Parameter estimation and uncertainty quantification of a fiber-reinforced concrete model by means of a multi-level Bayesian approach

Federico Ponsi [b], Elisa Bassoli [a],[*], Nicola Buratti [b], Loris Vincenzi [a]

[a] University of Modena and Reggio Emilia, Department of Engineering "Enzo Ferrari" (DIEF), Modena, Italy
[b] University of Bologna, Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), Bologna, Italy

ARTICLE INFO

ABSTRACT

The paper presents a procedure for the stochastic calibration of a cracked hinge model on the basis of an extensive experimental campaign performed on a large group of nominally identical fiber-reinforced specimens. The calibration is carried out in a multi-level Bayesian framework that allows to quantify and separate several uncertainty contributions affecting model parameters. Indeed, the variability in the experimental response for nominally identical specimens due to the material heterogeneity represents a significant uncertainty contribution as well as model error. The former can be quantified at the hyper-parameter level of the multi-level framework. The presented results highlight the good agreement of the numerical predictions with the experimental data and the superior performance of the multi-level framework compared to that of the classical single-level framework. We also perform analyses to explore the impact of the prior parameter model conditioned on hyper-parameters and assess the minimum number of specimen datasets needed to quantify the inherent variability of model parameters.

1. Introduction

Non-linear mechanical models have a critical role in nowadays engineering and research applications, in particular in the field of advanced cementitious materials. The parameters of these models [1,2] are usually estimated through calibration procedures based on inverse analysis [3]. For instance, for Fiber Reinforced Concretes (FRC) data from three-point bending tests (TPBT) on notched beams or from wedge splitting tests (WST) [4] can be used to this purpose. From a deterministic point of view, calibration is performed by minimizing the difference between some experimental data and the corresponding prediction provided by a model, for example the experimental force–CMOD (Crack Mouth Opening Displacement) curve obtained from a TPBT can be used to this purpose. The calibrated parameters are found by solving an optimization problem, based on the definition of a cost function measuring the discrepancy between numerical and simulated results. The parameters to calibrate can be associated to models with different degrees of complexity; for WSTs, Finite Element (FE) models are commonly used [5–7], while for TPBTs, both closed form solutions [8–12] and numerical simulations [7,13] have been developed. Among these models, the fictitious crack model proposed by Hillerborg [14] is probably one of the most adopted. However, other similar models can be found in literature, such as the cracked band model by Bazant and Oh [15].

The spectrum of optimization strategies and algorithms which have been adopted in literature for deterministic parameter calibration is wide, ranging from gradient-based [6,11] to non gradient-based algorithms [16,17]. The authors [18], for example, developed a smart strategy aimed at calibrating the parameters of models against multiple experimental results, proposing an application on TPBTs on FRC elements. The deterministic approach to model parameter calibration is focused only on the determination of the optimal parameter values, namely those minimizing the cost function of the optimization problem.

However, model calibration is subjected to several sources of uncertainties. In the literature it is frequent to distinguish between the uncertainty related to experimental data and the uncertainty related to the prediction model. The first one is commonly due to random measurement noise, systematic error caused by the measurement equipment or data post-processing. Uncertainty related to the adopted prediction model has been analyzed by many authors [19–21] that identified two main contributions: model form and model parameter uncertainty. The former is caused by deliberate or unintended assumptions and simplifications that are formulated to develop a model and are the result of a lack of knowledge or understanding of the true system [22]. On the other hand, uncertainty about the value of input parameters of the selected prediction model is defined as model parameter uncertainty. It can be associated to unknown properties of the model (e.g. regarding

materials or geometry) but also to inherent structural variability [23] that is an aleatory component of uncertainty. The latter is particularly relevant for the application under consideration. An example in the context of model calibration of constitutive laws is the variability of material mechanical parameters (elastic modulus and/or strength) due to material heterogeneity. This is particularly important in modeling the behavior of materials like FRC, because the effect of material heterogeneity and random distribution of fibers in the concrete matrix leads to a very high variability of the experimental force–CMOD curves. Indeed, it is very common that the results of several tests performed on FRC specimens with the same characteristics may significantly differ one from another. Another example from the dynamic identification field is the variability of structural properties such as mass, stiffness and boundary conditions due to changing environmental conditions [24]. This source of uncertainty is irreducible, and should not be confused with the uncertainty in the estimations of the model parameters, that can be reduced by acquiring new information. In many cases the extent of these uncertainties is not-negligible and the calibration should be treated as a statistical inference problem. This allows to estimate the effect of model and measurement uncertainties on both the calibrated parameters and the predictions of the calibrated model.

To the best of the authors' knowledge, works aimed at the application of statistical inference for the calibration of FRC models are limited. Buratti et al. [17] carried out the deterministic calibration of a cracked hinge model on the basis of the response of several groups of nominally identical fiber-reinforced concrete specimens and then computed the statistics of the calibrated parameters obtaining an high variability for each group. Cavalaro and Aguado [25] computed the coefficient of variation of the residual strength with reference to an extensive parametric study. Bolzon et al. [5] applied the Kalman Filter methodology to identify parameter and related uncertainties of a cohesive crack model with reference to concrete specimens subjected to WSTs. Dobrilla et al. [26] carried out the identification of fracture and bond parameters in reinforced concrete by means of Bayesian inference. Simwanda et al. [27] conducted a reliability analysis of ultra high-performance fiber-reinforced concrete beams exposed to fire after the stochastic calibration of their temperature-dependent thermal properties.

In order to expand the research in this field, the present work proposes the use of a Bayesian approach for model calibration. This approach is based on the Bayes' theorem for the expression of probability of a certain quantity conditioned on acquired information and on the interpretation of probability as relative plausibility given the current information level [28]. In this way, a rigorous probabilistic framework for model calibration based on probability logic can be exploited. Reference works about the application of Bayesian inference for model calibration have been developed by Beck and Katafygiotis [29], Beck [30] and Tarantola [31]. Relevant applications of Bayesian methods in civil engineering are reported in [32], including vibration-based model updating [33–36], structural health monitoring and damage assessment [21,37,38] and seismic attenuation prediction [39].

The novelty of the present paper is that Bayesian inference is performed in a hierarchical or multi-level framework [40,41] for the calibration of a FRC model. In this way, parameter uncertainty due to the inherent structural variability can be quantified in a rigorous probabilistic context and it can be also separated from other uncertainty sources. It represents an innovative approach in the field of FRC model calibration. The introduction of a hyper-parameter level in the updating framework allows the exchange of information among different datasets. Consequently, the inherent parameter variability can be quantified by the hyper-parameters while the remaining uncertainty contributions are embedded in the prediction error. Hierarchical Bayesian model updating has been successfully applied in the field of structural dynamics based on modal features [23,24,42] or time-domain responses for the identification of nonlinear hysteric models [43]. Moreover, hierarchical modeling has also been introduced in sparse Bayesian learning for structural damage identification [44,45].

In this work, the hierarchical framework is applied for the parameter estimation of a cracked hinge model. The calibration is performed on the basis of an extensive experimental campaign based on TPBTs on a significant number of specimen characterized by the same concrete class and fiber dosage. The results of the hierarchical framework are compared to those of the single-level framework, highlighting the advantages in terms of uncertainty quantification and propagation.

The paper is organized as follows. First, the experiments on the FRC specimens and their mechanical modeling are introduced in Sections 2 and 3. The Bayesian approach for parameter estimation is described in Section 4 with reference to both the single-level and the hierarchical framework, while the corresponding results are presented and discussed in Section 5. Finally, conclusions are drawn in Section 6.

2. Experimental tests

The experimental campaign is carried out on $150 \times 150 \times 600$ mm^3 prismatic specimens that are nominally identical, namely they are cast with the same concrete and fiber dosage. The concrete mix used is detailed in Table 1. The fibers adopted are polymeric, have a length of 42 mm and an equivalent diameter of 0.8 mm, i.e. they are Class II according to EN 14889-2. After casting, the prismatic specimens are cured for 26 days in water at 20 °C, then a 25 mm deep and 4 mm wide notch is cut at mid-span on each specimen, as prescribed by EN 14651. After notching, the specimens are cured in water for two days in order to be tested at 28 days. Before testing, a clip-on displacement transducer is installed at the bottom of the prismatic specimen, across the notch, in order to measure the Crack Mouth Opening Displacement (CMOD) and control the test. Tests are carried out with a MTS Landmark servo-hydraulic machine, connected to a MTS Flextest 40 controller. Closed loop PID is used to control the CMOD opening during the tests; as prescribed by EN 14651, a rate of 0.05 mm/min is set for CMOD < 0.1 mm which is then increased to 0.1 mm/min for larger CMOD values.

A total of 84 specimens are tested and they are classified with the acronym composed of the letter *T* followed by a number ranging from 1 and 84. For example, the first tested specimen is denoted as T1. The high number of tests performed is motivated by the interest in identifying the required number of specimens to achieve the convergence of the proposed calibration procedure, as discussed in Section 5.3. Specimens are produced in different batches, in each batch 6 prisms and 2 $150 \times 150 \times 150$ mm^3 cubes are cast. The compressive strength measured from cubes at 28 days is 59.2 MPa with a coefficient of variation of 0.11, no systematic difference is observed among specimens from different batches. A picture of the experimental setup is shown in Fig. 1a, while Fig. 1b presents all the force–CMOD curve measured during the tests. During tests no anomalies are noticed in the position of the crack, which always starts from the tip of the notch. Fig. 2a shows an example of the crack observed at the end of one of the bending tests. Furthermore, at the end of each test the prismatic specimen is split into two parts in order to observe the distribution of fibers; Fig. 2b shows the crack surfaces for two of the specimens tested. Also in this case no anomalies are observed. It is possible to observe a noticeable variability in the experimental curves, both in the peak and post-peak region, that is mainly due to the heterogeneity of the material and the random distribution of fibers inside the concrete matrix. This is the source of the inherent variability of calibrated model parameters, that will be quantified thanks to the hierarchical framework described in Section 4. The peak force and the residual strengths at CMOD values of 0.5, 1.5, 2.5 and 3.5 mm for all the 84 specimens are listed in Table A.5 for completeness. Mean values and standard deviations among all the specimens are also reported in the same table, highlighting the significant variability of the experimental response.
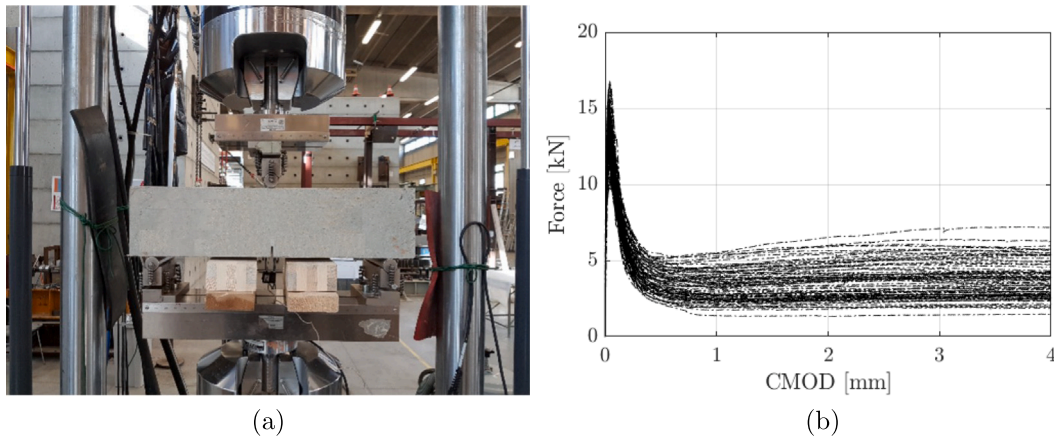
Fig. 1. (a) The experimental setup and (b) all the 84 force–CMOD experimental curves.

Fig. 2. (a) Crack observed at the end of one of the bending tests and (b) distribution of the fibers in the cross sections separated by the crack in the specimens T4 (left) and T6 (right). Fibers are marked with black dots.
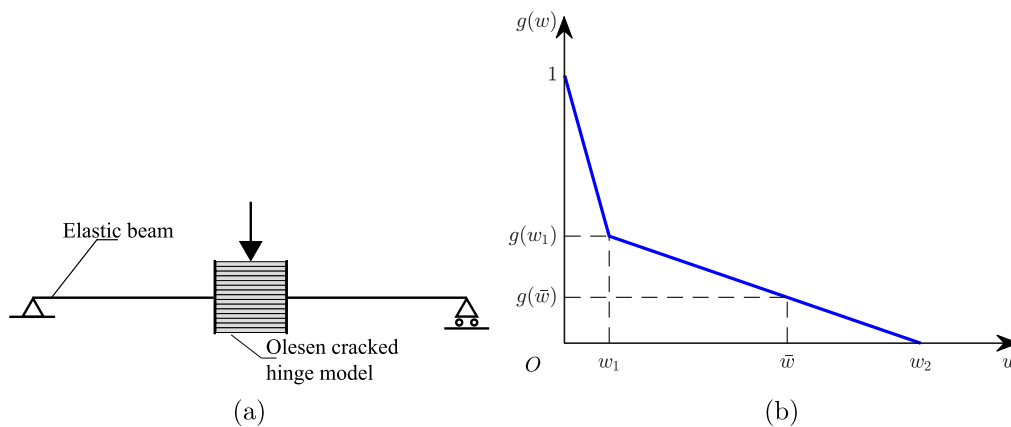

Fig. 3. (a) Schematic representation of the numerical model and (b) bi-linear stress–crack opening relationship.

3. Numerical modeling

The Bayesian procedure proposed in this paper can be applied to any parametric model, i.e. analytical, numerical etc., as far as it is possible to define a prediction error (Section 4). In this paper we adopt an analytical model adopted because of its very limited computational cost.

