Data-Centric Engineering (2026), 7: e46
doi:10.1017/dce.2026.10077

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Bayesian tendon breakage localization under model uncertainty using distributed fiber optic sensors

Daniel Andrés Arcones[1,2], Aeneas Paul[3], Martin Weiser[4], David Sanio[3], Peter Mark[3] and Jörg F. Unger[2]

[1]School of Engineering and Design, Technical University of Munich, Germany
[2]7.7 Modeling and Simulation, Bundesanstalt für Materialforschung und –prüfung, Germany
[3]Institute of Concrete Structures, Ruhr University Bochum, Germany
[4]Department of Modeling and Simulation of Complex Processes, Zuse-Institut Berlin, Germany
Corresponding author: Daniel Andrés Arcones; Email: daniel.andres@tum.de

Received: 09 April 2026 UTC; **Revised:** 13 July 2026 UTC; **Accepted:** 17 August 2026 UTC

Keywords: Bayesian updating; distributed fiber-optic sensors; identifiability; model-form uncertainty; tendon breakage

Abstract

This study develops a Bayesian, uncertainty-aware framework for tendon breakage localization in prestressed concrete members using high-resolution data from distributed fiber-optic sensors (DFOS). DFOS enables full-field monitoring of strain changes on the surface of prestressed concrete members due to such failure. A finite element model (FEM) of an experimental tendon-breakage test is constructed, and model parameters are calibrated probabilistically against DFOS measurements. To capture model-form uncertainty (MFU), stochastic perturbations are embedded directly into material parameters, enabling the joint inference of physical properties and MFU within a unified probabilistic framework. Gaussian Process surrogates are employed to efficiently emulate the nonlinear FEM response, supporting computationally tractable Bayesian inference. A ϕ-divergence-based influence analysis identifies the DFOS measurements that most strongly shape the posterior distributions, providing interpretable diagnostics of sensor informativeness and model adequacy. The calibrated parameters and embedded uncertainties are then transferred to a FEM of a full-scale structural configuration, enabling prediction of tendon breakage localization under realistic conditions. A separability analysis of the predictive strain distributions quantifies the identifiability of tendon breakage at varying depths, assessing the confidence with which different damage scenarios can be distinguished given the propagated uncertainties. Results demonstrate that the framework achieves robust parameter calibration, interpretable diagnostics, and uncertainty-informed damage detection, integrating experimental data, embedded MFU, and probabilistic modeling. By systematically propagating both experimental and model uncertainties, the approach supports reliable tendon breakage localization, informed decision-making, and optimal DFOS placement.

Impact Statement

This work demonstrates a practical framework for reliably localizing tendon breakage in prestressed concrete structures using high-resolution fiber-optic sensors combined with computational modeling. By embedding uncertainty directly into material and interface models, the approach allows predictions to be transferred to new structural configurations without additional testing, improving confidence in simulation-based assessments.

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Influence- and separability-based analyses identify which sensor measurements are most informative, guiding monitoring strategies and highlighting where models are most trustworthy. The methodology enhances transparency, reliability, and decision-making in structural health monitoring, offering industry-relevant tools for understanding complex behavior, optimizing sensor deployment, and assessing risk under uncertainty.

1. Introduction

Understanding the location of tendon breakage in prestressed concrete members is essential for assessing structural integrity (Sieradzki and Newman, 1987), guiding emergency interventions, and designing effective monitoring strategies within Structural Health Monitoring (SHM) (Farrar and Worden, 2007; Bergmeister et al., 2015; Richter et al., 2025). In recent years, SHM approaches have therefore gained popularity for extending the life of critical structures such as bridges (Becks et al., 2024; Herbers et al., 2024; Kang et al., 2025). Recent studies have explored both data-driven and probabilistic approaches for the assessment of prestressed concrete structures. Machine learning techniques were applied to predict the loss of prestressing forces (Zhang et al., 2023; Calò et al., 2025), while probabilistic data-driven frameworks were employed to evaluate the failure probability of prestressed concrete bridges considering tendon corrosion (Nettis et al., 2024). However, many data-driven approaches rely on representative training data and often provide limited physical interpretability and uncertainty quantification when extrapolated to configurations that differ from those used during model development. For tendon breakage, however, destructive experiments on operated structures cannot be performed, making direct calibration or validation of mechanical models against in situ tendon failures infeasible (Pirskawetz and Schmidt, 2023). As a result, model calibration must rely on controlled laboratory experiments, where tendon breakage can be safely induced and instrumented. The central challenge is therefore not only to calibrate a model to experimental data, but to do so in a way that allows reliable propagation of uncertainties from laboratory-scale tests to simulations of real structural systems, where tendon breakage localization is ultimately required.

In particular, tendon breakage produces abrupt and highly localized redistribution of strains that propagates along the member. These strain changes can be captured using a spatial grid of distributed fiber-optic sensors (DFOS) (Paul et al., 2024b), which provides high-resolution, full-field strain change profiles (Barrias et al., 2016). Such data contain rich information about load transfer mechanisms and bond–slip behavior, making DFOS a powerful sensing technology for tendon-related damage assessment. However, leveraging dense DFOS measurements for tendon breakage localization requires uncertainty-aware methodologies capable of reconciling measurement noise, nonlinear physics, and modeling errors, particularly when results must be transferred beyond the experimental configuration.

A natural framework for simulating the tendon–concrete interaction (Ayoub and Filippou, 2010; Abdelatif et al., 2017) and the resulting strain redistribution after breakage (Seiffert and Jensen, 2019; Paul et al., 2025) is provided by Finite element models (FEM). Yet, even detailed nonlinear FEMs are affected by model-form uncertainty (MFU) due to idealized constitutive laws, uncertain boundary conditions, and geometric simplifications. When calibration is performed solely against laboratory data, these deficiencies can lead to biased or overconfident predictions if not explicitly accounted for. Discrepancy-based approaches such as the Kennedy–O'Hagan (KOH) formulation (Kennedy and O'Hagan, 2001) can improve calibration fidelity in the experimental configuration (Arcones et al., 2024), but the inferred discrepancy terms are tied to the calibration setup and do not naturally transfer to new structural configurations, limiting their usefulness for real-world tendon breakage localization.

To overcome this limitation, embedded MFU is adopted, whereby uncertainty is absorbed directly into an enlarged stochastic parameterization of the model (Sargsyan et al., 2015; Sargsyan et al., 2019). For tendon-breakage problems, sensitivity studies indicate that the Young's modulus of the concrete matrix plays a dominant role in controlling the strain response (Paul et al., 2026). Treating this modulus as a stochastic variable allows MFU to be represented within the constitutive behavior itself, enabling Bayesian inference approaches to jointly estimate physical parameters and model uncertainty. This embedded representation produces probabilistic parameter descriptions that can be meaningfully

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

transferred from laboratory experiments to simulations of real structural systems without relying on non-transferable discrepancy corrections.

Bayesian calibration involving both deterministic and stochastic parameters requires repeated model evaluations, which are computationally prohibitive for detailed nonlinear FEMs. Gaussian Process (GP) surrogates are therefore employed to emulate FEM responses with quantified predictive uncertainty (Rasmussen, 2006), enabling efficient assimilation of full-field DFOS data and uncertainty propagation. Beyond calibration, the framework supports diagnostic analyses. A ϕ-divergence-based influence analysis (Weiss, 1996) identifies which DFOS measurements most strongly affect posterior distributions, providing insight into sensor informativeness and model adequacy. This data-centric diagnostic capability is essential for improving physical models in complex structural systems.

As demonstrated in previous studies on small-scale laboratory tests (Paul et al., 2025), structural-scale tests (de Domenico et al., 2026; Sträter et al., 2026), or tests on existing structures (Burger et al., 2024; Fiedler and Marx, 2026), breakage and subsequent re-anchoring of tendons represent a very localized type of damage. This allows the transfer of the embedded stochastic parameters into new FEM models representing more realistic structural configurations, such as T-beam cross-sections or varying tendon embedment depths. This enables the prediction of strain fields in scenarios where no experimental data are available. In particular, comparing predictive distributions across different tendon positions allows the assessment of *separability*—a measure of how distinguishable different structural configurations, realized by varying embedment depths of the tendons, are under the propagated uncertainty. This is also commonly known as structural identifiability (Bellman and Åström, 1970). Regions where predictive distributions overlap indicate sensor locations that provide limited information under the current propagated uncertainty, whereas regions with clear separation identify locations where measurements would be most informative for detecting tendon-related anomalies. This can be used for guiding potential sensor placement (Ostachowicz et al., 2019).

In this work, we present a fully uncertainty-aware workflow that integrates DFOS measurements, FEM modeling, Bayesian inference with embedded MFU, influence diagnostics, and uncertainty transfer from laboratory experiments to full-scale structural simulations. The framework addresses the fundamental SHM challenge of localizing tendon breakage in systems where direct experimental calibration is impossible, providing a rigorous and transferable basis for uncertainty-informed structural assessment.

2. Methodology

The overall workflow used in this study is summarized in Figure 1. The methodology combines surrogate modeling, Bayesian inference, uncertainty quantification, and influence analysis in a unified framework.

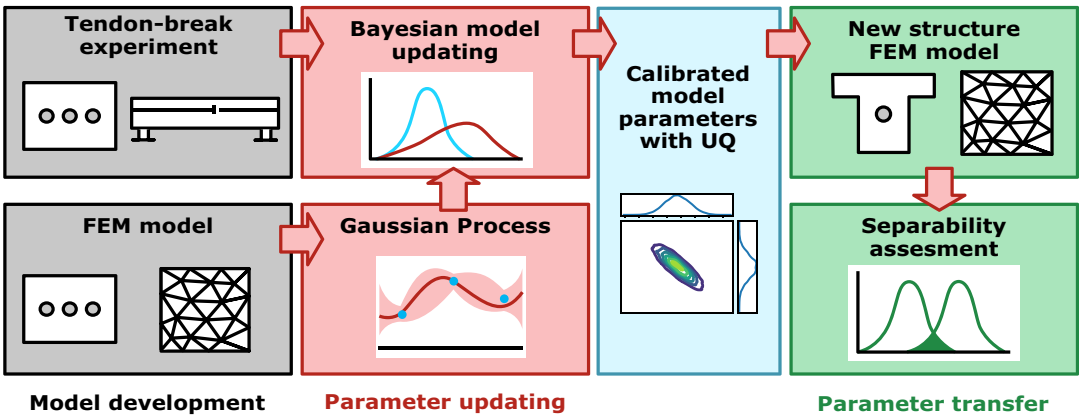


Figure 1. *Workflow for the parameter updating and transfer of the simulation of tendon breakage in pretensioned concrete.*

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

First, a computational model of the experimental setup is developed. It is approximated using a GP regressor in order to alleviate the cost of repeated high-fidelity simulations (Section 2.4). Using observational data together with the GP surrogate, the model parameters are inferred through a Bayesian calibration procedure that explicitly accounts for MFU (Section 2.1). Based on the resulting uncertainty estimates, an influence analysis is performed to assess the contribution of different data subsets and parameters to the inferred posterior distributions (Section 2.2). This analysis provides insight into dominant sources of discrepancy and supports targeted model improvements. Finally, the inferred parameters and their associated uncertainties are propagated to related prediction tasks in order to assess the robustness of the results under uncertainty (Section 2.3). The application of this methodology to the specific case study of tendon breakage in reinforced concrete is presented in a dedicated section, where the experimental setup, the high-fidelity model, and the corresponding results are described in detail.

2.1. Bayesian parameter updating under model-form uncertainty

2.1.1. Model calibration

We define the simulation-based forward model as

$$f : \mathcal{X}_f \times \Theta \subset \mathbb{R}^d \to \mathbb{R}^m, \quad (\mathbf{x}, \boldsymbol{\theta}) \mapsto f(\mathbf{x}, \boldsymbol{\theta}), \tag{1}$$

where $(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{X}_f \times \Theta \subset \mathbb{R}^d$ is a d-dimensional vector of input parameters composed of observation coordinates $\mathbf{x}$ and model parameters $\boldsymbol{\theta}$ and $f(\boldsymbol{\xi}) \in \mathbb{R}^m$ is the corresponding m-dimensional simulation output. For calibration, we have the observations dataset $\mathcal{Y} = (\mathbf{x}_i, y_i | i = 1, \ldots, N)$ or in matrix form $\mathcal{Y} = (X, \mathbf{y})$. Each observation relates to the simulation model f as

$$y = f(x, \boldsymbol{\theta}) + \delta_f(x) + \varepsilon_N, \tag{2}$$

where $\delta_f(\mathbf{x})$ is a discrepancy term between observations and predictions of f due to the MFU and ε_N is some prescribed noise, which will be treated as $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

Without MFU, the classical parameter updating problem consists of solving

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \|\mathbf{y} - f(X, \boldsymbol{\theta})\|, \tag{3}$$

where $\|\cdot\|$ is a norm to be defined, typically the Euclidean norm that leads to solving a minimization of the mean-squared error (MSE). Following the common principles of modularization (Bayarri et al., 2007), the training of the surrogate and the calibration of the parameters are performed sequentially and independently.

In the Bayesian framework, the main objective is to find the *posterior* probability distribution of $\boldsymbol{\theta}^*$ given the observations. We start by disregarding the discrepancy term $(\delta_f = 0)$ and imposing an error structure for the observation vector $\mathbf{y}$ such that it follows a multivariate normal distribution as $\mathbf{y} - f(X, \boldsymbol{\theta}) \sim \mathcal{MVN}(\mathbf{0}, \Sigma_\varepsilon)$, where Σ_ε is the covariance matrix equal to $\sigma_\varepsilon^2 I$ for independent normal errors. We obtain then the *likelihood* distribution as

$$\mathcal{L}(\boldsymbol{\theta}) \coloneqq \boldsymbol{\pi}(\mathbf{y} | \boldsymbol{\theta}) = (2\pi)^{-\frac{N}{2}} \det(\Sigma_\varepsilon)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y} - f(X, \boldsymbol{\theta})\|^2_{\Sigma_\varepsilon^{-1}} \right), \tag{4}$$

where $\|\mathbf{y} - f(X, \boldsymbol{\theta})\|^2_{\Sigma_\varepsilon^{-1}}$ denotes the Euclidean norm of the residuals weighted by the inverse of the covariance of the noise model. One option is directly substituting the model $f(X, \boldsymbol{\theta})$ by the mean of its approximation $\mathbf{m}_{\mathrm{GP}}(\boldsymbol{\theta}) = \mathbb{E}\left[\tilde{f}(X, \boldsymbol{\theta})\right]$, where $\tilde{f}(X, \boldsymbol{\theta})$ is the GP regression fitted for $f(X, \boldsymbol{\theta})$. The details on the GP implementation are provided in Section 2.4.

Applying Bayes' theorem, it is possible to obtain the posterior probability distribution of $\boldsymbol{\theta}$ from the likelihood $\mathcal{L}(\boldsymbol{\theta})$ and the prior $\pi(\boldsymbol{\theta})$. In particular,

$$\pi(\boldsymbol{\theta} | \mathbf{y}) \propto \pi(\mathbf{y} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \tag{5}$$

Sampling approaches such as those based on Markov Chain Monte Carlo (MCMC) draw realizations of the posterior distribution by evaluating samples of $\boldsymbol{\theta}$ at Equation 5.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

2.1.2. Model form uncertainty

Traditional Bayesian approaches do not treat MFU effectively, as the discrepancy term $\delta_f(x)$ is generally disregarded. Under additive Gaussian observation noise with diagonal covariance, Bayesian calibration reduces to a (weighted) least-squares problem, which, in the present case of a scalar diagonal covariance, coincides with ordinary least squares. As a result, the inferred posterior distribution $\pi(\boldsymbol{\theta}|\mathbf{y})$ concentrates around the least-squares estimate as the noise level decreases, and any unexplained variability in the observations is necessarily attributed to parameter uncertainty. Consequently, the posterior predictive distribution $f(X,\boldsymbol{\theta}|\mathbf{y})$ fails to reflect variability arising from structural model inadequacy, leading to overconfident predictions that do not capture the observed variability (Arcones et al., 2024).

One popular alternative is the framework from Kennedy and O'Hagan (2001), which implements $\delta_f(x)$ as a flexible parameterized function that compensates for the discrepancy, generally a GP. However, the $\delta_f(x)$ inferred within that framework cannot be propagated to other than systems or datasets than the one used for calibration. A promising alternative is the embedding of the uncertainty within the parameter formulation, as proposed in Sargsyan et al. (2019). The model with discrepancies is then reformulated as

$$y = f(x,\boldsymbol{\theta}) + \delta_f(x) + \varepsilon_N \approx f\left(x,\tilde{\boldsymbol{\theta}}\right) + \varepsilon_N, \tag{6}$$

where $\tilde{\boldsymbol{\theta}}$ is a stochastic extension of the parameter vector $\boldsymbol{\theta}$, endowed with a probability density $\pi_{\tilde{\theta}}\left(\tilde{\boldsymbol{\theta}}\right)$ defined on the parameter space Θ. This formulation implicitly accounts for the model discrepancy $\delta_f(x)$ through variability in the extended parameters. Introducing a probability distribution for $\tilde{\boldsymbol{\theta}}$ follows a hierarchical Bayesian construction and prevents excessive concentration of the posterior distribution of the affected parameters and their associated predictions. The resulting predictive uncertainty can be calibrated to provide a suitable representation of the MFU, provided that the extended model is sufficiently flexible to cover the range of the observations through variations in $\tilde{\boldsymbol{\theta}}$. The hyperparameters governing the density $\pi_{\tilde{\theta}}$ are inferred jointly with the original model parameters. In this paper, we follow the methodology developed in Andrés Arcones et al. (2026), defining explicitly such probability distributions as normal or log-normal, and adding their parameters to the inference. Therefore, $\boldsymbol{\theta}_{\text{ext}} = \{\boldsymbol{\theta}, \boldsymbol{\theta}_{\boldsymbol{\delta}}\}$, where $\boldsymbol{\theta}_{\boldsymbol{\delta}}$ represents the vector of parameters associated with the stochastic extension for the discrepancy $\tilde{\boldsymbol{\theta}}$.

