

Structural reliability updating on the basis of proof load testing and monitoring data

R. de Vries^{a,b,*}, E.O.L. Lantsoght^{a,c}, R.D.J.M. Steenbergen^{a,d}, M.A.N. Hendriks^{a,e},
M. Naaktgeboren^f

^a Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands

^b Reliable Structures, Netherlands Organization for Applied Scientific Research (TNO), Delft, the Netherlands

^c College of Sciences and Engineering, Universidad San Francisco de Quito (USFQ), Quito, Ecuador

^d Faculty of Engineering and Architecture, Ghent University, Ghent, Belgium

^e Department of Structural Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

^f Rijkswaterstaat, Ministry of Infrastructure and Water Management, Utrecht, the Netherlands

ARTICLE INFO

Keywords:

Bayesian updating
Laboratory testing
Load testing
Proof load
Reliability

ABSTRACT

As infrastructure continues to age and traffic levels intensify, there is a growing need for efficient methods to verify the reliability of many existing structures. Field testing offers the possibility to assess the current condition of a structure. Specifically, in a proof load test, substantial loads are applied to evaluate the structure's resistance to future loads that could compromise structural safety. However, to prevent excessive test loads and their potential damage, it is desirable to assess structural reliability by monitoring the response under more moderate loads. This study merges laboratory and in-situ testing results through a Bayesian update of the structural reliability after each successful load application. Two case studies are presented where laboratory testing on structurally similar elements and analytical modelling provide ample evidence to justify test load reductions of 20 % and 25 %. The proposed method offers a systematic framework to link the structure's response during testing to structural reliability and address the uncertainties in resistance, loads and measurements. Nonetheless, the representativeness of the data in terms of structural similarity and uncertainties related to measurements continue to be significant factors. Despite these challenges, incorporating monitoring data during proof load testing is expected to reduce target loads in most cases.

1. Introduction

Buildings and civil engineering works are expected to meet specific reliability requirements throughout their entire design life. Reliability assessment of an existing structure becomes relevant when the structure displays performance issues, the loads have significantly increased, or its original design life has passed. A design that may have been sufficient in the past may not be adequate today. Over time, degradation may have taken place, and the traffic loads have predominantly increased. Typically, assessing reliability requires in-depth information about the structural model, failure mechanisms, the description of loads and their combination. Moreover, for existing structures, assumptions regarding the uncertainties made in the past may no longer be true today as the knowledge about resistance and load models has evolved. Therefore, the original design reliability, based on prior knowledge, should be updated

to reflect the current state of the structure. In addition, the reliability requirements for the design of new structures are higher than those used for the assessment of existing structures, following from an economic motive [1,2]. In this article, comparisons will be made of the required target loads at reliability levels suitable for the assessment of existing structures

Inspections, structural assessments, and maintenance are essential to ensure sufficiently reliable bridges and viaducts. As the infrastructure ages and endures increased traffic loads and environmental challenges, accurate reliability assessment methods are needed to address these evolving conditions. In case no signs of deterioration are present, the typical desk study may confirm sufficient reliability if there are no reasons to suspect internal damage. However, wear is often present, and it is difficult to tell if it impairs structural reliability. Fortunately, tests can be carried out on the structure to gather supplementary data. Tests on reinforced concrete structures commonly entail measuring the

* Corresponding author at: Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands.

E-mail address: rein.devries@tudelft.nl (R. de Vries).

Nomenclature*List of abbreviations*

CDF	Cumulative distribution function
COV	Coefficient of variation
DIC	Digital image correlation
FEM	Finite element method
FORM	First-order reliability method
LHS	Latin hypercube sampling
LM1	Load model 1 (Eurocode)
MCMC	Markov chain Monte Carlo
MCS	Monte Carlo simulation
PLT	Proof load test
SORM	Second-order reliability method
WIM	Weigh-in-motion

List of symbols

C_{0Q}	Time-independent variability of traffic load
E	Load effect
E_{PL}	Load effect during proof load testing
E_s	Young's modulus of reinforcing steel
f_c	Concrete compressive strength
f_y	Yield stress of reinforcement
$F_X(\cdot)$	CDF of random variable X
G_{DL}	Dead load effect
G_{SDL}	Superimposed dead load effect
I	Indicator of performance
$m_{Q,PL}$	Mean proof load effect
m_X	Mean of resistance ratio
M	Moment, mean (sample)
M_G	Moment due to dead load
n	Sample size
N	Number of samples (MCS)

P	Load of an individual axle in proof load test
$p_X(\cdot)$	Probability density function of X
P_f	Failure probability
Q	Time-variant part of traffic load
$Q_{k,WIM}$	Characteristic traffic load following from WIM data
$Q_{k,LM1}$	Characteristic traffic load following from Eurocode LM1
Q_{PL}	Target load (proof load)
R	Resistance
s_X	Standard deviation of resistance ratio
S	Standard deviation (sample)
T	Student's t -distributed random variable
U	Standard normally distributed random variable
$\text{Var}(\cdot)$	Variance (operator)
V_X	Coefficient of variation of resistance ratio
$w_{\max}$	Maximum nominal crack width
$w_{\max,w}$	Maximum weighted nominal crack width
$\mathbf{x}$	Data vector (Bayesian)
X	Random variable, resistance ratio (R/E_{PL})
Z	Limit state function
Z_{PL}	Limit state function for proof load testing
α	Influence coefficients vector
β	Reliability index
ε_s	Steel strain
θ	Model parameters vector (Bayesian)
θ_E	Model uncertainty of the load effect
$\theta_{E,PL}$	Model uncertainty of the proof load effect
θ_R	Model uncertainty of the resistance
μ	Mean (population)
ν	Degrees of freedom (Student's t -distribution)
σ	Standard deviation (population)
$\Phi(\cdot)$	Standard normal CDF

geometry, drilling cores or scanning reinforcement. In some instances, elaborate setups that subject the structure to significant loads may be employed to test its resistance. The latter is referred to as *proof load testing*; by resisting a large load, the structure can prove to have sufficient resistance. However, applying the often large loads is resource-intensive and imposes a risk on the structure, equipment and personnel. To avoid excessively large loads, all relevant information about the structure should be considered, even when uncertain. Given this uncertainty, employing probabilistic techniques is necessary. Utilising all information, rather than often conservative design rules, can avoid excessive target loads. By rationally selecting the appropriate load level, proof load testing can effectively demonstrate the structure's resistance to anticipated future traffic loads [3–7].

The current research on the probabilistic substantiation of proof load testing aims to develop a comprehensive structural reliability updating framework [8]. This article presents a novel reliability updating method that integrates information from two distinct sources: (1) the survival of the applied load during the proof load test and (2) the data following from monitoring relevant indicators during the test, coupled to laboratory experiments giving the uncertain relation between these indicators and structural resistance [9]. Highly representative tests can significantly enhance the state of information, and thereby reduce resistance model uncertainty. Alternatively, tests on less representative specimens can be used, but will result in greater uncertainty. While the proof load is carefully increased in controlled increments to avoid unnecessary damage, the structural performance assessment may be based on indicators such as displacements and crack widths. These indicators, while not immediately indicative of overall structural health, are interpreted in the light of structural behaviour observed in laboratory experiments

on structural elements similar to those present in the proof loaded bridge. For example, strains may be interpreted via sectional analysis to identify a critical value. In this way, stop criteria can prevent unwanted damage [10], but the indicator value may also provide valuable information on structural performance.

To probabilistically interpret the information from indicators, first the theoretical background on structural reliability, statistical inference, and their relation to proof load testing is described. Then, the reliability updating method on the basis of proof load testing and monitoring data is presented. To illustrate the method's application, two case studies utilising laboratory measurements and analytical modelling are provided. The article rounds off with a discussion of the results, challenges, and conclusions.

2. Theoretical background

2.1. Structural reliability

The proposed method (Section 3) relies on the principle of structural reliability updating, which requires reliability assessment. A structural reliability assessment is based on a probabilistic model that includes a limit state function and the definition of random variables describing the load and resistance parameters and the modelling uncertainties. Reliability methods are used to calculate the reliability (or failure probability) of a structure or a structural component. The limit state function plays a central role as it expresses the boundary between safe and unsafe combinations of resistance and load effect. Negative values of the limit state function indicate failure, irrespective of the magnitude. An example of a typical limit state function involving just two random

variables is:

$$Z = g(\mathbf{X}) = g(R, E) = R - E \quad (1)$$

where R is the resistance and E is the load effect. The load effect may comprise the bending moment, shear force, axial load, and so on. Often, the combined effects of multiple loads are considered [11–13]. The state of information concerning the resistance and load effect varies between structures. Various sources of information can be utilised, ranging from literature to in-situ material testing and site-specific traffic data [8]. The Eurocode standards also allow for design assisted by testing and guidance is provided regarding the sampling process and statistical post-processing [14].

Commonly, the number of random variables is larger than just two, and the limit state function is more difficult to compute. There may be a very large number of random variables, and instead of having expressions in closed or analytic form available, complex finite element method (FEM) models may be needed to evaluate the limit state functions. For this reason, different reliability computation methods have been developed, and all have their pros and cons depending on the application. The calculation procedure for the proposed method (Section 3.5) makes use of the Markov chain Monte Carlo (MCMC) sampling [15], which in turn is based on Monte Carlo simulation (MCS). MCS is a straightforward method applicable to many reliability problems but is computationally expensive [16]. In an MCS, the random variables present in the problem formulation ($\mathbf{X}$) are repeatedly sampled and may be used to evaluate the limit state function. The probability of failure is obtained by calculating the fraction of failures that occur:

$$P_f = \frac{1}{N} \sum_{i=1}^N \mathbf{1}[g(\mathbf{x}_i) < 0] \quad (2)$$

where $\mathbf{1}[\cdot]$ denotes the indicator function; it is 1 when the condition within brackets is true and 0 otherwise. The number of samples is denoted by N , and the random vector containing the values of each sample is $\mathbf{x}_i$. The corresponding reliability index may be calculated via:

$$\beta = \Phi^{-1}(1 - P_f) = -\Phi^{-1}(P_f) \quad (3)$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution function (CDF) (i.e., mean = 0, standard deviation = 1). Reliability requirements are commonly formulated in terms of the reliability index and a reference period to account for the time-dependent nature of random processes. One way to unify reliability requirements for design and assessment is through the adoption of annual reliability targets [17, 18].

When the failure probability is small, many samples are required to estimate the reliability index accurately. Improved sampling methods [19–23] or other reliability methods are beneficial in such circumstances. After simplification, common reliability methods can be used in the proposed reliability updating method (Section 3.5). For example, using the first-order reliability method (FORM), a computationally efficient method with easily understandable output: the reliability index (β) and influence coefficients (α), indicating the relative importance of the random variables [24,25]. Further improvement of the reliability index is possible with the second-order reliability method (SORM) [26, 27] which utilises the second-order derivatives of the limit state function in the design point. After eigenvalue analysis or calculating determinants [28], they yield a correction factor to the FORM failure probability. Slightly different versions of the same idea were introduced by various authors, offering a small increase in accuracy [29,30].

2.2. Statistical inference

2.2.1. Principles

Deriving the statistical descriptions of the random variables within the limit state function is necessary to enable the reliability assessment.

These random variables represent uncertainties in material properties, loads, and modelling approaches. When tests are performed in the laboratory, the number of specimens is usually limited due to cost, time, and material availability constraints. Therefore, the number of tests and resulting data points is usually small and cannot be expected to capture the inherent variability fully. This also applies to calculating the resistance ratio distribution based on laboratory tests (Section 3.3.1). Both frequentist and Bayesian approaches can be used to infer the statistical descriptions and account for the limited number of tests. The frequentist approach solely uses the sample statistics of the observed data. The prediction distribution allows for the inclusion of statistical uncertainty due to small sample sizes and known and unknown standard deviations. On the other hand, the Bayesian approach also allows for incorporating prior knowledge to deliver posterior distributions that incorporate both data and subjective knowledge. The calculation of the posterior probability distributions typically requires numerical methods, such as Markov chain Monte Carlo (MCMC) sampling [15].