The behavior of the notched beams described in the previous section during the three-point bending test is simulated by combining the classical beam theory with the fictitious crack propagation model proposed by Olesen [8], which is a non-linear cracked hinge model based on the fracture mechanics concept of fictitious crack with a bi-linear stress–crack opening relationship. The part of the beam close to the propagating crack is considered as a layer of independent spring elements. These elements are formed by incremental horizontal strips and are attached at each end to a rigid boundary. Each rigid boundary is connected to an un-cracked beam element which behaves according to the classical beam theory (see Fig. 3a). The constitutive law of each

Table 1

Concrete mix used in the experimental campaign. Superplasticizer is added to the mix in order to achieve a slump of at least 150 mm.

Component	Dosage
CEM I 52.5 R	400 kg/m^3
Sand 0 mm–1 mm	885 kg/m^3
Sand 0 mm–5 mm	440 kg/m^3
Gravel 5 mm–15 mm	442 kg/m^3
Fibers	3.0 kg/m^3
Water/Cement ratio	0.45

horizontal strip in the fictitious crack is:

$$\sigma(w) = \begin{cases} E\varepsilon & \text{if } w = 0 \\ g(w) f_{ct} & \text{if } w > 0 \end{cases} \tag{1}$$

where E is the elastic modulus of concrete, ε is the elastic strain, f_{ct} is the tensile strength of concrete, w is the crack opening and $g(w)$ is, in this paper, a bi-linear curve defining the stress–crack opening relationship shown in Fig. 3b, that reads:

$$g(w) = \begin{cases} 1 - c_1 w & \text{if } 0 \le w \le w_1 \\ d_2 - c_2 w & \text{if } w_1 \le w \le w_2 \end{cases} \tag{2}$$

The parameters of the bi-linear curve are easily calculated as function of $w_1, g(w_1), w_2$ as $c_1 = (1 - g(w_1))/w_1$, $c_2 = g(w_1)/(w_2 - w_1)$ and $d_2 = g(w_1) + c_2 w_1$. The previous assumptions allow to formulate closed-form solutions for the moment–rotation relationship of the hinge and the force–CMOD curve of a three-point bending test, that are implemented in a MATLAB framework. To this aim, it is convenient to describe the hinge deformation by the rotation 2φ and the neutral axis depth y_0, that are related to the average curvature $\bar{\chi}$ and the average longitudinal strain $\bar{\varepsilon}$ as:

$$\bar{\chi} = \frac{2\varphi}{s}; \qquad \bar{\varepsilon}(y) = (y - y_0)\bar{\chi} \tag{3}$$

where s indicates the length of the nonlinear hinge which corresponds to the length of the nonlinear springs. The elongation $u(y) = s\bar{\varepsilon}(y)$ of a spring located at a depth y is given by the sum of the elastic elongation $(s\sigma(w(y))/E)$ and the crack opening $(w(y))$. By combining the previous relationship, the stress distribution along y can be expressed as:

$$\sigma(w(y)) = \left[2(y - y_0)\varphi - w(y)\right]\frac{E}{s} \tag{4}$$

The moment–rotation curve can be numerically obtained by varying the hinge rotation value inside a physically reasonable range and for each value by imposing the translation and rotational equilibrium of the hinge in order to compute the neutral axis depth y_0 and the bending moment. The stress distribution of the hinge changes through four distinct phases corresponding to $w = 0$, $0 < w < w_1$, $w_1 < w < w_2$ and $w > w_2$. For this reason, it is necessary to define transition points in order to correctly write the equilibrium equations. Transition points represent the rotation values causing the transition from a phase to the following one. Finally, the force corresponding to each moment value is obtained through beam equilibrium considerations. The complete formulation is described in [8].

4. Bayesian approach to model calibration

Model parameters obtained through inverse analysis are affected by different sources of uncertainty, such as errors and noise in the measure of an experimental data and uncertainty related to the prediction model. The main contributions to the last source of uncertainty are model form and model parameter uncertainty [20,21]. The latter can be associated to unknown properties of the model (e.g. regarding materials or geometry) but also to the inherent structural variability [23] induced by the aleatory of the system response. When the magnitude of measurement and/or model uncertainty is significant, the

deterministic approach to model calibration may not be the optimal strategy and the Bayesian approach is more suitable to quantify the effect of uncertainty on calibrated model parameters. Single-level and hierarchical Bayesian approach to model calibration are described in Sections 4.1 and 4.2, respectively. In the following, the vector collecting the model parameters subjected to calibration is denoted by $\mathbf{x}$.

4.1. Single-level Bayesian model updating

Bayes' theorem allows to infer posterior parameter distribution $p(\mathbf{x}|D)$ on the basis of a set of measured data D and the prior knowledge about the parameter values, represented by the prior distribution $p(\mathbf{x})$:

$$p(\mathbf{x}|D) = \frac{p(D|\mathbf{x})p(\mathbf{x})}{p(D)} \tag{5}$$

where $p(D)$ is the Bayesian evidence, a constant ensuring that the posterior distribution integrates to one, and it is computed as:

$$p(D) = \int_{\mathbf{I}} p(D|\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x} \tag{6}$$

where $\mathbf{I} \subset \mathbf{R}^n$ is the support of the prior distribution. The Bayesian evidence is usually employed for model class selection since it represents a trade-off between the ability of a model to fit the data and its complexity [30]. The dependency of all the terms of Eq. (5) on the stochastic model class (obtained by means of stochastic embedding of a deterministic mechanical model [46]) is here omitted but the reader could refer to [36] for all details.

$p(D|\mathbf{x})$ is the likelihood function representing the plausibility that the model class parameterized by $\mathbf{x}$ provides the measured data D. It reflects the contribution of data in the determination of the posterior distribution of parameters. The formulation of the likelihood function depends on the definition of the prediction error that in turn represents the discrepancy between the experimentally measured and numerically predicted features. In this study, the prediction error is defined based on the forces at eight reference values ($N = 8$) of CMOD, namely 0.03, 0.05, 0.1, 0.2, 0.5, 1.5, 2.5, and 3.5 mm. The corresponding force values derived from the experimental and numerical force–CMOD curve are collected in the vectors $\mathbf{F}^{(\text{ex})}$ and $\mathbf{F}^{(\text{num})}$, respectively. Consequently, the prediction error is modeled as a multivariate zero-mean Gaussian distribution:

$$e_F = \mathbf{F}^{(\text{ex})} - \mathbf{F}^{(\text{num})}(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_F) \tag{7}$$

Given the error definition of Eq. (7), the likelihood function values correspond to the probability density values of a multivariate Gaussian distribution with mean $\mathbf{F}^{(\text{ex})}$ and covariance matrix $\mathbf{\Sigma}_F$, computed for each vector $\mathbf{F}^{(\text{num})}(\mathbf{x})$, namely $p(D|\mathbf{x}) = \mathcal{N}\left(\mathbf{F}^{(\text{num})}(\mathbf{x})|\mathbf{F}^{(\text{ex})}, \mathbf{\Sigma}_F\right)$. The covariance matrix $\mathbf{\Sigma}_F$ summarizes the uncertainties in the prediction error, that in this formulation combines measurement and model error. The simplest approach involves the use of a diagonal covariance matrix with a single error parameter σ_F. An incorrect assumption regarding the prediction error may unfairly influence the Bayesian updating results. For this reason, it is possible to include the prediction error parameter in the updating process [21,47]. In this case, the vector $\mathbf{x}$ also includes the error parameter σ_F in addition to the model parameters.

When multiple datasets $\{D_i\}_{i=1}^{Nd}$ are available, with Nd representing the total number of datasets, Eq. (5) becomes:

$$p\left(\mathbf{x}|\{D_i\}_{i=1}^{Nd}\right) = \frac{p\left(\{D_i\}_{i=1}^{Nd}|\mathbf{x}\right)p(\mathbf{x})}{p\left(\{D_i\}_{i=1}^{Nd}\right)} \tag{8}$$

The likelihood function $p\left(\{D_i\}_{i=1}^{Nd}|\mathbf{x}\right)$ can be expressed as the product among the individual likelihood functions $p(D_i|\mathbf{x})$ under the assumptions of independence among different datasets, namely:

$$p\left(\{D_i\}_{i=1}^{Nd}|\mathbf{x}\right) = \prod_{i=1}^{Nd} p(D_i|\mathbf{x}) \tag{9}$$
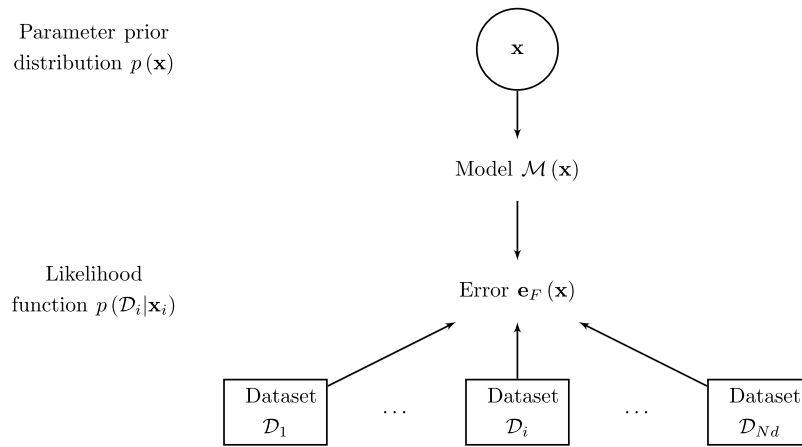
Fig. 4. Graphical representation of the single-level Bayesian framework.

In this work, the term dataset refers to the result of the experimental test on a single specimen, and corresponds to the set of force values at the aforementioned reference CMOD values. The schematic representation of the single-level framework in case of multiple datasets is depicted in Fig. 4 where inferred variables are denoted by a circle, while a rectangle is used for the observed quantities.

The computation of the joint posterior distribution $p\left(\mathbf{x}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ and of the marginal distributions is limited in practical applications by the need to solve high-dimensional integrals, for example the one of Eq. (6). For this reason, asymptotic approximations [29] or sampling methods [48] are generally employed, in which also an estimate of the Bayesian evidence $p(D)$ can be available. In this work, the sampling of the posterior distribution is performed through the Bayesian Annealed Sequential Importance Sampling (BASIS) algorithm [49], an improved version of the Transitional Markov Chain Monte Carlo algorithm [50] that has been largely diffused for Bayesian inference in structural engineering problems.

4.2. Hierarchical Bayesian model updating

The single-level updating has been successfully implemented for the identification of model parameters and model class selection in a variety of structural engineering problems. However, the covariance matrix of the posterior distribution $p\left(\mathbf{x}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ does not provide the total variability of updating model parameters [40]. This covariance matrix only represents the epistemic uncertainty, namely the uncertainty in the parameter mean estimation which will decrease with an increasing number of datasets [37,51,52]. The inherent variability of a parameter, that is the variability of the calibrated parameters among different experiments, cannot be quantified in a single-level framework.

Hierarchical or multi-level updating is an extension of the classical single-level framework and it allows the exchange of information among different datasets [41]. As the name suggests, the overall stochastic model is composed of different parameters hierarchically organized in different levels. The parameter-level is composed of the parameter vectors $\mathbf{x}_i$ ($i = 1, \ldots, Nd$), which contain also the error parameter $\sigma_{F,i}$. They are separately inferred from the datasets $\{\mathcal{D}_i\}_{i=1}^{Nd}$. Realizations of the parameters $\mathbf{x}_i$ can vary across each different dataset $\mathcal{D}_i$, thus the collection of all the parameter vectors of the hierarchical model is denoted as $\{\mathbf{x}_i\}_{i=i}^{Nd}$. The extra level that distinguishes the hierarchical framework is composed of the hyper-parameters, collected in the vector $\boldsymbol{\psi}$ and characterized in terms of their probability distribution. The inherent variability is quantified through the hyper-parameters, while all the other uncertainties are accounted for through the prediction error term. The authors believe that this approach is well suited to quantify the parameter variability of FRC models given the high scattering of the experimental response, as indicated in Section 2.

At the same time, this approach has not been investigated in the field of FRC model calibration and this work has been developed for this purpose. The graphical representation of the proposed hierarchical framework is depicted in Fig. 5.

The joint posterior distribution of parameters and hyper-parameters conditioned on the available datasets $\{\mathcal{D}_i\}_{i=1}^{Nd}$ can be obtained by means of the Bayes' theorem:

$$p\left(\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right) = C^{-1} p\left(\{\mathcal{D}_i\}_{i=1}^{Nd}|\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}\right) \times \left(p\left(\{\mathbf{x}_i\}_{i=1}^{Nd}|\boldsymbol{\psi}\right)p(\boldsymbol{\psi})\right) \tag{10}$$

where $p(\boldsymbol{\psi})$ denotes the prior distribution of hyper-parameters, $p\left(\{\mathbf{x}_i\}_{i=1}^{Nd},|\boldsymbol{\psi}\right)$ is the prior distribution of parameters conditioned on the hyper-parameters, C is the evidence of the hierarchical model class and $p\left(\{\mathcal{D}_i\}_{i=1}^{Nd}|\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}\right)$ is the joint likelihood function of the hierarchical model class. Given the independence of individual datasets and the dependence of $\mathbf{x}_i$ only on the dataset $\mathcal{D}_i$, the prior $p\left(\{\mathbf{x}_i\}_{i=1}^{Nd},|\boldsymbol{\psi}\right)$ and the likelihood function $p\left(\{\mathcal{D}_i\}_{i=1}^{Nd}|\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}\right)$ can be simplified as:

$$p\left(\{\mathbf{x}_i\}_{i=1}^{Nd}|\boldsymbol{\psi}\right) = \prod_{i=1}^{Nd} p\left(\mathbf{x}_i|\boldsymbol{\psi}\right) \tag{11}$$

$$p\left(\{\mathcal{D}_i\}_{i=1}^{Nd}|\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}\right) = \prod_{i=1}^{Nd} p\left(\mathcal{D}_i|\mathbf{x}_i\right) \tag{12}$$

$p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ and $p(\mathcal{D}_i|\mathbf{x}_i)$ are the prior distribution conditioned on the hyper-parameters and the likelihood of an individual dataset $\mathcal{D}_i$. We remark a substantial difference between the likelihood of the hierarchical framework (Eq. (12)) and the likelihood of the single-level framework in case of multiple datasets (Eq. (9)) because Eq. (12) considers a specific parameter vector $\mathbf{x}_i$ for each dataset $\mathcal{D}_i$. As happened in Section 4.1 for the single-level framework, the dependency of the terms in Eqs. (10)–(12) on the hierarchical model class $\{\mathbf{x}_i\}_{i=1}^{Nd}$ has been omitted. The same holds for the following discussion.