2.1.3. Uncertainty propagation

As $\tilde{\boldsymbol{\theta}}$ is a random variable, $f\left(x,\tilde{\boldsymbol{\theta}}\right)$ is also stochastic, which means that the probability distribution $\pi_{\tilde{\theta}}$ of $\tilde{\boldsymbol{\theta}}$ must be either sampled or propagated through f for its full description. As is common in embedded approaches—in contrast to hierarchical Bayesian methods—we opt for propagating $\tilde{\boldsymbol{\theta}}$ by using a Polynomial Chaos Expansion (PCE) (Sudret, 2021). For a given coordinate vector $\mathbf{x}$, we seek a PCE of the form:

$$f\left(\mathbf{x},\tilde{\boldsymbol{\theta}}\right) \approx \sum_{\boldsymbol{\alpha}\in \mathrm{A}} c_{\boldsymbol{\alpha}}(\mathbf{x})\Psi_{\boldsymbol{\alpha}}\left(\tilde{\boldsymbol{\theta}}\right), \tag{7}$$

where $\left\{\Psi_{\boldsymbol{\alpha}}\left(\tilde{\boldsymbol{\theta}}\right)\right\}$ is a multivariate orthogonal polynomial basis with respect to the probability measure of $\tilde{\boldsymbol{\theta}}$, $\boldsymbol{\alpha}\in\mathbb{N}_0^n$ is a multi-index defining the total degree of each multivariate polynomial, $\mathrm{A}\subset\mathbb{N}_0^n$ is the finite set of indices used in the expansion (e.g., corresponding to polynomials up to total degree p), $c_{\boldsymbol{\alpha}}(\mathbf{x})\in\mathbb{R}$ are the PCE coefficients, which are functions of the input coordinate $\mathbf{x}$. The polynomials Ψ_α are chosen to be orthonormal following Askey's scheme based on the input distributions of $\tilde{\boldsymbol{\theta}}$ (Xiu and Karniadakis, 2002) such that $\langle\Psi_{\boldsymbol{\alpha}},\Psi_{\boldsymbol{\alpha}}\rangle = 1$ with $\langle\Psi_{\boldsymbol{\alpha}},\Psi_{\boldsymbol{\alpha}}\rangle = \int_{\mathbb{R}^n}\Psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\theta})\pi_{\tilde{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta}$.

To approximate $c_{\boldsymbol{\alpha}}(\mathbf{x})$, we use a pseudo-spectral projection approach based on Gaussian quadrature with Q quadrature points, such that the coefficients of the PCE are obtained as

$$c_{\boldsymbol{\alpha}}(\mathbf{x}) = \sum_{q=1}^{Q} w^{(k)} f\left(\mathbf{x},\boldsymbol{\theta}^{(q)}\right)\Psi_{\boldsymbol{\alpha}}\left(\boldsymbol{\theta}^{(q)}\right), \tag{8}$$

where $\left\{\boldsymbol{\theta}^{(q)}, w^{(q)}\right\}_{q=1}^{Q}$ are the quadrature nodes and weights associated with $\pi_{\tilde{\theta}}$. For a collection of input points $X = \left\{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(N)}\right\}$, the model is evaluated at each $\mathbf{x}^{(j)}$ for all quadrature nodes $\boldsymbol{\theta}^{(q)}$. A single model simulation yields

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$f\left(X,\boldsymbol{\theta}^{(q)}\right)=\begin{bmatrix} f\left(\mathbf{x}^{(1)},\boldsymbol{\theta}^{(q)}\right) \\ f\left(\mathbf{x}^{(2)},\boldsymbol{\theta}^{(q)}\right) \\ \vdots \\ f\left(\mathbf{x}^{(N)},\boldsymbol{\theta}^{(q)}\right) \end{bmatrix}\in\mathbb{R}^{N}, \tag{9}$$

which allows constructing all $c_{\boldsymbol{\alpha}}\left(\mathbf{x}^{(j)}\right)$ simultaneously for only Q evaluations of the model. As full model evaluations f are typically computationally expensive, evaluations of the mean $\mathbf{m}_{\mathrm{GP}}$ of the surrogate model $\tilde{f}$ are used instead. The PCE provides a full surrogate model for the stochastic response $\mathbf{m}_{\mathrm{GP}}\left(\tilde{\boldsymbol{\theta}}\right)=\mathbb{E}\left[\tilde{f}\left(X,\tilde{\boldsymbol{\theta}}\right)\right]$ across all input locations $X=\left\{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(N)}\right\}$.

The approximation is expressed as

$$\mathbf{m}_{\mathrm{GP}}\left(\tilde{\boldsymbol{\theta}}\right)\approx\sum_{\boldsymbol{\alpha}\in\mathrm{A}}\mathbf{c}_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}\left(\tilde{\boldsymbol{\theta}}\right), \tag{10}$$

where each $\mathbf{c}_{\boldsymbol{\alpha}}\in\mathbb{R}^{N}$ is a vector of PCE coefficients evaluated at all input points, such as

$$\mathbf{c}_{\boldsymbol{\alpha}}=\begin{bmatrix} c_{\boldsymbol{\alpha}}\left(\mathbf{x}^{(1)}\right) \\ c_{\boldsymbol{\alpha}}\left(\mathbf{x}^{(2)}\right) \\ \vdots \\ c_{\boldsymbol{\alpha}}\left(\mathbf{x}^{(N)}\right) \end{bmatrix}. \tag{11}$$

From this expansion, the mean at the observations $\mathbf{m}_{\mathrm{PCE}}(\boldsymbol{\theta})\in\mathbb{R}^{N}$ and the variances at observations $\boldsymbol{\sigma}^{2}_{\mathrm{PCE}}\in\mathbb{R}^{N}$ of the model response over X are obtained directly as components of the surrogate:

$$\mathbf{m}_{\mathrm{PCE}}(\boldsymbol{\theta}_{\mathrm{ext}})=\mathbb{E}\left[\mathbf{m}_{\mathrm{GP}}\left(\tilde{\boldsymbol{\theta}}\right)\right]\approx\mathbf{c}_{\mathbf{0}}, \tag{12}$$

$$\boldsymbol{\sigma}^{2}_{\mathrm{PCE}}(\boldsymbol{\theta}_{\mathrm{ext}})=\mathbb{V}\left[\mathbf{m}_{\mathrm{GP}}\left(\tilde{\boldsymbol{\theta}}\right)\right]\approx\sum_{\boldsymbol{\alpha}\neq\mathbf{0}}\mathbf{c}^{2}_{\boldsymbol{\alpha}}\langle\Psi_{\boldsymbol{\alpha}},\Psi_{\boldsymbol{\alpha}}\rangle, \tag{13}$$

where the square $\mathbf{c}^{2}_{\boldsymbol{\alpha}}$ is computed component-wise, and therefore the covariance matrix is $\Sigma_{\mathrm{PCE}}=\boldsymbol{\sigma}^{2}_{\mathrm{PCE}}(\boldsymbol{\theta}_{\mathrm{ext}})I$, assuming independent outputs. The last expressions follow immediately from the orthogonality of the polynomial basis. Further statistical moments can be extracted, but once the PCE is constructed, it is possible to approximate the full statistical behavior of $f\left(X,\tilde{\boldsymbol{\theta}}\right)$, without requiring any additional sampling or postprocessing.

Since the response of $\tilde{f}\left(X,\tilde{\boldsymbol{\theta}}\right)$ is stochastic and depends on the extrinsic parameters $\boldsymbol{\theta}_{\mathrm{ext}}$, one can identify those parameters for which the induced variance of $f\left(X,\tilde{\boldsymbol{\theta}}\right)$ provides a quantitative measure of MFU. To do so, we adapt the likelihood term defined in Equation 4 to include the variance of the prediction under the assumption that the response follows a normal distribution. The resulting likelihood is

$$\mathcal{L}_{\mathrm{IN}}(\boldsymbol{\theta}_{\mathrm{ext}})=(2\pi)^{-\frac{N}{2}}\det(\Sigma_{\mathrm{PCE}})^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\|\mathbf{y}-\mathbf{m}_{\mathrm{PCE}}(\boldsymbol{\theta}_{\mathrm{ext}})\|^{2}_{\Sigma^{-1}_{\mathrm{PCE}}}\right). \tag{14}$$

In a given iteration of an MCMC loop, samples of $\boldsymbol{\theta}_{\mathrm{ext}}$ are drawn; then $\tilde{\boldsymbol{\theta}}$ is defined based on them; the PCE for $\tilde{f}\left(X,\tilde{\boldsymbol{\theta}}\right)$ is constructed using evaluations of the forward model at a set of quadrature points; the statistical moments of the PCE $\mathbf{m}_{\mathrm{PCE}}$ and Σ_{PCE} are computed; and finally the likelihood $\mathcal{L}_{\mathrm{IN}}(\boldsymbol{\theta}_{\mathrm{ext}})$ is evaluated to obtain a sample of the posterior distribution $\pi(\boldsymbol{\theta}_{\mathrm{ext}}|\,\mathbf{y})$.

2.2. *Influence analysis*

2.2.1. *ϕ-divergence with full posterior*

Let $Y=(y_1,\ldots,y_n)$ denote the whole dataset of observations. Let S denote a nonempty proper subset of indices $\{1,\ldots,n\}$ and let its complement be $S^{c}=\{1,\ldots,n\}\setminus S$. We denote by $Y_S=\{Y_i:i\in S\}$ and

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$Y_{S^c} = \{Y_i : i \in S^c\}$ the corresponding subvectors of observations. We investigate the effect of removing S from the dataset on the posterior distribution of the model parameters. Assuming that $\pi(\boldsymbol{\theta} \mid Y)$ has global support, we define the posterior distributions

$$\pi(\boldsymbol{\theta} \mid Y_S) \propto \pi(\boldsymbol{\theta}) \pi(Y_S \mid \boldsymbol{\theta}) \quad \text{and} \quad \pi(\boldsymbol{\theta} \mid Y_{S^c}) \propto \pi(\boldsymbol{\theta}) \pi(Y_{S^c} \mid \boldsymbol{\theta}). \tag{15}$$

The *influence* of S can be quantified using a ϕ-divergence (Weiss, 1996). The ϕ-divergence between the two probability density functions $\pi(\boldsymbol{\theta} \mid Y)$ and $\pi(\boldsymbol{\theta} \mid Y_{S^c})$ is defined as

$$D_\phi(S) = \int \phi\left(\frac{\pi(\boldsymbol{\theta} \mid Y_{S^c})}{\pi(\boldsymbol{\theta} \mid Y)}\right) \pi(\boldsymbol{\theta} \mid Y) \, d\boldsymbol{\theta}, \tag{16}$$

where $\phi : \mathbb{R}_0^+ \to \mathbb{R}$ is a convex function satisfying $\phi(1) = 0$. A common choice for ϕ is the reverse Kullback–Leibler (KL) divergence, obtained by setting $\phi(\cdot) = -\log(\cdot)$. For simplification, we will denote by $D_\phi(S)$ the ϕ-divergence of the posterior distribution obtained with the full data Y and after excluding the subset S, using the reverse KL divergence.

This formulation has the advantage that it does not require recomputing the posterior distribution for each subset Y_{S^c}. Instead, it can be evaluated directly from posterior samples and their likelihood values (Weiss, 1996; Zhu et al., 2012). The data vector $Y = (Y_S, Y_{S^c})$ fulfills the chain rule of probability

$$\pi(Y \mid \boldsymbol{\theta}) = \pi(Y_S \mid Y_{S^c}, \boldsymbol{\theta}) \pi(Y_{S^c} \mid \boldsymbol{\theta}). \tag{17}$$

Therefore, we define the ratio of likelihoods between using all data Y and using only Y_{S^c} as

$$\pi_S(\boldsymbol{\theta}) = \frac{\pi(Y \mid \boldsymbol{\theta})}{\pi(Y_{S^c} \mid \boldsymbol{\theta})} = \pi(Y_S \mid Y_{S^c}, \boldsymbol{\theta}). \tag{18}$$

Then, applying Bayes' law and Equation 18,

$$\pi(\boldsymbol{\theta} \mid Y_{S^c}) = \frac{\pi(Y_{S^c} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int \pi(Y_{S^c} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{19}$$

$$= \frac{\pi(Y \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) [\pi_S(\boldsymbol{\theta})]^{-1}}{\int \pi(Y \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) [\pi_S(\boldsymbol{\theta})]^{-1} d\boldsymbol{\theta}} \tag{20}$$

$$= \frac{\pi(\boldsymbol{\theta} \mid Y) \pi(Y) [\pi_S(\boldsymbol{\theta})]^{-1}}{\int \pi(\boldsymbol{\theta} \mid Y) \pi(Y) [\pi_S(\boldsymbol{\theta})]^{-1} d\boldsymbol{\theta}} \tag{21}$$

$$= \frac{\pi(\boldsymbol{\theta} \mid Y) [\pi_S(\boldsymbol{\theta})]^{-1}}{\int \pi(\boldsymbol{\theta} \mid Y) [\pi_S(\boldsymbol{\theta})]^{-1} d\boldsymbol{\theta}} \tag{22}$$

$$= \frac{\pi(\boldsymbol{\theta} \mid Y) [\pi_S(\boldsymbol{\theta})]^{-1}}{\mathbb{E}_{\boldsymbol{\theta} \mid Y}\left[\pi_S(\boldsymbol{\theta})^{-1}\right]}, \tag{23}$$

where $\mathbb{E}_{\boldsymbol{\theta} \mid Y}$ denotes expectation with respect to the posterior distribution $\pi(\boldsymbol{\theta} \mid Y)$. Substituting this expression into the definition of $D_\phi(S)$ and simplifying yields the following expression for the influence of S

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$D_\phi(S) = \int \phi\left(\frac{[\pi_S(\boldsymbol{\theta})]^{-1}}{\mathbb{E}_{\boldsymbol{\theta}|Y}\left[\pi_S(\boldsymbol{\theta})^{-1}\right]}\right)\pi(\boldsymbol{\theta}\,|\,Y)d\boldsymbol{\theta} = \mathbb{E}_{\boldsymbol{\theta}|Y}\left[\phi\left(\frac{[\pi_S(\boldsymbol{\theta})]^{-1}}{\mathbb{E}_{\boldsymbol{\theta}|Y}\left[\pi_S(\boldsymbol{\theta})^{-1}\right]}\right)\right]. \tag{24}$$

Finally, substituting the influence function $\phi(u) = -\log(u)$ we get

$$D_\phi(S) = \log \mathbb{E}_{\boldsymbol{\theta}|Y}\left[\pi_S(\boldsymbol{\theta})^{-1}\right] + \mathbb{E}_{\boldsymbol{\theta}|Y}[\log \pi_S(\boldsymbol{\theta})]. \tag{25}$$

This influence measure can be computed for different sets S to evaluate the relative influence of some observations over others on the inferred posterior. See Appendix B.1 for a stable estimation of the global influence from posterior samples.

2.2.2. ϕ-divergence on the marginal posterior using kernel density estimation

We are interested in extending the formulation of D_ϕ of Equation 25 for a single parameter component θ_j. Since $\pi_S(\boldsymbol{\theta})$ is not a probability density, marginalization must be defined with respect to the joint posterior distribution. We therefore consider the conditional posterior expectation of $\pi_S(\boldsymbol{\theta})$ given θ_j. The posterior expectation conditional on the value of θ_j is

$$\pi_S^{(j)}(\theta_j) = \mathbb{E}_{\boldsymbol{\theta}|Y}\left[\pi_S(\boldsymbol{\theta})\,|\,\theta_j\right] = \int \pi_S(\theta_j, \theta_{-j}) p\left(\theta_{-j}|\theta_j, Y\right) d\theta_{-j}, \tag{26}$$

where θ_{-j} denotes all components except θ_j. An alternative would be to marginalize the perturbed $\pi(Y_{S^c}\,|\,\boldsymbol{\theta})$ and reference posteriors $\pi(Y\,|\,\boldsymbol{\theta})$ with respect to θ_{-j} and compute a divergence directly between the resulting marginal distributions of θ_j.

In practice, we approximate $\pi_S^{(j)}(\theta_j)$ from a posterior sample $\{\boldsymbol{\theta}^{(i)}\}_{i=1}^N$ using a one-dimensional weighted kernel density in the θ_j coordinate (Hastie et al., 2009). Let $K : \mathbb{R} \to \mathbb{R}^+$ be a kernel and $h > 0$ the bandwidth. We define the Gaussian kernel $K_h(u) = \frac{1}{h}K\left(\frac{u}{h}\right)$ with $K(u) = (2\pi)^{-1/2}e^{-u^2/2}$. Using sample values $\theta_j^{(i)}$ and weights $w_i \propto \pi_S(\boldsymbol{\theta}^{(i)})$ computed from full-sample likelihoods, we compute the Nadaraya–Watson estimator (Nadaraya, 1964; Watson, 1964) of the conditional expectation as

$$\widehat{\pi}_S^{(j)}(\theta_j) = \frac{\sum_{i=1}^N w_i K_h\left(\theta_j - \theta_j^{(i)}\right)}{\sum_{i=1}^N K_h\left(\theta_j - \theta_j^{(i)}\right)}, \tag{27}$$

where the bandwidth h is chosen following Scott's rule (Scott, 2008). We then insert this smoothed conditional ratio into the ϕ-divergence influence functional. The marginal influence for component j is

$$D_\phi^{(j)}(S) = \log \mathbb{E}_{\theta_j|Y}\left[\widehat{\pi}_S^{(j)}(\theta_j)^{-1}\right] + \mathbb{E}_{\theta_j|Y}\left[\log \widehat{\pi}_S^{(j)}(\theta_j)\right]. \tag{28}$$

See Appendix B.2 for a stable estimation of the marginal influence from posterior samples using the KDE approach.

2.2.3. ϕ-divergence on the marginal posterior using fixed-mean parameter estimators

It is also of interest to compute the marginal influence on one parameter when the others are fixed to their estimated values. In this case, this marginal influence is approximated by fixing them at the mean of their posterior samples. We first compute such mean of θ_{-j} as

$$\overline{\theta}_{-j} = \mathbb{E}_{\boldsymbol{\theta}|Y}\left[\theta_{-j}\right] \approx \frac{1}{N}\sum_{i=1}^N \theta_{-j}^{(i)}. \tag{29}$$

The fixed-mean perturbation as a function of θ_j is the pointwise evaluation

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$\pi_S^{(\mathrm{fix})}(\theta_j) = \pi_S(\theta_j, \overline{\theta}_{-j}) = \frac{\pi(Y|\theta_j, \overline{\theta}_{-j})}{\pi(Y_{S^c}|\theta_j, \overline{\theta}_{-j})}. \tag{30}$$

Using this plug-in perturbation, we define the fixed-mean influence

$$D_\phi^{(j),(\mathrm{fix})}(S) = \log \mathbb{E}_{\theta_j|Y}\left[\pi_S^{(\mathrm{fix})}(\theta_j)^{-1}\right] + \mathbb{E}_{\theta_j|Y}\left[\log \pi_S^{(\mathrm{fix})}(\theta_j)\right]. \tag{31}$$

See Appendix B.3 for a stable estimation of the marginal influence from posterior samples using the fixed-mean approach.