2.2.2. Prediction distribution

Given n observations of an unknown random variable X , the prediction distribution describes the probability distribution of the next-to-be-observed value X_{n+1} . Generally, two situations are distinguished: one where the standard deviation is known, and the other where it must be estimated from the data. If the observations X_i come from the same normally distributed population, are independent and identically distributed (i.i.d.), it follows that [31,32]:

$$U = \frac{X_{n+1} - M}{\sigma\sqrt{1 + 1/n}} \sim \mathcal{N}(0, 1) \quad (4)$$

where U is a standard normally distributed random variable, $M = (X_1 + X_2 + \dots + X_n)/n$ is the sample mean, and σ is the population standard deviation. Only when data is observed, random variable M becomes a realisation of the sample mean (denoted with lowercase m). The normalising term follows from considering the variance of the difference in the numerator, i.e. $\text{Var}(X_{n+1} - M) = \text{Var}(X_{n+1}) + \text{Var}(M) = \sigma^2 + \sigma^2/n = \sigma^2(1 + 1/n)$. Taking its square root gives the denominator in Eq. (4). Intuitively, it may be understood as the standard deviation following from random variable X_{n+1} and the, typically smaller, standard deviation of the sample mean of values 1 to n . Solving Eq. (4) for X_{n+1} gives:

$$X_{n+1} = M + U\sigma\sqrt{1 + 1/n} \quad (5)$$

where the right-hand side may be interpreted as the prediction distribution of X and follows a normal distribution. It should be realised that X_{n+1} will not actually follow the prediction distribution, but Eq. (5) allows for incorporating the uncertainty about the mean, similar to the posterior predictive distribution in Bayesian inference (Section 2.2.3). The ‘penalty’ incurred by estimating the population mean with the sample mean is contained in the increased standard deviation and diminishes with an increasing number of observations n .

In case the standard deviation of X is not known and needs to be inferred from the data as well, Student’s t -distribution emerges [33]. The t -distribution is wider than the normal distribution, reflecting the more significant uncertainty when the standard deviation is unknown. If, once again, the i.i.d. observations X_i from the same normally distributed population are considered, it follows that [31]:

$$T = \frac{X_{n+1} - M}{S\sqrt{1 + 1/n}} \sim t_{\nu=n-1} \quad (6)$$

where T is a (standard) t -distributed random variable with $\nu = n - 1$ degrees of freedom and S is the sample standard deviation including Bessel’s bias correction for the variance:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2} \quad (7)$$

The same normalisation consideration holds for the denominator in Eq. (6) as in Eq. (4). Solving Eq. (6) for X_{n+1} gives:

$$X_{n+1} = M + TS\sqrt{1 + 1/n} \quad (8)$$

where the right-hand side may be interpreted as the prediction distribution of X and follows a shifted and scaled t -distribution. Compared to Eq. (5), another ‘penalty’ is introduced by using the random variable T instead of U , because the t -distribution is typically wider. In some cases, an intermediate approach is followed where n is artificially increased when (prior) information suggests a similar standard deviation as calculated from the dataset. In addition, knowledge about the type of distribution may also be incorporated. It is typical to assume a lognormal distribution for material properties and, thus, apply the above procedure to log-transformed values [14,34].

2.2.3. Bayesian inference

The Bayesian inference process enables the integration of prior knowledge to improve the statistical description of data. It is a systematic process in which *prior* beliefs are updated according to the available information (or evidence, or data), resulting in *posterior* beliefs. In a Bayesian inference context, usage of the t -distribution is equivalent to adopting a non-informative prior for the parameters of the normal distribution. However, in many cases, a non-informative prior is too generic and does not adequately represent prior knowledge or constraints known about the parameters, leading to overly conservative posterior estimates. In such scenarios, incorporating more informative priors by leveraging expert judgement and historical data leads to more realistic outcomes [35]. Bayesian inference can be used to update beliefs about important parameters in the structural reliability analysis. The reliability of the structure may be re-evaluated each time new information becomes available, resulting in Bayesian reliability updating [36–38]. In a sequential updating, scheme the posterior distribution following from the previously acquired data is used as a prior for the next iteration [39,40].

In Bayesian inference, the distribution of a random variable X itself may be updated, or the parameters of its assumed model. In the latter hierarchical model, the distribution parameters are modelled as random variables with specific distributions as well. The prior distributions may then be specified by providing values for the hyperparameters, e.g. the mean and standard deviation of the mean of X . In this scenario, the model parameters are the mean and standard deviation of X , collected in $\theta = (\mu_X, \sigma_X)$, and may be updated through Bayes’ theorem:

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta)p(\theta)}{\int p(\mathbf{x}|\theta)p(\theta) d\theta} \quad (9)$$

where θ is a vector containing the model parameters (random variables), $\mathbf{x}$ is a vector containing the data, $p(\theta|\mathbf{x})$ is the posterior distribution, $p(\mathbf{x}|\theta)$ is the likelihood of observing the data given the model parameters, $p(\theta)$ is the prior distribution, and the denominator is called the marginal likelihood and acts as a normalising constant. If prior information about the parameters to be inferred is not available, a non-informative prior may be used. In the typical case with the mean and the standard deviation as model parameters, the non-informative prior is obtained by the following (improper) probability density functions [41]:

$$p(\mu_X) = 1 \quad (10a)$$

$$p(\sigma_X) = 1/\sigma_X \quad (10b)$$

from which the prior distribution follows as $p(\theta) = p(\mu_X)p(\sigma_X) = 1/\sigma_X$. Because the resulting prior distribution is improper, it cannot be sampled directly. Instead, a calculation procedure is required that draws its random values in a different way, such as Markov chain Monte Carlo (MCMC) [42,43] or Importance Sampling [19]. When the often multi-dimensional posterior distribution $p(\theta|\mathbf{x})$ has been obtained, it can be

used to update the probability distribution of X itself, i.e. the posterior predictive distribution or Bayes’ distribution [44]:

$$p_X^{\text{Bayes}}(x) = \int p_X(x|\theta)p(\theta|\mathbf{x}) d\theta \quad (11a)$$

$$F_X^{\text{Bayes}}(x) = \int F_X(x|\theta)p(\theta|\mathbf{x}) d\theta \quad (11b)$$

where $p_X(\cdot)$ denotes the probability density function and $F_X(\cdot)$ the cumulative distribution function of random variable X .

2.3. Reliability following from proof load testing

In proof load testing, the so-called *target load* plays a central role, as it is specifically chosen to simulate today’s and future usage conditions and account for a degree of uncertainty. The magnitude of the target load controls the desired reliability of the structure, as higher reliability demands require higher loads during testing. If a structure withstands the target load without signs of distress [45], it is deemed sufficiently reliable for continued operation after the test. Selecting the target load based on reliability considerations may be done in different ways, depending on how much information is available about the structure. For a reliability assessment of a bridge or viaduct, at minimum, a statistical description of expected traffic loads is required. By assuming that the resistance of the tested bridge or viaduct is at least equal to the permanent load effects and the target load effect ($R \geq G + Q_{\text{PL}}$), the limit state function, including model uncertainties, may be written as [5]:

$$Z = \theta_{E,\text{PL}}Q_{\text{PL}} - \theta_E C_{0Q}Q \quad (12)$$

where $\theta_{E,\text{PL}}$ is the model uncertainty of the load effect pertaining to the proof load testing situation (Section 3.2), Q_{PL} is the target load, θ_E is the model uncertainty of the load effect for the regular traffic load situation (correlated with $\theta_{E,\text{PL}}$), C_{0Q} accounts for the time-independent variability of the traffic load, and Q is the time-variant part of the traffic load. Evaluating the limit state function in Eq. (12) is referred to as the *lower-bound approach*, as it provides the most conservative estimate of the posterior resistance distribution [46]. A comparison with target loads obtained through this relatively straightforward method is presented in the case studies (Sections 4 and 5).

Alternatively, the distribution function of the resistance may be explicitly considered. This procedure effectively leads to truncating the left tail of the distribution to exclude the possibility that the resistance is lower than the load effect produced during the test [47]. However, this truncation is not abrupt but rather gradual, owing to some uncertainty about the actual load effect created by the applied load [48]. In principle, this resistance relates to the specific loading position and method of application. On a structural level, the effect of the applied load on a particular component or cross-section is more valuable. Therefore, the limit state, in principle, considers a resistance and load effect, not the applied load itself. The update of the resistance distribution may also be achieved via the application of Bayes’ theorem, and an indicator likelihood function providing the value 0 for resistances lower than the target load and 1 otherwise. With the proof load effect described by a random variable, this procedure also results in an appropriate posterior distribution for the resistance. The prior distribution may then be formulated using a mean value based on the mean annual traffic load and a relatively large coefficient of variation to reflect the large degree of uncertainty about the resistance. When a resistance distribution is available, possibly updated by in-service proven strength, it is also possible to evaluate the reliability during the proof load testing situation [46].

3. Reliability updating method

3.1. Reliability updating using two information sources

After each load increment in the proof load test, the estimation of the resistance on the basis of measurements forms the first source of information (Fig. 1, point 1). The proposed method relates the observed in-situ response, through the use of indicator values, to the response observed in laboratory tests on similar structural elements or derived from suitable analytical models. The indicator value is measured in situ and on representative specimens in the lab. For example, strains derived from horizontal displacement measurements at the bottom of a beam or slab can serve as critical indicators for structural performance. By utilising the same indicators in both laboratory and in-situ testing, it is possible to determine the mean and standard deviation of the structure's resistance. Instead of a mean and standard deviation, the prediction distribution of R may be established via other statistical approaches, including Bayesian inference (Section 2.2).

If a structure can withstand a specific proof load, it also means that its resistance (R) is equal to or greater than the load effect during the test (Section 2.3). Considering the uncertainty of the proof load test, the gradual truncation of the resistance is the second source of information (Fig. 1, point 2). The information from the two sources is processed in the presented order and allows for variable load increments to determine the reliability of the structure. Generally, as the applied load increases, the structural reliability also tends to increase. A flowchart outlining the steps in the proposed proof load testing assessment method is provided in Fig. 2.

3.2. Probabilistic model for reliability updating

Structural reliability may be assessed by evaluating a limit state function (Section 2.1). The primary limit state considers the situation in which the regular traffic loads act on the bridge. The proposed function aligns with the guidelines from the Probabilistic Model Code [34] and fib Bulletin 80 [49]:

$$Z = \theta_R R - \theta_E (G_{DL} + G_{SDL} + C_{0Q} Q) \quad (13)$$

where θ_R is the uncertainty associated with resistance calculation, R is the resistance, θ_E is the model uncertainty of the load effect calculation, G_{DL} is the dead load effect, G_{SDL} is the superimposed dead load effect, C_{0Q} is the time-invariant part of the live load effect, and Q is the time-variant part of the live load effect (i.e. traffic load effect). The reliability of the bridge prior to proof load testing could be evaluated if distributions were assigned to all random variables, including R . Because of the low information state, the distribution of R would incorporate a large variability and thus result in low reliability. Instead of employing a conventional structural resistance model, the resistance R will be estimated by combining in-situ measurements with insights gained from laboratory experiments or analytical modelling. This also affects how the variability of θ_R , the uncertainty associated with

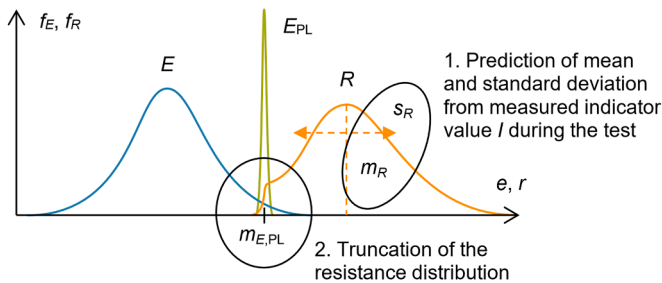


Fig. 1. General principle of updating the resistance distribution from two sources of information.

resistance calculation, should be quantified (Section 3.4).