Our goal is to obtain the marginal posterior distribution of the hyper-parameters $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ and of the parameters $p\left(\mathbf{x}_i|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$, which can be theoretically obtained by marginalizing the joint posterior distribution of parameters and hyper-parameters $p\left(\{\mathbf{x}_i\}_{i=1}^{Nd},\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ of Eq. (10). Because of the complexity associated with directly evaluating the joint posterior distribution, alternative methodologies can be adopted for their computation [40,41,43,53,54], which in most cases allow to avoid the computation of the joint posterior distribution. We adopt the one proposed by Wu et al. [55], that involves an efficient approximation based on the use of Importance Sampling [56] and it is outlined in Section 4.2.1.
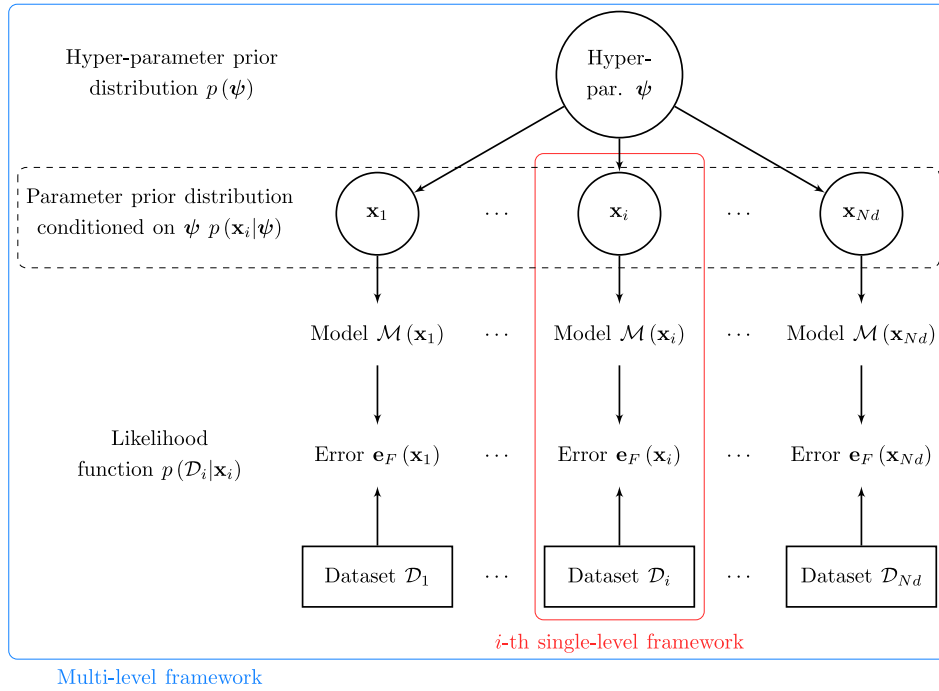
Fig. 5. Graphical representation of the hierarchical Bayesian framework.

4.2.1. Numerical sampling in the hierarchical framework

The procedure for sampling marginal posterior distribution of the hyper-parameters and parameters is summed up through the flowchart represented in Fig. 6. In step 1, single-level Bayesian model updating (Eq. (5)) is performed separately for each dataset $\mathcal{D}_i$ by sampling the posterior distribution $p\left(\mathbf{x}_i | \mathcal{D}_i\right)$. A predefined number Ns of samples is drawn for each dataset by means of the BASIS algorithm introduced in Section 4.1 and also an estimate of the Bayesian evidence $p\left(\mathcal{D}_i\right)$ is computed. Then, step 2 aims to calibrate the hyper-parameters whose posterior distribution $p\left(\boldsymbol{\psi} | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)$ is expressed through the Bayes' theorem as:

$$p\left(\boldsymbol{\psi} | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right) = \frac{p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd} | \boldsymbol{\psi}\right) p\left(\boldsymbol{\psi}\right)}{p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)} \qquad (13)$$

where $p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd} | \boldsymbol{\psi}\right)$ is the likelihood for $\boldsymbol{\psi}$, $p(\boldsymbol{\psi})$ is the prior of the hyper-parameters $\boldsymbol{\psi}$ and $p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)$ is the evidence for $\boldsymbol{\psi}$, a constant ensuring that the posterior distribution $p\left(\boldsymbol{\psi} | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)$ integrates to one. Its exact computation is characterized by the same issues highlighted for the single-level Bayesian procedure in Section 4.1. For this reason, the natural logarithm of the evidence, log-evidence in the following, is numerically estimated by the BASIS sampling algorithm as a by-product.

The computation of the likelihood $p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd} | \boldsymbol{\psi}\right)$ needs the repeated evaluations of the likelihood $p\left(\mathcal{D}_i | \mathbf{x}_i\right)$ for each dataset $\mathcal{D}_i$, that represents the largest contribution to the total computational effort of the whole procedure. The idea is to approximate $p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd} | \boldsymbol{\psi}\right)$ by using Importance Sampling and exploiting the samples of the posterior distribution $p\left(\mathbf{x}_i | \mathcal{D}_i\right)$ for each dataset (drawn in step 1), as:

$$p\left(\left\{\mathcal{D}_i\right\}_{i=1}^{Nd} | \boldsymbol{\psi}\right) \approx \prod_{i=1}^{Nd} \frac{p\left(\mathcal{D}_i\right)}{Ns} \sum_{k=1}^{Ns} \frac{p\left(\mathbf{x}_i^{(k)} | \boldsymbol{\psi}\right)}{p\left(\mathbf{x}_i^{(k)}\right)} \qquad (14)$$

where $\mathbf{x}_i^{(k)}$ is the kth sample of the proposal distribution $p\left(\mathbf{x}_i | \mathcal{D}_i\right)$, with $k = 1, 2, \ldots, Ns$, and provided that Ns is sufficiently large. Moreover, $p\left(\mathcal{D}_i\right)$ is the evidence of the single-level Bayesian procedure, estimated

in the first step for each dataset $\mathcal{D}_i$, $p\left(\mathbf{x}_i^{(k)} | \boldsymbol{\psi}\right)$ and $p\left(\mathbf{x}_i^{(k)}\right)$ are the values of the priors $p\left(\mathbf{x}_i | \boldsymbol{\psi}\right)$ and $p\left(\mathbf{x}_i\right)$ for the sample $\mathbf{x}_i^{(k)}$. $p\left(\mathbf{x}_i\right)$ is defined in the first step and it is assumed as an uniform distribution in the following. Hence, even if, in the general case, $p\left(\mathbf{x}_i^{(k)}\right)$ depends on k and on the dataset $\mathcal{D}_i$, in this case it is a constant since the same bounds of the uniform distribution are adopted for all the datasets. On the contrary, the prior $p\left(\mathbf{x}_i^{(k)} | \boldsymbol{\psi}\right)$ typically follows a Gaussian distribution, which is assumed in the following.

The final step 3 aims to drawn sample from $p\left(\mathbf{x}_i | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)$. On the basis of the dependency assumptions previously mentioned and using Importance Sampling, it can be written as:

$$p\left(\mathbf{x}_i | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right) \approx \frac{p\left(\mathcal{D}_i | \mathbf{x}_i\right)}{Ns} \sum_{k=1}^{Ns} \frac{p\left(\mathbf{x}_i | \boldsymbol{\psi}^{(k)}\right)}{p\left(\mathcal{D}_i | \boldsymbol{\psi}^{(k)}\right)} \qquad (15)$$

where $\boldsymbol{\psi}^{(k)}$ is the kth sample of $p\left(\boldsymbol{\psi} | \left\{\mathcal{D}_i\right\}_{i=1}^{Nd}\right)$ and the values of the likelihood $p\left(\mathcal{D}_i | \boldsymbol{\psi}^{(k)}\right)$ are known from the previous step. Moreover, $p\left(\mathbf{x}_i | \boldsymbol{\psi}^{(k)}\right)$ is the value of the prior distribution $p\left(\mathbf{x}_i | \boldsymbol{\psi}\right)$ whose parameters are defined by the sample $\boldsymbol{\psi}^{(k)}$.

As reported in Wu et al. [55], the variance of results is strictly connected with the closeness of the integrand and of proposal distribution. Solutions to reduce the variance are the choice of a prior $p\left(\mathbf{x}_i\right)$ very close to $p\left(\mathbf{x}_i | \boldsymbol{\psi}\right)$ or the use of a large number N_s of posterior samples. Given the difficulty of knowing the appropriate prior parameters at the beginning of the procedure, the simplest solution to obtain good accuracies is to choose a large number of samples N_s. The reader could refer to [55,57] for the complete description of the procedure.

5. Results

In the present section, model parameters of the cracked hinge model described in Section 3 are calibrated on the basis of the Bayesian framework of Section 4. Model parameters to be calibrated are the tensile strength of concrete f_{ct} and the parameters of the bi-linear stress–crack opening relationship in Eq. (2). Concerning the parametrization of the bi-linear stress–crack opening curve for inverse analysis, it can be convenient to consider the parameter w_1, $g\left(w_1\right)$ and $g\left(\bar{w}\right)$ (see Fig. 3b) rather than c_1, c_2 and d_2, where $\bar{w}$ is a predefined value of

1. Single-level Bayesian updating of parameters $\mathbf{x}_i$

For $i =, 1, 2, \ldots, N_d$

2. Updating of hyper-parameters ψ

3. Refined updating of parameters $\mathbf{x}_i$
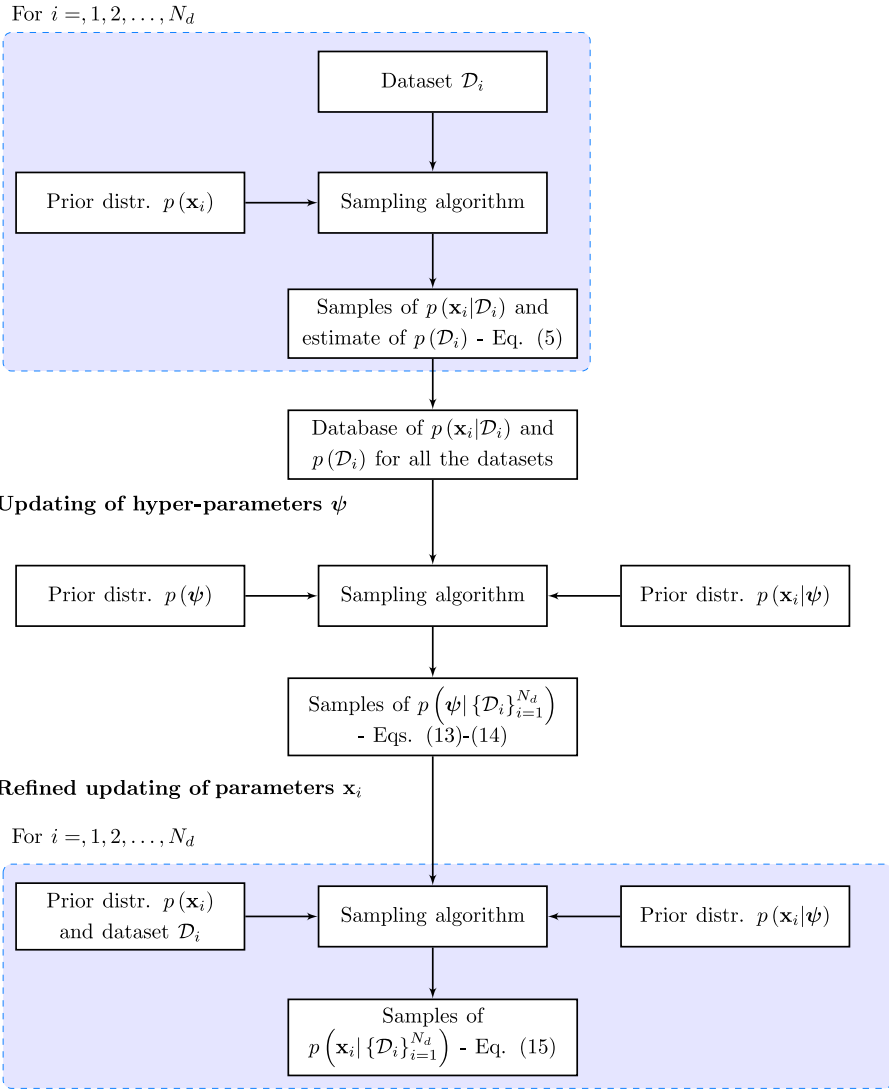
For $i =, 1, 2, \ldots, N_d$

Fig. 6. Flowchart of the procedure for hierarchical model updating.

crack opening that has been reached during the experimental tests, in this work $\bar{w}$=3.5 mm. In fact, it is more reasonable to calibrate $g(\bar{w})$ rather than w_2, that corresponds to a stress equal to zero. Indeed, w_2 is commonly larger than the maximum CMOD reached in a test and its calibration can result in an ill-posed problem. Consequently, the second line of Eq. (2) is rewritten as:

$$g(w) = g(w_1) - \frac{g(w_1) - g(\bar{w})}{\bar{w} - w_1}(w - w_1) \qquad \text{if } w_1 \leq w \leq w_2 \qquad (16)$$

This parametrization directly involves physical quantities that are easier to interpret and to bound. Indeed, if inappropriate bounds are selected for the parameters, nonlinear constraint are necessary in order to keep the physical significance of the model, see [8] for further details.