2.3. Uncertainty propagation, separability and identifiability

Unlike classical approaches, the embedded formulation enables the propagation of quantified MFU through an extended parameter set. Let g be a real-valued map, analogous to f, defined as

$$g : \mathcal{X}_g \times \Theta_g \times \Lambda_g \subset \mathbb{R}^q \to \mathbb{R}^m, \quad (\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \mapsto g(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda}), \tag{32}$$

where $\mathbf{x} \in \mathcal{X}_g$ denotes the input coordinates at which the model response is evaluated, $\boldsymbol{\theta} \in \Theta_g \subseteq \Theta$ is the vector of parameters previously calibrated using model f, and $\boldsymbol{\lambda} \in \Lambda_g$ is a vector of additional parameters that are not identified in the calibration stage. In contrast to $\boldsymbol{\theta}$, the parameters $\boldsymbol{\lambda}$ are quantities of direct interest that must be inferred from measurement data in the physical system, for example, damage-related state variables.

The model g represents a downstream prediction task in which the calibrated parameters $\boldsymbol{\theta}$ are reused under a different modeling context. When the parameters $\boldsymbol{\theta}$ are calibrated with embedded uncertainties, they are represented by random variables $\tilde{\boldsymbol{\theta}}$, rendering the model response $g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda})$ inherently stochastic. For each fixed value of $\boldsymbol{\lambda}$, the uncertainty in $\tilde{\boldsymbol{\theta}}$ induces a predictive distribution of the model output. The central objective is to assess how this propagated uncertainty affects the ability to distinguish between different candidate values of $\boldsymbol{\lambda}$ based on the resulting predictive distributions.

In this work, the model g is a FEM of a reinforced concrete beam used to investigate whether measurements from a SHM system can be linked to corresponding values of the model-state parameters $\boldsymbol{\lambda}$. Specifically, $\boldsymbol{\lambda}$ encodes the location of a tendon failure, parameterized through the depth of the tendon break. The predictive distributions of the structural response for different values of $\boldsymbol{\lambda}$ are compared against observed strain patterns to evaluate how accurately the damaged tendon can be identified in the presence of embedded MFU. As in the calibration stage, a PCE surrogate is employed to approximate the stochastic model response, following Equations 7–13. This surrogate enables efficient propagation of the uncertainty in $\tilde{\boldsymbol{\theta}}$ for each candidate value of $\boldsymbol{\lambda}$, yielding the predictive distributions required for subsequent identifiability and influence analyses.

Using the calibrated parameter set $\boldsymbol{\theta}$, we simulate system responses $g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda})$ for various hypothesized values of $\boldsymbol{\lambda}$. Running these simulations over a range of possible breakage locations generates a family of predictive distributions, each representing the expected system response for a particular damage scenario. These predictive distributions provide the basis for defining detection thresholds within the monitoring framework, which can be tailored to specific spatial locations $\mathbf{x}$ in the structure. A diagram of the proposed workflow is represented in Figure 2.

However, the MFU embedded in $\tilde{\boldsymbol{\theta}}$ leads to overlapping output distributions for different values of $\boldsymbol{\lambda}$. When the predictive distributions for $g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda}_1)$ and $g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda}_2)$, corresponding to $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}_1 + \Delta\boldsymbol{\lambda}$, are not statistically distinguishable, it becomes impossible to unambiguously infer the change in $\boldsymbol{\lambda}$ from the observed data. This is the essence of the identifiability problem. Therefore, a central objective is to quantify the minimal detectable perturbation $\Delta\boldsymbol{\lambda}_{\min}$ such that the induced change in system response exceeds the uncertainty envelope defined by the calibrated model. This enables a principled definition of detection thresholds that are both sensitive and robust, grounded in the probabilistic structure of the FEM response. The identifiability of a change in the parameter vector $\boldsymbol{\lambda}$ depends not only on the magnitude of the perturbation $\Delta\boldsymbol{\lambda}$, but also on the current operating point $\boldsymbol{\lambda}_0$ and the direction of change (i.e., whether

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

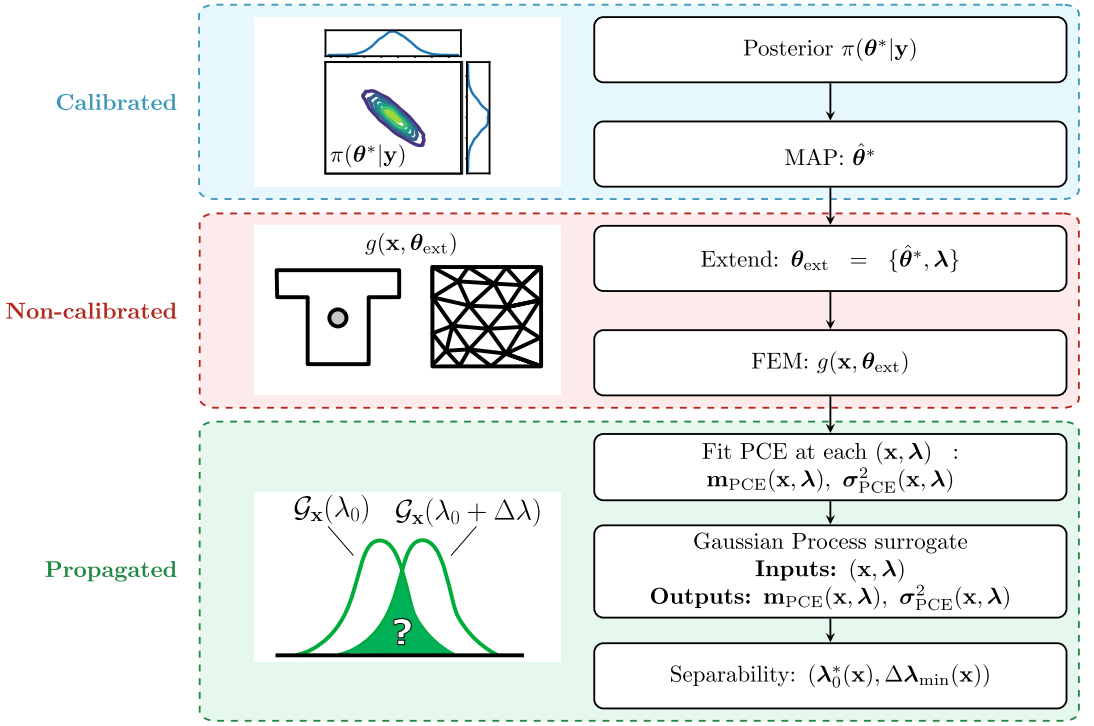


Figure 2. *Flow diagram of uncertainty propagation and identifiability analysis. The calibrated parameters $\boldsymbol{\theta}^*$ are extended with a set of non-calibrated parameters $\boldsymbol{\lambda}$, propagated through the FEM model g, approximated via PCEs, and used to train GP surrogates for separability analysis.*

$\boldsymbol{\lambda}_0 \pm \Delta\boldsymbol{\lambda}$). Due to the nonlinear and probabilistic nature of the forward model $g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda})$, the propagated response distributions may differ significantly depending on both $\boldsymbol{\lambda}_0$ and the sign of the perturbation.

The identifiability analysis is conducted locally at each point $\mathbf{x}$ in the spatial domain. For a given $\mathbf{x}$, the objective is to determine the smallest perturbation $\Delta\boldsymbol{\lambda}$ that leads to a statistically distinguishable change in the model response, as well as the corresponding parameter configuration $\boldsymbol{\lambda}_0^*$ for which this minimal change is most difficult to detect. Let the response distribution at $\mathbf{x}$ be defined as $\mathrm{G}_{\mathbf{x}}(\boldsymbol{\lambda}) = g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda})$. The chosen identifiability criterion is based on the principle of non-overlap of the confidence intervals at 95%. We define the spatially-localized maximin optimization problem as

$$(\boldsymbol{\lambda}_0^*(\mathbf{x}), \Delta\boldsymbol{\lambda}_{\min}(\mathbf{x})) = \arg\max_{\boldsymbol{\lambda}_0 \in \Lambda_g} \min_{\Delta\boldsymbol{\lambda} > 0} \|\Delta\boldsymbol{\lambda}\| \text{ subject to } \begin{cases} \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda}_0)] \cap \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda}_0 + \Delta\boldsymbol{\lambda})] = \varnothing, \\ \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda}_0)] \cap \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda}_0 - \Delta\boldsymbol{\lambda})] = \varnothing. \end{cases} \tag{33}$$

Here, $\Delta\boldsymbol{\lambda}_{\min}(\mathbf{x})$ is the smallest detectable parameter shift at point $\mathbf{x}$ and $\boldsymbol{\lambda}_0^*(\mathbf{x})$ is the least favorable configuration, that is, the point in parameter space where distinguishability is hardest under the confidence interval criterion. This formulation ensures that detection thresholds are defined locally and conservatively, based on the worst-case identifiability scenario at each spatial location. In practical terms, solving this problem enables spatially varying sensitivity maps to be constructed, guiding sensor placement and interpretation of system responses under monitoring conditions.

Due to the high computational cost of evaluating the FEM model $g(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\lambda})$, a surrogate modeling strategy is adopted to make the identifiability optimization tractable. The surrogate captures the dependence of the distribution of g on the parameters $\boldsymbol{\lambda}$, accounting for the uncertainty introduced by the parameters $\tilde{\boldsymbol{\theta}}$. For each pair $(\mathbf{x}, \boldsymbol{\lambda})$, the distribution $\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda}) = g(\mathbf{x}, \tilde{\boldsymbol{\theta}}, \boldsymbol{\lambda})$ is approximated using a PCE with $\tilde{\boldsymbol{\theta}}$ as

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

input, yielding the conditional mean $\mathbf{m}_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda})$ and variance $\boldsymbol{\sigma}^2_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda})$ of the model response as formulated in Equations 7–13. These statistics are then modeled using two independent GP regressors, one for the mean and one for the variance, whose implementation details will be reviewed in Section 2.4. The training dataset for the GP surrogates is constructed as follows: a set of spatial locations $X = [\mathbf{x}_i]_{i=1}^{N_x}$ is selected as a subset of the mesh nodes used in the FEM simulation. Then, a discrete grid $\{\boldsymbol{\lambda}_j\}_{j=1}^{N_\lambda} \subset \Lambda_g$ is chosen to cover the admissible domain Λ_g of the uncertain parameters. For each training pair $(\mathbf{x}_i,\boldsymbol{\lambda}_j)$, the corresponding PCE is constructed to compute the mean and variance of $\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda})$, which are then used to train the GPs. This dataset can be formulated as

$$\mathcal{D}_{\mathrm{train}} = \left\{ (X,\boldsymbol{\lambda}_i), \left(\mathbf{m}_{\mathrm{PCE}}(X,\boldsymbol{\lambda}_i), \boldsymbol{\sigma}^2_{\mathrm{PCE}}(X,\boldsymbol{\lambda}_i)\right) \middle| i = 1,\ldots,N_\lambda \right\}. \tag{34}$$

Only $N_\lambda \cdot N_Q$ evaluations of g are required, where N_Q is the number of quadrature points for the PCE. Once trained, the GPs provide efficient predictions $\mathbf{m}_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda})$ and $\boldsymbol{\sigma}^2_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda})$ at any $(\mathbf{x},\boldsymbol{\lambda})$ pair within the domain of interest. The 95% confidence intervals for the response distributions are then approximated as

$$\mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}}(\boldsymbol{\lambda})] \approx \left[\tilde{\mathbf{m}}_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda}) - 1.96\tilde{\boldsymbol{\sigma}}^2_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda}), \tilde{\mathbf{m}}_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda}) + 1.96\tilde{\boldsymbol{\sigma}}^2_{\mathrm{PCE}}(\mathbf{x},\boldsymbol{\lambda})\right], \tag{35}$$

where $\tilde{\mathbf{m}}_{\mathrm{PCE}}$ and $\tilde{\boldsymbol{\sigma}}^2_{\mathrm{PCE}}$ are the statistical moments estimated by their respective GPs. The variance introduced by the GP is disregarded for simplicity, as $\mathcal{D}_{\mathrm{train}}$ is sufficiently dense. The confidence interval approximation of Equation 35 can be directly plugged into the optimization problem in Equation 33. The domain Λ_g is discretized for the maximization part of the optimization problem, choosing the same fixed grid $\{\boldsymbol{\lambda}_j\}_{j=1}^{N_\lambda}$ as used in the GP training, thereby improving the evaluation efficiency. The minimization problem for $\Delta\boldsymbol{\lambda}$ is solved using the Nelder–Mead algorithm, while the maximization of $\boldsymbol{\lambda}_0$ is performed by comparison in the grid. This formulation enables efficient computation of the minimal detectable changes across the spatial domain, making it feasible to generate full-field maps of identifiability under uncertainty.

If separability is not achieved within a predefined maximum perturbation $\Delta\lambda_{\max}$, the optimization problem in Equation 33 is infeasible. Even in that case, we can extract some insight from partial solutions. To address this, we first check separability at $\lambda \pm \Delta\lambda_{\max}$ for each candidate λ. If separability is already satisfied at this level, then the optimization proceeds as usual. Otherwise, we fix $\Delta\lambda = \Delta\lambda_{\max}$ and quantify the degree of indistinguishability by computing the overlap between the two distributions. The overlap integral is computed to measure the amount of shared probability mass between the two distributions corresponding to λ and $\lambda \pm \Delta\lambda_{\max}$ as

$$\begin{aligned} O(\mathbf{x},\lambda;\Delta\lambda_{\max}) \equiv \frac{1}{2}\Bigg(& \int_{-\infty}^{\infty} \min\left\{\mathcal{G}_{\mathbf{x}}(\lambda),\, \mathcal{G}_{\mathbf{x}}(\lambda + \Delta\boldsymbol{\lambda}_{\mathbf{max}})\right\} \mathrm{d}y \\ & + \int_{-\infty}^{\infty} \min\left\{\mathcal{G}_{\mathbf{x}}(\lambda),\, \mathcal{G}_{\mathbf{x}}(\lambda - \Delta\boldsymbol{\lambda}_{\mathbf{max}})\right\} \mathrm{d}y \Bigg). \end{aligned} \tag{36}$$

where y represents the output of the (i.e., the support of $\mathcal{G}_{\mathbf{x}}$) over which the probability mass is defined. In practice, $\mathcal{G}_{\mathbf{x}}$ is modeled as a normal random variable defined by the PCE, and the integrals are calculated by numerical integration in the range of $\tilde{\mathbf{m}}_{\mathrm{PCE}} \pm 4\tilde{\boldsymbol{\sigma}}^2_{\mathrm{PCE}}$, truncating the distribution everywhere else.

From the collection of overlaps across the parameter grid, summary statistics can be extracted as

$$\begin{aligned} O_{\min}(\mathbf{x}) &= \min_{\lambda \in \Lambda_g} O(\mathbf{x},\lambda;\Delta\lambda_{\max}), \\ O_{\max}(\mathbf{x}) &= \max_{\lambda \in \Lambda_g} O(\mathbf{x},\lambda;\Delta\lambda_{\max}), \\ R_O(\mathbf{x}) &= O_{\max}(\mathbf{x}) - O_{\min}(\mathbf{x}). \end{aligned} \tag{37}$$

These values provide complementary information: $O_{\min}$ indicates the best-case separability, $O_{\max}$ the worst-case indistinguishability, and R_O quantifies the variability of overlap across Λ_g. The complete procedure is summarized in Algorithm 1.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Algorithm 1 Minimal detectable change with separability check at $\Delta\boldsymbol{\lambda}_{\max}$

Require: Spatial nodes $X=\{\mathbf{x}_i\}_{i=1}^{N_x}$, parameter grid $\Lambda_g=\{\lambda_j\}_{j=1}^{N_\lambda}$, GP surrogates for $\tilde{\mathbf{m}}_{\mathrm{PCE}},\tilde{\sigma}^2_{\mathrm{PCE}}$, maximum perturbation $\Delta\lambda_{\max}$, optimizer $\mathcal{O}$.
Ensure: For each $\mathbf{x}_i$: $\left(\boldsymbol{\lambda}_0^*(\mathbf{x}_i),\Delta\boldsymbol{\lambda}_{\min}(\mathbf{x}_i)\right)$, overlap map $O\left(\mathbf{x}_i,\lambda_j\right)$, and $(O_{\min},O_{\max},R_O)$.

for each spatial node $\mathbf{x}_i\in X$ **do**
  for each parameter $\lambda_j\in\Lambda_g$ **do**
    Check separability at $\Delta\lambda_{\max}$:

$$\text{Separable} \leftarrow \begin{cases} \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}_i}(\lambda_j)]\cap\mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}_i}(\lambda_j+\Delta\lambda_{\max})] \stackrel{?}{=} \varnothing, \\ \mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}_i}(\lambda_j)]\cap\mathrm{CI}_{0.95}[\mathcal{G}_{\mathbf{x}_i}(\lambda_j-\Delta\lambda_{\max})] \stackrel{?}{=} \varnothing. \end{cases},$$

  end for
  if Separable$\forall\lambda_j\in\Lambda_g$ **then**
    {Perform optimization}
    Solve Eq. (33) with optimizer $\mathcal{O}$ to obtain $\left(\boldsymbol{\lambda}_0^*(\mathbf{x}_i),\Delta\boldsymbol{\lambda}_{\min}(\mathbf{x}_i)\right)$.
  else
    {Fix to $\Delta\lambda_{\max}$ and compute overlap}
    Set $\Delta\lambda^*\left(\mathbf{x}_i,\lambda_j\right)\leftarrow\Delta\lambda_{\max}$.
    for each parameter $\lambda_j\in\Lambda_g$ **do**
      Compute overlap $O\left(\mathbf{x}_i,\lambda_j\right)$ using Eq. (36).
      Compute statistics $O_{\min},O_{\max},R_O$ with Eq. (37).
    end for
  end if
end for
return Maps of $\boldsymbol{\lambda}_0^*$, $\Delta\boldsymbol{\lambda}_{\min}$, O, $O_{\min},O_{\max},R_O$.