Instead of directly assigning a distribution to resistance R , it will be expressed as a function of the proof load effect via the factor X . Each time the structure withstands a new load level in a proof load test, the distribution function of R can be updated to reflect the information obtained via measurements, in addition to the truncation (Section 3.1). The revised distribution of R can be expressed as the product of a resistance ratio X and the load effect produced during the last successful load test cycle ($R = X E_{PL}$). Given a certain observed indicator value I , the distribution of X may be obtained – as determined before proof load testing (Section 3.3). During a proof load test, the traffic load will be absent, and instead, the test load is present. Thus, the limit state function describing the proof load testing situation is:

$$Z_{PL} = \theta_R X E_{PL} - E_{PL} = (\theta_R X - 1) E_{PL} \propto \theta_R X - 1 \quad (14)$$

where $E_{PL} = \theta_E (G_{DL} + G_{SDL}) + \theta_{E,PL} Q_{PL}$ with Q_{PL} denoting the load effect created by the proof load and $\theta_{E,PL}$ its corresponding model uncertainty. By assuming that the proof load is withstood, it directly follows that $Z_{PL} > 0$ and thus $\theta_R X > 1$. The conditionality can be satisfied in a Bayesian updating process by obtaining the joint posterior distribution of $\theta = (\theta_R, X)$ as:

$$p(\theta | Z_{PL} > 0) \propto p(Z_{PL} > 0 | \theta) p(\theta) \quad (15)$$

where the likelihood $p(Z_{PL} > 0 | \theta)$ acts as an indicator function, or potential [50,51], and $p(\theta)$ is the prior probability. In a Monte Carlo simulation, this process involves differentiating between samples that either withstand or fail the proof load test (Section 3.5). After this update, the marginal distributions should not be sampled independently because the interdependence of variable combinations significantly influences the outcomes. Returning to the original traffic load situation, Eq. (13), with $R = X E_{PL}$ and denoting the updated variables as θ_R' and X' gives:

$$Z = \theta_R' X' [\theta_E (G_{DL} + G_{SDL}) + \theta_{E,PL} Q_{PL}] - \theta_E (G_{DL} + G_{SDL} + C_{0Q} Q) \quad (16)$$

and can be used to evaluate the structural reliability after a successful test cycle.

The chosen probabilistic formulation accounts for model uncertainties in both the load effect caused by regular loads (θ_E) and the load effect specific to the load applied in the proof load test ($\theta_{E,PL}$). The statistical characterisation of these uncertainties is specific to the application. In particular, the model uncertainty pertaining to the proof load can address various factors, such as the method of load application, the number and configuration of tested positions and lanes, and the considered failure mode. It should be realised that the model uncertainties θ_E and $\theta_{E,PL}$ are likely correlated, given that the same mathematical principles and models are employed to calculate both load effects.

3.3. Distribution of the resistance ratio (X)

3.3.1. Using laboratory test data

When laboratory data are available, the relationship between measurements or indicator values and the resistance ratio ($X = R/E$) can be inferred from these tests. The laboratory tests should be conducted on similar elements and in a configuration comparable to the in-situ proof load test. The laboratory measurements are processed for each load step to analyse the resistance ratio distribution as the indicator values increase. Each specimen has a resistance (R) that corresponds to the load effect at the moment that the limit state is reached (failure). During each load step, the load effect (E) can be calculated, resulting in a corresponding resistance ratio. Typically, an estimation of the self-weight is required to calculate the load effect from both permanent and applied loads (E). This procedure results in a resistance ratio versus indicator value (I) curve for each specimen (for example, see Fig. 5 in Section 4.2).

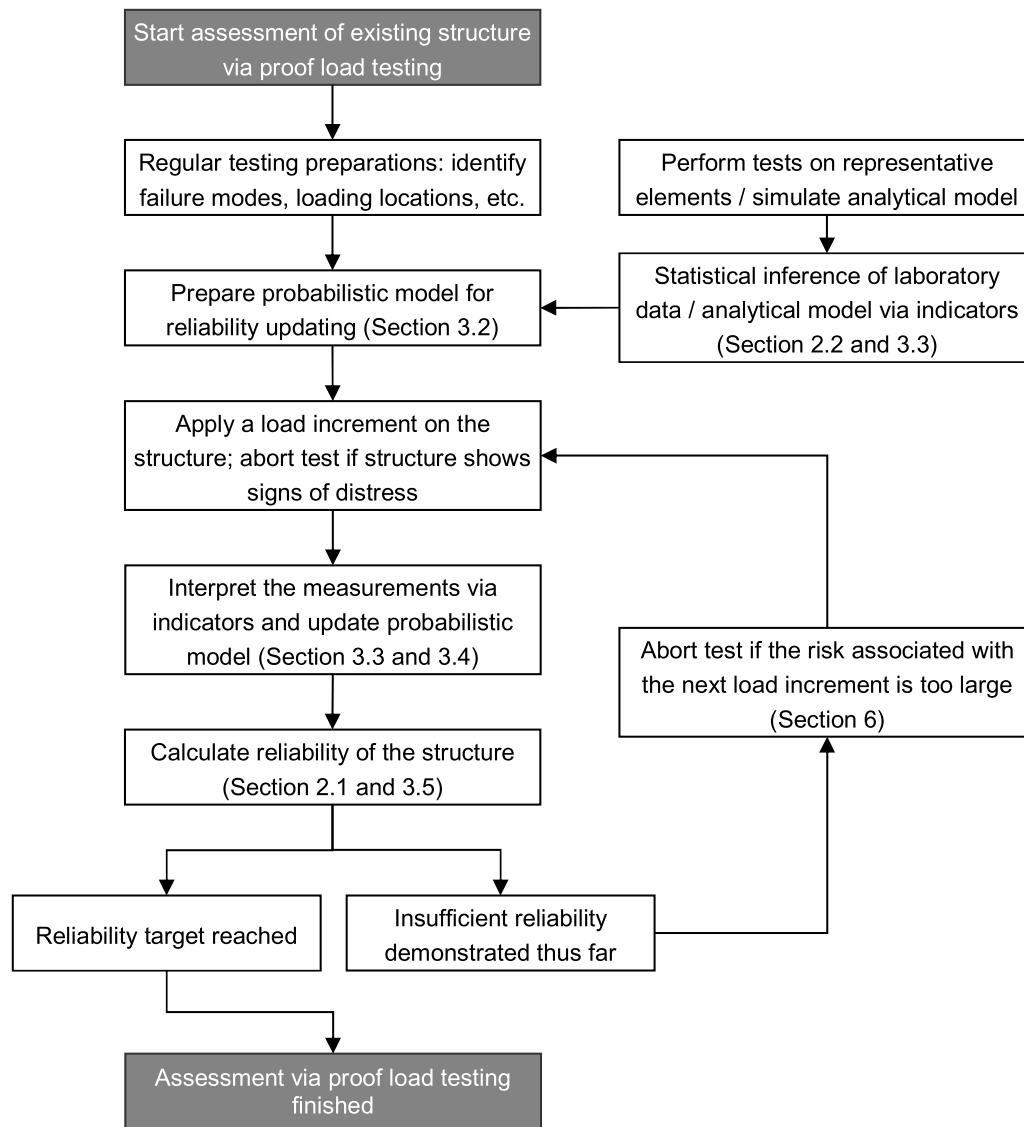


Fig. 2. Flowchart outlining the steps in the proposed proof load testing assessment method.

The resistance ratio is modelled as a random variable due to the inherent uncertainties in both the resistance and the indicator value at a given load level. Moreover, the uncertainty typically decreases as the indicator's value increases. To reduce noise and erratic responses, the maximum indicator value observed up to each load step is used. The same post-processing practice should be followed during in-situ tests for consistency. When the specimen has failed, the resistance ratio is 1, indicating that the resistance is at least equal to the current load effect but not higher.

Once the resistance ratio curves have been obtained for the laboratory tests, statistical post-processing may be performed to infer the statistical description for a range of indicator values. Given a certain indicator value, the data points (X_i) are obtained as the resistance ratios from each specimen. Interpolation of the resistance ratio curves is required to obtain intermediate values. Then, the data points can be analysed using established sample testing methodologies (Section 2.2). Because the prediction distribution of X is used, the actual variation is more significant than indicated by the standard deviation alone. Bayesian inference can also be employed instead of the prediction equation. However, this may be impractical since the inference process needs to be repeated for a range of indicator values. A better approach would also include the trend of model parameters with varying indicator

values combined with a nuanced treatment of measurement error (Section 3.4). An application of the prediction equation and an exponential model for the trend in the resistance ratio is provided in Section 4.2.

3.3.2. Using an analytical model

In cases where laboratory measurements are unavailable, computer simulations can be utilised as an alternative. Instead of calculating the typical design resistance, the aim is to determine the resistance ratio distribution, which cannot be directly obtained through conventional methods. A regular resistance model is developed, but the parameters are random variables. To obtain the statistical distribution of the resistance ratio, it is necessary to integrate over all random variables included in the resistance model. There are several methods to accomplish this integration. The most straightforward method is Monte Carlo simulation (MCS) [16]. However, often a complex numerical (FEM) model is used for the resistance calculation. In these cases, the application of Latin hypercube sampling (LHS) can be beneficial as this method allows for a more efficient representation of the random space with fewer samples [23,52].

By using LHS, numerous resistance ratio versus measurement value (indicator) curves are generated. These curves may then be statistically

analysed for a range of indicator values. By assuming a normal distribution, a mean and standard deviation curve of the resistance ratio versus indicator value can be produced. Because the random space is directly integrated, there is no need to account for statistical uncertainty. Only if the number of simulations is small (e.g. fewer than 30) an approach similar to the post-processing of laboratory experiments should be followed (Section 3.3.1). The additional uncertainty inherent in analytical modelling, as compared to physical testing, can be reflected in a larger coefficient of variation for the resistance model uncertainty (θ_R ; see Section 3.4). An application of this approach, in which analytical modelling of the bending resistance is performed, is presented in Section 5.

3.4. Resistance model and measurement uncertainty

In the described probabilistic framework, θ_R accounts for uncertainty in the resistance calculation introduced by the proposed method. This uncertainty is small when the laboratory specimens closely resemble the actual structure or when the mechanical model has been validated for accuracy. Conversely, significant uncertainty is expected if the structures differ substantially, the mechanical model is overly simplistic, measurement errors are significant, or there are inconsistencies in data post-processing. Practical applications of the proposed method will provide valuable insights into appropriate model uncertainty values as the application in real-world scenarios can highlight the difference between laboratory and in-situ observations [53].

When dealing with a small number of laboratory tests, statistical methods like Student's t -distribution or Bayesian inference effectively account for the inherent statistical uncertainty and variability in the data. The statistical uncertainty would be incorporated directly in the resistance R , thus separated from the modelling uncertainty θ_R . Student's t -distribution is particularly useful when the sample size is small, and the standard deviation is unknown. It is wider than the normal distribution, reflecting the increased uncertainty that comes with fewer laboratory measurements. Bayesian inference, on the other hand, offers a flexible way of incorporating prior knowledge and can account for measurement noise as well. To capture the noise, the data points may be viewed under an assumed distribution for the measurement error by adjusting the likelihood calculation. Assuming the likelihood model (X) and noise (ε) are both normally distributed and independent, their combined variance may be used to define a substitute random variable. The likelihood of observing the data points may then be calculated using the probability density function of the substitute random variable rather than X directly [54,55].

Measurement errors can significantly influence results, especially with small values, such as minor crack widths or small strains. As the magnitude of the values increases, the relative impact of measurement errors typically decreases. In addition, the error also depends on the parameter being measured and whether it can be directly measured or must be inferred. Typically, larger measurement errors are anticipated when estimating crack widths using digital image correlation (DIC) compared to direct strain measurements. Measuring strains on a concrete surface is more susceptible to errors than taking strain measurements directly on the reinforcement. The moderate load values in the proposed method will typically result in small indicator values, and thus, the large uncertainty should be appropriately accounted for. While Bayesian methods are well-suited for treating noise, other analytical approaches can also enhance model accuracy by incorporating physically expected trends with increasing indicator values, thereby providing a nuanced understanding of the data (see, for instance, Section 4.2).

3.5. Calculation procedure

In order to compute the structural reliability after a successful proof load test, the knowledge of surviving the applied load needs to be

incorporated, see Eq. (15). To obtain the posterior distribution $p(\theta | Z_{PL} > 0)$, several calculation methods can be employed. In the direct Bayesian Monte Carlo (BMC) method, the prior distributions are directly sampled. Each sample is used in the simulation to determine if the random values (θ) result in the survival of the proof load test (likelihood evaluation). During the simulation, all parameter sets that produce the desired outcome are stored and collectively describe the posterior distribution [56,57]. In Markov chain Monte Carlo (MCMC), a Markov chain is constructed to obtain samples from the posterior distribution. Algorithms like Metropolis-Hastings [42,43] and Gibbs sampling [58] can generate such sequences (Section 2.2.3). In this work, the MCMC method is adopted because of its versatility. Owing to its Monte Carlo nature, the chain's current state is directly used to evaluate the reliability of the structure with the posterior distribution. For each chain state, $\theta = (\theta_R, X)$, the sample is supplemented by random realizations of the remaining random variables to evaluate the limit state function, i.e. Eq. (16).