The calibration will be carried out in both the single-level (Section 5.1) and hierarchical (Section 5.2) framework, highlighting differences between them and the advantages of the hierarchical framework.

5.1. Results for the single-level updating

The calibration is first performed in the single-level framework described in Section 4.1. The posterior distribution of the 5 parameters

f_{ct}, w_1, $g(w_1)$, $g(\bar{w})$ and σ_F are inferred on the basis of Eq. (5) and the likelihood function defined by Eq. (9) for the case of multiple datasets. The joint prior distribution $p(\mathbf{x})$ of the updating parameter is obtained by considering the a priori independence of parameters unless for $g(w_1)$ and $g(\bar{w})$, since it is needed that $g(w_1) \geq g(\bar{w})$. The prior distribution of each parameter is described in Table 2. Note that the parameter $g(\bar{w})$ can assume negative values during the calibration. This means that the coordinate w_2 corresponding to $g(w_2)$=0 is lower than the predefined value of $\bar{w}$. In this case, negative values of $g(w)$ are not considered in the bi-linear curve, and the values for $w > w_2$ are set to zero.

The results of the single-level updating are summed up in the last columns of the same table. As reported by other authors [40,52], the model parameter standard deviations do not represent the inherent parameter variability but the uncertainty in the mean value estimation, that reduces with an increasing dimension of datasets (for instance, in the present case, increasing the number N of forces at prescribed values of CMOD used to compute the prediction error in Eq. (7)). Indeed, the standard deviations listed in Table 2 are very small, one or more orders of magnitude lower than the corresponding mean values. The comparison between experimental and numerical force–CMOD curves
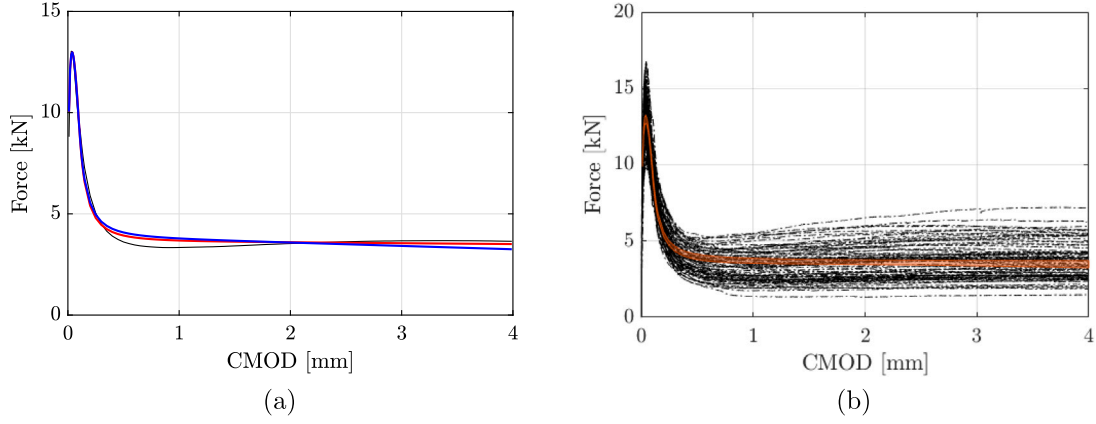
Fig. 7. Representation of experimental and numerical curves: (a) mean experimental curve (black) compared to the mean numerical predicted curve of the single-level (blue) and hierarchical (red) framework; (b) variability of the experimental curves (black dashed lines) and 99% uncertainty bounds for numerical predictions (red region) obtained by considering the model parameter uncertainty in the single-level framework. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
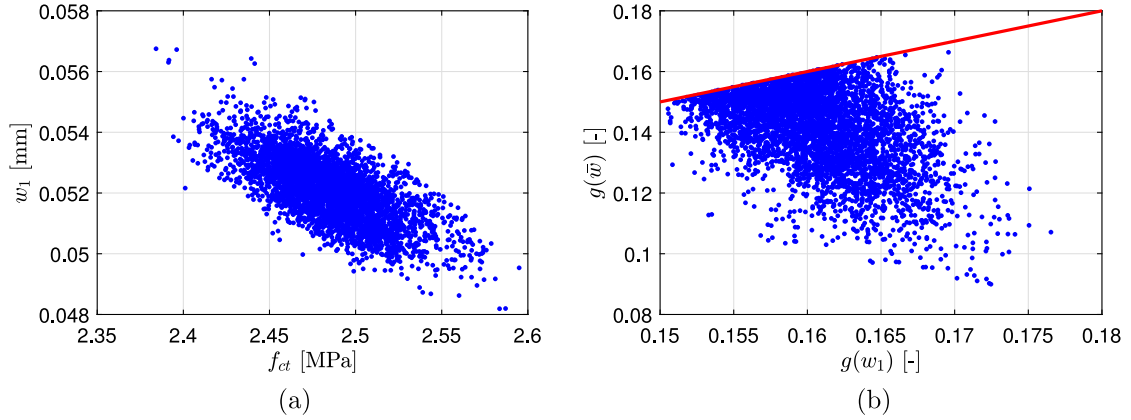

Fig. 8. Scatter plot of the most correlated model parameters in the single-level framework: (a) f_{ct} versus w_1 and (b) $g(w_1)$ versus $g(\bar{w})$.

is presented in Fig. 7. Fig. 7a shows the high correspondence between the experimental mean curve and the numerical mean curve, while Fig. 7b shows the 99% uncertainty bounds for model predictions obtained considering only the model parameter uncertainty. The bounds clearly do not cover the variability of the experimental response. It is worth noticing that larger prediction bounds can be obtained by adding the calibrated prediction error σ_F to the deterministic model response. In this case, two main drawbacks are highlighted. First, the bounds overestimate the experimental variability since the prediction error includes all the possible error sources in the single-level framework. Second, the inherent parameter variability is not directly quantified and it cannot be propagated to different model predictions.

As concerns the correlation among the updated parameters, the strongest correlations are observed for the pairs f_{ct}-w_1 and $g(w_1)$-$g(\bar{w})$, whose scatter plots are represented in Fig. 8. The corresponding correlation coefficients are -0.69 (f_{ct}-w_1, Fig. 8a) and -0.44 ($g(w_1)$-$g(\bar{w})$, Fig. 8b). From the last plot, it is easy to view the influence of the prior distribution imposing $g(w_1) \geq g(\bar{w})$. The red line represents the boundary line of equation $g(w_1) = g(\bar{w})$. The hypothesis of a posteriori uncorrelated parameters is not appropriate in this case. The absolute values of the correlation coefficients among all the other pairs of updated parameters are less than 0.3.

5.2. Results for the hierarchical updating

The calibration is thus carried out according to the hierarchical framework of Section 4.2. As regards the prior distribution $p(\mathbf{x}_i|\boldsymbol{\psi})$

Table 2
Prior $p(\mathbf{x})$ and posterior $p\left(\mathbf{x}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ distributions of the updating parameters in the single-level framework. a, b: left and right bounds of a univariate uniform distribution. μ_{po}, σ_{po}: mean and standard deviation of the posterior distribution $p\left(\mathbf{x}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ (see Eq. (8)).

Parameter	Units	Prior dist.	a	b	μ_{po}	σ_{po}
f_{ct}	MPa	Uniform	0.50	5	2.49	0.03
w_1	mm	Uniform	0.02	0.30	0.05	$1 \cdot 10^{-3}$
$g(w_1)$	–	Uniform[a]	0.01	0.99	0.16	$4 \cdot 10^{-3}$
$g(\bar{w})$	–		-0.50	0.99	0.14	0.01
σ_F	kN	Uniform	0.01	10	1.18	0.03

[a] Coupled distribution since $g(w_1) \geq g(\bar{w})$.

of parameters $\mathbf{x}_i$ conditioned on the hyper-parameters $\boldsymbol{\psi}$, Gaussian distributions are chosen for all the five parameters involved. Hence the hyper-parameter vector $\boldsymbol{\psi}$ has dimension equal to 10 and it is composed of the pairs mean-standard deviation of each of the five Gaussian distributions. Hyper-parameters are characterized by uniform prior distributions ($p(\boldsymbol{\psi}) \sim \mathbf{U}(a, b)$) and their posterior distributions are sampled by means of the approximated method described in Section 4.2.1. Table 3 sums up the characteristic of prior $p(\boldsymbol{\psi})$ and posterior distribution $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ for each hyper-parameter. Note that the symbols μ and σ used in first column of the table denote the hyper-parameters $\boldsymbol{\psi}$ related to the five updating parameters. $\mu_{f_{ct}}$, μ_{w_1}, $\mu_{g(w_1)}$, $\mu_{g(\bar{w})}$ and μ_{σ_F} represent the mean values of the updating parameters, while $\sigma_{f_{ct}}$, σ_{w_1}, $\sigma_{g(w_1)}$, $\sigma_{g(\bar{w})}$ and σ_{σ_F} quantify the inherent variability of the updating
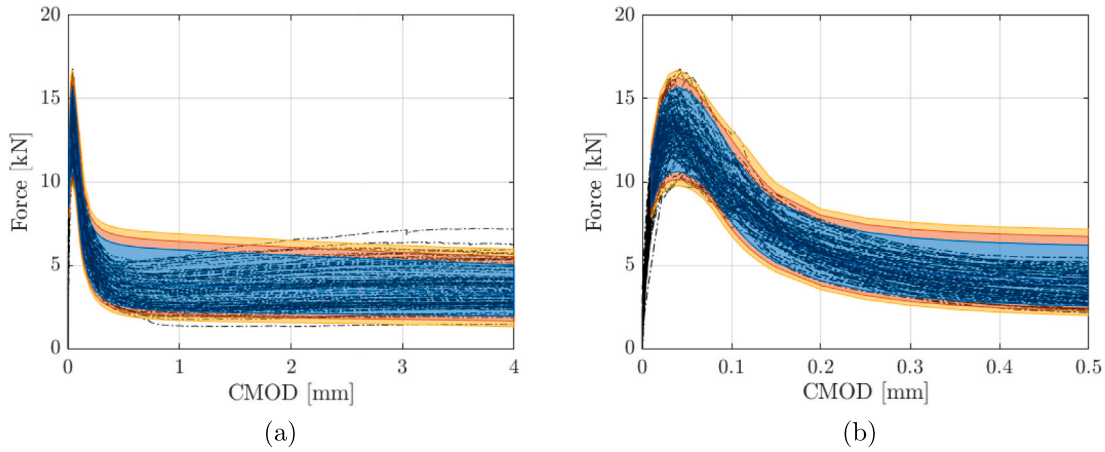
Fig. 9. Comparison between experimental curves (black dashed lines) and 99.9% (yellow), 99% (red) and 95% (blue) uncertainty bounds for numerical predictions obtained by considering the parameter inherent variability in the hierarchical framework. (a) entire curve, (b) zoom on the peak region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

Prior $p(\boldsymbol{\psi})$ and posterior $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$ distributions of the hyper-parameters in the hierarchical framework. a, b: left and right bounds of a univariate normal distribution. μ_{po}, σ_{po}: mean and standard deviation of the posterior distribution $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$.

Hyper-par.	Units	Prior dist.	a	b	μ_{po}	σ_{po}
$\mu_{f_{ct}}$	MPa	Uniform	0.50	5	2.49	0.03
$\sigma_{f_{ct}}$	MPa	Uniform	0.50	3	0.30	0.02
μ_{w_1}	mm	Uniform	0.02	0.30	0.05	$8\cdot10^{-4}$
σ_{w_1}	mm	Uniform	$1\cdot10^{-3}$	0.2	$7\cdot10^{-3}$	$8\cdot10^{-4}$
$\mu_{g(w_1)}$	–	Uniform	0.01	0.99	0.17	$5\cdot10^{-3}$
$\sigma_{g(w_1)}$	–	Uniform	0.01	0.60	0.05	$5\cdot10^{-3}$
$\mu_{g(\bar{w})}$	–	Uniform	−0.50	0.99	0.10	$6\cdot10^{-3}$
$\sigma_{g(\bar{w})}$	–	Uniform	0.01	0.60	0.04	$9\cdot10^{-3}$
μ_{σ_F}	kN	Uniform	0.01	3	0.59	0.02
σ_{σ_F}	kN	Uniform	0.01	1.5	0.07	0.02

parameters. This source of uncertainty is definitely not negligible since they correspond to coefficient of variations with respect to the mean value of 11.9%, 12.8%, 28.7%, 43.9% and 12.6%, respectively.

Conversely, μ_{po} and σ_{po} denote the mean and standard deviation of samples of the hyper-parameter posterior distribution $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{Nd}\right)$. σ_{po} represents the uncertainty in the hyper-parameter mean estimate, that is at least one order of magnitude lower than the mean estimate for most of the hyper-parameters. The mean predicted force–CMOD curve exhibits a close similitude with the mean experimental curve and with the mean predicted curve of the single-level updating (Fig. 7a).