2.4. Gaussian Process surrogate model

We consider the previously defined simulation f defined in Equation 1 as the forward model to be calibrated, mapping model parameters to observable system responses. Throughout this section, we construct a GP surrogate $\tilde{f}$ to emulate the input–output behavior of the model. Let $\{\boldsymbol{\xi}_i\in\Omega\subset\mathbb{R}^d, i=1,\ldots,n\}$ denote a set of design points. Evaluations of the model at these points yield the training dataset

$$\mathcal{D}=\{(\xi_i,f(\boldsymbol{\xi}_i)):i=1,\ldots,n\},\quad \Xi=\left[\boldsymbol{\xi}_1^\top\ \cdots\ \boldsymbol{\xi}_n^\top\right]\in\mathbb{R}^{d\times n},\quad \mathbf{f}=[f(\boldsymbol{\xi}_1)\ \ \cdots\ \ f(\boldsymbol{\xi}_n)]\in\mathbb{R}^{m\times n}.$$

For each output component, a GP prior is placed as

$$\tilde{f}(\boldsymbol{\xi})\sim\mathcal{GP}(\mu(\boldsymbol{\xi}),k(\boldsymbol{\xi},\boldsymbol{\xi}')) \tag{38}$$

where $\mu:\Omega\to\mathbb{R}$ is the prior mean and $k:\Omega\times\Omega\to\mathbb{R}$ is a positive semi-definite covariance kernel (Rasmussen, 2006). Since the training outputs are standardized to zero mean for numerical stability, we take $m(\boldsymbol{\xi})\equiv 0$.

Let $\Xi^*=\left[\boldsymbol{\xi}_1^*,\ldots,\boldsymbol{\xi}_n^*\right]^\top$ be a set of test points with corresponding outputs $\mathbf{f}^*=\tilde{f}(\Xi^*)$. The GP prior implies the joint distribution

$$\begin{bmatrix}\mathbf{f}\\ \mathbf{f}^*\end{bmatrix}\sim\mathcal{N}\left(\mathbf{0},\begin{bmatrix}K(\Xi,\Xi) & K(\Xi,\Xi^*)\\ K(\Xi^*,\Xi) & K(\Xi^*,\Xi^*)\end{bmatrix}\right) \tag{39}$$

with $K(\Xi,\Xi')=\left[k\left(\boldsymbol{\xi}_i,\boldsymbol{\xi}_j'\right)\right]$. Denoting the vector of the kernel evaluation for a given test point $\boldsymbol{\xi}^*$ with respect to the training points $\boldsymbol{\xi}_i$ as $\mathbf{k}(\boldsymbol{\xi}^*)=[k(\boldsymbol{\xi}^*,\boldsymbol{\xi}_1),\ldots,k(\boldsymbol{\xi}^*,\boldsymbol{\xi}_n)]^\top$, the predictions for such point are

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$\mathbb{E}\left[\tilde{f}(\boldsymbol{\xi}^*)\right] = \mathbf{k}(\boldsymbol{\xi}^*)^\top K(\Xi,\Xi)^{-1}\mathbf{f}, \tag{40}$$

$$\mathbb{V}\left[\tilde{f}(\boldsymbol{\xi}^*)\right] = \mathbf{k}(\boldsymbol{\xi}^*,\boldsymbol{\xi}^*) - \mathbf{k}(\boldsymbol{\xi}^*)^\top K(\Xi,\Xi)^{-1}\mathbf{k}(\boldsymbol{\xi}^*). \tag{41}$$

The chosen kernel will be a radial basis function (RBF) with added white noise:

$$k(\xi_i,\xi_j) = \sigma_f^2 \exp\left(-\frac{\|\xi_i - \xi_j\|^2}{2\ell^2}\right) + \sigma_n^2\delta_{ij}, \tag{42}$$

where σ_f^2 is the signal variance, ℓ is the correlation length scale, σ_n^2 is the white noise variance, and δ_{ij} is the Kronecker delta, which is 1 if $i=j$ and 0 otherwise. Other kernels are possible but do not present any apparent advantages due to the expected smoothness of the target response surface.

Training a GP regressor as a surrogate model is reduced to solving the minimization problem

$$\boldsymbol{\theta}_{\mathrm{GP}}^* = \arg\min_{\boldsymbol{\theta}_{\mathrm{GP}}} -\log \pi(\mathbf{f}|\Xi,\boldsymbol{\theta}_{\mathrm{GP}}), \tag{43}$$

where $\boldsymbol{\theta}_{\mathrm{GP}} = \left\{\sigma_f^2, \ell, \sigma_n^2\right\}$ denotes the set of hyperparameters of the kernel, and $\pi(\mathbf{f}|\Xi,\boldsymbol{\theta}_{\mathrm{GP}})$ is the probability density function of the GP evaluated at the training outputs $\mathbf{f}$, which follow a multivariate normal distribution. This corresponds to maximizing the log marginal likelihood of the training data under the GP prior. The optimization is performed using gradient-based methods, as implemented in scikit-learn's `GaussianProcessRegressor` (Pedregosa et al., 2011), which follows the standard algorithmic framework described in Rasmussen (2006). All models are trained using automatic hyperparameter tuning via marginal likelihood maximization with multiple restarts to avoid poor local optima.

GP constitute an appropriate surrogate modeling framework for the present application due to their sample efficiency and strong approximation capabilities for smooth, deterministic response surfaces. Their kernel-based structure enables accurate emulation of complex simulation outputs from a limited number of high-fidelity evaluations (Forrester et al., 2008). We employ GP surrogates in two distinct ways. The first approach approximates the response surfaces with respect to a set of parameters $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$ at a fixed set of spatial coordinates $X = \{\mathbf{x}_i\}_{i=1}^N$. The second approach constructs a single GP surrogate defined jointly over both coordinates and parameters.

In the first case, the input to the GP is restricted to the parameter vector, that is, $\boldsymbol{\xi} = \{\boldsymbol{\theta}\}$. For each spatial coordinate $\mathbf{x}_i \in X$, a separate GP is trained and later evaluated as

$$\tilde{f}(\mathbf{x}_i,\boldsymbol{\theta}) \sim \mathcal{GP}(m(\boldsymbol{\theta}), k(\boldsymbol{\theta},\boldsymbol{\theta}')), \tag{44}$$

using the training dataset $\mathcal{D} = \left\{\left(\boldsymbol{\theta}_j, f\left(\mathbf{x}_i,\boldsymbol{\theta}_j\right)\right)\right\}_{j=1}^n$. Since a single simulation produces N observations (one per coordinate in X), only n simulations are required to train the N independent GPs, one for each fixed coordinate. The surrogate can therefore interpolate only in the parameter space Θ, while predictions must be evaluated at the fixed coordinate set X. This formulation is suitable when all quantities of interest are measured or required at predefined observation locations. This first formulation is used in the calibration of the parameters of the simulation model f based on experiments. In general, including a GP in the calibration process generates an additional uncertainty that should be considered in the likelihood. In some cases, it can be exploited for more efficient surrogate building (Perrin et al., 2025). However, in the current paper we validate the GP, showing that the variances in the validation dataset can be neglected, and use it as a mean response surface.

In the second case, the GP input includes both the coordinate and parameter vectors, $\boldsymbol{\xi} = \{\mathbf{x},\boldsymbol{\theta}\}$, with $\boldsymbol{\xi} \in \Omega = \mathcal{X} \times \Theta \subset \mathbb{R}^d$, where $\mathcal{X}$ and Θ denote the coordinate and parameter spaces, respectively. A single GP is trained on this joint domain, enabling predictions across both spatial coordinates and parameter values. This formulation is particularly suitable when predictions are needed on a dense or variable spatial grid, such as for computing full-field system responses. This second formulation is used for the propagation of the uncertainties in the FEM model g, which will be used for additional parameter identification.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

3. Application to a model for tendon breakage in reinforced concrete structures

The methodology outlined in the former section 2 establishes the foundation for an uncertainty-embedded calibration. This framework is applied to a tendon-breakage experiment in pretensioned concrete instrumented with DFOS (see Section 3.1.2) on the surface, allowing the measurement of strain changes with high spatial resolution. This section provides a comprehensive description of the experimental setup and the accompanying FEM. It also presents the resulting calibration, influence, and identifiability analysis. The objective of the identifiability assessment is to analyze how accurately can we distinguish the breakage of a tendon at different depths of a concrete specimen considering measurement noise and model uncertainty.

3.1. Experimental investigation

3.1.1. Experimental setup

The experimental investigation was executed on a pretensioned concrete beam of rectangular cross section ($L \cdot b \cdot h = 2000 \cdot 300 \cdot 200$ [mm^3]). In a stressing bed, a prestress of $\sigma_{p,0} = 755$ N/mm^2 was applied to three prestressing wires with smooth surfaces ($\varnothing_p = 9.4$ mm, St 1375/1570) and transferred to the beam after hardening of the concrete (28 d). A normal strength concrete was employed. Based on three tested cubes, a mean compressive strength of $f_{cm,cube} = 55.0$ N/mm^2 was measured. The installation of a plastic tube between the designated breakage point of the tendon and the concrete surface ensured the accessibility of the tendon after pouring of the concrete. The location was initially shifted in the x-direction to ensure full re-anchoring on at least one side of the breakpoint. However, DFOS measurements conducted on the tendon during the experiment showed that complete re-anchorage was attained on both sides of the breakage (Paul et al., 2024a). After application of the measurement equipment (details see Section 3.1.2), the beam was transferred to the test rig, which is illustrated in Figure 3a. The breakage of the centric tendon (at $x = 800$ mm, $y = z = 0$, cf. Figure 3b) was then induced using a drill. The complete separation was visually confirmed through the previously installed plastic tube and terminated the experimental investigation.

3.1.2. Instrumentation

DFOS (Speck et al., 2019; Clauß et al., 2021) were applied to the concrete surface (Janiak et al., 2023; Paul et al., 2024a) to instrument the test specimen. The measurement principle of DFOS can be roughly described as follows: A light beam is emitted into the sensor core ($\varnothing \approx 5$ µm; see Figure 4a), which possesses a variable refractive index along the sensor due to micro-inclusions. This variation enables identification of each measurement section through a distinctive "fingerprint." The backscatter is

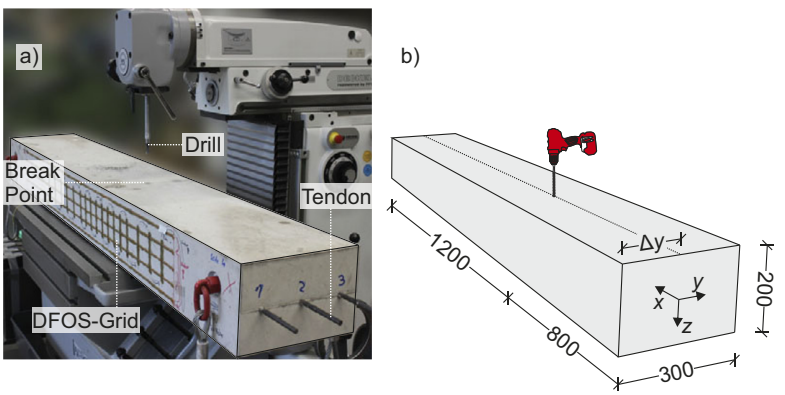


Figure 3. *Setup of the experimental investigation: specimen with drill and applied grid of distributed fiber optical sensors (a), sketch of the specimen with corresponding geometrical properties (b).*

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

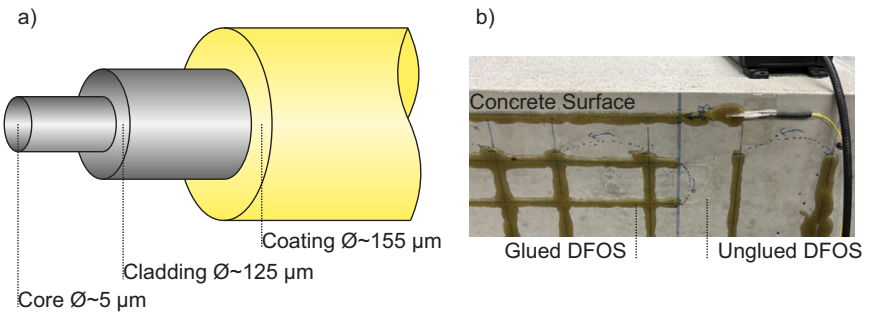


Figure 4. *Distributed fiber-optic sensor, as used in the instrumentation of the specimen: general structure of a sensor (a), sensor grid applied to the concrete surface (b).*

converted into the frequency domain by discrete Fourier transformation. Changes in strain $\Delta\varepsilon$ cause alterations in the wavelength (Samiec, 2011; Barrias et al., 2016).

A variety of sensor configurations are available. The main distinction between these sensors lies in their outer sensor layer (*coating*, see Figure 4a), which exerts a substantial influence on the transfer of strain from the sample to the sensor. Sensors with a polyimide, acrylate, or nylon coating are popular. Rigid coatings, such as polyimide (Young's modulus of $E_{\text{Polyimide}} = 400E_{\text{Acrylate}}$ (Chapeleau and Bassil, 2021)), are recommended for high-precision measurements (Herbers et al., 2023). The increased risk of breakage and sensor loss exhibited by stiff coatings in the presence of cracks can be disregarded here, as the induced damage to the prestressing steel usually does not result in concrete cracks (Sträter et al., 2024).

After the transfer of the prestressing force and removal of the formwork, the DFOS was applied to the side surfaces of the concrete structure using a two-component epoxy-based adhesive as shown in Figure 4b. The application was executed parallel and orthogonal to the tendon axis as a 2D-grid, following a meandering pattern in which a single fiber traversed varying height levels. Loops were left unglued. The outer points of each layer of the grid were distinctly designated to their respective positions on the fiber strand according to the *touch-to-locate method* (Konertz et al., 2019). The interpolation of strains between measurement points and layers was achieved through bilinear interpolation (Paul et al., 2025). The measurements were performed throughout the entire experiment with a measuring point spacing of 2.6 mm and a frequency of 1 Hz.

3.2. Numerical model

A numerical model of the experimentally investigated specimen was developed by means of the FEM. For this purpose, the commercial software *ABAQUS* was used. The concrete volume was modeled with cylindrical recesses corresponding to the positions of the tendons. Both the tendon and the concrete were discretized using three-dimensional solid elements with linear shape functions. A uniform element size was used in regions remote from the tendons, while progressive mesh refinement was applied toward the prestressing steel. This approach is consistent with established FEM practices for reinforced and prestressed concrete structures (van Meirvenne et al., 2018). The analysiswas performed using a static step-wise procedure, with loads applied incrementally under quasi-static conditions. Details are provided in the upcoming subsections.

3.2.1. Constitutive laws

The nonlinear Concrete Damage Plasticity (CDP) constitutive law, developed by Lubliner et al. (1989) and improved by Lee and Fenves (1998) was employed for the concrete. It combines isotropic, non-associative plasticity with damage parameters to consider the degradation of the material. The effective stress $\overline{\boldsymbol{\sigma}}$ was calculated by means of the elasticity matrix $\boldsymbol{D}_0^{el}$, the total strain tensor $\boldsymbol{\varepsilon}$, and the plastic strain tensor $\boldsymbol{\varepsilon}^{pl}$

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$\overline{\boldsymbol{\sigma}} = \boldsymbol{D}_0^{el} : \left(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{pl}\right). \tag{45}$$

For the computation of the Cauchy stress tensor $\boldsymbol{\sigma}$, the scalar degradation factor $(1-d)$ was additionally employed

$$\boldsymbol{\sigma} = (1-d)\overline{\boldsymbol{\sigma}}. \tag{46}$$

The computation of d depends on $\overline{\boldsymbol{\sigma}}$ and a set of two hardening variables $\tilde{\boldsymbol{\varepsilon}}^{pl}$. These hardening variables refer to equivalent plastic strains and are introduced to differentiate between behavior under compression (crushing failure; $\tilde{\varepsilon}_c^{pl}$) and tension (cracking failure; $\tilde{\varepsilon}_t^{pl}$). In addition to the degradation of the elastic stiffness, these also control the evolution of the yield function (Jankowiak and Lodygowski, 2005)

$$F\left(\overline{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{pl}\right) \leq 0 \text{with} : \tilde{\boldsymbol{\varepsilon}}^{pl} = \begin{bmatrix} \tilde{\varepsilon}_c^{pl} \\ \tilde{\varepsilon}_t^{pl} \end{bmatrix}. \tag{47}$$

$\overline{\boldsymbol{\sigma}}$ was decomposed into one invariant describing the volume change (hydrostatic stress $\overline{p}$)

$$\overline{p} = -\frac{1}{3}\text{trace}(\overline{\boldsymbol{\sigma}}) \tag{48}$$

and one component describing the shape change (deviatoric stress $\overline{\mathbf{S}}$)

$$\overline{\boldsymbol{S}} = \overline{\boldsymbol{\sigma}} + \overline{p}\mathbf{I} \tag{49}$$

with $\mathbf{I}$ as the identity tensor. $\overline{\boldsymbol{S}}$ was then transformed into the invariant von Mises equivalent stress $\overline{q}$

$$\overline{q} = \sqrt{1.5\left(\overline{\boldsymbol{S}} : \overline{\boldsymbol{S}}\right)}. \tag{50}$$

The yield function F was expressed in terms of $\overline{p}$, $\overline{q}$, and the Macauley Bracket $\langle x \rangle = 0.5(|x| + x)$ as

$$F\left(\overline{p}, \overline{q}, \tilde{\boldsymbol{\varepsilon}}^{pl}\right) = \frac{1}{1-\alpha}\left(\overline{q} + 3\alpha\overline{p} + \beta\left(\tilde{\boldsymbol{\varepsilon}}^{pl}\right)\langle\overline{\sigma}_{max}\rangle - \gamma\langle -\overline{\sigma}_{max}\rangle\right) - \overline{\sigma}_c\left(\tilde{\varepsilon}_c^{pl}\right) \geq 0. \tag{51}$$

Its shape was defined by α, β, and γ, which are calculated according to

$$\alpha = \frac{(\sigma_{b0}/\sigma_{c0}) - 1}{2(\sigma_{b0}/\sigma_{c0}) - 1}, \tag{52}$$

$$\beta = \frac{\overline{\sigma}_c\left(\tilde{\varepsilon}_c^{pl}\right)}{\overline{\sigma}_t\left(\tilde{\varepsilon}_t^{pl}\right)}(1-\alpha) - (1+\alpha), \tag{53}$$

$$\gamma = \frac{3(1-K_c)}{2K_c - 1}. \tag{54}$$

The ratio of initial equibiaxial to uniaxial compressive yield stress $\frac{\sigma_{b0}}{\sigma_{c0}}$ was used to compute α (Eq. (52)). The function β considers the effective stress $\overline{\boldsymbol{\sigma}}$ computed by means of the hardening variables $\tilde{\varepsilon}_c^{pl}$ and $\tilde{\varepsilon}_t^{pl}$ (eq. (53)), while γ is derived from K_c, which represents the ratio of the second stress invariant of the tensile meridian to that of the compressive meridian (eq. (54)). $\frac{\sigma_{b0}}{\sigma_{c0}} = 1.16$ and $K_c = \frac{2}{3}$ are commonly used (van Meirvenne et al., 2018). $\overline{\sigma}_{max}$ denotes the maximum eigenvalue of $\overline{\boldsymbol{\sigma}}$ (Jankowiak and Lodygowski, 2005).