A computationally attractive alternative is using the SORM (Section 2.1), which can account for the non-linearity present in the limit state function. Because the survival condition, $Z_{PL} > 0$, cannot be incorporated directly, an approximation must be made. This is achieved by introducing a substitute random variable $Y = \theta_R X$, which follows Student's t -distribution when X is described by a t -distribution and a lognormal distribution otherwise. The variances of the original random variables may be combined such that $\text{Var}(Y) = \text{Var}(\theta_R) + \text{Var}(X)$. Then, the distribution of Y is left-truncated to the set $[1, \infty)$ to impose the survival of the proof load test. Applying this alternative procedure results in an error in the reliability index of approximately 0.1 within the range of common target values (as experienced in the case studies, Sections 4 and 5).

4. Shear resistance assessment supported by laboratory tests

4.1. Description

In order to illustrate the practical application of the method proposed in Section 3, the reliability of a hypothetical shear-critical reinforced concrete slab bridge is considered. The case exemplifies older Dutch slab bridges that lack shear reinforcement. For simplicity, the slab is designed to match the exact width of a single traffic lane (Fig. 3). Normally, a slab bridge would include several lanes, along with sidewalks and railings. The single-lane slab bridge, assumed to experience heavy traffic primarily from trucks, represents a relatively conservative scenario.

Deep beams representing sections of such a slab were tested in the laboratory to evaluate their shear resistance. The reliability of the bridge under consideration can be assessed using the resistance data and the measurements obtained from load tests. However, because the bridge is fictional and no actual in-situ measurements were performed, these values must be estimated to demonstrate the application of the proposed method. The laboratory measurements employed in this case study were initially designed to examine the shear behaviour of reinforced concrete beams lacking shear reinforcement [59]. The tests are a continuation of the study into the parameters that play a role in the transition between flexural and shear failure of reinforced concrete beams without reinforcement [60]. For this case study, H-variants (H121, H401, H403, H404, H602) from the test series were selected because their dimensions correspond to those of the studied concrete slab. The strips, or deep beams, tested in the laboratory had a length of 9 m, a width of 0.3 m and a height of 1.2 m. They were subjected to a load via a single jack positioned near the midpoint of the span, leading to shear failure near the supports (see Fig. 4). Given the specified lane width of 3.6 m, the slab comprises $3.6 / 0.3 = 12$ strips.

Because the experiments have already been conducted, providing a resistance distribution based on the test results (i.e. five V_u values) would be directly possible. Supplemented by the knowledge that the

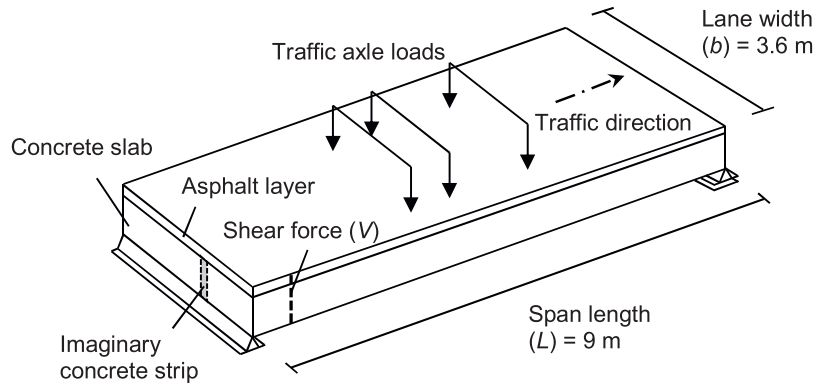


Fig. 3. Hypothetical reinforced concrete slab used as a case study [9].

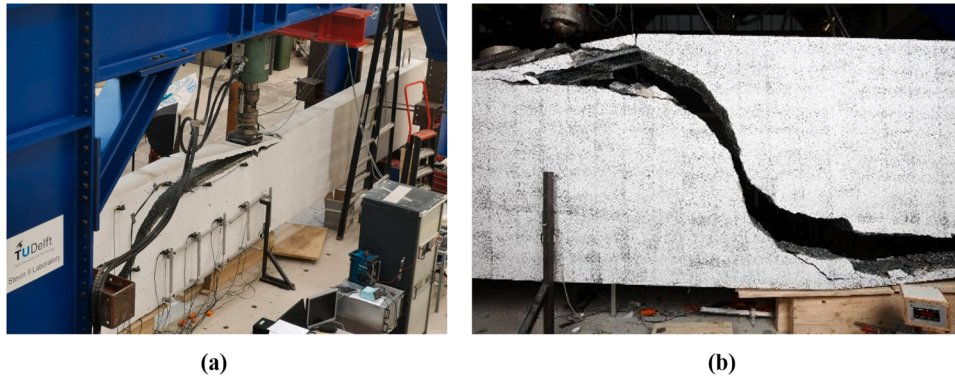


Fig. 4. Photos of (a) the test set-up and (b) the typical crack pattern at failure [61].

structure survives the proof load, this information would give an alternative strategy for reliability updating. However, this alternative procedure does not account for the in-situ observations. By using these observations, a distinction can be made between structures performing well, exhibiting small crack widths, and those performing poorly, showing larger crack widths. In addition, the use of a resistance ratio allows laboratory results from similar structural elements to be applied to other structures that are not entirely identical. However, the validity of this approach and the increase in resistance model uncertainty become relevant factors (Section 3.4).

4.2. Laboratory data post-processing

The digital image correlation (DIC) method was used to monitor and analyse the development of cracks in the concrete during the tests [62]. DIC allows for the determination of the nominal crack width at various locations. The nominal crack width results from the consolidation of several smaller cracks that are closely clustered, forming a single significant crack. The concept of virtual gauge length which is used in DIC, requires the area captured in photographs to be sufficiently large to facilitate the nominal crack width calculations. In this study, the optimal gauge length was found to be $0.8d$, where d represents the depth from the top of the beam to the centre of the longitudinal reinforcement [63]. At each selected location, the nominal crack width is calculated at the level of the reinforcement. To assign greater importance to the cracks near the supports, a weighted crack width calculation is introduced, multiplying the nominal crack width by the factor Vd/M , where V and M refer to the shear force and bending moment at the specific location, respectively. Both the maximum nominal crack width ($w_{\max}$) and the maximum weighted nominal crack width ($w_{\max,w}$) are computed. These maximum values refer to the (weighted) crack widths occurring at any point along the beam's length. Finally, the resistance ratio $X = R/E_{PL}$

$= V_u/V$ for each indicator value was plotted against each load step (Fig. 5).

The data was post-processed to derive the sample mean and standard deviation for each indicator value (Fig. 6). It may be observed that the standard deviation is generally smaller for the maximum *weighted* nominal crack width indicator ($w_{\max,w}$), suggesting it as the preferable metric for subsequent modelling and field testing. Data points where the weighted nominal crack is less than 0.08 mm displayed noticeably high mean and standard deviation values. The high values are likely due to noise affecting the DIC measurements at very small displacements. Generally, the mean and standard deviation exhibited some variability across different values of the indicator $w_{\max,w}$. An exponential model was applied to the data to address the erratic behaviour (Fig. 6):

$$m_{V_u/V}(w_{\max,w}) = \begin{cases} 1.5\exp(-3w_{\max,w}) + 0.88 & w_{\max,w} < 0.85 \text{ mm} \\ 1 & w_{\max,w} \geq 0.85 \text{ mm} \end{cases} \quad (17a)$$

$$s_{V_u/V}(w_{\max,w}) = \begin{cases} 0.53\exp(-2.1w_{\max,w}) - 0.05 & w_{\max,w} < 1.1 \text{ mm} \\ 0 & w_{\max,w} \geq 1.1 \text{ mm} \end{cases} \quad (17b)$$

Eqs. (17a) and (17b) effectively capture the trend towards a mean value of 1 and a standard deviation of 0 as the load increases. It also becomes clear that for values of $w_{\max,w} \geq 1.1$ mm, the method offers no advantages compared to using the lower-bound approximation [5] as the resistance ratio becomes a deterministic value of 1. It should be noted that measurement uncertainty is not explicitly considered here (Section 3.4) since the possibly underestimated uncertainty (low standard deviation) is corrected by the adopted exponential model for small crack width values (Fig. 6). In addition, the use of Student's *t*-distribution is equivalent to the assumption of a non-informative prior for the mean and standard deviation. If Bayesian inference was performed,

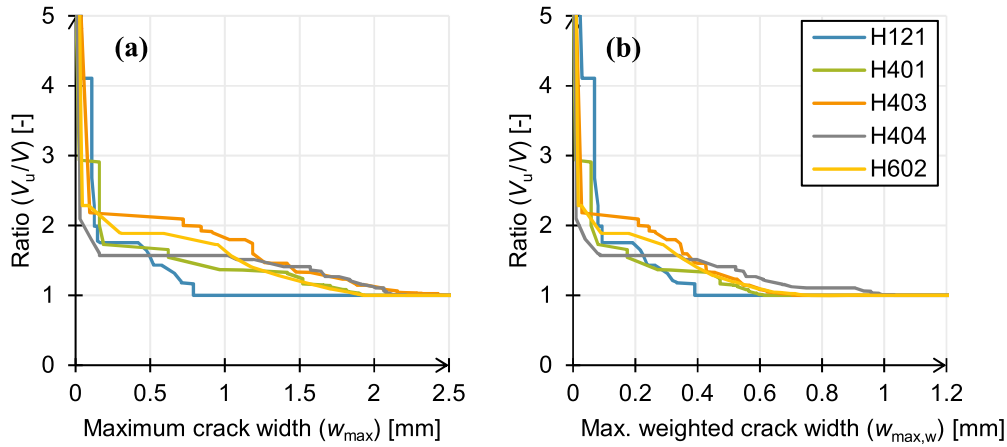


Fig. 5. Shear resistance ratio versus (a) the maximum nominal crack width and (b) the maximum weighted nominal crack width [9].

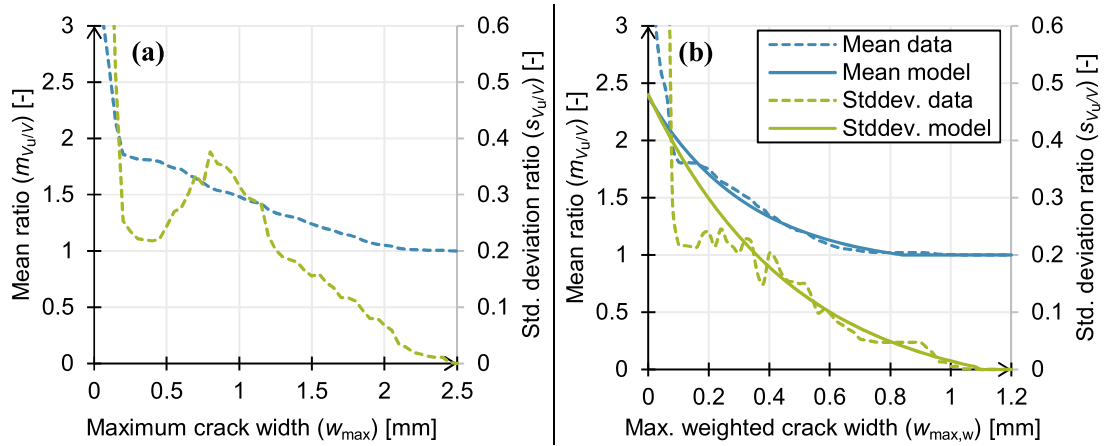


Fig. 6. Mean and standard deviation of the shear resistance ratio versus (a) the maximum nominal crack width and (b) the maximum weighted nominal crack width.

low-informative priors would be used, leading to somewhat lower dispersion.

4.3. Assumed load testing results

The bridge considered in this case study is theoretical and has not been tested. However, indicator values that would normally be acquired during testing are needed to apply the proposed method. Meaningful target loads can be based on the characteristic value of the traffic load multiplied by a certain factor. Given a target load, the tests performed in the laboratory may also be used to estimate likely indicator values (Table 1). In real-world applications, this step would not be required. The discussion (Section 6) describes the sensitivity to the assumed load testing results. It should be noted that measurements performed on the in-situ structure will already include the superimposed dead load (G_{SDL})

thus the measurements will begin from a different starting value. To compensate for this difference, the values should be increased by the crack widths provided in Table 1 for the case $G_{DL} + G_{SDL}$. Although the value is small in the current application, this step is important to align the measurements between the laboratory tests and the structure being monitored during the load test.