The variability of predictions is investigated in Fig. 9. The uncertainty bounds of the numerical predictions have been computed by accounting for only the inherent parameter variability quantified by the hyper-parameter vector $\boldsymbol{\psi}$. Hence, the prediction error σ_F has not been included. The bounds clearly cover the variability of the experimental curve both in the peak (Fig. 9b) and post peak (Fig. 9a) regions. Only few portions of the experimental curves are not covered by the uncertainty bounds because of the error committed by the numerical model in the reproduction of the experimental response.

The effect of the introduction of the hierarchical framework on the parameter samples can be analyzed thanks to Fig. 10. The blue bars represent the histogram of the model parameter for all the 84 datasets obtained in the first phase of the method, namely after that the single-level Bayesian model updating has been performed for each dataset. The red bars represent the histogram of the model parameter obtained at the end of the hierarchical framework, namely when the prior distribution conditioned on the hyper-parameters $p(\mathbf{x}_i|\boldsymbol{\psi})$ has been included in the sampling process. At the end of the hierarchical updating, the samples are clearly more concentrated in a narrow range

with respect to first phase. This effect is more pronounced for w_2 and $g(\bar{w})$ (see Figs. 10b and 10d) but it is visible also for f_{ct} and $g(w_2)$ (see Figs. 10a and 10c). Moreover, the samples are very close to the probability distributions defined by the corresponding pair of hyper-parameters, depicted with a black line in the same figures.

The hierarchical framework has the clear advantage of allowing the direct quantification of the parameter inherent variability, as proved by the previous results. Moreover, the calibration of a single experimental test can be also refined. Two representative examples are shown in Fig. 11. The 95% uncertainty bounds of model predictions for specimen T79 and T81 are compared. The red region represents the bounds for the single-level updating, while the blue one represents the bounds for the hierarchical framework. The exchange of information among different datasets carried out in the hierarchical framework allows to reduce the uncertainty bounds in cases where the prediction error of the single-level framework $\sigma_{F,S}$ is high, as for the case of T81 (Fig. 11b). On the other hand, when the prediction error $\sigma_{F,S}$ is not excessively high, the reduction obtained thanks to the hierarchical framework is negligible, as for T79 (Fig. 11a). For a more comprehensive analysis, all the datasets have been considered. The reductions of the prediction error mean value from the single-level to the hierarchical framework ($\Delta\sigma_F = \sigma_{F,S} - \sigma_{F,H}$) have been compared. The reduction $\Delta\sigma_F$ is represented in function of the prediction error mean value of the single-level updating $\sigma_{F,S}$ in Fig. 12. $\Delta\sigma_F$ is approximately linearly proportional to $\sigma_{F,S}$, namely the reduction is the greater the higher the initial value of the prediction error in the single-level framework. It is interesting to note that $\Delta\sigma_F$ can be negative, hence a linear increment of the prediction error mean value can be performed in the hierarchical framework if $\sigma_{F,S}$ is lower than a threshold value, 0.56 kN in this case.

5.3. Effect of the dataset size on the parameter variability

The results of Section 5.2 refer to the calibration based on the overall available datasets (84 specimens). The authors have also analyzed the convergence of hyper-parameters according to a varying number of datasets, ranging from 1 to 84. First, the force–CMOD curves obtained by the experimental tests of the specimens T1-T84 are randomly shuffled obtaining a random sequence of 84 curves. Then, the hyper-parameter posterior distribution $p\left(\boldsymbol{\psi}|\{\mathcal{D}_i\}_{i=1}^{n}\right)$ is inferred according to a number n of datasets ranging from 1 to 84, by selecting the first n curves of the generated random sequence. Results of this analysis are represented in Fig. 13. As concerns the mean value of the 5 parameters, very few datasets (namely 3-4) are needed in order to have stability for the mean value in the case of w_1 (Fig. 13b), $g(w_1)$
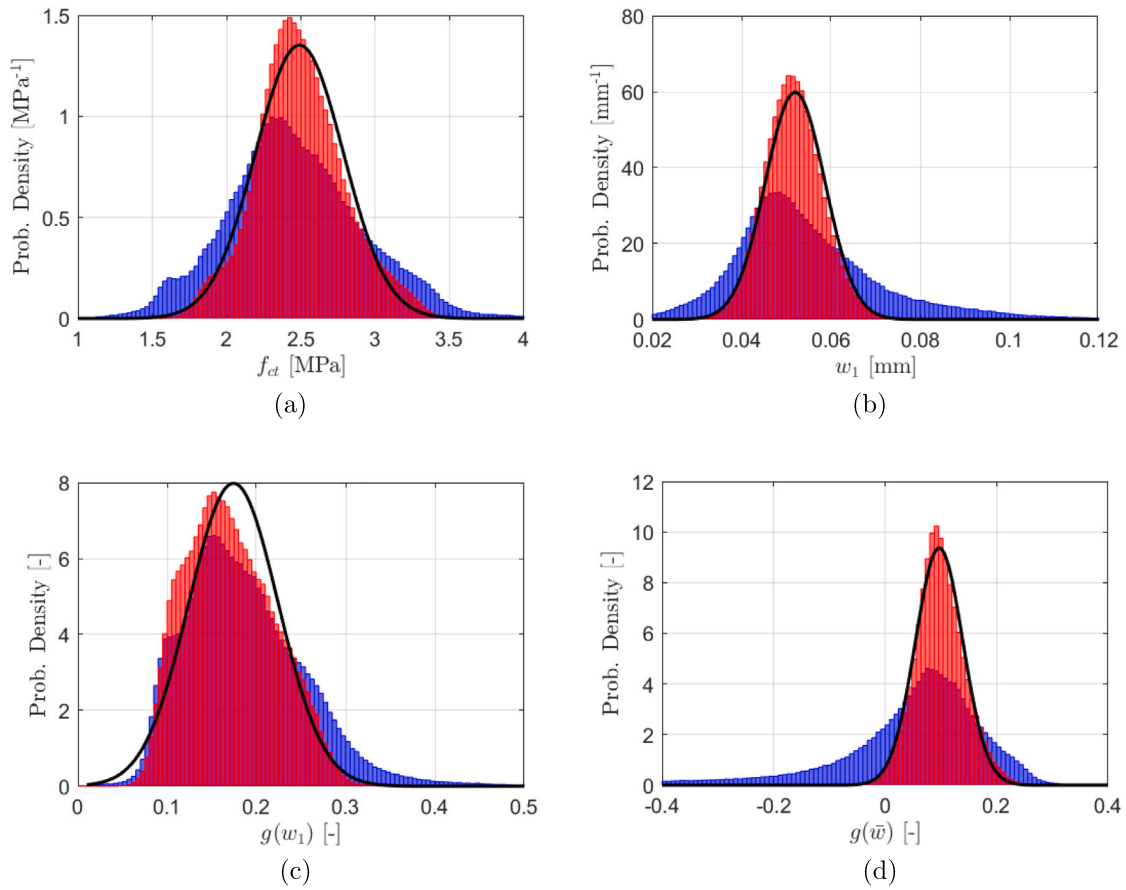
Fig. 10. Histogram of the model parameter for all the 84 datasets: (a) f_{ct}, (b) w_1, (c) $g(w_1)$ and (d) $g(\bar{w})$ in the first stage of the hierarchical framework (blue bars) and in the final stage (red bars). Black line: probability distribution of the corresponding prior $p(\mathbf{x}_i|\boldsymbol{\psi})$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
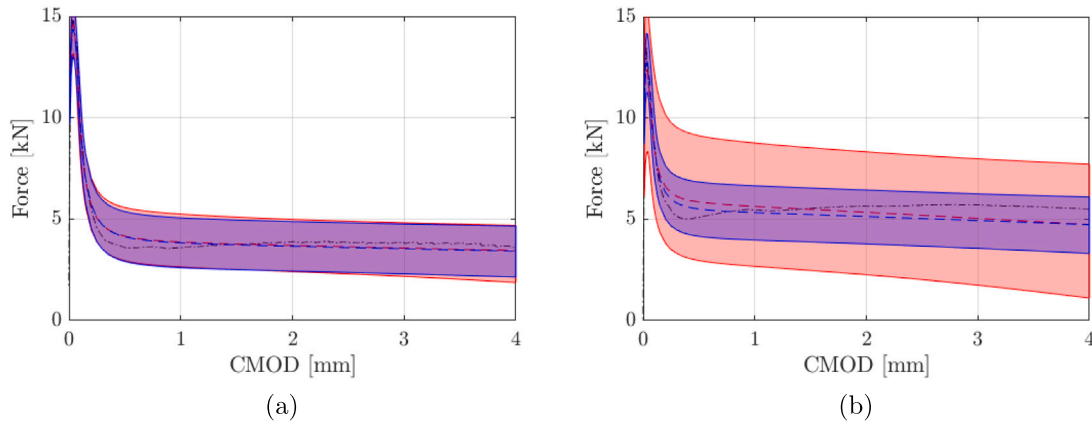

Fig. 11. 95% uncertainty bounds obtained for the (a) T79 and (b) T81 specimen according to the single-level (red region) and hierarchical (blue region) framework. Black line: experimental curve; red and blue dashed line: numerical mean curve of the single-level and hierarchical framework, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fig. 13c) and $g(\bar{w})$ (Fig. 13d). For f_{ct} (Fig. 13a) and σ_F (Fig. 13e) an higher number, approximately 10, is needed.

As concerns the parameter variability, a minimum number of datasets equal to 16 is required to reach quite stable uncertainty bounds. This number is the result of the convergence analysis described in the following. First, the values of the hyper-parameters $\sigma_{f_{ct}}$, σ_{w_1}, $\sigma_{g(w_1)}$ and $\sigma_{g(\bar{w})}$ updated on the basis of an increasing number of datasets, i.e. from 1 to 84, are collected in specific vectors. These vectors are then normalized by dividing each of their components by the mean value of the vector. Finally, a single vector $\boldsymbol{\eta}$ representing

a global index for all the hyper-parameters is defined by adding the squared normalized vectors. The minimum number of datasets is determined as the point where the global index last exceeds a threshold equal to 1.3 times the value of $\boldsymbol{\eta}$ calculated for 84 specimens.

5.4. Alternative parametrizations and prior distributions

The results of Section 5.2 refer to the parametrization of the bi-linear stress-crack opening curve in terms of f_{ct}, w_1, $g(w_1)$ and $g(\bar{w})$
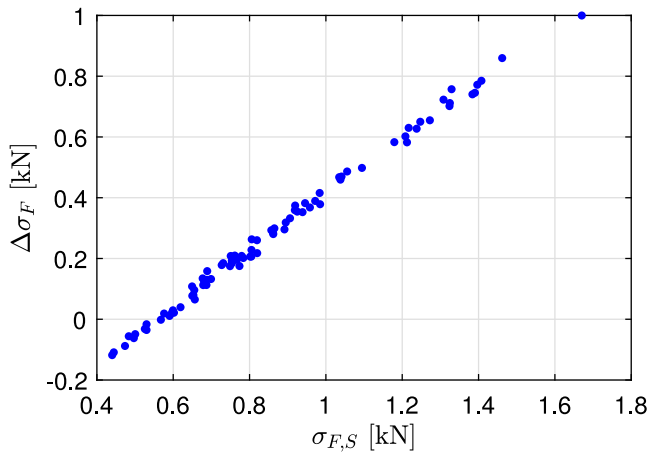
Fig. 12. Reduction $\Delta\sigma_F$ of the prediction error mean value for the 84 datasets in function of the prediction error mean value $\sigma_{F,S}$ determined in the single-level updating.

Table 4
Log-evidence values for different parametrizations of the model and different prior probability models.

		Parametrization	
		PAR. A	PAR. B
Prior model $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$	Normal	−919	−1058
	Uniform	−1125	−1161
	Lognormal	−926	−1071
	Beta	−755	−850

for $\bar{w} = 3.5$ mm. In the current section the authors investigate how the choice of alternative parametrizations and prior probability models of parameters conditioned on hyper-parameters $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ can affect the results of the hierarchical framework. By referring to the previous parametrization as PAR. A, two alternative parametrizations are analyzed: PAR. B that has the same first three parameters of PAR.A but the forth is replaced by the value $g\left(\bar{w}\right)$ at $\bar{w} = 1$ mm instead of $\bar{w} = 3.5$ mm. The parametrization PAR. C replaces the forth parameter with Δw representing the difference between w_2 and w_1 (that is $w_2 = w_1 + \Delta w$).

Performing the updating with PAR. C revealed a clear difficulty, namely the very low sensitivity of the likelihood function $p\left(D_i|\mathbf{x}_i\right)$ to the parameter Δw. We consider as an example the calibration based on the experimental response of the specimen T40, analyzing how the choice of the parameter bounds for Δw affects the calibration results. Fig. 14 compares the posterior distribution of Δw for the range [0.5; 15] mm (Fig. 14a) and [0.5; 150] mm (Fig. 14b). Except for small values of Δw, both distributions are nearly uniform, highlighting how a variation of Δw does not cause a substantial modification of the likelihood function. Furthermore, increasing the upper limit of the parameter range produces an increase in samples that have high Δw values, without being able to find a result independent of the imposed upper and lower bounds. The reason could be sought in the way the experimental tests were carried out. It has been already pointed out that the experimental tests have been stopped at CMOD = 4 mm and that w_2 is commonly larger than 4 mm, hence it is more reasonable to calibrate the stress reduction for CMOD values reached in the experimental tests. The low sensitivity towards Δw makes the calibration problem ill-conditioned, for this reason the parametrization PAR.C is discarded and no other results regarding this parametrization will be presented in the following.