For non-associative plasticity, the plastic flow $\dot{\boldsymbol{\varepsilon}}^{pl}$ resulted from the derivative of the flow potential function

$$\dot{\boldsymbol{\varepsilon}}^{pl} = \dot{\lambda}\frac{\partial G(\overline{\boldsymbol{\sigma}})}{\partial\overline{\boldsymbol{\sigma}}} \text{with} : G(\overline{p}, \overline{q}) = \sqrt{\epsilon f_{ctm}\tan(\psi)^2 + \overline{q}} - \overline{p}\tan(\psi). \tag{55}$$

The Drucker–Prager hyperbolic function was used for this purpose. It also takes $\overline{p}$ and $\overline{q}$ as input parameters and is defined by the eccentricity ϵ, the mean tensile strength of the concrete f_{ctm} and the

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

dilation angle in the $\overline{p}-\overline{q}$-plane ψ. Commonly used values are $\epsilon = 0.1$ mm and $\psi = 30°$ (van Meirvenne et al., 2018).

The CDP distinguishes between compressive and tensile loading. For both loading paths, the formulation developed by Krätzig and Pölling (2004) was used, employing the parameters suggested by Birtel and Mark (2006). Details regarding the formulations can be derived from their work. Under uniaxial compressive stress, ductile behavior with pronounced nonlinear hardening and a softer drop in resistance was captured. As shown in Figure 5a, the first branch of the stress–strain relationship was linear-elastic $\sigma_{c(1)}(\varepsilon_c)$ (for $\sigma_c < 0.4 f_{cm}$). The second branch $\sigma_{c(2)}(\varepsilon_c)$ of the stress–strain relation represented the nonlinear hardening phase (for $0.4 f_{cm} \leq \sigma_c \leq f_{cm}$). It reflected progressive microcracking and damage accumulation in the cement matrix, accompanied by stress redistribution and aggregate interlock, which together enable a continued increase in load-carrying capacity despite a gradual reduction in stiffness (van Mier et al., 1987). The third branch described the behavior after the stress exceeded the mean compressive strength f_{cm}. It was governed by a progressive loss of load-bearing capacity due to crack formation.

Under tension (Figure 5b), the model depicted brittle behavior, which occured after the mean tensile strength f_{ctm} was exceeded. Subsequently, for $\sigma_t > f_{ctm}$, σ_t was calculated from the crack width w, which according to Hillerborg (1983) was computed as $w = l_e \cdot \varepsilon_t^{in}$ (with the element length l_e).

The constitutive law of the steel was linear-elastic, since the prestress applied to the tendon σ_{p0} in the experiments did not exceed the yield strength of St 1375/1570. The constitutive law was therefore described using only the Young's Modulus E_p and Poisson Ratio ν_p.

3.2.2. Contact formulation

The pretensioning of a tendon leads to a reduction in its diameter $\varnothing_p$. The release of the prestressing force during a tendon breakage hence results in an increase of $\varnothing_p$ to its initial value (see Figure 5c). To appropriately model this so-called Hoyer effect (Briere et al., 2013), the tendon was modeled as a volume element. Contact between steel and concrete was prescribed parallel and orthogonal to the concrete–

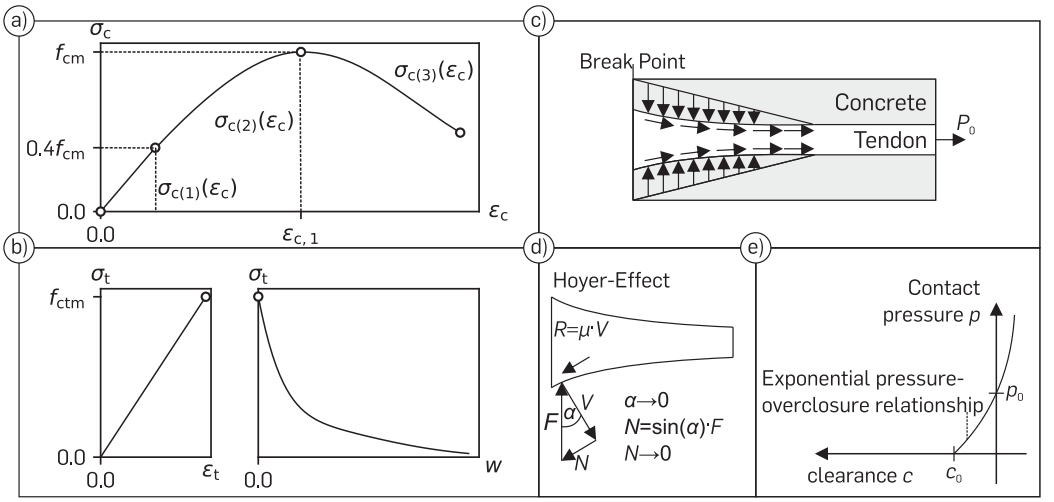


Figure 5. *Specifications of the computational model: behavior of concrete under uniaxial compression (a), behavior of concrete under uniaxial tension (b), close-up of the interface at the breakage point (c), local forces at the steel-concrete interface (d), contact specification normal to the interfaces by an exponential relationship between clearance c and contact pressure p, according to ABAQUS [Dassault Systèmes Simulia Corporation (2025)] (e).*

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

tendon interface (Figure 5d). Normal to the interface, an exponential relationship provided in *ABAQUS* was used. The contact pressure p was computed from the clearance between the interfaces c using an exponential function. It was defined by the clearance at which contact pressure is initiated (c_0) and the contact pressure p_0 when $c = 0$, as visualized in Figure 5e. The employed contact formulation allowed a realistic load transfer, while concurrently permitting slip and decoupling as a consequence of local damage.

Parallel to the interface, contact was modeled by employing Coulomb's friction law through the friction coefficient μ.

3.2.3. Simulation steps and boundary conditions

The simulation of the tendon breakage consists of three subsequent steps (cf. Table 1). The boundary conditions are set accordingly. Symmetry with respect to yz-plane at the breakage point ($x = 0$) was used, and only one part of the beam (free end at $l = 1200$ mm, cf. Figure 3) was modeled. In the symmetry plane ($x = 0$), a corresponding boundary condition ($u_x = \theta_y = \theta_z = 0$) was applied to the beam (concrete and tendons). It should be noted that solid elements do not possess rotational degrees of freedom, which reduces the symmetry conditions to $u_x = 0$. Vertical displacements were prohibited ($u_z = 0$) through all simulation steps (1–3) at the bottom of the beam ($z = 100$ mm), while displacements in y (u_y) were not constrained.

To implement the prestress to the tendons, a predefined stress field ($\sigma_{p,0} = 755$ N/mm^2) in the x-direction was initially applied to them. During the first simulation step (*Initial*), axial displacements of the tendons were prohibited at their ends ($x = 0$ and $x = l$). Then, the initial boundary conditions were altered in the two subsequent simulation steps: In step 2 (*Loading*), the transfer of the prestress $\sigma_{p,0}$ from the tendons to the concretewas simulated by releasing the fixation of u_x at the free end ($x = l$) for all tendons. In simulation step 3 (*Break*), the breakage was initiated by releasing the boundary condition of the central tendon at $x = 0$. The computation of the strain difference between the simulation steps, as shown in Figure 6 (right), corresponded to the measurements taken in the experiment. Table 1 summarizes the sequence of loading steps and the related boundary conditions.

3.3. Experimental results

The axial strains (x-direction) measured on the concrete surface with DFOS after the breakage of the tendon $\Delta\varepsilon_{c,x}$ are visualized in Figure 6 (left). The regions between the sensor strands were interpolated using bilinear interpolation (Paul et al., 2025). Positive strain changes of up to 25 µm/m were measured in the proximity of the breakage point (marked with a black cross). As the distance from the breakage in x-direction increased, strains exhibited a decreasing trend, reaching a value of 0 µm/m at approximately 350 mm. Furthermore, a slight change in the strain field was observed in the z-direction. The strain changes exhibited a maximum at the bottom edge ($z = 0.5h$, $\Delta\varepsilon_{c,x} \approx 25$ µm/m) while smaller changes were measured at the upper edge ($z = -0.5h$, $\Delta\varepsilon_{c,x} \approx 18$ µm/m). This was attributed to the full-surface support of the specimen on its bottom surface ($z = 100$ mm) during the experiment and corresponding friction.

It is important to note that the DFOS were applied after the induction of the prestressing force (after step 2), resulting in the specimen being compressed at the reference time t_0. Due to the tendon breakage (at t_1), the compression diminished around the breakage point and increased again toward the free end because of

Table 1. *Boundary conditions for different simulation steps of the computational model*

Simulation step	Bottom surface ($z = 0.5h$)	Central tendon cross section ($x = 0$)	Concrete cross section ($x = 0$)	Tendons free end ($x = l$)
Initial (1)	$u_z = 0$	$u_x = 0$	$u_x = 0$	$u_x = 0$
Loading (2)	$u_z = 0$	$u_x = 0$	$u_x = 0$	[–]
Breakage (3)	$u_z = 0$	[–]	$u_x = 0$	[–]

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

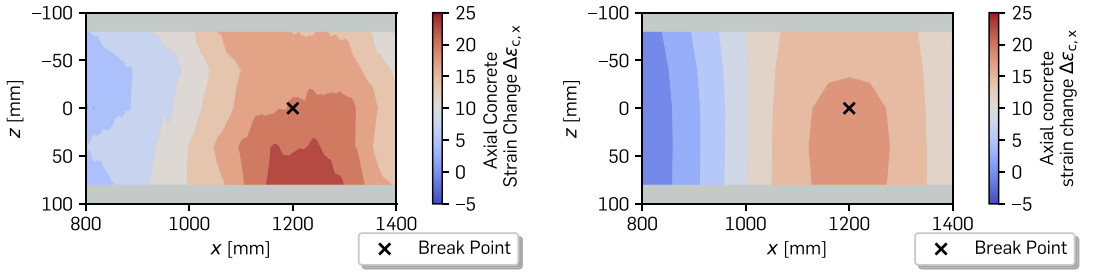


Figure 6. *Strain field determined in the direction of the tendon axis $\Delta\varepsilon_{c,x}$. (Left) Experimental, result from a broken tendon; (right) computational result showing strain changes between simulation steps 2 and 3.*

the re-anchorage of the tendon in the surrounding concrete (Hegger and Bertram, 2010). The measured tensile strain changes were therefore calculated as ($\Delta\varepsilon_{c,x} = \varepsilon_{c,x}(t_1) - \varepsilon_{c,x}(t_0)$).

3.4. Numerical results

Figure 6 (right) shows the strain changes coming from the simulation model presented in Section 3.2 at simulation step 3. The strain changes were computed at positions corresponding to those of the sensors on the test specimen (Section 3.1.2). To ensure better comparability with the measurements, the x-coordinates of the strain field were shifted to match those of the experimental investigation. Local tensile strains of up to 20 µm/m were observed around the breakage point. They decreased with increasing distance in x-direction. As observed in the experimentthe simulation also computed slightly greater strain changes at the bottom of the beam ($z = 100$ mm; 20 µm/m) than at the top ($z = -100$ mm; 17 µm/m). This was attributed to the implementation of the boundary condition on the bottom surface (see Table 1). Overall, the calculated strain field was slightly smoother than the experimental one, as the influence of measurement noise and minor inaccuracies in sensor localization during the touch-to-locate process was mitigated. Nonetheless, correspondence between the numerical simulation and the experiment was underscored by the coefficient of determination, $R^2 = 0.914$, and the correlation coefficient, $r = 0.957$.

3.5. Inverse problem definition

For the inverse problem, the parameters to be inferred are $\boldsymbol{\theta} = [E_{cm}, p_0, c_0, \mu]$. All the other parameters detailed in Sections 3.1 and 3.2, required for the numerical model, were set to their default values as identified from literature or material testing (see Table 2). The geometry was set to reproduce the experimental setup described in Section 3.1.

The observations used for the calibration were a discretization of the observations of the DFOS setup described in Section 3.1. Five rows of measurements were observed at heights $z \in \{-80, -40, 0, 40, 80\}$ mm starting at the tendon breakage point toward the end of the beam ($x = 0$) with a distance of 40 mm, equivalent to the vertical spacing of the DFOS. At a distance of approximately 400 mm from the break point, the measured strain changes became marginal (see Figure 6), resulting in 11 measurement points per height and accordingly 55 measurement points overall.

The predictions of the FEM model were used as output to be calibrated against. These predictions provided the difference in tensile strain $\Delta\varepsilon_{c,x}$ before and after breakage at the DFOS location. The predictions were collected at the equivalent locations in the model as the observations. The aim was to calibrate the model parameters such that the FEM model could reliably reproduce the observations.

3.6. Gaussian Process training

The GP surrogate model was trained using four input parameters: E_{cm}, c_0, p_0, and μ. These parameters were selected following a sensitivity analysis by Paul et al. (2026), which demonstrated their dominant

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Table 2. *Model parameters, descriptions, units, and default values*

Characteristic	Parameter	Description	Unit	Default
Concrete	E_{cm}	Young's modulus of concrete	MPa	Calibrated
	ν_c	Poisson's ratio of concrete	–	0.20
	σ_{b0}/σ_{c0}	Ratio of initial equibiaxial to uniaxial compressive yield stress	–	1.16
	K_c	Ratio of second stress invariants of tensile and compressive meridians	–	2/3
	ϵ	Eccentricity of CDP	mm	0.1
	f_{ctm}	Mean tensile strength of concrete	MPa	4.1
	ψ	Dilation angle of CDP	°	30.0
	η	Viscosity Parameter	–	0.0005
Steel	E_p	Young's modulus of steel	MPa	196000
	ν_p	Poisson's ratio of steel	–	0.30
Contact	μ	Friction coefficient	–	Calibrated
	c_0	Clearance at which contact pressure is initiated	mm	Calibrated
	p_0	Contact pressure at clearance$c = 0$	MPa	Calibrated
Solver	ζ	Numerical damping for solver stability	–	0.0002
Error	σ	Sensor noise standard deviation	–	0.01

Note: Parameters calibrated in $\boldsymbol{\theta}$ are indicated.

influence on the system response, while the remaining parameters exhibited negligible effects. The output quantity of interest was the difference in the strain field $\Delta\varepsilon_c$, extracted at N_{points} spatial locations over the surface of the finite element (FEM) model corresponding to the DFOS setup. This resulted in an output vector of dimension N_{points} for each simulation run.

A total of $N_{\text{samples}} = 100$ training points were generated using Latin Hypercube Sampling (LHS), ensuring good coverage of the multidimensional input space. The corresponding FEM simulations provided the training outputs. The GP was implemented with a RBF kernel combined with a white-noise kernel, using the `scikit-learn` framework. Inputs were normalized to the range of 0 to 1 in the space of the training samples. Outputs were normalized jointly as well to [0,1] over the whole dataset of training points. Table 3 summarizes the ranges of the input parameters and the kernel hyperparameters.

The hyperparameters were optimized by maximizing the log-marginal likelihood. The noise variance σ_n^2 converged to the minimum admissible value ($\sigma_n^2 = \sigma_{n,\min}^2$), which reflects the high consistency between the surrogate and the training data. One GP was trained for each of the observation points, leading to different optimal hyperparameters for each case.

Table 3. *Input parameter ranges and Gaussian Process hyperparameter bounds*

Parameter	Lower bound	Upper bound
c_0 [mm]	0.2012	0.7988
p_0 [MPa]	2.008	5.992
E_{cm} [MPa]	$2.702e+04$	$3.898e+04$
μ [-]	0.202	1.198
Length scalesℓ_i	0.01	100.0
Signal varianceσ_f^2	0.001	1000.0
Noise varianceσ_n^2	$1e-07$	0.1

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

The predictive performance of the trained GP model was assessed using an independent validation set of N_{val} samples, generated through LHS within the same parameter ranges as the training data. The FEM simulations corresponding to these validation points served as the reference values for comparison. The evaluation was based on standard regression metrics, including the coefficient of determination (R^2), the root mean square error (RMSE), the mean absolute error (MAE), the maximum error, the normalized RMSE (NRMSE) in percentage, and coverage statistics based on the absolute Z-value. The results were validated on 49 random samples on the training domain and are summarized in Table 4.

The obtained R^2 value close to unity confirmed that the GP surrogate accurately reproduced the FEM reference outputs across the parameter space. The RMSE and MAE values were small relative to the magnitude of the strains ε, indicating low overall prediction error. Additionally, the predictive uncertainty estimates provided by the GP were consistent with the residuals observed in the validation set. The distribution of the normalized prediction errors remained within the 95% confidence bounds for the majority of the validation points, demonstrating that the model not only achieved high accuracy but also provided reliable uncertainty quantification. These results validated the suitability of the GP surrogate for representing the input–output mapping of the FEM model in the DFOS configuration.

3.7. Parameter updating

The Bayesian calibration was carried out using the affine-invariant ensemble sampler implemented in `emcee`, with $N_{\text{walkers}} = 20$ walkers evolved for $N_{\text{iter}} = 10000$ iterations. The parameter space was explicitly bounded to remain within the domain of validity of the GP surrogate. In the presence of residual model–data discrepancy and strong parameter bounds, the resulting posterior could exhibit multiple disconnected modes, some corresponding to low-probability regions near the surrogate limits. In such cases, ensemble samplers could converge to distinct local probability wells (Hou et al., 2012).

To address this issue, we applied a likelihood-based clustering and pruning procedure after burn-in and at the end of sampling, designed to retain the dominant posterior mode while discarding walkers trapped in low-probability regions. The full algorithmic formulation is provided in Appendix A; here, we summarize the key principles. After burn-in, the mean log-posterior of each walker i was computed over the final fraction α of its trajectory

$$\ell_i = \frac{1}{T}\sum_{t=1}^{T} \ln \pi_{i,t}, \qquad T = \lfloor \alpha, T_{\text{burn}} \rfloor. \tag{56}$$

The values $\{\ell_i\}_{i=1}^{N_{\text{walkers}}}$ were sorted in ascending order, and clusters were identified by detecting large jumps in consecutive differences

$$d_k = \ell_{(k+1)} - \ell_{(k)}, \qquad k = 1, \ldots, N_{\text{walkers}} - 1, \tag{57}$$

where a jump was declared whenever

Table 4. *Training and validation metrics for both GP emulators*

Metric	Train	Validation
RMSE	0.1627	0.4409
MAE	0.1127	0.2738
R2	0.9991	0.9940
Max Error	0.8214	1.9970
NRMSE (%)	0.67	1.74
Mean |z|	0.6277	0.7481
|z| < 2 (%)	94.98	93.21
|z| > 0.5 (%)	43.48	52.39

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$d_k > \gamma \,\text{Median}\{d_k\}_{k=1}^{N_{\text{walkers}}-1}, \tag{58}$$

with $\gamma = 5$. These jumps partitioned the ensemble into disjoint clusters in log-posterior space. The cluster corresponding to the largest coherent posterior mass was retained, while walkers belonging to lower-probability clusters were pruned. Pruned walkers were replaced by new initial states obtained via convex combinations of retained walkers, ensuring that the ensemble remained within the high-probability region and inside the GP training domain. The corrected ensemble was then evolved for an additional $N_{\text{iter}} = 10000$ iterations. At the end of the full chain, the same clustering criterion was reapplied to obtain the final posterior ensemble, without resampling. The likelihood function was Gaussian and evaluated using the GP surrogate mean predictions and the observed DFOS data supposing a Gaussian noise variance $\sigma = 0.01$. GP hyperparameters were fixed during MCMC sampling at the optimized values obtained in Section 3.5.