4.4. Probabilistic model and reliability analysis

To enable the update of structural reliability, a probabilistic model specific to the considered structure is required. The mean and coefficient of variation of the random variables in Table 2 are based on the Probabilistic Model Code [34] and fib Bulletin 80 [49]. The coefficient of variation used for model uncertainty of the shear resistance ($V_{\theta R} = 0.15$) is deemed appropriate for structures without shear reinforcement [64].

Table 1
Expected indicator readings given a proof load.

Test load level	Loads acting on structure	Expected shear force [kN]	Maximum crack width, weighted by position ($w_{max,w}$) [mm]					
			H121	H401	H403	H404	H602	Average
-	G_{DL}	30.0	0	0	0	0	0	0
-	$G_{DL} + G_{SDL}$	34.9	0.004	0.002	0.002	0.002	0.010	0.004
1	$G + 1.0 Q_k$	73.0	0.069	0.058	0.019	0.010	0.016	0.034
2	$G + 1.2 Q_k$	80.6	0.069	0.058	0.023	0.010	0.029	0.038
3	$G + 1.4 Q_k$	88.2	0.069	0.071	0.026	0.016	0.058	0.048
4	$G + 1.6 Q_k$	95.9	0.069	0.135	0.158	0.032	0.087	0.096
5	$G + 1.8 Q_k$	103	0.074	0.174	0.244	0.052	0.256	0.160

Table 2

Overview of random variables in the limit state function.

Var.	Description	Distribution	Mean	COV
θ_R	Model uncertainty of the resistance	Lognormal	1	0.15
X	Resistance to current load effect ratio	Student's t	(varies)	(varies)
θ_E	Model uncertainty of the load effect	Lognormal	1	0.10
G_{DL}	Dead load effect	Normal	356 kN	0.05
G_{SDL}	Superimposed dead load effect	Normal	59.3 kN	0.10
C_{OQ}	Time-independent uncertainty of the traffic load, including dynamic effects	Lognormal	1.1	0.10
Q	Traffic load effect, annual maximum	Gumbel	390 kN	0.035
$\theta_{E,PL}$	Model uncertainty of the proof load effect; correlation $\rho(\theta_E, \theta_{E,PL}) = 0.7$	Lognormal	1	0.10
Q_{PL}	Load effect achieved by proof load	Normal	(varies)	0.02

The statistical description of the traffic load effect (Q) is based on weigh-in-motion (WIM, [65]) measurement data recorded in the Netherlands. This data was analysed to provide extreme value distributions for the annual maximum bending moment at midspan and shear force near the supports of a single-span structure [5]. The corresponding 1000-year characteristic load effect is obtained as $Q_{k,WIM} = F_Q^{-1}(0.999)$ according to the Eurocode where $F_Q^{-1}(\cdot)$ indicates the inverse CDF of Q . The mean test load levels ($m_{Q,PL}$) are related to the characteristic load effect via a factor (1.0–1.8) in the following analyses. The mean and standard deviation of the resistance ratio for each load level is calculated using Eqs. (17a) and (17b) based on the average indicator values from Table 1. The corresponding coefficient of variation is obtained as $V_X = s_X/m_X$ (Table 3). Finally, by making use of the probabilistic model specified in Table 2, reliability calculations are performed using the method outlined in Section 3. The calculated reliability indices are provided in Table 4.

The results presented in Table 4 clearly show that the proposed method provides higher annual reliability values compared to the lower-bound approach, i.e. Eq. (12). This difference signifies the advantage of integrating additional information from indicators that reflect structural performance. The value of this additional information is particularly significant when the test loads are relatively low, illustrated by an increase in the reliability index of about 1.5 with moderate loads. In addition, Table 4 provides the reliability during the proof load test, indicating the risk associated with applying the target load. A reliability index of 1.8, corresponding to a failure probability of approximately 0.036, suggests the structure is indeed likely to survive the applied load. The reliability during testing is also updated with each increment in the load level, enabling continuous assessment of risk during the test (Section 6).

Table 3Mean and coefficient of variation of the resistance ratio (X) for each load test cycle.

WIM char. load factor ($m_{Q,PL}/Q_{k,WIM}$)	Mean PL effect, strip ($m_{Q,PL,s}$) [kN]	Mean PL effect, lane ($m_{Q,PL}$) [kN]	Indicator value ($w_{max,w}$) [mm]	Mean of resistance ratio (m_X)	COV of resistance ratio (V_X)
1.0	38.1	457	0.034	2.23	0.198
1.2	45.7	549	0.038	2.22	0.198
1.4	53.3	640	0.048	2.18	0.197
1.6	61.0	732	0.096	2.00	0.191
1.8	68.7	823	0.0160	1.81	0.182

Table 4

Calculated reliability indices given different levels of loading, during and after proof load testing (PLT).

WIM characteristic load factor ($m_{Q,PL}/Q_{k,WIM}$)	LM1 characteristic load factor ($m_{Q,PL}/Q_{k,LM1}$)	Reliability during PLT	Annual reliability after successful PLT	
			Proposed method	Lower-bound
1.0	0.78	1.80	2.99	0.53
1.2	0.93	1.79	3.77	1.89
1.4	1.09	1.78	4.55	3.04
1.6	1.24	1.71	5.23	4.03
1.8	1.40	1.61	5.83	4.85

5. Bending resistance assessment supported by a calculation model

5.1. De Beek viaduct

In this case study, proof load levels relating to the bending resistance of the reinforced concrete slab viaduct De Beek in the Netherlands will be examined. The viaduct from 1963 passes over the A67 highway near Ommel, but the viaduct itself is part of the secondary road network. The highway traffic passes beneath the viaduct's two central spans, each with a length of 15.4 m. Additional spans on either side have a length of 10.8 m. The cross-section height varies parabolically along the longitudinal direction from 0.47 m to 0.87 m, measured in the heart of the deck. The bridge is constructed as a continuous slab and thus experiences support moments above its middle three supports (Fig. 7). The sidewalks, added during a later phase of construction, are assumed to offer no significant contribution to the structural resistance (but can influence the stiffness). A pilot proof load test of the first span has already been performed for both bending and shear mechanisms, but only bending will be considered in this case study. Thus, in contrast to the previous case study (Section 4), in-situ test results and measurements are available [66]. A calculation model for the bending resistance will be used to interpret the measurements during testing because laboratory tests on similar bridge decks are not available.

In 2015, an inspection and assessment of the bridge took place and it was concluded that the bridge had insufficient resistance [67]. During the inspection, cracks at the bottom of the concrete slab were found. Afterwards, the detected cracks were filled to prevent water ingress. Later that year, further investigation and a pilot proof load test were performed, confirming that the resistance of span 1 was sufficient for two traffic lanes, both calculated and tested. For span 2, the resistance was only deemed sufficient if plastic redistribution would be allowed to take place, but this typically results in unwanted cracking. The bridge inspection and assessment led to a traffic restriction that reduced the original two lanes to one central lane, only allowing traffic from one direction at a time [68].

5.2. Traffic load effect

The viaduct was designed for two lanes, with traffic in opposite directions. The width of the carriageway is 7.44 m. After subtracting the width of the curbs, the remaining lane width is 3.5 m for each direction. To obtain the most accurate statistical description of the load effect, additional traffic simulations were performed using WIM data recorded in the Netherlands. In this way, the specific continuous slab configuration with traffic moving in opposite directions can be taken into account. The viaduct crosses the A67 highway but is part of the secondary road network. However, only WIM data recorded at highway locations is available in the Netherlands. Using the highway measurements likely results in overestimating the true load effect for this particular location. However, because the viaduct is situated in a rural area, tractors and slurry tanks also make use of it. Although the number of heavy

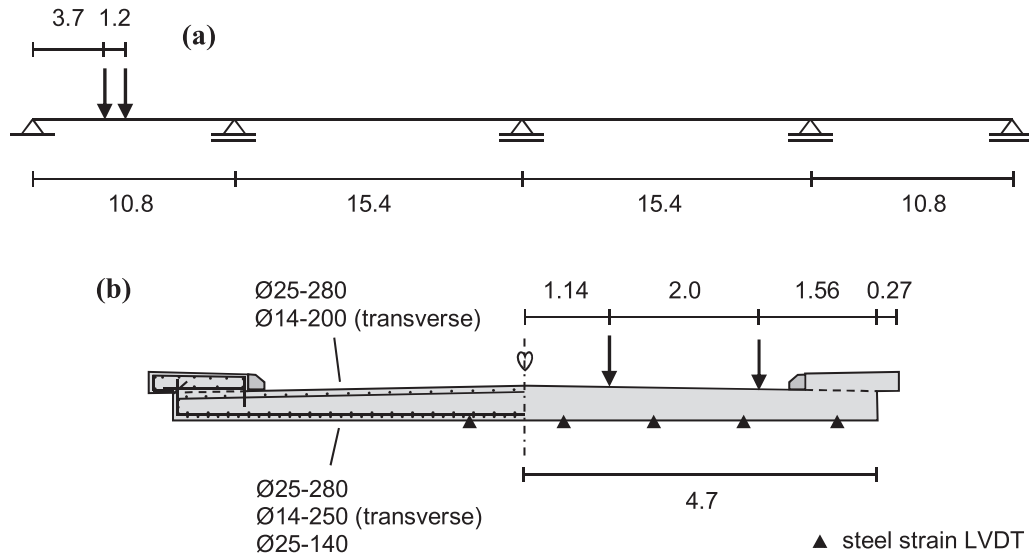


Fig. 7. Schematic (a) side view and (b) cross-section of the De Beek viaduct with the bending proof load test location indicated by the arrows. Reinforcement diameters and spacings provided in mm, other measurements provided in m.

agricultural vehicles is small, they may exert relatively large axle loads on the viaduct. Therefore, no reduction of the traffic load effect following the highway WIM data was applied.

From the WIM measurements, only the first lane data was used in the traffic simulations (i.e. where the trucks predominantly drive). The simulation uses a time-discretisation method in which the vehicle axle loads are placed on the bridge according to the time they were originally recorded. The linear load effect, i.e. the bending moment, was calculated through the superposition of the load effects caused by each individual axle. This calculation may be effectively performed using influence lines, which express the load effect at a certain location, given a unity axle load (Fig. 8). The influence lines were derived using a 1D finite element analysis utilising many small Euler-Bernoulli beam elements to account for the varying deck height in the longitudinal direction.

By analysing the bending moment caused by permanent and traffic loads it follows that the maximum combined load effect can be expected around $x = 4.5$ m. To identify the difference in the bending moment following from one lane and two lanes in opposite directions, both situations were analysed. The load effect in the two-lane configuration is not simply the one-lane load effect multiplied by a factor 2 because it is very unlikely that, in both directions, a heavy truck is present at the same time. However, the variability in the load effect increases considerably, as shown by the more gradually descending right tail of the distribution in the two-lane configuration (Fig. 9). The Gumbel

distribution for the weekly maxima was found by fitting the right tail of the weekly extremes for the WIM highway locations (A16L, A27L, A50L and A67L) directly. The yearly distributions we found by modifying the Gumbel distribution location parameter as $\mu_{yr} = \mu_{wk} + \ln(52)\beta$ where 52 is the number of weeks in the year, and β is the scale parameter of the Gumbel distribution (which remains unchanged). The assessment of the first span will be performed for the two-lane configurations, for which the annual load effect is found to follow a Gumbel distribution with a mean value of 1301 kNm and a variation coefficient of 0.058.