The focus is then moved to analyzing the influence of different prior models $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ for PAR.A and PAR.B. Four different prior models, namely normal, uniform, lognormal and Beta, are considered. In the analyzed cases, the five parameter distributions are described by the same prior model. For the lognormal case, all the parameters have

positive ranges of variation except for $g\left(\bar{w}\right)$ that can assume negative values. For this reason, the lognormal distribution is chosen to describe the variation of $1 − g\left(\bar{w}\right)$.

The comparison between different prior models $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ is performed according to the value of the Bayesian evidence for the hyper-parameter likelihood, introduced as the constant at the denominator of Eq. (13). The evidence can be used in problems of model class selection since it represents a trade-off between the average ability of a model to fit the data and its complexity [30]. If the prior probability of competing models is the same, a larger evidence value implies an higher plausibility in the choice of the corresponding model according to the available experimental data. The log-evidence values for the parametrization PAR.A and PAR.B and for the different prior probability models $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ are listed in Table 4. The influence of $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ on the log-evidence value is distinctive and two general considerations can be formulated. If the parametrization PAR.A and PAR.B are separately analyzed, it is evident how uniform and beta prior models are associated to the lowest and the highest log-evidence values, respectively. Normal and lognormal prior models have intermediate values. On the other hand, comparing the two parametrizations for a fixed prior model $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$, we observe that PAR.A has always larger log-evidence values.

Parameter inherent variability obtained with different parametrizations and prior models are shown in Fig. 15. These plots have been realized thanks to the Matlab tool developed by [58]. The shaded area represents the [25; 75] % percentile range, the horizontal line indicates the median while the white dot indicates the mean. Uniform prior distributions are very different from the other represented distribution. Moreover, since uniform prior has the lowest log-evidence values, the considerations that are expressed in the following refer to the other prior models. The distributions of f_{ct} (Fig. 15a), w_1 (Fig. 15b) and $g\left(w_1\right)$ (Fig. 15c) have similar central trend indices and [25; 75] % percentile ranges. In general, a large discrepancy between distributions related to PAR.A or PAR.B is not observed, except for the case of Beta prior model and PAR.B where the parameter variability is different from the others. For the parameters $g\left(\bar{w}\right)$ (Fig. 15d) and σ_F (Fig. 15e), distributions related to PAR.A have noticeable differences compared to those of PAR.B. As concerns $g\left(\bar{w}\right)$, it is due to the fact that the pre-assigned crack opening $\bar{w}$ is different for PAR.A and PAR.B, hence it is expected the $g\left(\bar{w}\right)$ is on average larger at 1 mm than at 3.5 mm. The distribution of the prediction error σ_F depends on the sensitivity of the likelihood function to the calibration parameters, that varies for PAR.A and PAR.B in function of the value of $\bar{w}$.

Finally, the probabilistic features of the numerical predictions have been investigated for the two prior models $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$ that allow to have the highest log-evidence values for PAR.A, namely Beta and normal model. Fig. 16 shows the comparison between the probability density functions (pdf) of the numerical predictions and of the experimental responses. The pdfs of the numerical predictions are estimated via a normal kernel smoothing function [59]. The comparison is presented for two representative cases, namely the force distribution at CMOD equal to 0.05 mm (Fig. 16a) and 3.5 mm (Fig. 16a). In both cases, the pdfs of the numerical predictions obtained with a normal and beta prior model are very similar. No remarkable differences between the two prior models are noticed for the force distributions at different CMOD values. For the force distribution at 0.05 mm, there is a high accordance between the pdfs of the numerical and experimental quantities. As the force distribution at 3.5 mm is concerned, some differences are noticed for values of force larger than 5 kN. In the authors' opinion, these differences can be imputed to the model error rather than an inappropriate prior model $p\left(\mathbf{x}_i|\boldsymbol{\psi}\right)$. Indeed, the few experimental force samples in the range [5.0; 7.0] kN are not well fitted by the numerical pdfs because of the simulation inaccuracies of the numerical model. The empirical cumulative distributions of the same quantities are reported in Fig. 17. According to the probability densities, the cumulative distributions of the force at CMOD equal to 0.05 mm are very close to the experimental distribution (see Fig. 17a). The numerical and experimental force distribution at 3.5 mm, instead, are more distant in the range [5.0; 7.0] kN (see Fig. 17b).
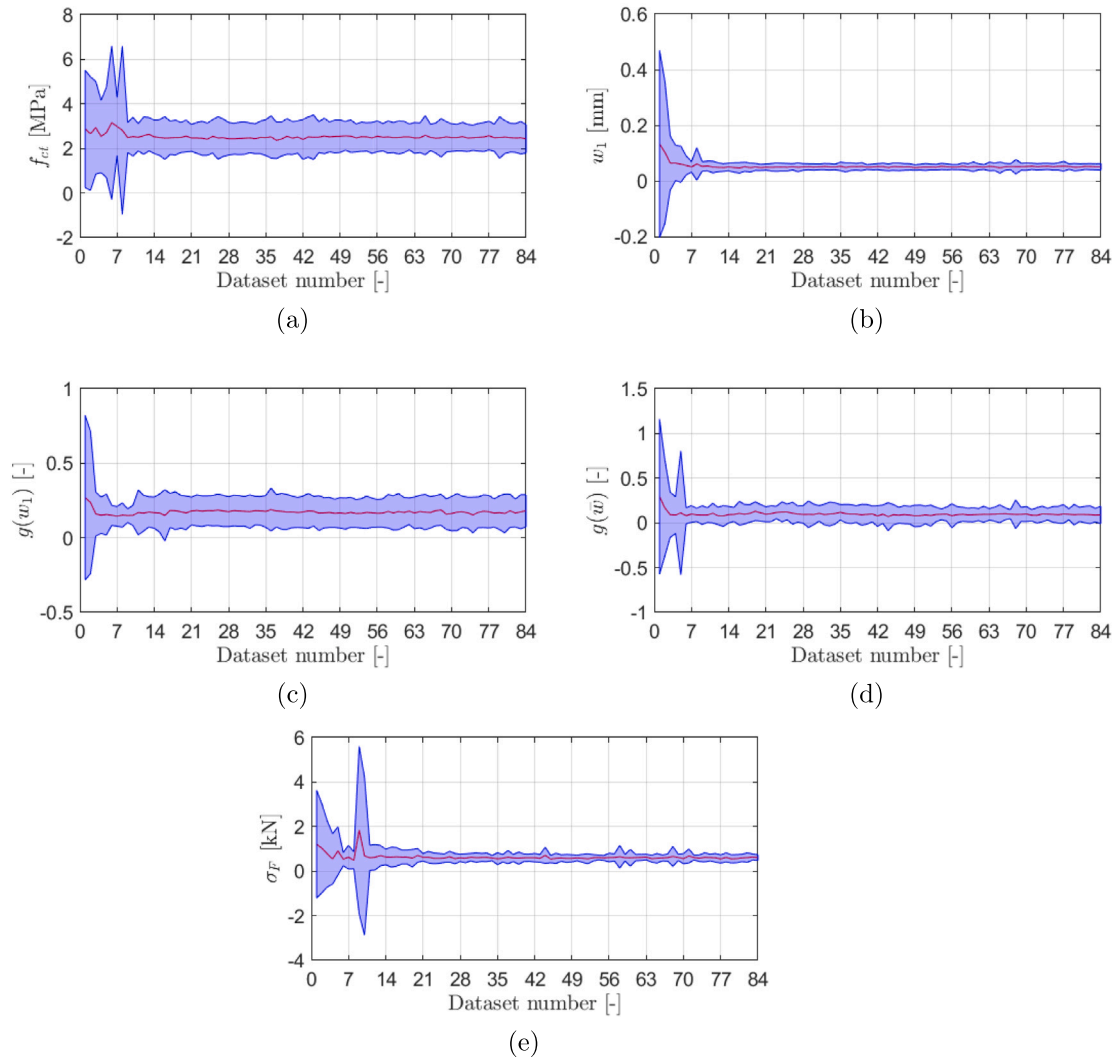
Fig. 13. Convergence of hyper-parameters pairs for (a) f_{ct}, (b) w_1, (c) $g(w_1)$, (d) $g(\bar{w})$ and (e) σ_F in function of the number of available datasets. Red line: mean value; blue region: 95% uncertainty bounds. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
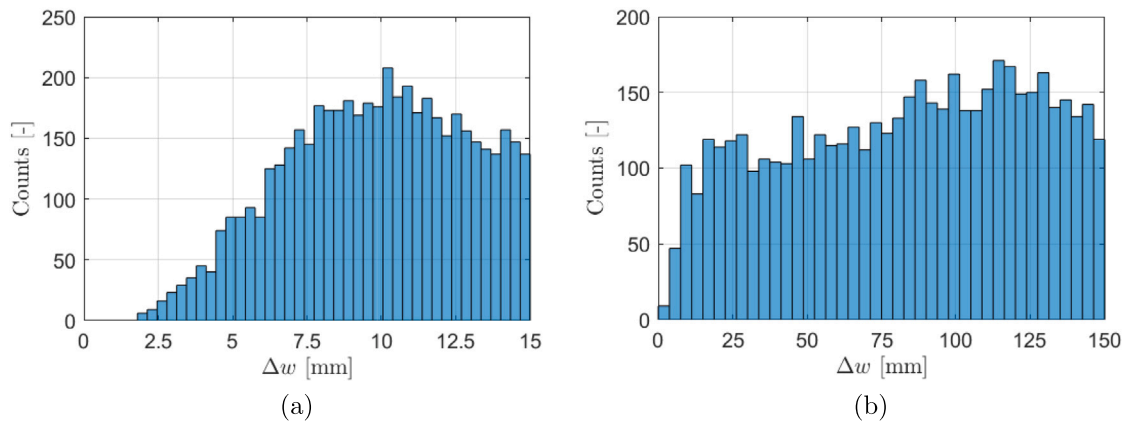

Fig. 14. Comparison between the posterior distributions of Δw for the parameter range (a) [0.2; 15] mm and (b) [0.2; 150] mm.

6. Conclusions

The present work proposed a Bayesian hierarchical (multi-level) framework to quantify the inherent variability of calibrated parameters of mechanical models, which can be relevant in structural engineering problems involving materials with highly variable properties such as

FRCs. The proposed procedure was applied to the calibration of the parameters of an analytical cracked-hinge model, considered as case study, based on the results of an experimental campaign performed on a large group of nominally identical FRC specimens.

The obtained results revealed the advantages of the hierarchical framework in contrast to single-level updating. In fact, the hierarchical
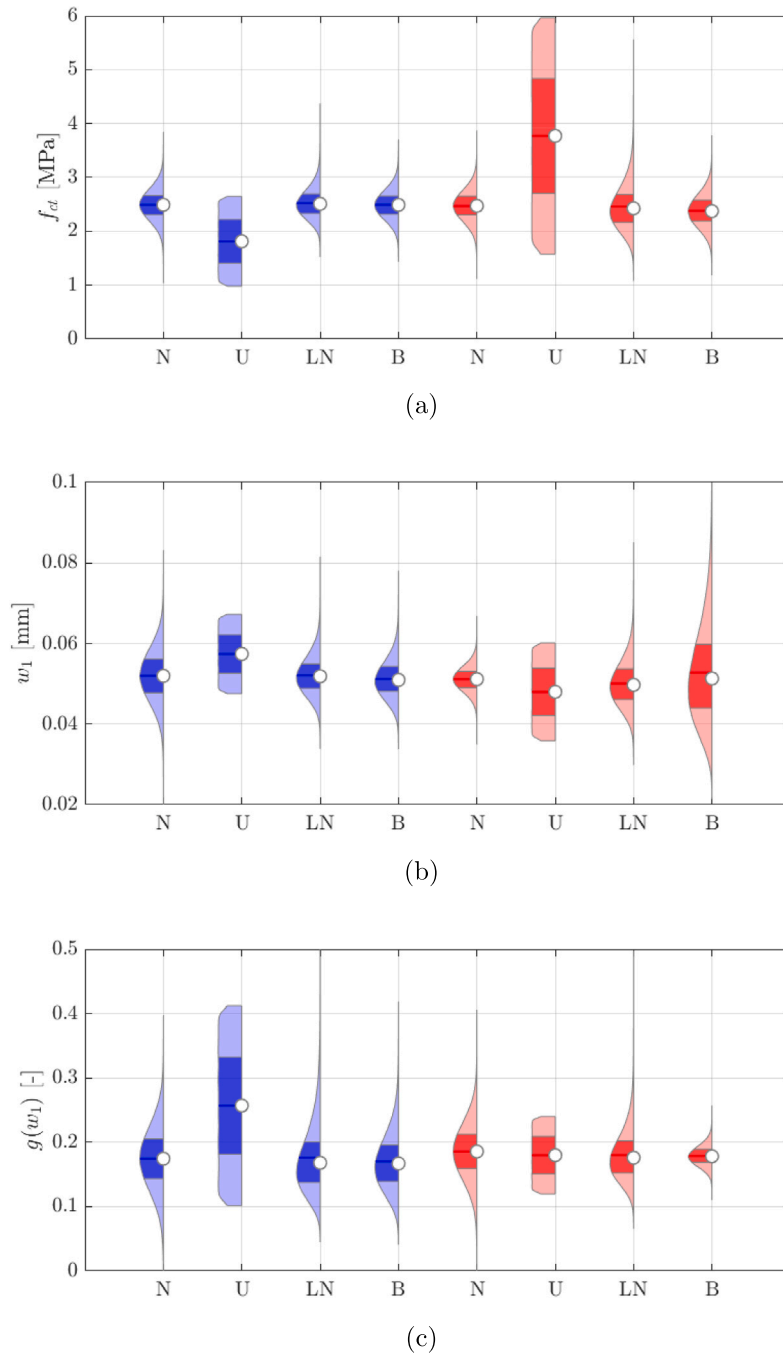
(a)

(b)

(c)

Fig. 15. Parameter inherent variability for different prior models and for the parametrization PAR.A (blue) and PAR.B (red). (a) f_{ct}, (b) w_1, (c) $g(w_1)$, (d) $g(\bar{w})$ and (e) σ_F. N: normal, U: uniform, LN: lognormal, B: Beta. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

approach allows a direct quantification of the inherent variability of parameters. Moreover, also a calibration based on a single experimental test can be refined by exploiting the information related to the other tests. In fact, the exchange of information among different datasets allows to reduce the uncertainty bounds in cases where the prediction error of the single-level framework is high.

