $-\beta_1$, in [11] its estimators have been collected for each data sample and a lognormal distribution have been fitted to them, its parameters are summarized in Table 6.

In order to formalize the prior information on the location and scatter of the fatigue limit, the probability distribution introduced in [20] has been employed:

$$f = k\sigma_v^{-[v'+\delta(n')+1]} \exp\left\{-\frac{1}{2\sigma_v^2} [v' s'^2 + n'(\mu_v - m')^2]\right\} \quad (8)$$

where s' is the prior expectation of the standard deviation, v' is the number of degrees of freedom for s' , m' is the prior expectation of the mean value and n' is the number of degrees of freedom for m' . If $n'=0$, the function $\delta(n')=0$, otherwise $\delta(n')=1$. In case of vague information, the degree of freedom should be zero

whereas a value between 20 and 40 can be assigned if experience leads to almost deterministic knowledge.

The parameters have been estimated in [11] and are reported in Table 6. The estimators of the model parameters and the coefficient of variation and the skewness related to their distributions are reported in Table 7. The value of the log-likelihood function at the MLE is 100.26. This is lower compared to the value obtained by the LRM, when runouts were not included. Instead, it is significantly higher than the value of the log-likelihood function obtained by inferring the complete dataset using the LRM and the MLM. The MLE and the non-informative Bayesian estimator are almost equal to each other and also the estimated coefficients of variation are close. The differences are

and the uncertainty affecting the parameter p are significantly reduced.

It should also be noted that with respect to the determination of quantiles of the fatigue life considering the epistemic uncertainty, it is possible that in the Monte Carlo procedure negative values of the standard deviation of the fatigue limit and the parameter p might be sampled if the results of the Wald statistic are employed. For this reason, when the quantiles of the fatigue life are computed by nested sampling the aleatory and the epistemic uncertainty estimated by the Wald statistic, the parameters σ , σ_v and p cannot be negative. Thus, a truncation scheme should be applied in order to limit the sampling to the domain of the model parameter.

The highest difference in the fatigue life predicted in the different frequentist scenarios is expected to be in the transition from the finite to the infinite life region. By comparing the predicted 5% quantiles of the fatigue limit distribution, the maximum difference obtained is less than 1% for the frequentist methods and using the Wald statistic the predicted 5% quantile is 25.27MPa. IBI predicts 26.17MPa, which is 1.16 times higher than the NBI. Despite the Bootstrap allows a better estimation of the epistemic uncertainty, the fatigue life predicted by the 5% quantile does not significantly differ by that one estimated using the MLM and the Wald statistic, even at relatively low stress ranges, see Figure 4.

4. Conclusions

By the analyses performed in the present work, the following conclusions can be drawn:

- (1) In order to include epistemic uncertainty in the estimation of the fatigue life distribution by performing nested sampling of epistemic and aleatory uncertainty in a Monte-Carlo simulation, the correlation between the model parameters should be estimated. Thus, the

Table 7. Results of the LRM fitting the “large sample” data. The estimators of the model parameters and the COV are reported for each method employed.

	Model Parameters					
	β_0	β_1	σ	μ_v	σ_v	p
	Estimators					
MLM	1.11E+01	-2.78E+00	1.32E-01	1.54E+00	7.27E-02	4.74E-01
NBI	1.09E+01	-2.70E+00	1.29E-01	1.52E+00	8.98E-02	6.70E-01
IBI	1.12E+01	-2.82E+00	1.34E-01	1.54E+00	7.36E-02	4.31E-01
	Coefficient of Variation [%] / Skewness					
Wald	2.88 / -0.05	5.17 / 0.05	5.87 / -0.00	1.47 / -0.01	28.20 / 0.02	38.21 / 0.08
Bootstrap	3.42 / 0.30	6.11 / 0.25	7.29 / 0.26	1.88 / -0.83	31.40 / -0.34	46.45 / 1.62
NBI	3.15 / -0.33	5.61 / 0.26	6.59 / -0.09	2.19 / -1.16	24.14 / 0.48	44.05 / 2.19
IBI	2.08 / -0.11	3.72 / 0.07	5.58 / 0.28	0.83 / -0.18	14.19 / 0.61	27.93 / 0.58

addressed to the small skewness of the distributions that never exceeded the interval ± 2 , except for the parameter p , which is also the parameter more affected by uncertainty. As result of the IBI, both the skewness

linear regression theory cannot be employed to estimate the uncertainty underlying the model parameters.