Two updating approaches were considered. In the first approach, the parameter vector was defined as $\boldsymbol{\theta} = [E_{cm}, p_0, c_0, \mu]$. In the second approach, the parameter vector was augmented with $\sigma_{E_{cm}}$, where E_{cm} was interpreted as the mean of a lognormal distribution and $\sigma_{E_{cm}}$ as its standard deviation, such that $\tilde{E}_{cm} \sim \mathcal{LN}(E_{cm}, \sigma_{E_{cm}})$. The extended parameter $\tilde{E}_{cm}$ was propagated through the GP surrogate using a PCE of degree 2, as described in Section 2.1. Table 5 presents the prior distributions and bounds for the parameters to be calibrated. The likelihood function was adapted accordingly, as shown in Equation 14. The choice of augmenting E_{cm} was motivated by its high sensitivity, as it was identified as the primary driver of variations in the model response. The uncertainty represented through $\sigma_{E_{cm}}$ was intended to combine the effect of potential model discrepancies, regardless of their physical origin, even when these discrepancies were not directly associated with an incorrect representation of E_{cm}. The resulting uncertainty should therefore be interpreted as a vehicle for representing MFU and generating predictive distributions consistent with the observed measurements. Nevertheless, the embedded approach was inherently limited by the selected embedding parameter and the range of response variability that could be reproduced through its stochastic augmentation. In particular, when discrepancies arose from mechanisms not explicitly represented in the model and whose effects could not be approximated through variations in E_{cm} (e.g., bond–slip behavior or tendon–concrete interaction effects), part of the model discrepancy remained unaccounted for. Therefore, the transferability and effectiveness of the embedded MFU formulation depended on the ability of the chosen parameter to reproduce the dominant patterns of discrepancy observed in the response (Table 5).

The comparison between the non-embedded and embedded uncertainty calibration results is summarized in Table 6, while the calibration performance metrics are reported in Table 7. The corresponding trace and pair plots of the posterior samples are shown in Figure 7, and the predictive comparisons against the experimental observations are presented in Figure 8. Without embedding the uncertainty, the confidence intervals (CIs) were generated exclusively from the prescribed noise σ. Despite this, the resulting CIs fail to encompass the experimental observations, indicating the presence of model-form uncertainties not captured by the GP variance alone. As a consequence, the calibrated parameters weredriven toward the boundaries of their prior ranges, reflecting an artificial compensation for the missing model discrepancy. This behavior indicated that the calibration in this configuration was not

Table 5. *Priors for parameters used in MCMC calibration*

Parameter	Prior type	Parameters / bounds
E_{cm}	Lognormal	$\mu = 33000\,\text{MPa}, \sigma = 3300\,\text{MPa}[25200, 37050]$
$\sigma_{E_{cm}}$	Uniform	[0.25,7.41]
p_0	Uniform	[2.1,5.7]
c_0	Uniform	[0.21,0.76]
μ	Uniform	[0.21,1.14]

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

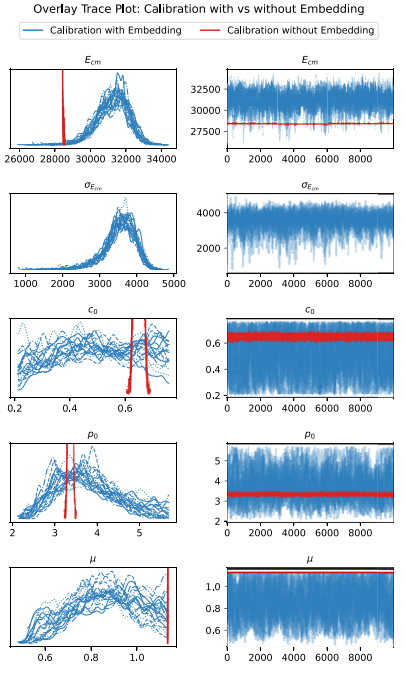
Table 6. *Predictive statistics (residual-based and |Z|-values) for the non-embedded and embedded uncertainty calibration cases*

	Residual-based				**|Z|-values**						
Case	Mean	RMSE	Median	MAD	Mean	Std. dev.	Median	MAD	$\|Z\| > 2$ (%)	$\|Z\| < 0.5$ (%)	Coverage (95%)
Non-UQ	0.23	**1.64**	0.16	1.66	1.87	1.49	1.52	1.22	40	**16**	60
UQ	**0.27**	1.67	**0.11**	**1.35**	**0.94**	**0.79**	**0.61**	**0.68**	**13**	42	**87**

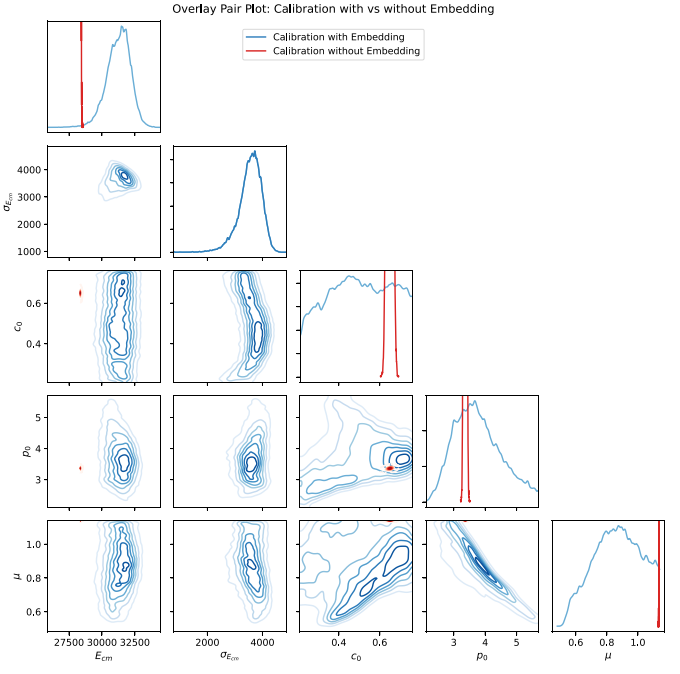
Note: Best values in bold.

Table 7. *Comparison of posterior summaries acros calibrations with and without embedding*

	Without embedding		With embedding	
Parameter	Mean (std)	95% CI	Mean (std)	95% CI
E_{cm}	28368.30 (18.35)	[28350.47, 28417.44]	31244.27 (894.13)	[29336.98, 32803.68]
$\sigma_{E_{cm}}$	–	–	3548.81 (416.91)	[2581.66, 4216.89]
p_0	3.36 (0.03)	[3.30, 3.43]	3.77 (0.73)	[2.56, 5.37]
c_0	0.65 (0.01)	[0.63, 0.67]	0.50 (0.15)	[0.23, 0.75]
μ	1.14 (0.00)	[1.14, 1.14]	0.87 (0.15)	[0.58, 1.12]



(a) Trace plot of MCMC calibration. Without (in red) and with embedding (in blue).



(b) Pair plot of MCMC calibration. Without (in red) and with embedding (in blue). PDF comparison in the diagonal.

Figure 7. *Comparison with and without embedding of MCMC calibration for the parameters E_{cm}, $\sigma_{E_{cm}}$, p_0, c_0, and μ. The left subfigure shows trace plots, and the right shows parameter pair plots.*

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

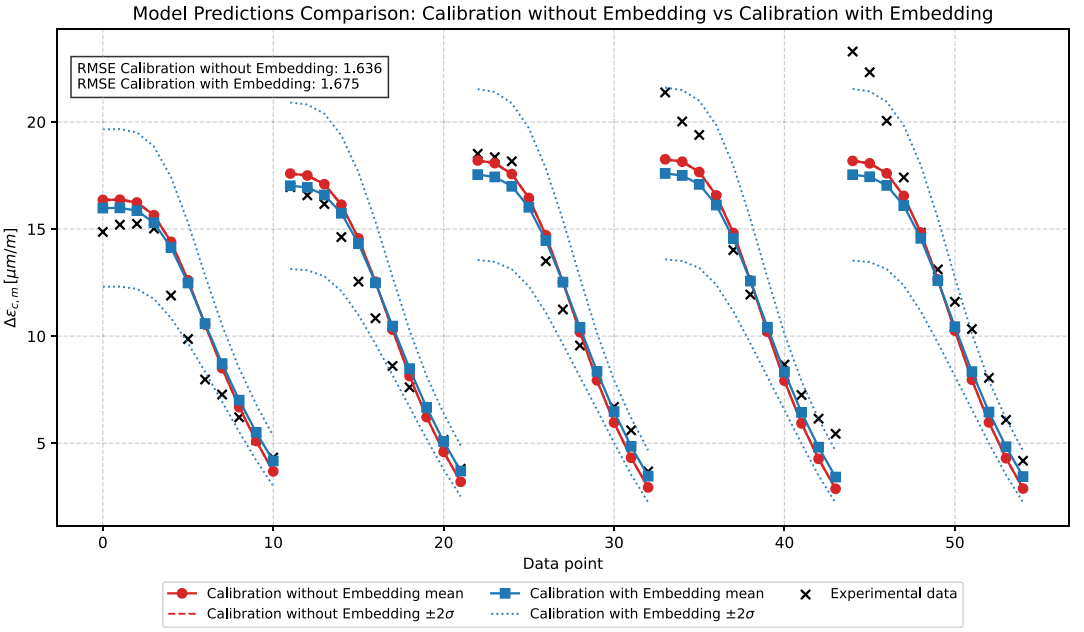


Figure 8. *Predicted confidence intervals at* 95% *of the posterior predictive for* $\widehat{\boldsymbol{\theta}}$ *with and without embedding. The confidence interval in the calibrated model without embedding depends only on the prescribed noise* σ.

physically meaningful, as reflected by the value of μ tending toward its maximum allowable value and E_{cm} approaching its minimum. In this case, the predictive variance did not constitute an actual quantification of MFU in the simulation model, but rather a by-product of uncertainty in the GP surrogate, which became less accurate near the boundaries of the domain due to the sampling scheme used to generate the training dataset.

When the uncertainty was embedded through the stochastic representation of E_{cm}, the inferred variance was directly associated with the material stiffness. This embedding inflated the CIs and improved their coverage of the observational data. The posterior distributions of the parameters were observed to be less concentrated, as a result of the flatter likelihood induced by the inclusion of variance in the quantification. Nevertheless, these posteriors converged toward more reasonable values than the extreme estimates obtained in the model without embedding. In particular, the inferred Young's modulus of approximately 31000 MPa, with an associated uncertainty of 3500 MPa, lay within the expected range of experimentally observed values. Additionally, a general lack of influence of d, together with a clear correlation between p_0 and μ, was observed in the posterior behavior. The Z-values, defined as the normalized residuals between observations and model predictions, confirmed these findings. They showed a clear increase in coverage (87% with versus 60% without embedding) and a decrease in outliers (13% of absolute Z-values exceeding 2 with versus 40% without embedding), despite an increase in the percentage of the dataset for which the discrepancy was smaller than the predictive intervals (42% with versus 16% non-without embedding). Nevertheless, some observations still fell outside the 95% predictive intervals, indicating that part of the MFU remains unaccounted for, even after embedding. Moreover, the predicted variance wasvsignificantly larger than the observed residual in more than one third of the predictions, suggesting a general overestimation of the predictive uncertainty due to the remaining discrepancies.

To further investigate the remaining discrepancies observed in the predictions, an influence analysis was performed on the model calibrated with embedded uncertainty. The predictions, representing the change in normal strains after breaking the tendon on the DFOS, still did not fully cover the experimental observations.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

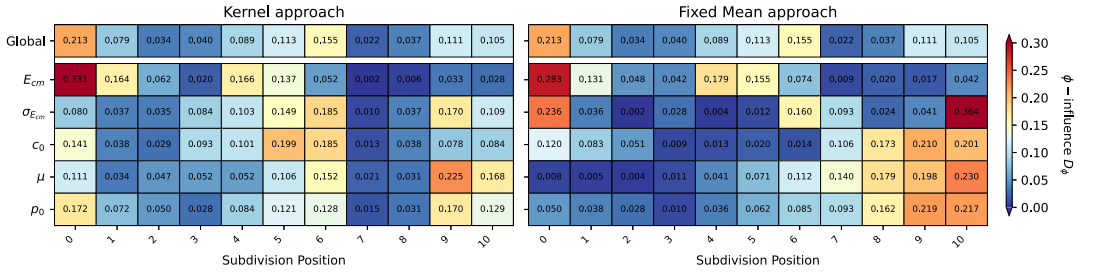


Figure 9. *Influence analysis on the model calibrated with embedded parameters. At the top for both sides, $\widehat{D}_\phi(S_i)$ for the full posterior for each S_i. On the left, influence $\widehat{D}_\phi^{(j)}(S_i)$ of each S_i on the marginal posterior of the parameters $\theta^{(j)}$ using KDE smoothing. On the right, influence $\widehat{D}_\phi^{(j),(\mathrm{fix})}(S_i)$ of each S_i on the marginal posterior of the parameters $\theta^{(j)}$ fixing the means of the other parameters. Influence values have been normalized to add to 1 in their respective case to enable the direct comparison.*

Since the sensor data originated from five different DFOS located at distinct vertical positions z, the observations were grouped by their distance to the tendon break, forming subsets S_i. Each subset S_i thus represented measurements obtained at a similar structural region with respect to the failure point.

The influence of each subset S_i on the posterior distribution of the parameters was quantified using ϕ-divergences, specifically the reverse Kullback–Leibler (KL) divergence between the posterior obtained using all observations for calibration and only a subset of them (see Section 2.2). This approach allowed assessing of how the removal of each subset affected the posterior, thereby identifying which experimental groups exerted the largest influence on the calibrated model. The resulting influence metrics are illustrated in Figure 9. The top plots display the global influence $\widehat{D}_\phi(S_i)$ for each subset on the full posterior, while the lower panels show, respectively, the marginal influences $\widehat{D}_\phi^{(j)}(S_i)$ and $\widehat{D}_\phi^{(j),(\mathrm{fix})}(S_i)$ for each parameter $\theta^{(j)}$, obtained using two alternative regularization approaches. All influence values were normalized to sum to one within each case, enabling direct comparison between subsets and parameter-wise influences.

The global influence analysis revealed a dominant contribution from the first subset, corresponding to the portion of the DFOS located closest to the tendon break. This region coincided with the observations least well captured by the predictive intervals, confirming that the local response near the failure point drove the largest posterior changes. A secondary effect was also observed in the sixth to seventh subdivision, affected by the increased discrepancies in the first set of observations (data points 7 and, to a lesser extent, 6).

The marginal influence analyses highlighted that the parameter E_{cm} was primarily driven by observations located close to the tendon breaking point. These measurements corresponded to the region exhibiting the largest model–data discrepancies and the highest sensitivity to calibration, indicating that variations in E_{cm} induced the most significant changes in the predicted strain response in this zone. The embedded variance parameter $\sigma_{E_{cm}}$ was also strongly affected in the fixed-point influence analysis, suggesting that part of the local discrepancy was being absorbed through an increased variability of the embedded stiffness rather than through systematic shifts in the mean response. A similar, albeit weaker, effect was observed for the last observation, which also displayed residual discrepancy at the identified optimum.

Two complementary influence formulations were considered to disentangle these effects. In the kernel-based influence analysis, the parameters were marginalized, such that the influence measure reflected the sensitivity of the joint posterior distribution after integrating out parameter dependencies. In this formulation, part of the discrepancy attributed to a given parameter could be compensated by correlated variations in other parameters, leading to a more spatially homogeneous influence pattern for $\sigma_{E_{cm}}$. In contrast, the fixed-mean influence analysis evaluated sensitivity by conditioning on fixed values of all parameters except the one under investigation. This isolates the direct contribution of individual

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

parameters to the posterior change and was therefore more sensitive to localized discrepancies that could not be alleviated through compensating parameter interactions.

The KDE-based analysis provided a global view of how observations influenced the inferred uncertainty when all parameter interactions were accounted for, while the fixed-mean analysis reveals which parameters were directly responsible for reconciling local discrepancies in the predictions. The latter was particularly relevant in the context of parameter transfer, where mean values are typically propagated to simulations of real structures. Taken together, the two influence measures indicated that the observed discrepancies near the breakage region stem from a localized model deficiency, which was partially absorbed through increased embedded stiffness variability when parameter interactions were allowed.

3.8. Identifiability assessment of parameters of interest

The final step of the analysis aimed to assess the identifiability of parameters of interest when the calibrated and uncertainty-embedded model was transferred to a full-scale, realistic structure. The parameters inferred from the calibration with uncertainty quantification, specifically the stochastic Young's modulus $\tilde{E}_{cm}$, were propagated to a new FE model representing a structural component with a different cross-section, based on the dimensions of a real bridge. This transfer enabled the propagation of both epistemic and aleatory uncertainties obtained from the experimental calibration into a context that more closely resembled a real engineering application.

The motivation for this analysis arose from the intended use of the model in SHM of prestressed concrete bridges. In a realistic setting, the calibrated parameters from laboratory-scale experiments would be employed to interpret field measurements, with the objective of diagnosing potential failures in the prestressing tendons. In particular, if a tendon breaks, the resulting strain field on the concrete surface changed in a way that depended on the depth of the breakage a (Paul et al., 2024b). Identifying which tendon had failed therefore corresponded to solving an inverse problem: given a new set of observed strains, determine the value of a (i.e., the depth of the broken tendon) that most likely produced them. Such identification was only feasible if the predictive distributions corresponding to different tendon depths were sufficiently distinct. The separability analysis presented here thus evaluated whether the embedded model uncertainty allowed discrimination between potential tendon breakage depths.