Furthermore, specific analyses have been carried out to determine the minimum number of specimens needed for an accurate assessment of inherent parameter variability. With the dataset considered, results

indicate that 16 specimens are sufficient for this purpose. This may allow to reduce time and costs in future experimental campaigns aimed at characterizing the behavior of similar materials.

The main practical application of the proposed procedure is probabilistic assessment of structural safety and the calibration of partial safety factors used in design. In fact, a proper calibration of the inherent uncertainty of material parameters is essential for uncertainty propagation and therefore for structural reliability assessment.
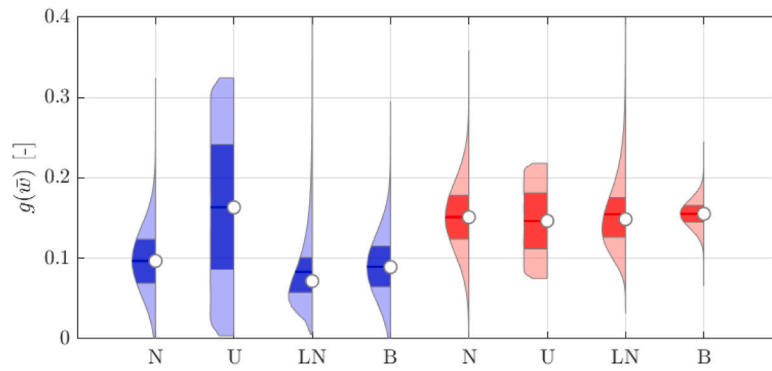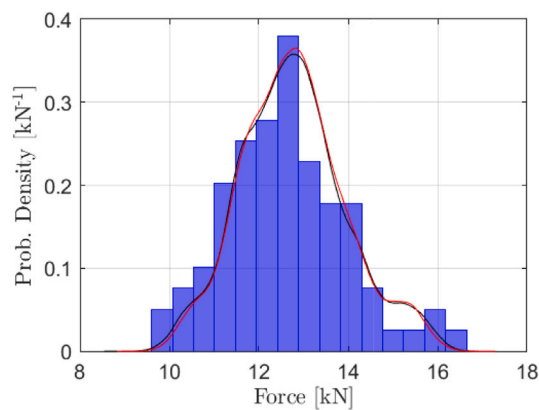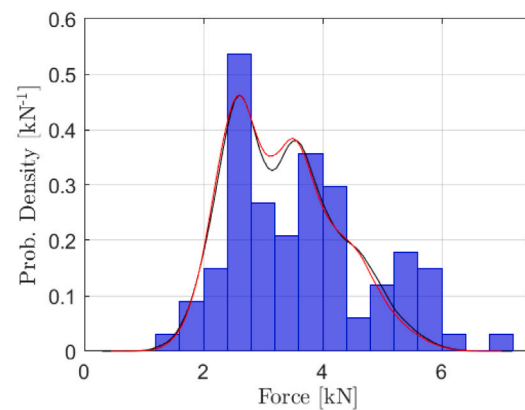
(d)
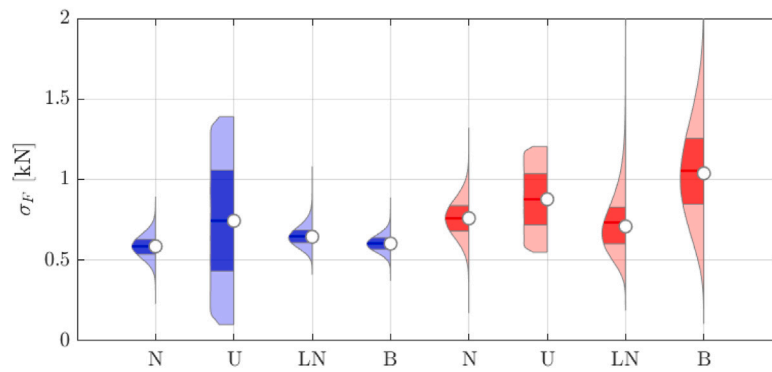
(e)

Fig. 15. (*continued*).

(a)

(b)

Fig. 16. Comparison between the probability densities of the numerical predictions and the experimental responses (blue bars) at CMOD values of (a) 0.05 mm and (B) 3.5 mm. Black line: normal prior model; red line: Beta prior model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Federico Ponsi: Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Elisa Bassoli:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Nicola Buratti:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Loris Vincenzi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, 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.

Data availability

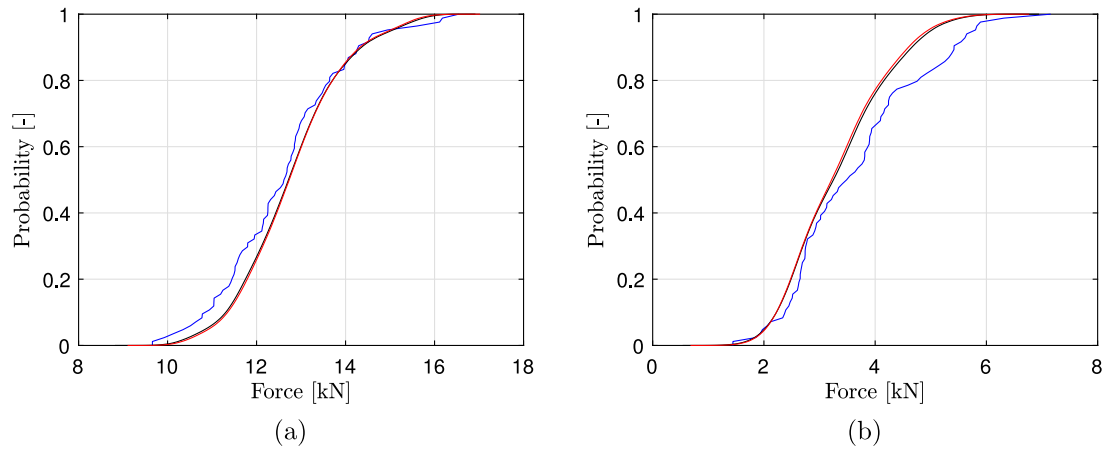Data will be made available on request.

Fig. 17. Comparison between the cumulative distributions of the numerical predictions and the experimental responses (blue line) at CMOD values of (a) 0.05 mm and (b) 3.5 mm. Black line: normal prior model; red line: Beta prior model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix A

See Table A.5.

Table A.5
Peak load and residual strength at CMOD values of 0.5, 1.5, 2.5 and 3.5 mm for all the 84 specimens together with the related statistics.

Specimen	F_{peak} [kN]	$F_{0.5}$ [kN]	$F_{1.5}$ [kN]	$F_{2.5}$ [kN]	$F_{3.5}$ [kN]
T1	14.95	3.88	3.64	3.94	3.94
T2	12.38	3.06	2.48	2.66	2.75
T3	12.28	3.12	2.63	2.73	2.68
T4	13.22	4.50	4.38	4.74	4.93
T5	12.97	2.37	1.78	1.91	1.93
T6	11.23	4.17	4.25	4.49	4.39
T7	13.10	3.91	4.08	4.23	4.24
T8	12.78	4.31	4.36	4.71	4.59
T9	14.08	3.93	3.84	4.06	4.02
T10	13.12	4.13	3.27	3.31	3.24
T11	13.14	4.70	4.66	5.12	5.32
T12	12.32	3.84	3.56	3.77	3.55
T13	13.76	5.42	5.78	5.95	5.89
T14	13.63	4.79	5.16	5.47	5.42
T15	13.38	3.37	3.62	4.06	4.15
T16	11.71	4.44	4.61	5.02	5.15
T17	11.84	4.53	4.25	4.69	4.75
T18	12.18	4.19	3.89	3.90	3.81
T19	13.39	3.98	3.64	3.89	3.9
T20	14.11	4.90	4.80	5.01	5.04
T21	14.59	4.10	3.97	4.10	3.9
T22	14.43	2.17	2.02	2.28	2.39
T23	15.63	3.70	4.11	4.40	4.31
T24	16.63	2.68	2.35	2.67	2.88
T25	16.21	3.17	2.67	2.82	2.95
T26	16.29	4.61	4.79	5.22	5.42
T27	14.06	5.18	6.10	6.78	7.16
T28	15.67	3.47	3.38	3.76	3.89
T29	13.45	4.99	4.53	4.98	5.25
T30	12.13	2.99	2.31	2.46	2.62
T31	14.92	2.66	2.48	2.89	3.09
T32	15.88	2.35	2.27	2.53	2.65
T33	15.26	3.63	3.48	3.90	4.17
T34	15.60	2.74	2.52	2.84	3.02
T35	13.69	3.00	2.85	3.02	3.21
T36	15.52	2.42	2.12	2.24	2.37
T37	13.54	3.85	4.43	4.79	4.82
T38	13.26	4.41	4.87	5.57	5.83
T39	14.36	4.59	4.59	4.41	4.26
T40	16.27	4.76	5.67	6.27	6.32
T41	15.00	4.89	4.50	5.07	5.37
T42	12.16	2.83	2.43	2.55	2.66
T43	14.02	3.94	3.38	3.46	3.49

Table A.5 (*continued*).

Specimen	F_{peak} [kN]	$F_{0.5}$ [kN]	$F_{1.5}$ [kN]	$F_{2.5}$ [kN]	$F_{3.5}$ [kN]
T44	13.44	4.03	3.85	3.80	3.93
T45	12.13	2.98	1.91	1.95	1.86
T46	12.99	4.29	3.85	4.07	4.24
T47	12.64	3.60	2.86	2.91	2.93
T48	13.70	3.19	2.32	2.42	2.47
T49	12.82	2.95	2.61	2.71	2.74
T50	11.10	2.70	2.45	2.66	2.77
T51	12.77	3.64	3.45	3.64	3.69
T52	12.75	3.04	3.12	3.29	3.34
T53	10.26	4.44	3.83	3.78	3.76
T54	13.34	2.14	1.34	1.38	1.44
T55	13.05	3.28	3.47	3.75	3.80
T56	11.91	3.27	2.47	2.52	2.51
T57	13.90	4.03	3.58	3.67	3.41
T58	13.41	4.67	4.16	4.08	4.1
T59	10.82	2.38	1.88	1.96	1.97
T60	11.99	3.93	3.82	4.00	4.09
T61	10.51	2.81	2.40	2.56	2.65
T62	14.78	2.73	2.44	2.68	2.78
T63	14.13	3.69	3.08	3.14	3.13
T64	12.18	2.73	1.95	1.97	2.12
T65	11.60	2.39	2.20	2.49	2.69
T66	13.38	3.20	2.95	3.13	3.32
T67	14.19	2.32	2.00	2.08	2.06
T68	13.77	3.06	2.39	2.60	2.63
T69	13.94	3.24	2.66	2.57	2.52
T70	16.78	2.92	2.81	2.96	3.02
T71	12.90	4.36	4.94	5.32	5.65
T72	12.64	3.00	2.82	2.75	2.74
T73	14.22	2.61	2.47	2.73	2.91
T74	13.41	2.57	2.21	2.33	2.34
T75	12.68	3.43	2.73	2.74	2.74
T76	12.74	5.08	5.36	5.59	5.55
T77	12.59	5.23	5.66	6.01	5.80
T78	13.72	3.65	3.77	3.76	3.64
T79	13.46	5.11	5.53	5.69	5.61
T80	12.55	2.78	2.65	2.58	2.6
T81	14.82	3.60	3.75	3.86	3.81
T82	12.95	2.70	2.35	2.42	2.45
T83	15.21	3.36	3.01	3.07	3.13
T84	13.91	3.54	3.40	3.79	3.85
Mean	13.50	3.61	3.42	3.62	3.67
Std. dev.	1.40	0.85	1.10	1.19	1.20

References

[1] A. Nour, B. Massicotte, R. de Montaignac, J.-P. Charron, Development of an inverse analysis procedure for the characterisation of softening diagrams for FRC beams and panels, Constr. Build. Mater. 94 (2015) 35–44, http://dx.doi.org/10.1016/j.conbuildmat.2015.06.049, URL https://www.sciencedirect.com/science/article/pii/S0950061815300064.

[2] S.J. Stephen, B. Raphael, R. Gettu, S. Jose, Determination of the tensile constitutive relations of fiber reinforced concrete using inverse analysis, Constr. Build. Mater. 195 (2019) 405–414, http://dx.doi.org/10.1016/j.conbuildmat.2018.11.014, URL https://www.sciencedirect.com/science/article/pii/S0950061818326795.

[3] M. Tanaka, G.S. Dulikravich, Inverse Problems in Engineering Mechanics, Elsevier, 1998.

[4] S.P. Shah, A. Carpinteri, Fracture mechanics test methods for concrete : report of Technical Committee 89-FMT Fracture Mechanics of Concrete, Test Methods, RILEM (the International Union of Testing and Research Laboratories for Materials and Structures), 1991.

[5] G. Bolzon, R. Fedele, G. Maier, Parameter identification by Kalman filter of a cohesive crack model, Comput. Methods Appl. Mech. Engrg. 191 (2002) 2847–2871.