5.3. Proof load test and measurements

In November 2015, proof load tests were performed on the first span of the viaduct De Beek using load levels described by relevant Dutch guidelines and standards [69–71]. The load was applied as the Eurocode LM1 tandem with two axles at a distance of 1.2 m with a wheelbase of 2 m and a wheel print size of $0.4 \text{ m} \times 0.4 \text{ m}$. The axle load was varied to achieve a load effect corresponding to various load levels (Table 5). The corresponding load effects caused by the tandem axle loads are calculated using the influence lines derived in Section 5.2. The bending moment caused by the LM1 tandem is obtained as $M = 3.75 P$ where P is the load of an individual axle. The load test situation mimics the situation in which a heavy vehicle passes the bridge in a two-lane configuration. Due to the eccentric placement of the load, the bending moment

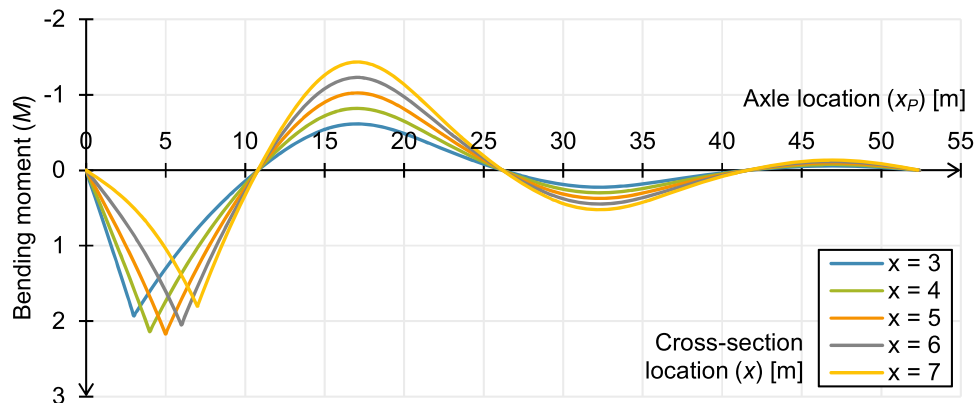


Fig. 8. Influence lines for the bending moment considering various cross-section locations, i.e., the moment for a cross-section located at x , resulting from a unit axle load ($P = 1$) positioned at x_p .

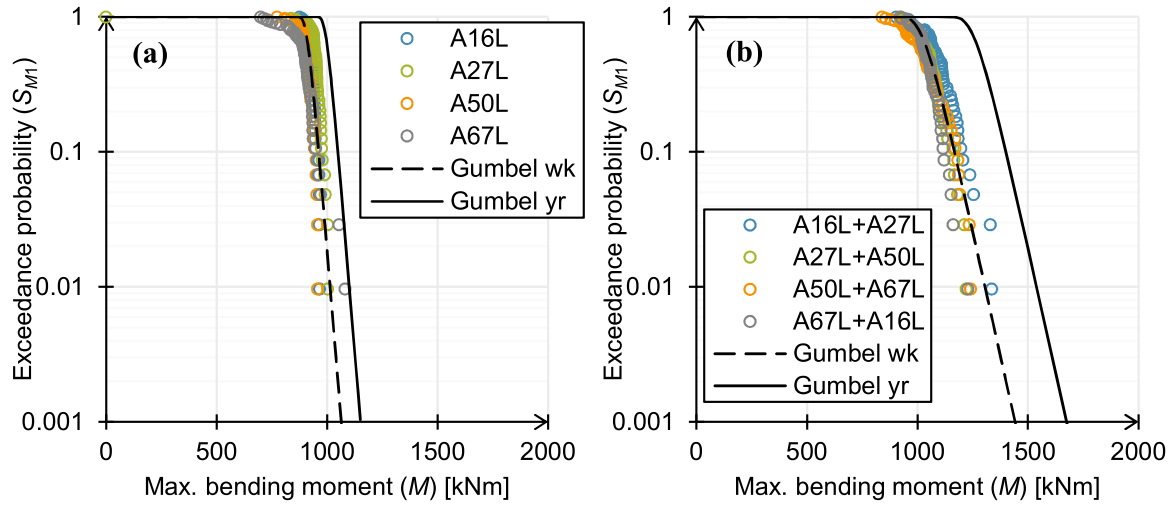


Fig. 9. Extreme value distributions of the maximum load effect in the first span for the (a) one-lane and (b) two-lane configurations.

Table 5

Overview of test loads, load effects and measured steel strains.

Cycle	Description	P_{tot} [kN]	P [kN]	M [kNm]	ϵ_s [10^{-6}]
1	Unfactored tandem	600	300	1125	165
2	Service level	1000	500	1875	340
3	RBK usage	1400	700	2625	570
4	Intermediate	1700	850	3188	790
5	Eurocode design	1750	875	3281	830

varies in the transverse direction of the deck. In this case study, the deck will be treated as a beam for simplicity, but the relative increase of strains near the edge will be taken into account (see Section 5.4.1).

During the test, various measurements were performed, including displacements, concrete surface strains and reinforcement strains. At a number of locations, the concrete cover was removed to inspect the reinforcement and attach strain gauges [68]. If the reinforcement cannot be accessed directly, stop-criteria expressed in concrete surface strains may prove beneficial [72]. In this case study, the maximum steel strain measured at 0.86 m from the edge of the deck (see Fig. 7) is of primary interest since yielding of the reinforcement is undesired for durability reasons. In the longitudinal direction, the strains are measured at a distance of 4.05 m from the support and are considered representative of the section between the axle loads (3.7–4.9 m), as the moment remains nearly constant throughout this range.

5.4. Analytical resistance model

5.4.1. Modified sectional analysis

If no tests have been performed on representative specimens, it is possible to use an analytical model to interpret measurements during the test. The resistance model may range from simple to complex, e.g. cross-sectional analysis [73,74], strip model [75,76], linear finite element model, non-linear finite element model [77,78]. In the considered case, the proof load test has already been performed, and therefore, the measurements can be used to correct a relatively simple sectional analysis model. This calculation is performed by gradually incrementing the strains across the section according to a linear strain distribution and finding the compression zone depth for which the axial force and bending moments in the section are in balance. The result is often presented as a moment-curvature (M - κ) diagram. In this case study, the bottom steel strain (ϵ_s) is utilised instead of the curvature (κ) because the steel strain was measured during testing.

For the sectional analysis of the De Beek bridge deck, a width of $b = 9.4$ m is used, along with an equivalent deck height of $h = 0.47$ m at

the location where the maximum moment is deemed to occur ($x = 4.5$ m, Section 5.2). The reinforcement, as schematically presented in Fig. 7, listed from bottom to top, corresponds to the longitudinal reinforcement areas $A_{s,bot,1} = 3506$ mm²/m, $A_{s,bot,2} = 1753$ mm²/m and $A_{s,top} = 1753$ mm²/m. Thorenfeldt's model [79] is used for the concrete compressive stress-strain curve, and no concrete tensile strength contribution is assumed. The analysis is first performed using mean values for the geometry and material properties (Table 6) to establish the model correction.

Analysis of the measurement data reveals that the recorded strains must be adjusted to compare the numerically obtained response with the observed measurements. If a linear response up to and between load cycles 1 and 2 is assumed, the microstrain increment corresponding to $1875 - 1125 = 750$ kNm is $340 - 165 = 175$ (Table 5). Then, a microstrain of $1125/750 \cdot 175 = 263$ would be expected for the first cycle, which is about 100 higher than recorded. However, the corrected model is non-linear, as a result of Eq. (18), and the required adjustment has been determined as 85 iteratively.

The moment caused by the self-weight of the deck and the superimposed loads is estimated as $M_G = 747$ kNm [66]. The corrected cross-sectional analysis model results in a corresponding bottom steel strain $\epsilon_{sG} = 154 \cdot 10^{-6}$. The calculated steel strains need to be lowered compared to the original model to align with the measured response (Fig. 10). This increase in stiffness could be caused by the contribution of the sidewalks. In addition, the measurements show a gradual decrease in stiffness that is not present in the modelled response. Therefore, a power law is included to increase the steel strain proportionally to the reinforcement's yield strain (ϵ_y ; mean value $\epsilon_{ym} = f_{ym} / E_{sm} = 1417 \cdot 10^{-6}$). The adopted expression for the modified steel strain is:

$$\epsilon_s^* = [c_1 + c_2 (\epsilon_s / \epsilon_y)^{c_3}] \epsilon_s \quad (18)$$

Table 6

Overview of the random variables in the mechanical model.

Var.	Description	Distribution	Mean	COV
c	Concrete cover thickness	Gamma	30 mm	0.17
h	Height of the cross-section	Lognormal	470 mm	0.10
E_s	Young's modulus of reinforcing steel	Lognormal	205 GPa	0.02
f_c	Concrete compressive strength	Log- t ($n = 6$)	57.5 MPa ^a	0.10 ^a
f_y	Yield stress of the reinforcement	Log- t ($n = 3$)	290.5 MPa ^a	0.034 ^a
c_2	Non-linearity of moment-strain relation	Lognormal	0.41	0.10

^a Sample statistics are reported. The log- t prediction distribution accounts for the small number of samples and will effectively lead to higher variability.

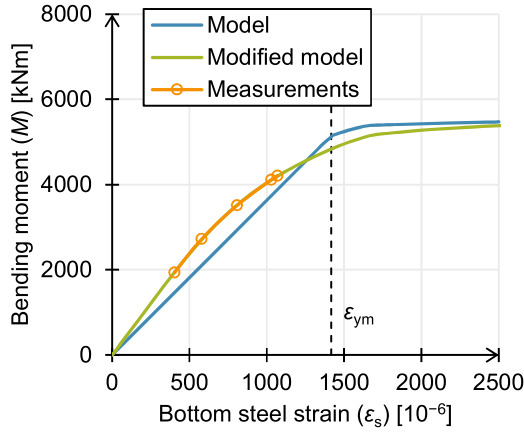


Fig. 10. Comparison of the moment-strain diagram following from modelling and measurements.

where c_1 accounts for the difference in stiffness, c_2 controls the degree of non-linearity and c_3 specifies the shape of the curve. An optimal fit for this case study is obtained with $c_1 = 0.75$, $c_2 = 0.41$ and $c_3 = 4.1$. As a result, the modified model predicts a lower (local) yield moment, but the ultimate resistance remains the same. In this way, the model accounts for eccentric loading with increased strains towards the side of the bridge deck where the load is applied.

5.4.2. Sampling of the mechanical model

Since no laboratory measurements are available, the uncertainty regarding the structural response must be included in an alternative manner. To this end, the mechanical model can be set up using random variables for the parameters (Table 6). The distribution types and coefficients of variation (COV) follow fib [49] and JCSS [34]. A larger COV value is adopted for the height of the cross-section since it is an equivalent value for the actual height, which varies along the longitudinal and transverse directions. The steel area of the reinforcement is not included because its variability is minimal. Material testing was performed on the concrete and the reinforcing steel [67]. It is expected that concrete compressive strength and yield stress of the reinforcement would follow a lognormal distribution when many measurements are available. Therefore, the prediction distribution, Eq. (8), is used to describe the logarithm of these material properties. The resulting log- t distributed random variables for the concrete compressive strength (f_c) and the yield stress of the reinforcement (f_y) are:

$$f_c = \exp\left(4.05 + 0.106 T_{\nu=6-1} \sqrt{1 + 1/6}\right) \text{ [MPa]} \quad (19a)$$

$$f_y = \exp\left(5.76 + 0.0346 T_{\nu=3-1} \sqrt{1 + 1/3}\right) \text{ [MPa]} \quad (19b)$$

where random variable T_ν follows Student's t -distribution with ν degrees of freedom. The sample mean and standard deviation values have been calculated from the log-transformed measurement data (i.e. taking the natural logarithm of each value). The uncertainty about the non-linearity of the modified sectional analysis model is included in the probabilistic description by using a lognormal distribution for the parameter c_2 in Eq. (18), which controls the degree of non-linearity.

Latin hypercube sampling [23,52] was used to simulate responses according to the probabilistic model (Table 6). Compared to the more common Monte Carlo simulation [16], Latin hypercube sampling better represents the output distribution when using a small number of samples. Although the mechanical model is not computationally demanding in this case study, a more complex (FEM) model will require considerable computation time. The calculated moment-strain curves and the resistance ratios for 100 samples are displayed in Fig. 11. Because the yielding of reinforcement is undesired, the yield moment resistance (M_y) is used instead of the ultimate bending resistance (M_u) to calculate the ratio $X = M_y/M$. The sample yield moment is obtained by calculating the sample yield stress ($\epsilon_y = f_y / E_s$) and subsequently interpolating the moment-strain curve to find the corresponding moment. In Fig. 12, the calculated mean and standard deviation of the resistance ratio are plotted against the bottom steel strain. In contrast to the approach in Section 4.2, the statistics of the resistance ratio X can be directly obtained from the curves. The mean and standard deviation relations are discretised using a strain interval of $50 \cdot 10^{-6}$ and can be linearly interpolated in the subsequent reliability analysis.