To perform this assessment, the stochastic Young's modulus $\tilde{E}_{cm}$, characterized by its mean and variance from the calibration, was treated as a random variable and propagated through the FE model for each tendon depth a. This yields predictive distributions of the strain response $\varepsilon \sim \mathcal{G}_{\mathbf{x}}(a)$ at spatial points $\mathbf{x}$, quantifying how uncertainty in material and bond parameters translated into uncertainty in the observable strain field.

The upscaled structural model represented a symmetric T-shaped beam of length l, shown in Figure 10. Its cross-section was defined by the web height $h_w = 1000$ mm, web width $t_w = 1200$ mm, flange thickness $t_f = 250$ mm, and flange width $b_f = 2250$ mm. The geometric parameters remain constant for all analyses. Since the study focused on the breakage of a single tendon and the concrete remained uncracked, the intact tendons were not explicitly modeled. The position of the tendon within the cross-section was defined by $(\Delta y, \Delta z)$, where Δy denoted the embedding depth with respect to the evaluated concrete surface and corresponded to the control parameter a. The strain response was evaluated along the outer surface of the web at $y = 0.5t_w$.

The central question of this study was whether the propagated uncertainty from the calibrated parameters still permitted distinguishable predictions for different values of a. If the predictive distributions for different tendon depths overlapped significantly, then new observations would not allow reliable inference of which tendon had failed. Conversely, well-separated predictive distributions implied that the distributed surficial strain measurement could be effectively used to identify the failure location. Hence, the separability of the predictive strain distributions provided a quantitative measure of the model's diagnostic capability and defined a metric for the informativeness of potential sensor locations.

Direct evaluation of the FE model for all a values and sensor points would be computationally prohibitive. To overcome this limitation, the propagation of $\tilde{E}_{cm}$ was performed using a PCE of degree 2, yielding the mean $\tilde{\mathbf{m}}_{\mathrm{PCE}}$ and standard deviation $\tilde{\boldsymbol{\sigma}}_{\mathrm{PCE}}$ fields at each spatial point. Two GP surrogate

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

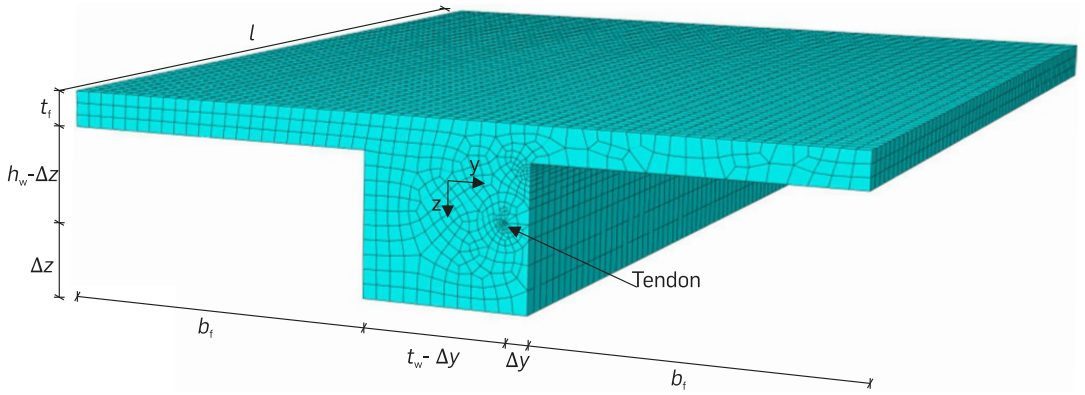


Figure 10. *Computational model of an upscaled T-beam used to investigate the separability of predictions for different tendon depths.*

models were trained to emulate the PCE outputs: one for the mean field and one for the standard deviation. Each GP takes as input $(\mathbf{x}, a)$ and outputs either $\tilde{\mathbf{m}}_{\mathrm{PCE}}$ or $\tilde{\boldsymbol{\sigma}}_{\mathrm{PCE}}$. Both employed an isotropic Matern kernel with a white-noise component, with inputs and outputs scaled to [0,1]. The optimized hyperparameters of the GP surrogates are listed in Table 8. The input dataset was composed of x, z and a in a grid $[5 \times 15 \times 10]$ divisions, from which 2500 data points were taken randomly for training the GPs and the remaining 5000 were used for validation. Validation metrics are presented in Table 9.

The separability of the predictive distributions associated with different tendon depths a was quantified using the GP surrogates following Algorithm 1. Two complementary statistics were employed. The first is a Min- Δ metric, which measures the smallest distinguishable difference in tendon depth Δa such that the corresponding predictive distributions at a given sensor location remained separable under the propagated uncertainty. Smaller values of Δa indicated higher sensitivity of the sensor to changes in a. The second is a PDF-overlap metric, which quantifies the degree of overlap between predictive probability density

Table 8. *Hyperparameter summary for mean and std GP surrogates (input parameter a)*

	Bounds	Mean GP	Std GP
ℓ	[0.01,100.0]	0.5561	0.2561
σ_f^2	[0.001,10000.0]	780.3	0.6031
σ_n^2	$[1e-08, 0.01]$	1.373e–05	7.549e–05

Table 9. *Validation metrics for GP mean and std emulators*

Metric	Mean GP	Std GP
RMSE	0.1639	0.0282
MAE	0.05257	0.00938
R2	0.9978	0.9938
Max Error	3.772	0.4837
NRMSE (%)	0.23	0.45
Mean |z|	0.3125	0.2445
|z| < 2 (%)	97.70	97.51
|z| > 0.5 (%)	13.09	10.19

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

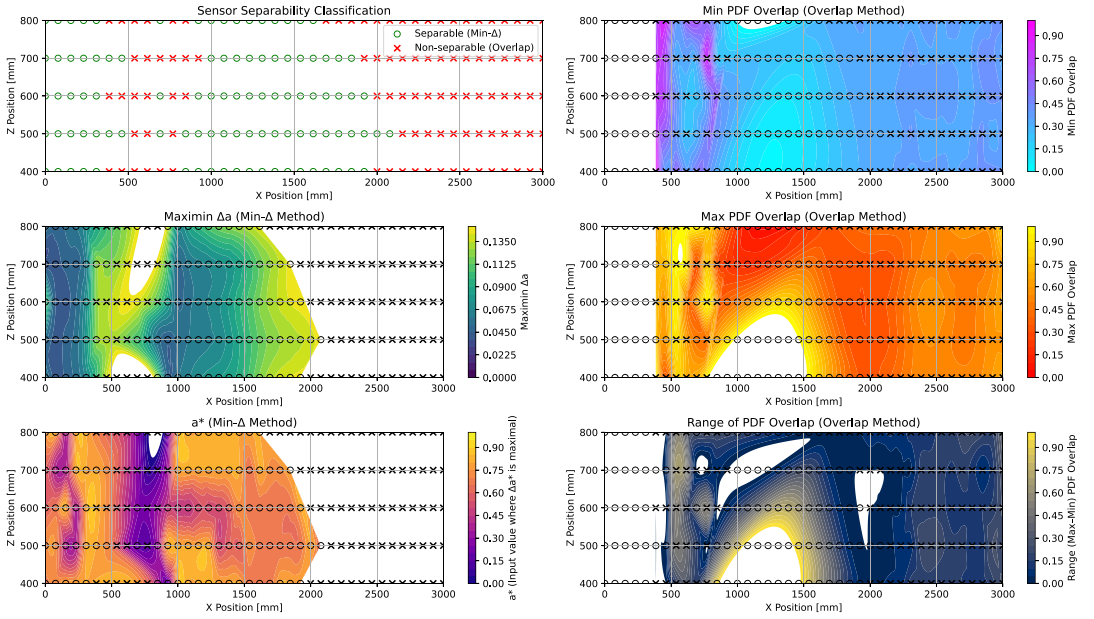


Figure 11. *Comparison of sensor separability and overlap metrics for two classification approaches. Each subplot shows the spatial distribution of virtual sensors in the x–z plane, with separable sensors (∘) identified by the Min-Δ method and non-separable sensors (×) identified by the PDF-overlap method. The position of the tendon breakage is at $x = 0$ mm and $z = 600$ mm. (Top left) Sensor separability classification map. (Middle left) Maximin Δa field (Min-Δ method). (Bottom left) a^* field indicating the tendon depth where Δa^* is maximal. (Top right) Minimum PDF overlap field. (Middle right) Maximum PDF overlap field (Overlap method). (Bottom right) Range of PDF overlap summarizing overlap variability. Interpolation between applicable sensors is performed using cubic splines.*

functions associated with different values of a; low overlap implies high discriminative power, while high overlap indicates limited ability to distinguish tendon breakage depths.

The optimization required to compute these statistics was performed using the Nelder–Mead method. The resulting spatial fields of separability and overlap are shown in Figure 11. The top-left panel presented a classification of virtual sensor locations in the x-z plane on the concrete surface ($y = t_w/2$) into separable (∘) and non-separable (×) regions based on the combined criteria of the two metrics. The middle-left panel showed the spatial distribution of the maximin Δa values obtained with the Min- Δ method, highlighting regions where small changes in tendon depth could be reliably distinguished. The bottom-left panel reported the value a^* at which Δa is maximized, indicating which tendon depth was most influential at each sensor location. The top-right panel displayed the minimum PDF overlap across all considered tendon depths, identifying sensor locations that provide the strongest discrimination. The middle-right panel showed the maximum overlap, revealing regions where predictive distributions remained highly overlapping regardless of a. The bottom-right panel summarized the range of PDF overlap, providing a measure of variability in discriminative power across depth scenarios.

The results of the separability analysis revealed four distinct regions along the beam that corresponded to different identifiability regimes. Close to the breakage point ($x < 450$ mm), the predictive distributions for different tendon depths a were well separated, indicating that the model response was highly sensitive to the breakage position and that the propagated material-form uncertainty did not obscure this sensitivity. In this region, the strain response was dominated by the local stress redistribution caused by the tendon break, which led to measurable and distinguishable strain patterns across different tendon depths.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Moving away from the breakage zone, a region of poor separability was observed (450 mm $< x <$ 900 mm). Here, the predictive distributions corresponding to different a values exhibited significant overlap. This behavior was attributed to the interaction between the geometric eccentricity of the broken tendon, which induces a bending moment around the vertical axis z in the structure, and the uncertainty embedded in the stochastic Young's modulus $\tilde{E}_{cm}$. The superposition of these two effects generated similar strain patterns for different tendon depths, thereby reducing the discriminatory power of the model in this intermediate region.

Further along the beam (900 mm $< x <$ 1900 mm), the separability increased again, as the bending-induced strains stabilized and the relative effect of the tendon position on the overall strain field becomes more distinct. However, this trend does not persist indefinitely: far from the breakage point ($x >$ 1900 mm), the strain changes become too small to be distinguishable within the propagated uncertainty, and the predictive distributions once again overlapped. This spatial pattern confirmed that the propagation of the material-form uncertainty through the calibration framework provided valuable diagnostic information. It identifies regions where predictions were robust and distinguishable, and conversely, where model predictions should not be trusted for inverse inference. Consequently, this information could directly guide sensor placement strategies in practice, ensuring that measurements were taken in regions of maximal separability when the number of available sensors was limited. This assessment was specific to the calibrated model, a different setup would generate other predictive distributions and affect separability.

It should be noted that the present separability assessment was conducted on a single point-sensor basis, evaluating each spatial location independently. Considering multiple sensor locations simultaneously or a DFOS, and accounting for their spatial correlation, would enable a more comprehensive and tighter assessment of separability, potentially improving the identification of the breakage depth. This extension was left for future work. Additionally, the possible error introduced by transferring the calibrated parameters from the experimental setup to the upscaled model was not been explicitly quantified, as such transfer errors were inherently model-dependent and could not be rigorously evaluated within the current framework.

4. Conclusions

This work presented a comprehensive, uncertainty-aware workflow for tendon breakage localization in prestressed concrete members, integrating laboratory-scale tendon breakage experiments and DFOS measurements, high-fidelity FEM, surrogate modeling via GPs, and Bayesian inference with embedded MFU. The framework addressed the fundamental challenge that direct calibration on real structures is infeasible, as destructive tendon-breakage tests can hardly be performed in situ. By first calibrating models under controlled laboratory conditions and embedding uncertainties probabilistically, the workflow enabled transferable, trustworthy predictions of tendon breakage locations and depths in realistic structural configurations such as bridges.

The starting point of the workflow was a detailed FEM representation of the tendon-breakage experiment, capable of reproducing the nonlinear structural response from re-anchoring and strain redistribution following a tendon breakage. To enable efficient probabilistic calibration, the computationally intensive FEM simulations were replaced by GP surrogate models that emulated the structural response across the parameter space while retaining accuracy in the regions most informed by the data. This surrogation step provided the foundation for the subsequent Bayesian calibration, in which the parameters governing material and bond behavior were inferred using the MCMC sampler. A key achievement of this study lay in the explicit quantification of model form uncertainty through parameter embedding. By interpreting the effective Young's modulus E_{cm} as a stochastic variable with an inferred variance $\sigma_{E_{cm}}$, the calibration gained the ability to represent structural and modeling discrepancies probabilistically. This embedding not only improved the statistical consistency between model predictions and observations but also provided a physically interpretable mechanism for uncertainty propagation, thereby increasing the reliability and trustworthiness of the inferred parameters. Building upon the calibrated model, an influence analysis based on ϕ-divergence (using Kullback–Leibler divergence) was

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

conducted to quantify the impact of specific data subsets on the posterior distributions. This diagnostic step allowed the identification of regions and parameters most responsible for model discrepancy, most notably, the sensitivity of the posterior to measurements near the tendon breakage. The insights gained from this analysis could directly guide the improvements to the model, targeting the area around the tendon break.

The final stage of the workflow demonstrated the transferability and predictive power of the calibrated and uncertainty-embedded model. By propagating the stochastic parameters to a realistic structural configuration, the study assessed the identifiability of an additional parameter, for example, the depth of a broken tendon, through a separability analysis of the propagated predictive distributions. This analysis confirmed that the embedded MFU preserved sufficient separability near the breakage point to enable tendon-depth identification, while also revealing spatial regions where predictions are unreliable due to overlapping uncertainties. These findings provided a rational basis for optimizing sensor placement and for assessing confidence in inverse identification tasks. In summary, the developed framework achieves a seamless integration of experimental evidence, physics-based simulation, surrogate modeling, and probabilistic reasoning. The resulting methodology formed a comprehensive, data-informed approach for uncertainty-aware model calibration and validation in structural applications, and established a foundation for extending such techniques to large-scale engineering systems.

Although the proposed workflow demonstrated a reliable integration of experimental data, surrogate modeling, and Bayesian uncertainty embedding, several limitations must be acknowledged. The calibration strongly depended on the quality and representativeness of the experimental DFOS measurements, which were specific to the laboratory setup. Consequently, the inferred posterior distributions and embedded uncertainty reflected the conditions and boundary effects of that configuration. When transferring the parameters to different geometries or scales, these dependencies might not be fully preserved, introducing unquantifiable epistemic uncertainty. The use of GP surrogates and the embedding of MFU through a single stochastic variable $\tilde{E}_{cm}$ constituted additional simplifications. While these approaches enabled efficient calibration and propagation, they assumed that the main discrepancies can be captured by the variability in the effective stiffness, neglecting spatial or multi-parameter sources of model error. Similarly, both the influence and separability analyses were performed under independence assumptions, that is, considering each sensor or subset separately, thus not fully capturing correlations that could enhance discriminative power. These limitations highlighted that, although the framework provided a systematic and trustworthy foundation for uncertainty-aware model calibration and interpretation, future developments should aim at multi-source uncertainty representation, adaptive surrogate refinement, and joint sensor analysis to further improve generalization and reliability.

Future developments of the proposed framework will focus on extending its capability to capture more complex and realistic sources of uncertainty and to enhance its diagnostic resolution. Incorporating correlations between sensor measurements and spatially distributed observations (Mu et al., 2022) will enable multivariate separability and influence analyses, providing tighter and more reliable identification of parameter changes. Adaptive surrogate modeling strategies (Semler and Weiser, 2023), such as active learning or multi-fidelity GP, could be employed to refine the representation of highly nonlinear responses while maintaining computational efficiency. Moreover, the extension of the embedding concept to multiple MFU parameters, potentially coupled with spatial variability fields, would allow a more physically consistent quantification of discrepancies (Strong et al., 2015). Finally, integrating experimental design optimization into the workflow could guide sensor placement and data acquisition strategies, maximizing the information gain for future structural monitoring or model updating campaigns (de Souza Mello et al., 2024; Orozco et al., 2025; Zhang et al., 2025; Bakeer et al., 2025). In this context, the separability analysis presented in this study could provide a quantitative basis for sensor layout design. Rather than relying exclusively on deterministic sensitivity metrics, sensor placement strategies could prioritize locations that maximize the separability of potential damage states while accounting for both model and measurement uncertainties. Such an approach may lead to more robust monitoring systems and more reliable damage localization in operational bridge environments, where uncertainty levels are typically higher than those encountered under laboratory conditions. Uncertainty from the effect of

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

unknown loads and environmental conditions could enter as additional stochastic terms in the forward model evaluated for the separability assessment, allowing for a seamless integration of those more complex scenarios. Additionally, the current formulation is most directly applicable to the optimization of the placement of individual sensors. Extending the framework to multisensor systems, particularly distributed sensing technologies such as DFOS, would require explicitly accounting for the covariance structure among measurements. In such cases, separability should be evaluated based on the joint response of the sensor network, as correlations between measurements can significantly influence both identifiability and damage-localization performance.

Data availability statement. Replication data and code can be found at Zenodo: https://doi.org/10.5281/zenodo.18713387.

Acknowledgments. The authors used ChatGPT to assist with drafting of this article, specially to improve text clarity and grammar. GPT-4 and GPT-5 were accessed from https://chatgpt.com/ and used without modification between July 2025 and February 2026. All scientific content, ideas, and interpretations originated from the authors. Claude Sonnet 4.5 was used through the Github Copilot interface for assisted programming in generating the code required to obtain the results. All the content was thoroughly reviewed by the authors.

Author contribution. Conceptualization: M.W., D.S., P.M., J.F.U.; Data curation: D.A.A., A.P.; Data visualization: D.A.A., A.P.; Formal analysis: D.A.A., A.P.; Funding acquisition: M.W., D.S., P.M., J.F.U.; Investigation: D.A.A., A.P.; Methodology: D.A.A., A.P.; Project administration: M.W., D.S., P.M., J.F.U.; Resources: P.M., J.F.U.; Review and editing: M.W., D.S., P.M., J.F.U.; Software: D.A.A., A.P.; Supervision: M.W., D.S., P.M., J.F.U.; Writing original draft: D.A.A., A.P.. All authors approved the final submitted draft.