[6] J.-K. Kim, Y. Lee, S.-T. Yi, Fracture characteristics of concrete at early ages, Cem. Concr. Res. 34 (3) (2004) 507–519, http://dx.doi.org/10.1016/j.cemconres.2003.09.011.

[7] S.H. Kwon, Z. Zhao, S.P. Shah, Effect of specimen size on fracture energy and softening curve of concrete: Part II. Inverse analysis and softening curve, Cem. Concr. Res. 38 (8) (2008) 1061–1069, http://dx.doi.org/10.1016/j.cemconres.2008.03.014.

[8] J. Olesen, Fictitious crack propagation in fiber-reinforced concrete beams, J. Eng. Mech. 127 (3) (2001) 272–280.

[9] L. Ostergaard, J. Olesen, Comparative study of fracture mechanical test methods for concrete, Fract. Mech. Concr. Struct. 1 (2004) 455–462.

[10] J. Barros, V. Cunha, A. Ribeiro, J. Antunes, Post-cracking behaviour of steel fibre reinforced concrete, Mater. Struct. (Mater. Constr.) 38 (275) (2005) 47–56, http://dx.doi.org/10.1617/14058.

[11] J.L.A.d. e Sousa, R. Gettu, Determining the tensile stress-crack opening curve of concrete by inverse analysis, J. Eng. Mech. 132 (2) (2006) 141–148, http://dx.doi.org/10.1061/(ASCE)0733-9399(2006)132:2(141).

[12] K. Chiranjeevi Reddy, K.V. Subramaniam, Analysis for multi-linear stress-crack opening cohesive relationship: Application to macro-synthetic fiber reinforced concrete, Eng. Fract. Mech. 169 (2017) 128–145, http://dx.doi.org/10.1016/j.engfracmech.2016.11.015.

[13] M. Alberti, A. Enfedaque, J. Gálvez, E. Reyes, Numerical modelling of the fracture of polyolefin fibre reinforced concrete by using a cohesive fracture approach, Composites B 111 (2017) 200–210, http://dx.doi.org/10.1016/j.compositesb.2016.11.052.

[14] A. Hillerborg, Analysis of fracture by means of the fictious crack model, particularly for fiber-reinforced concrete, Int. J. Cem. Compos. Lightweight Concr. 2 (4) (1980) 177–184.

[15] Z. Bazant, B. Oh, Crack band theory for fracture of concrete, Matér. Constr. 16 (1983) 155–177, http://dx.doi.org/10.1007/BF02486267.

[16] F. Tin-Loi, N. Que, Identification of cohesive crack fracture parameters by evolutionary search, Comput. Methods Appl. Mech. Engrg. 191 (49) (2002) 5741–5760, http://dx.doi.org/10.1016/S0045-7825(02)00483-8.

[17] N. Buratti, C. Mazzotti, M. Savoia, Post-cracking behaviour of steel and macro-synthetic fibre-reinforced concretes, Constr. Build. Mater. 25 (5) (2011) 2713–2722.

[18] L. Vincenzi, F. Ponsi, E. Bassoli, N. Buratti, A computationally efficient procedure for calibrating model parameters of multiple specimens, Constr. Build. Mater. 411 (2024) 134757, http://dx.doi.org/10.1016/j.conbuildmat.2023.134757, URL https://www.sciencedirect.com/science/article/pii/S0950061823044781.

[19] M.I. Friswell, J.E. Mottershead, Finite Model Updating in Structural Dynamics, Kluwer Academic Publishers, Dordrecht, 1995.

[20] M. Kennedy, A. O'Hagan, Bayesian calibration of computer models, J. R. Stat. Soc. Ser. B 63 (2001) 425–464, http://dx.doi.org/10.1111/1467-9868.00294.

[21] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, Mech. Syst. Signal Process. 56–57 (2015) 123–149.

[22] J. Legault, R. Langley, J. Woodhouse, Physical consequences of a nonparametric uncertainty model in structural dynamics, J. Sound Vib. 331 (25) (2012) 5469–5487, http://dx.doi.org/10.1016/j.jsv.2012.07.017, URL https://www.sciencedirect.com/science/article/pii/S0022460X12005433.

[23] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Accounting for modeling errors and inherent structural variability through a hierarchical Bayesian model updating approach: An overview, Sensors 20 (14) (2020) http://dx.doi.org/10.3390/s20143874, URL https://www.mdpi.com/1424-8220/20/14/3874.

[24] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, J. Sound Vib. 374 (2016) http://dx.doi.org/10.1016/j.jsv.2016.03.022.

[25] S.H.P. Cavalaro, A. Aguado, Intrinsic scatter of FRC: an alternative philosophy to estimate characteristic values, Mater. Struct. 48 (11) (2015) 3537–3555, http://dx.doi.org/10.1617/s11527-014-0420-6.

[26] S. Dobrilla, H.G. Matthies, A. Ibrahimbegovic, Considerations on the identifiability of fracture and bond properties of reinforced concrete, Internat. J. Numer. Methods Engrg. 124 (17) (2023) 3662–3686, http://dx.doi.org/10.1002/nme.7289, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.7289, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.7289.

[27] L. Simwanda, A.J. Babafemi, N.D. Koker, C. Viljoen, Bayesian calibration and reliability analysis of ultra high-performance fibre reinforced concrete beams exposed to fire, Struct. Saf. 103 (2023) 102352, http://dx.doi.org/10.1016/j.strusafe.2023.102352, URL https://www.sciencedirect.com/science/article/pii/S0167473023000395.

[28] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge University Press, 2003.

[29] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, J. Eng. Mech. 124 (1998) 455–461.

[30] J.L. Beck, Bayesian system identification based on probability logic, Struct. Control Health Monit. 17 (2010) http://dx.doi.org/10.1002/stc.424.

[31] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, 2005.

[32] K.-V. Yuen, Bayesian Methods for Structural Dynamics and Civil Engineering, John Wiley & Sons, 2010.

[33] E. Zhang, P. Feissel, J. Antoni, A comprehensive Bayesian approach for model updating and quantification of modeling errors, Probab. Eng. Mech. 26 (4) (2011) 550–560, http://dx.doi.org/10.1016/j.probengmech.2011.07.001, URL http://www.sciencedirect.com/science/article/pii/S0266892011000439.

[34] B. Goller, J. Beck, G. Schuëller, Evidence-based identification of weighting factors in Bayesian model updating using modal data, J. Eng. Mech. 138 (2012) 430–440, http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0000351.

[35] L. Ierimonti, I. Venanzi, N. Cavalagli, F. Comodini, F. Ubertini, An innovative continuous Bayesian model updating method for base-isolated RC buildings using vibration monitoring data, Mech. Syst. Signal Process. 139 (2020) 106600, http://dx.doi.org/10.1016/j.ymssp.2019.106600, URL http://www.sciencedirect.com/science/article/pii/S0888327019308210.

[36] F. Ponsi, E. Bassoli, L. Vincenzi, Bayesian and deterministic surrogate-assisted approaches for model updating of historical masonry towers, J. Civ. Struct. Health Monit. 12 (2022) 1469–1492, http://dx.doi.org/10.1007/s13349-022-00594-0.

[37] M.W. Vanik, J.L. Beck, S.K. Au, Bayesian probabilistic approach to structural health monitoring, J. Eng. Mech. 126 (7) (2000) 738–745.

[38] Y. Huang, C. Shao, B. Wu, J. Beck, H. Li, State-of-the-art review on Bayesian inference in structural system identification and damage assessment, Adv. Struct. Eng. 22 (2018) 136943321881154, http://dx.doi.org/10.1177/1369433218811540.

[39] K.-V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, Struct. Saf. 32 (5) (2010) 338–346, http://dx.doi.org/10.1016/j.strusafe.2010.03.011, Probabilistic Methods for Modeling, Simulation and Optimization of Engineering Structures under Uncertainty in honor of Jim Beck's 60th Birthday. URL https://www.sciencedirect.com/science/article/pii/S0167473010000305.

[40] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, Mech. Syst. Signal Process. 64–65 (2015) 360–376, http://dx.doi.org/10.1016/j.ymssp.2015.03.026, URL http://www.sciencedirect.com/science/article/pii/S0888327015001545.

[41] J.B. Nagel, B. Sudret, A unified framework for multilevel uncertainty quantification in Bayesian inverse problems, Probab. Eng. Mech. 43 (2016) 68–84, http://dx.doi.org/10.1016/j.probengmech.2015.09.007, URL https://www.sciencedirect.com/science/article/pii/S0266892015300266.

[42] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Hierarchical Bayesian modeling framework for model updating and robust predictions in structural dynamics using modal features, Mech. Syst. Signal Process. 170 (2022) 108784, http://dx.doi.org/10.1016/j.ymssp.2021.108784, URL https://www.sciencedirect.com/science/article/pii/S0888327021010967.

[43] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Nonlinear model updating through a hierarchical Bayesian modeling framework, Comput. Methods Appl. Mech. Engrg. 392 (2022) 114646, http://dx.doi.org/10.1016/j.cma.2022.114646, URL https://www.sciencedirect.com/science/article/pii/S0045782522000494.

[44] Y. Huang, J.L. Beck, H. Li, Hierarchical sparse Bayesian learning for structural damage detection: Theory, computation and application, Struct. Saf. 64 (2017) 37–53, http://dx.doi.org/10.1016/j.strusafe.2016.09.001, URL https://www.sciencedirect.com/science/article/pii/S0167473016300662.

[45] R. Hou, X. Wang, Q. Xia, Y. Xia, Sparse Bayesian learning for structural damage detection under varying temperature conditions, Mech. Syst. Signal Process. 145 (2020) 106965, http://dx.doi.org/10.1016/j.ymssp.2020.106965, URL https://www.sciencedirect.com/science/article/pii/S0888327020303514.

[46] G.C. Goodwin, M.E. Salgado, A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models, Internat. J. Adapt. Control Signal Process. 3 (4) (1989) 333–356, http://dx.doi.org/10.1002/acs.4480030405, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/acs.4480030405, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/acs.4480030405.

[47] C. Papadimitriou, Bayesian updating of weight values uncertainties in weighted residuals methods for model updating and response predictions, in: Proceedings of the XXIV International Conference on Sound and Vibration in Engineering and III International Conference on Uncertainty in Structural Dynamics, Leuven, Belgium, 2010, pp. 5207–5222.

[48] S. Brooks, A. Gelman, G. Jones, X.-L. Meng, Handbook of Markov Chain Monte Carlo, CRC Press, 2011.

[49] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian Annealed Sequential Importance Sampling: An Unbiased Version of Transitional Markov Chain Monte Carlo, ASCE-ASME J. Risk Uncertain. Engrg. Syst. Part B Mech. Engrg. 4 (1) (2017) 011008, http://dx.doi.org/10.1115/1.4037450, arXiv:https://asmedigitalcollection.asme.org/risk/article-pdf/4/1/011008/6072407/risk_004_01_011008.pdf.

[50] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, J. Eng. Mech. 133 (2007) http://dx.doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816).

[51] H. Sohn, K.H. Law, A Bayesian probabilistic approach for structure damage detection, Earthq. Eng. Struct. Dyn. 26 (12) (1997) 1259–1281.

[52] I. Behmanesh, B. Moaveni, Probabilistic identification of simulated damage on the dowling hall footbridge through Bayesian finite element model updating, Struct. Control Health Monit. 22 (2014) http://dx.doi.org/10.1002/stc.1684.

[53] P. Congdon, Applied Bayesian Hierarchical Methods, Chapman and Hall/CRC, 2010.

[54] J.B. Nagel, B. Sudret, Hamiltonian Monte Carlo and borrowing strength in hierarchical inverse problems, ASCE-ASME J. Risk Uncertain. Eng. Syst. A 2 (3) (2016) B4015008, http://dx.doi.org/10.1061/AJRUA6.0000847, URL https://ascelibrary.org/doi/abs/10.1061/AJRUA6.0000847, arXiv:https://ascelibrary.org/doi/pdf/10.1061/AJRUA6.0000847.

[55] S. Wu, P. Angelikopoulos, J.L. Beck, P. Koumoutsakos, Hierarchical Stochastic Model in Bayesian Inference for Engineering Applications: Theoretical Implications and Efficient Approximation, ASCE-ASME J. Risk Uncertain. Engrg. Syst. Part B Mech. Engrg. 5 (1) (2018) 011006, http://dx.doi.org/10.1115/1.4040571, arXiv:https://asmedigitalcollection.asme.org/risk/article-pdf/5/1/011006/6386801/risk_005_01_011006.pdf.

[56] T. Kloek, H.K. van Dijk, Bayesian estimates of equation system parameters: An application of integration by Monte Carlo, Econometrica 46 (1) (1978) 1–19.

[57] A. Economides, G. Arampatzis, D. Alexeev, S. Litvinov, L. Amoudruz, L. Kulakova, C. Papadimitriou, P. Koumoutsakos, Hierarchical Bayesian uncertainty quantification for a model of the red blood cell, Phys. Rev. Appl. 15 (2021) 034062, http://dx.doi.org/10.1103/PhysRevApplied.15.034062, URL https://link.aps.org/doi/10.1103/PhysRevApplied.15.034062.

[58] B. Bechtold, Violin plots for matlab, 2016, http://dx.doi.org/10.5281/zenodo.4559847, https://github.com/bastibe/Violinplot-Matlab.

[59] A. Bowman, A. Azzalini, Applied smoothing techniques for data analysis: The kernel approach with S-plus illustrations, in: Oxford Statistical Science Series, OUP Oxford, 1997, URL https://books.google.it/books?id=7WBMrZ9umRYC.