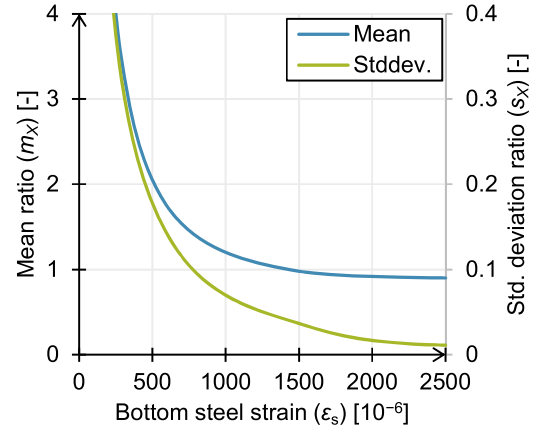


Fig. 12. Mean and standard deviation of the simulated resistance ratio.

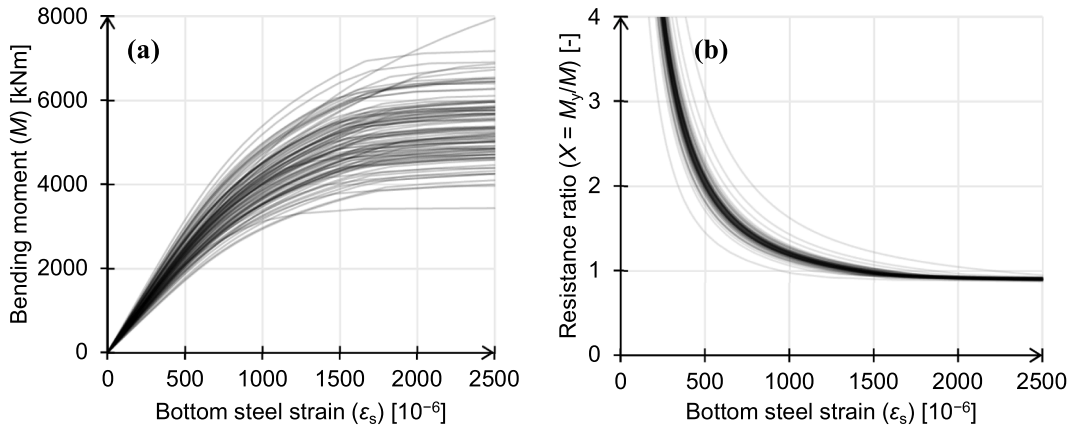


Fig. 11. Simulation result displayed as moment-strain (a) and resistance ratio-strain curves (b).

5.5. Probabilistic model and reliability analysis

The probabilistic model described in Section 3.2 is now employed to perform the reliability analysis. The mean values and variation coefficients of the random variables are specified in Table 7 and are based on the Probabilistic Model Code [34] and fib Bulletin 80 [49]. A statistical description of the traffic load effect was obtained in Section 5.2. The statistics for the resistance ratio depend on the indicator value – as discussed in Section 5.4.2. In this case, the indicator is the steel strain in the bottom reinforcement near the edge of the deck (ϵ_s). In contrast to the previous case study (Section 4), the resistance ratio (X) follows a normal distribution, not Student's t -distribution, because the resistance model is based on a stochastic model. The coefficient of variation used for the model uncertainty of the resistance ($V_{\theta R} = 0.15$) is twice the value generally used for bending [64]. This increased value reflects the greater uncertainty associated with the use of a basic sectional analysis model for a bridge deck subjected to eccentric loading (Section 5.4.1).

For each of the load levels used in the proof load test, the mean and coefficient of variation of the resistance ratio are calculated (Table 8). The load levels may also be related to the WIM and LM1 characteristic traffic load effects to enable comparisons with the previous case study (Section 4). For each of these load levels, the reliability during and after surviving the proof load test has been calculated (Table 9). Similar to the previous case study, using the indicator data results in markedly higher reliability indices at low load levels than the lower-bound method, i.e. Eq. (12). The reliability during testing is relatively high but decreases rapidly with higher test loads, underlining the need for risk assessment during the proof load test between load steps (Section 6).

6. Discussion

In the data post-processing of the shear resistance case study (Section 4.2), the model deviates from the calculated values for $w_{\max,w} < 0.08$ mm. This region is important for the subsequent reliability calculations and greatly influences the outcomes. As described in Section 3.4, measurement errors play a significant role with small indicator values and various strategies are discussed to account for them. The power-law model was chosen because analytical resistance models show a similar decreasing trend (Section 5.4). Further data inspection reveals very high resistance ratios in the small-value region with a positive distribution skew. Thus, the adoption of a conservative model in this region effectively mitigates measurement-error issues. Additional calculations using Bayesian inference to statistically describe the resistance ratio resulted in higher reliability indices (up to 0.5), thereby confirming the effectiveness of the chosen approach.

Attention must be given to ensuring the similarity between the in-

Table 7
Overview of random variables in the limit state functions.

Var.	Description	Distribution	Mean	COV
θ_R	Model uncertainty of the resistance	Lognormal	1	0.15
X	Resistance to current load effect ratio	Normal	(varies)	(varies)
θ_E	Model uncertainty of the load effect	Lognormal	1	0.10
G_{DL}	Dead load effect	Normal	604 kNm	0.05
G_{SDL}	Superimposed dead load effect	Normal	143 kNm	0.10
C_{OQ}	Time-independent uncertainty of the traffic load, including dynamic effects	Lognormal	1.1	0.10
Q	Traffic load effect, annual maximum	Gumbel	1301 kNm	0.058
$\theta_{E,PL}$	Model uncertainty of the proof load effect; correlation $\rho(\theta_E, \theta_{E,PL}) = 0.7$	Lognormal	1	0.10
Q_{PL}	Load effect achieved by proof load	Normal	(varies)	0.02

Table 8

Mean and coefficient of variation of the resistance ratio (X) for each load test cycle.

Cycle	Description	Measured steel strain [10^{-6}]	Indicator value (ϵ_s) [10^{-6}]	Mean of ratio (m_X)	COV of ratio (V_X)
1	Unfactored tandem	165	419	2.42	0.090
2	Service level	340	594	1.76	0.082
3	RBK usage	570	824	1.36	0.068
4	Intermediate	790	1044	1.17	0.056
5	Eurocode design	830	1084	1.15	0.054

Table 9

Calculated reliability indices given different levels of loading, during and after proof load testing (PLT).

WIM characteristic load factor ($m_{Q,PL}/Q_k$, WIM)	LM1 characteristic load factor ($m_{Q,PL}/Q_{k,LM1}$)	Reliability during PLT	Annual reliability after successful PLT	
			Proposed method	Lower-bound
0.67	0.46	4.78	3.58	-1.70
1.12	0.76	3.18	3.99	1.95
1.57	1.06	1.79	5.00	4.12
1.91	1.29	0.91	5.78	5.27
1.96	1.33	0.80	5.92	5.40

situ tested structure and the results from laboratory tests or analytical models. In the shear resistance case study, it is assumed that the response of the bridge slab is similar to the response of the strips tested in the laboratory. However, still a relatively large coefficient of variation was assigned to the model uncertainty to account for the remaining differences (e.g., transverse load distribution, boundary conditions, edge and size effects). With analytical models, it is crucial to evaluate whether they can reliably provide correct indicator values (such as crack widths, strains, etc.), as these models often primarily consider the resistance of structural members. Experience with the proposed method in practice can establish suitable model uncertainty values.

In reality, the tested structure may exhibit a very different response than assumed in both case studies. If the structure's condition is above average, smaller indicator values are anticipated and therefore the reliability increases. Conversely, for structures that show large indicator values under small loads, lower reliability is expected. The results include reliability indices that will likely be observed in practical applications, but may differ from case to case. The same holds for the relation between the target load, the statistical description of the traffic load effect and the load effect calculated from standards. For example, the required target load for the bending resistance assessment of the viaduct is relatively low when expressed by the unfactored LM1 characteristic load (Section 5.5). This result stems from the detailed statistical analysis of the traffic load effect, indicating that the LM1 characteristic load may be rather conservative for the two-lane situation with opposite driving directions.

The proposed method also provides the opportunity to calculate the reliability during the proof load test. Although no standards or guidelines specify minimum reliability levels for load testing, case-specific risk analysis can help determine optimal values. Such an analysis can take into account the load-testing conditions (e.g., absence of traffic, restricted areas), allowing for informed decisions on whether to continue with higher load levels or stop the test. This approach offers a risk-optimal alternative to pre-determined stop criteria but requires additional analysis and calculation effort.

The value of the current study lies in the proposed method for proof load testing rather than the calculated reliability values and target loads. Future research and practical applications can further refine the

framework, providing valuable insights into model uncertainties and measurement techniques. Real-world validation is recommended to establish the robustness of the adopted models and data post-processing procedures. Large-scale implementation would benefit from research into generally applicable indicators, resistance ratio curves, identification of application criteria and possibly (FEM) modelling guidelines.

7. Conclusions

The proposed proof load testing method involves updating the resistance distribution using two sources of information: (1) the observed in-situ response, which is related to the resistance via indicator values and associated resistance ratios, and (2) the survival of the applied load during the proof load test. Although the prediction of the resistance on the basis of the in-situ response is associated with considerable uncertainty, this information is still valuable and can be accounted for probabilistically. A Bayesian procedure for updating the resistance distribution given the two information sources is presented, along with a method for calculating the posterior reliability using the Markov chain Monte Carlo (MCMC) technique. In addition, the probability of failure during the test is calculated using the most current data, offering a risk-optimal alternative to pre-determined stop criteria.

The application of the proposed method was demonstrated through two case studies. The first case study considered a hypothetical shear-critical concrete bridge using laboratory data, while the second examined the bending resistance of an existing viaduct in the Netherlands using an analytical model. These case studies illustrated how the framework could be applied in real-world scenarios, and the potential gains in reliability when monitoring data is included in the assessment. The two case studies in which in-situ measurement data was included allowed for test load reductions of 20 % and 25 %. The proposed method's versatility was highlighted by using laboratory experiments in the first case study and an analytical model in the second to establish the relationship between indicator values and resistance ratios. In cases with complex mechanical behaviour where tests or analytical models are less representative of the in-situ structure, smaller test load reductions should be expected due to the increased uncertainty.

Advanced probabilistic models and calculation techniques were utilised to account for uncertainties in resistance, material properties, load effects, and measurement errors. Laboratory data post-processing, including the consideration of weighted crack widths, provided insight into the resistance ratios for different target loads. The impact of model uncertainty and measurement errors was addressed, particularly for small indicator values. Alternatively, analytical models can be used to derive similar insights, ensuring the model closely matches the in-situ conditions.

The suggested method offers a more comprehensive and accurate approach to evaluating existing infrastructure using proof load testing. Using in-situ measurements, the procedure also enables the calculation of failure probability during the test, allowing for risk-based decisions on whether to proceed or stop. Practical applications of the method can determine whether similar reductions in test loads, as found in the current study, are feasible. If so, proof load testing can become more economically attractive and less time-consuming, minimising the traffic disruption involved in bridge testing.

CRedit authorship contribution statement

Hendriks M.A.N.: Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Steenbergen R.D.J.M.:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Naaktgeboren M.:** Writing – review & editing, Funding acquisition, Conceptualization. **Lantsoght E.O.L.:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Funding acquisition, Data curation, Conceptualization. **de Vries R.:**

Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to express their gratitude and sincere appreciation to the Dutch Ministry of Infrastructure and Water Management (Rijks-waterstaat) for financing the research work. In addition, the fruitful discussions with Sonja Fennis of Rijkswaterstaat, Ton Vrouwenvelder of TNO and Yuguang Yang from Delft University of Technology have been of great help.

Data Availability

Data will be made available on request.