- (2) If the number of test data is relatively small, the Bayesian Inference employing the

Random-Walk Metropolis-Hastings algorithm determines a Markov-Chain appears to be sensitive to the scale parameter of the candidate distribution. This resulted in difficulties in estimating the posterior distribution of the model parameters.

- (3) The inference performed using the 6PRFLM determines a higher value of the log-likelihood function than the LRM when the “large sample” dataset is employed.
- (4) For the cover plate steel detail, the fatigue life predicted in the transition from finite to infinite life region is affected by large uncertainty. This is due to the small amount of test data executed at relatively small stress range levels. It has been found that by employing a prior distribution for the parameters of the fatigue limit distribution this uncertainty decreases but remains still high.
- (5) Despite the Bootstrap, different from the Wald statistic, allows any shape of the distribution of the model parameters, the difference of the 5% quantile of the fatigue life obtained considering the epistemic uncertainty for the 6PRFLM fitting the “large sample” data is negligible. Thus, for relatively large samples inferred with complex models such as the 6PRFLM, the computational effort spent to perform the Bootstrap estimation of the epistemic uncertainty does not seem to lead to a better prediction.

The authors would like to acknowledge the Dutch infrastructural asset owners Rijkswaterstaat and ProRail and the research organization TNO for their support.

References

- [1] C. Schneider and S. Maddox, “Best practice guide on statistical analysis of fatigue data,” 2003.
- [2] M. B. A. Van Asselt, *Perspectives on Uncertainty and Risk - the PRIMA Approach to Decision Support*. 2000.
- [3] M. H. Faber, “On the Treatment of Uncertainties and Probabilities in Engineering Decision Analysis,” *J. Offshore Mech. Arct. Eng.*, vol. 127, no. 3, p. 243, Aug. 2005.
- [4] J. Fisher, P. Albrecht, B. Yen, and D. Klingerman, “Effects of weldments on the fatigue strength of steel beams-transverse stiffness and attachments, September 1972,” *Fritz Lab. Reports*, Jan. 1972.
- [5] R. Slockbower, “Fatigue resistance of full scale cover-plated beams.,” *Theses Diss.*, Jan. 1976.
- [6] O. H. Basquin, “The exponential law of endurance tests,” *Proc ASTM*, vol. 10, pp. 625–630, 1910.
- [7] J. Schijve, “Statistical distribution functions and fatigue of structures,” *Int. J. Fatigue*, vol. 27, no. 9, pp. 1031–1039, 2005.
- [8] F. G. Pascual and W. Q. Meeker, “Estimating Fatigue Curves with the Random Fatigue-Limit Model,” *Technometrics*, vol. 41, no. 4, pp. 277–290, 1999.
- [9] Luca D’Angelo and Alain Nussbaumer, “Estimation of fatigue S-N curves of welded joints using advanced probabilistic approach. Under review: IJFATIGUE-D-16-00685,” *Int. J. Fatigue*, vol. 97, pp. 98–113, 2017.
- [10] L. D’Angelo, M. Rocha, A. Nussbaumer, and E. Brühwiler, “S-N-P Fatigue Curves Using Maximum Likelihood: Method for fatigue resistance curves with application to straight and welded rebars,” *Eurosteel 2014*, no. 1, pp. 1–8, 2014.
- [11] D. Leonetti, J. Maljaars, and H. H. (Bert. Snijder, “Fitting fatigue test data with a novel S-N curve using frequentist and Bayesian inference,” *Int. J. Fatigue*, 2017.
- [12] Y. Pawitan, *In all likelihood: statistical modelling and inference using likelihood*. Clarendon Press, 2001.
- [13] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses,” *Ann. Math. Stat.*, vol. 9, no. 1, pp. 60–62, 1938.
- [14] P. Diaconis and B. Efron, “Computer-intensive methods in statistics,” *Sci. Am.*, vol. 248, no. 5, pp. 116–130, 1983.
- [15] I. Babuska, Z. Sawlan, M. Scavino, B. Szabó, and R. Tempone, “Bayesian inference and model comparison for metallic fatigue data,” Dec. 2015.
- [16] W. K. Hastings, “Monte carlo sampling methods using Markov chains and their applications,” *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
- [17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of State Calculations by Fast Computing Machines,” *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, Jun. 1953.
- [18] W. M. Bolstad, *Understanding Computational Bayesian Statistics*. 2012.
- [19] G. Sedlacek, A. Hobbacher, J. B. Schleich, A. Nussbaumer, S. J. Maddox, and J. Brozzetti, “First draft of the background document prEN 1993-1-9.”
- [20] JCSS, “Background documentation Eurocode 1 (ENV 1991) - Part 1: Basis of design.,” 1996.