References

- [1] Steenbergen RDJM, Vrouwenvelder ACWM. Safety philosophy for existing structures and partial factors for traffic loads on bridges. *Heron* 2010;55(2): 123–39.
- [2] Vrouwenvelder T, Scholten N. Assessment criteria for existing structures. *Struct Eng Int* 2010;20(1):62–5.
- [3] Faber MH, Val DV, Stewart MG. Proof load testing for bridge assessment and upgrading. *Eng Struct* 2000;22:1677–89.
- [4] Casas JR, Gómez JD. Load rating of highway bridges by proof-loading. *KSCSE J Civ Eng* 2013;17(3):556–67.
- [5] De Vries R, et al. Proof load testing method by the American association of state highway and transportation officials and suggestions for improvement. *Transp Res Rec* 2023.
- [6] Frangopol DM, et al. *Reliability-based analysis and lifecycle management of load tests*. Book chapter. In: Lantsoght EOL, editor. Load testing of bridges: Proof load testing and the future of load testing; 2019. p. 263–92.
- [7] Lantsoght EOL. Load testing of bridges: Proof load testing and the future of load testing. In: Frangopol DM, editor. *Structures and Infrastructures*. London, UK: CRC Press / Balkema - Taylor & Francis Group; 2019.
- [8] De Vries R. Reliability assessment of existing reinforced concrete bridges and viaducts through proof load testing. *Proceedings of the 11th International Conference on Bridge Maintenance, Safety and Management (IABMAS)*, Barcelona, Spain; 2022.
- [9] De Vries R. Structural reliability updating using monitoring data from in-situ load testing and laboratory test results. *Proceedings of the 13th International Conference on Bridge Maintenance, Safety and Management (IABMAS)*, Copenhagen, Denmark; 2024. p. 409–417.
- [10] Zarate Garnica GI, Lantsoght EOL. Stop criteria for proof load testing of reinforced concrete structures. *Proc 2021 Sess 13th fib Int PhD Symp Civ Eng* 2021:195–202.
- [11] Ditlevsen O, Madsen HO. *Book. Structural reliability methods*. John Wiley & Sons; 1996.
- [12] Madsen HO, Krenk S, Lind NC. *Methods of structural safety*. Englewood Cliffs, New Jersey: Prentice Hall; 1986. p. 403.
- [13] Der Kiureghian A. *Structural and system reliability*. Cambridge University Press; 2022.
- [14] CEN, Eurocode 0: Basis of structural design. Standard, EN 1990+A1+A1/C2:2019, European Committee for Standardization (CEN), Brussels, Belgium; 2019.
- [15] Wasserman L. *All of statistics: A concise course in statistical inference*. Book, Springer; 2004.
- [16] Metropolis N, Ulam S. The Monte Carlo method. *J Am Stat Assoc* 1949;44(247): 335–41.
- [17] De Vries R, Steenbergen RDJM, Maljaars J. Annual reliability requirements for bridges and viaducts. *Heron* 2023;68(2).
- [18] Melhem M.M., Caprani C.C., Stewart M.G., Zhang S. Bridge Assessment Beyond the AS 5100 Deterministic Methodology Research Report AP-R617-20, Austroads, Sydney, Australia; 2020.
- [19] Kloek T, Van Dijk HK. Bayesian estimates of equation system parameters: an application of integration by Monte Carlo. *Econometrica* 1978;46(1):1–19.
- [20] Bjerager P. Probabilistic integration by directional simulation. *J Eng Mech* 1988; 114(8):1285–302.
- [21] Waarts PH. Structural reliability using finite element analysis - An appraisal of DARS: Directional Adaptive Response surface Sampling. Delft University of Technology; 2000.
- [22] Moustapha M, Marelli S, Sudret B. Active learning for structural reliability: survey, general framework and benchmark. *Struct Saf* 2022;96:102174.

- [23] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 1979;21(1):239–45.
- [24] Hasofer AM, Lind NC. Exact and invariant second-moment code format. *J Eng Mech Div* 1974;100(1):111–21.
- [25] Rackwitz R, Fiessler B. Structural reliability under combined random load sequences. *Comput Struct* 1978;9: 498–494.
- [26] Fiessler B, Neumann HJ, Rackwitz R. Quadratic limit states in structural reliability. *Eng Mech* 1979;105(4):661–76.
- [27] Breitung K. Asymptotic approximations for multinormal integrals. *Eng Mech* 1984; 110(3):357–66.
- [28] Phoon KK. Numerical recipes for reliability analysis – a primer. in *Reliability-Based Design in Geotechnical Engineering: Computations and Applications*. Taylor & Francis; 2008. p. 75.
- [29] Hohenbichler M, et al. New light on first- and second-order reliability methods. *Struct Saf* 1987;4:267–84.
- [30] Tvedt, L., Two second order approximations to the failure probability. Technical report, RDIV/20-004-83, Det Norske Veritas; 1983.
- [31] Geisser S. Predictive inference: An introduction. New York: Chapman & Hall; 1993.
- [32] Box GE, Hunter JS, Hunter WG. *Statistics for Experimenters: Design, Innovation, And Discovery*. 2nd edition. Hoboken, New Jersey, USA: Wiley-Interscience; 2005.
- [33] Gosset WS. The probable error of a mean. *Biometrika* 1908;6(1):1–25.
- [34] JCSS ; 2015, Available from: (<https://www.jcss-lc.org/jcss-probabilistic-model-code/>). Probabilistic model code(<https://www.jcss-lc.org/jcss-probabilistic-model-code/>).
- [35] Ditlevsen O, Vrouwenvelder A. "Objective" low informative priors for Bayesian inference from totally censored Gaussian data. *Struct Saf* 1994;16:175–88.
- [36] Straub D, Papaioannou I. Bayesian updating with structural reliability methods. *Eng Mech* 2015;141(3).
- [37] Ghose MK, Rajagopalan S. Bayesian reliability assessment for discrete data—a case study. *Reliab Eng* 1985;27–36.
- [38] Peterka V. Bayesian system identification. *Automatica* 1981;17(1):41–53.
- [39] Lauritzen SL. Sequential Bayesian updating – BS2 statistical inference, lectures 14 and 15. Lect Notes, Univ Oxf 2009.
- [40] Alam J, et al. Sequential Bayesian updating for time-variant reliability analysis of ageing structures. *Mech Syst Signal Process* 2023;204(110774).
- [41] Jeffreys H. *Theory of probability*, 3rd ed Book, Oxford University Press; 1961.
- [42] Metropolis N, et al. Equations of state calculations by fast computing machines. *Chem Phys* 1953;21(6):1087–91.
- [43] Hastings WK. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 1970;57(1):97–109.
- [44] Benjamin JR, Cornell CA. *Probability, statistics, and decision for civil engineers*. Book, McGraw-Hill; 1970.
- [45] AASHTO The manual for bridge evaluation Standard, 3rd Edition. Washington, D. C., USA; 2018.
- [46] De Vries R, et al. Time-dependent reliability assessment of existing concrete bridges with varying knowledge levels by proof load testing. *Struct Infrastruct Eng* 2023;20(7–8):1053–67.
- [47] Lin TS, Nowak AS. Proof loading and structural reliability. *Reliab Eng* 1984;8: 85–100.
- [48] Bröske H. Structural test design with value of information PhD Thesis (Report 401), DTU, Kongens Lyngby, Denmark; 2018.
- [49] fib Partial factor methods for existing concrete structures. Bulletin 80, Recommendation, Task Group 3.1, Fédération internationale du béton, Lausanne, Switzerland; 2016.
- [50] Abril-Pla O, et al. PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Comput Sci* 2023;9.
- [51] Lauritzen SL, et al. Independence properties of directed Markov fields. *Networks* 1990;20:491–505.
- [52] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf* 2003;81:23–69.
- [53] Sykora M, Holický M. Assessment of uncertainties in mechanical models. *Appl Mech Mater* 2013;378:13–8.
- [54] MacKay DJC. *Bayesian methods for adaptive models*. Pasadena, California, U.S.A.: California Institute of Technology; 1991.
- [55] Bishop CM. *Pattern recognition and machine learning*. New York, U.S.A.: Springer; 2006.
- [56] Hornberger GM, Spear RC. Eutrophication in peel inlet—I. The problem-defining behavior and a mathematical model for the phosphorus scenario. *Water Res* 1980; 14(1):29–42.
- [57] Dilks DW, Canale RP, Meier PG. Development of Bayesian Monte Carlo techniques for water quality model uncertainty. *Ecol Model* 1992;62(1):149–62.
- [58] Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell* 1984;6(6).
- [59] Yang Y, Van der Ham HWM, Naaktgeboren M. Shear capacity of RC slab structures with low reinforcement ratio – An experimental approach. Lisbon, Portugal: fib Symposium; 2021.
- [60] Koekkoek R.T., Yang Y. Measurement report on the transition between flexural and shear failure of RC beams without shear reinforcement Delft University of Technology, Stevin Report 25.5-16-04, Delft, The Netherlands; 2016.
- [61] Zhang F. Acoustic emission-based indicators of shear failure of reinforced concrete structures without shear reinforcement. The Netherlands: Delft University of Technology, Delft; 2022.
- [62] Jones EM, Iadicola MA. A good practices guide for digital image correlation. *International Digital Image Correlation Society*; 2018.
- [63] Zarate Garnica G.L., Vries R.De, Lantsoght E.O.L. Analysis report of reinforced concrete slabs for stop criteria Stevin Report 25.5-22-02, Delft University of Technology, Delft, The Netherlands; 2022.
- [64] Sykora M, et al. Uncertainties in resistance models for sound and corrosion-damaged RC structures according to EN 1992-1-1. *Material and Structures* 2015;48 (10):3415–30.
- [65] FHWA, Weigh-in-motion pocket guide - Part 1: WIM technology, data acquisition, and procurement guide. Federal Highway Association, FHWA-PL-18-015, Washington D.C., USA; 2018.
- [66] Koekkoek R.T., Analysis report for the assessment of viaduct De Beek by proof loading Stevin Report 25.5-16-01, Delft University of Technology, Delft, The Netherlands; 2016.
- [67] Iv-Infra INPA120717-17-RAP-0043, Papendrecht, The Netherlands (in Dutch). 51H-304-01 - De Beek - Herberekening brugdek (Recalculation bridge deck) Report; 2015.
- [68] Lantsoght EOL, et al. Pilot proof-load test on viaduct De Beek: case study. *J Bridge Eng* 2017;22(12).
- [69] NEN Beoordeling van de constructieve veiligheid van een bestaand bouwwerk bij verbouw en afkeuren - Grondslagen (Assessment of the structural safety of an existing structure for renovation and rejection - Basic requirements). Standard, NEN 8700:2011, Nederlands Normalisatie-instituut, Delft, The Netherlands; 2011.
- [70] RWS Richtlijnen beoordeling kunstwerken - Beoordeling van de constructieve veiligheid van een bestaand kunstwerk bij verbouw, gebruik en afkeur (Guidelines for assessing structures - Assessment of the structural safety of an existing structure for renovation, use and rejection). Rijkswaterstaat, Report, Version 1.1, 27 May; 2013.
- [71] CEN Standard, EN 1991-2+C1:2015, European Committee for Standardization (CEN), Brussels, Belgium, Eurocode 1: Actions on structures - Part 2: Traffic loads on bridges; 2015.
- [72] Lantsoght EOL, et al. Stop criteria for flexure for proof load testing of reinforced concrete structures. *Front Built Environ* 2019;5.
- [73] Mörsch E. *Der Eisenbetonbau: seine Theorie und Anwendung* (Reinforced concrete: theory and practice), Konrad Wittwer, Stuttgart, Germany; 1908.
- [74] Wight JK, MacGregor JG. *Reinforced concrete: Mechanics and design*. 6th edition. New Jersey, USA: Pearson Education, Inc; 2012.
- [75] Lantsoght EOL, et al. Extended strip model for reinforced concrete slabs under concentrated loads. *Acids Struct J* 2017;114(2):565–74.
- [76] Lantsoght E.O.L., Ospina C.E., Alexander S.D.B. and , Punching capacity of spread footings using ACI 318-19 and the strip model ACI Special Issue SP-357: Punching shear of concrete slabs: insights from new materials, tests, and analysis methods, American Concrete Institute, Farmington Hills, USA; 2023.
- [77] De Borst R, et al. *Non-linear finite element analysis of solids and structures*. 2nd edition. West Sussex, UK: John Wiley & Sons, Ltd; 2012.
- [78] Lantsoght EOL, et al. Optimizing finite element models for concrete bridge assessment with proof load testing. *Front Built Environ* 2019;5.
- [79] Thorenfeldt E., Tomaszewicz A., Jensen J.J. and , Mechanical properties of high-strength concrete and applications in design Proceedings of the Symposium on the Utilization of High-Strength Concrete, Stavanger, Norway; 1987.