Funding statement. This research was supported by the German Research Foundation (DFG) - Project number 461030501 in the special focus program SPP 2388/1 “Hundert plus – Verlängerung der Lebensdauer komplexer Baustrukturen durch intelligente Digitalisierung” (Hundred Plus – Extending the service life of complex building structures through intelligent digitalization) in the subprojects B04: Pattern detection of internal prestressing wire breaks on concrete surfaces (project number 501774158) and C07: Data driven model adaptation for identification of digital twins of bridges (project number 501811638).

Competing interest. The authors declare none.

Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

Abdelatif AO, **Owen JS and Hussein MFM** (2017) Modeling and parametric study of the reanchorage of ruptured tendons in bonded posttensioned concrete. *Journal of Structural Engineering 143*(12), 04017162. https://doi.org/10.1061/(asce)st.1943-541x.0001898.

Arcones DA, **Weiser M**, **Koutsourelakis P-S and Unger JF** (2024) Model bias identification for bayesian calibration of stochastic digital twins of bridges. *Applied Stochastic Models in Business and Industry*. https://doi.org/10.1002/asmb.2897.

Arcones DA, **Weiser M**, **Koutsourelakis P-S and Unger JF** (2026) Embedded model form uncertainty quantification with measurement noise for bayesian model calibration. *Data-Centric Engineering 7*. https://doi.org/10.1017/dce.2025.10035.

Ayoub A and Filippou FC (2010) Finite-element model for pretensioned prestressed concrete girders. *Journal of Structural Engineering 136*(4), 401–409. https://doi.org/10.1061/(asce)st.1943-541x.0000132.

Bakeer T, **Herbers M and Marx S** (2025) Sensor informativeness, identifiability, and uncertainty in bayesian inverse problems for structural health monitoring. *arXiv*: https://arxiv.org/abs/2511.16628, November.

Barrias A, **Casas JR and Villalba S** (2016) A review of distributed optical fiber sensors for civil engineering applications. *Sensors 16*(5), 748. https://doi.org/10.3390/s16050748; Available at https://www.mdpi.com/1424-8220/16/5/748.

Bayarri MJ, **Berger JO**, **Paulo R**, **Sacks J**, **Cafeo JA**, **Cavendish J**, **Lin C-H and Tu J** (2007) A framework for validation of computer models. *Technometrics 49*(2), 138–154. https://doi.org/10.1198/004017007000000092.

Becks H, **Lippold L**, **Winkler P**, **Moeller M**, **Rohrer M**, **Leusmann T**, **Anton D**, **Sprenger B**, **Kähler P**, **Rudenko I**, **Arcones DA**, **Koutsourelakis P-S**, **Unger JF**, **Weiser M**, **Petryna Y**, **Schnellenbach-Held M**, **Lowke D**, **Wessels H**, **Lenzen A**, **Zabel V**, **Könke C**, **Claßen M and Hegger J** (2024) Neuartige Konzepte für die Zustandsüberwachung und -analyse von Brückenbauwerken – Einblicke in das Forschungsvorhaben SPP100+/Novel concepts for the condition monitoring and analysis of bridge structures – Insights into the spp100+ research project. *Bauingenieur 99*(10), 327–338. https://doi.org/10.37544/0005-6650-2024-10-63.

Bellman R and Åström KJ (1970) On structural identifiability. *Mathematical Biosciences 7*(3–4), 329–339. https://doi.org/10.1016/0025-5564(70)90132-x.

Bergmeister K, **Mark P**, **Österreicher M**, **Sanio D**, **Heek P**, **Krawtschuk A**, **Strauss A and Ahrens MA** (2015) Innovative Monitoringstrategien für Bestandsbauwerke. In *BetonKalender*. https://doi.org/10.1002/9783433603406.ch7.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Birtel V and Mark P.(2006) Parameterised finite element modelling of RC beam shear failure. *ABAQUS users conference*. Bochum: Ruhr-University Bochum; 2006. p. 95–108.

Briere V, **Harries KA**, **Kasan J and Hager C** (2013) Dilation behavior of seven-wire prestressing strand – The Hoyer effect. *Construction and Building Materials 40*, 650–658. https://doi.org/10.1016/j.conbuildmat.2012.11.064.

Burger H, **Betz P**, **Richter B**, **Herbers M**, **Schramm N**, **Diers J**, **Schacht G**, **Lingemann J**, **Marx S and Fischer O** (2024) Untersuchungen zur Verbundverankerung von durchtrennten Spanngliedern beim Brückenrückbau. *Beton- und Stahlbetonbau 119*(10), 753–763. https://doi.org/10.1002/best.202400046.

Calò M, **Ruggieri S**, **Buitrago M**, **Nettis A**, **Adam JM and Uva G** (2025) An ml-based framework for predicting prestressing force reduction in reinforced concrete box-girder bridges with unbonded tendons. *Engineering Structures 325*, 119400. https://doi.org/10.1016/j.engstruct.2024.119400.

Chapeleau X and Bassil A (2021) A general solution to determine strain profile in the core of distributed fiber optic sensors under any arbitrary strain fields. *Sensors 21*(16), 5423. https://doi.org/10.3390/s21165423 https://www.mdpi.com/1424-8220/21/16/5423.

Clauß F, **Ahrens MA and Mark P** (2021) A comparative evaluation of strain measurement techniques in reinforced concrete structures. *Structural Concrete 22*(5), 2992–3007. https://doi.org/10.1002/suco.202000706.

Dassault Systèmes Simulia Corporation (2025) Abaqus theory guide.

de Domenico D, **Shamsaddinlou A**, **Maugeri N**, **Longo P**, **Messina D**, **Recupero A**, **La Mazza D**, **Longo M**, **Darò P and Mancini G** (2026) Modal identification of prestressed concrete beams with artificial damage on prestressing strands. *Procedia Structural Integrity 78*, 65–72. https://doi.org/10.1016/j.prostr.2025.12.009.

de Souza Mello FM, **Pereira JLJ and Gomes GF** (2024) Multi-objective sensor placement optimization in shm systems with kriging-based mode shape interpolation. *Journal of Sound and Vibration 568*, 118050. https://doi.org/10.1016/j.jsv.2023.118050.

Farrar CR and Worden K (2007) An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365*(1851), 303–315. https://doi.org/10.1098/rsta.2006.1928.

Fiedler M and Marx S (2026) Probabilistic estimation of reliable sensor spacing for wire break detection in prestressed concrete using acoustic emission technique. *Structural Concrete*. https://doi.org/10.1002/suco.70672.

Forrester AIJ, **Sóbester A and Keane AJ** (2008) *Engineering Design Via Surrogate Modelling: A Practical Guide*. Chichester: Wiley. https://doi.org/10.1002/9780470770801.

Hastie T, **Tibshirani R and Friedman J** (2009) *The Elements of Statistical Learning*. New York: Springer. https://doi.org/10.1007/978-0-387-84858-7.

Hegger J and Bertram G (2010) Verbundverhalten von vorgespannten Litzen in UHPC. *Beton- und Stahlbetonbau 105*(6), 379–389. https://doi.org/10.1002/best.201000019.

Herbers M, **Bartels J-H**, **Richter B**, **Collin F**, **Ulbrich L**, **Al-Zuriqat T**, **Geck CC**, **Naraniecki H**, **Hahn O**, **Jesse F**, **Smarsly K and openLAB SM** (2024) Eine Forschungsbrücke zur Entwicklung eines digitalen Brückenzwillings. *Beton- und Stahlbetonbau*. https://doi.org/10.1002/best.202300094.

Herbers M, **Richter B**, **Gebauer D**, **Classen M and Marx S** (2023) Crack monitoring on concrete structures: Comparison of various distributed fiber optic sensors with digital image correlation method. *Structural Concrete 24*(5), 6123–6140. https://doi.org/10.1002/suco.202300062.

Hillerborg A (1983) Analysis of one single crack. In F.H. Wittmann, *Fracture Mechanics of Concrete (Developments in Civil Engineering)*, pp. 223–249. United States: Elsevier. Available at https://portal.research.lu.se/en/publications/analysis-of-one-single-crack.

Hou F, **Goodman J**, **Hogg DW**, **Weare J and Schwab C** (2012) An affine-invariant sampler for exoplanet fitting and discovery in radial velocity data. *The Astrophysical Journal 745*(2), 198. https://doi.org/10.1088/0004-637x/745/2/198.

Janiak T, **Becks H**, **Camps B**, **Classen M and Hegger J** (2023) Evaluation of distributed fibre optic sensors in structural concrete. *Materials and Structures 56*(9), 1–18. https://doi.org/10.1617/s11527-023-02222-9; Available at https://link.springer.com/article/10.1617/s11527-023-02222-9.

Jankowiak T and Lodygowski T (2005) Identification of parameters of concrete damage plasticity constitutive model. *Foundations of Civil and Environmental Engineering 6*, 53–69. Available at https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-bpp1-0059-0053.

Kang C, **Arcones DA**, **Becks H**, **Beetz J**, **Blankenbach J**, **Claßen M**, **Degener S**, **Eisermann C**, **Göbels A**, **Hegger J**, **Hermann R**, **Kähler P**, **Peralta P**, **Petryna Y**, **Schnellenbach-Held M**, **Schulz O**, **Smarsly K**, **Sönmez MF**, **Sprenger B**, **Unger JF**, **Vassilev H**, **Weiser M and Marx S** (2025) Intelligente digitale Methoden zur Verlängerung der Nutzungsdauer der Nibelungenbrücke. *Beton- und Stahlbetonbau*. https://doi.org/10.1002/best.70070.

Kennedy MC and O'Hagan A (2001) Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(3), 425–464. https://doi.org/10.1111/1467-9868.00294.

Konertz D, **Löschmann J**, **Clauß F and Mark P** (2019) Faseroptische Messung von Dehnungs- und Temperaturfeldern. *Bauingenieur 94*(07-08), 292–300. https://doi.org/10.37544/0005-6650-2019-07-08-70.

Krätzig WB and Pölling R (2004) An elasto-plastic damage model for reinforced concrete with minimum number of material parameters. *Computers & Structures 82*(15-16), 1201–1215. https://doi.org/10.1016/j.compstruc.2004.03.002.

Lee J and Fenves GL (1998) Plastic-damage model for cyclic loading of concrete structures. *Journal of Engineering Mechanics 124*(8), 892–900. https://doi.org/10.1061/(ASCE)0733-9399(1998)124:8(892).

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Lubliner J, **Oliver J**, **Oller S and Oñate E** (1989) A plastic-damage model for concrete. *International Journal of Solids and Structures 25*(3), 299–326. https://doi.org/10.1016/0020-7683(89)90050-4.

Mu H-Q, **Liang X-X**, **Shen J-H and Zhang F-L** (2022) Analysis of structural health monitoring data with correlated measurement error by bayesian system identification: Theory and application. *Sensors 22*(20), 7981. https://doi.org/10.3390/s22207981.

Nadaraya EA (1964) On estimating regression. *Theory of Probability & Its Applications 9*(1), 141–142. https://doi.org/10.1137/1109020.

Nettis A, **Nettis A**, **Ruggieri S and Uva G** (2024) Corrosion-induced fragility of existing prestressed concrete girder bridges under traffic loads. *Engineering Structures 314*, 118302. https://doi.org/10.1016/j.engstruct.2024.118302.

Orozco JAP, **Chamoin L**, **Cortial J**, **de Buhan M and Soulier B** (2025) Optimal placement of distributed optic fiber sensors with the modified constitutive relation error. *Mechanical Systems and Signal Processing 238*, 113206. https://doi.org/10.1016/j.ymssp.2025.113206.

Ostachowicz W, **Soman R and Malinowski P** (2019) Optimization of sensor placement for structural health monitoring: A review. *Structural Health Monitoring 18*(3), 963–988. https://doi.org/10.1177/1475921719825601.

Paul A, **Sanio D and Mark P** (2024a) Detection of internal tendon breaks by fiber-optical measurements – Influence of the re-anchoring behavior. *Procedia Structural Integrity 64*, 1287–1294. https://doi.org/10.1016/j.prostr.2024.09.199.

Paul A, **Sanio D and Mark P** (2024b) Monitoring tendon breaks in concrete structures at different depths using distributed fiber optical sensors. *e-Journal of Nondestructive Testing 29*(7). https://doi.org/10.58286/29598.

Paul A, **Sanio D and Mark P** (2025) Detektion innerer Spannstahlbrüche durch faseroptische Messungen an Betonoberflächen. *Beton- und Stahlbetonbau* (120). https://doi.org/10.1002/best.202400100.

Paul A, **Sanio D and Mark P** (2026) Sensitivity study on accuracy and relevant parameters in numerical modeling of internal tendon breaks. *Structural Concrete*. https://doi.org/10.1002/suco.70619.

Pedregosa F, **Varoquaux G**, **Gramfort A**, **Michel V**, **Thirion B**, **Grisel O**, **Blondel M**, **Prettenhofer P**, **Weiss R**, **Dubourg V**, **Vanderplas J**, **Passos A**, **Cournapeau D**, **Brucher M**, **Perrot M and Duchesnay E** (2011) Scikit-learn: Machine learning in python. *Journal of Machine Learning Research 12*, 2825–2830.

Perrin G, **Marco RJD**, **Soize C and Funfschilling C** (2025) Estimating intractable posterior distributions through gaussian process regression and metropolis-adjusted langevin procedure. *arXiv*arxiv: https://arxiv.org/abs/2506.13336, June.

Pirskawetz SM and Schmidt S (2023) Detection of wire breaks in prestressed concrete bridges by acoustic emission analysis. *Developments in the Built Environment 14*, 100151. https://doi.org/10.1016/j.dibe.2023.100151.

Rasmussen CE (2006) *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.

Richter B, **Will E**, **Herbers M and Marx S** (2025) Detection of prestressing wire breaks in post-tensioned concrete structures using distributed fiber optic strain sensing. *Journal of Civil Structural Health Monitoring*. https://doi.org/10.1007/s13349-025-01018-5.

Samiec D (2011) Verteilte faseroptische Temperatur- und Dehnungsmessung mit sehr hoher Ortsauflösung. *Photonik* (6), 34–37.

Sargsyan K, **Najm HN and Ghanem R** (2015) On the statistical calibration of physical models. *International Journal of Chemical Kinetics 47*(4), 246–276. https://doi.org/10.1002/kin.20906.

Sargsyan K, **Huan X and Najm HN** (2019) Embedded model error representation for Bayesian model calibration. *International Journal for Uncertainty Quantification 9*(4), 365–394. https://doi.org/10.1615/int.j.uncertaintyquantification.2019027384.

Scott DW (2008) *Multivariate Density Estimation Theory, Practice, and Visualization*. New York: John Wiley & Sons, Incorporated.

Seiffert A and Jensen A (2019) A practical approach for modeling tendon and wire failures for model-based damage detection of prestressed concrete bridges. *In 5th International Conference on Smart Monitoring, Assessment and Rehabilitation of Civil Structures*, volume *25*, Potsdam, August. e-Journal of Non-destructive Testing. https://www.ndt.net/?id=24910.

Semler P and Weiser M (2023) Adaptive gaussian process regression for efficient building of surrogate models in inverse problems. *Inverse Problems 39*(12), 125003. https://doi.org/10.1088/1361-6420/ad0028.

Sieradzki K and Newman RC (1987) Stress-corrosion cracking. *Journal of Physics and Chemistry of Solids 48*(11), 1101–1113. https://doi.org/10.1016/0022-3697(87)90120-x.

Speck K, **Vogdt F**, **Curbach M and Petryna Y** (2019) Faseroptische Sensoren zur kontinuierlichen Dehnungsmessung im Beton. *Beton- und Stahlbetonbau 114*(3), 160–167. https://doi.org/10.1002/best.201800105.

Sträter N, **Clauß F**, **Ahrens MA and Mark P** (2024) Detection of tendon breaks in prestressed concrete structures using coda wave interferometry. *Structural Concrete*. https://doi.org/10.1002/suco.202400680.

Sträter N, **Clauß F**, **Ahrens MA and Mark P** (2026) Identification of invisible damage in prestressed concrete bridges—A demonstrator case using coda wave interferometry. *Structural Concrete*. https://doi.org/10.1002/suco.70674.

Strong M, **Oakley JE**, **Brennan A and Breeze P** (2015) Estimating the expected value of sample information using the probabilistic sensitivity analysis sample. *Medical Decision Making 35*(5), 570–583. https://doi.org/10.1177/0272989x15575286.

Sudret B (2021) Polynomial chaos expansions in 90 minutes. [Conference presentation]. 4th National Conference on Multi-disciplanry Design, Analysis and Optimization (NCMDAO 2021), ETH Zurich. https://doi.org/10.3929/ethz-b-000508852

van Meirvenne K, **de Corte W**, **Boel V and Taerwe L** (2018) Non-linear 3d finite element analysis of the anchorage zones of pretensioned concrete girders and experimental verification. *Engineering Structures 172*, 764–779. https://doi.org/10.1016/j.engstruct.2018.06.065.

van Mier JGM, **Reinhard HW and van der Vlugt BW** (1987) Ergebnisse dreiachsiger verformungsgesteuerter Belastungsversuche an Beton. *Bauingenieur* (62), 353–361.

Watson GS (1964) Smooth regression analysis. *Sankhhyā: The Indian Journal of Statistics, Series A 26*(4), 359–372.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

Weiss R (1996) An approach to bayesian sensitivity analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(4), 739–750. https://doi.org/10.1111/j.2517-6161.1996.tb02112.x.

Xiu D and Karniadakis GE (2002) The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing 24*(2), 619–644. https://doi.org/10.1137/s1064827501387826.

Zhang H, **Guo Q-Q and Xu L-Y** (2023) Prediction of long-term prestress loss for prestressed concrete cylinder structures using machine learning. *Engineering Structures 279*, 115577. https://doi.org/10.1016/j.engstruct.2022.115577.

Zhang C, **Chun Q and Lin Y** (2025) Optimal sensor placement and structural health monitoring methods of ancient stone bridges based on an improved genetic algorithm: Taking lugou bridge as an example. *Measurement 241*, 115680. https://doi.org/10.1016/j.measurement.2024.115680.

Zhu H, **Ibrahim JG**, **Cho H and Tang N** (2012) Bayesian case influence measures for statistical models with missing data. *Journal of Computational and Graphical Statistics 21*(1), 253–271. https://doi.org/10.1198/jcgs.2011.10139.

A. Clustering and pruning of walker ensembles

After the burn-in phase of an ensemble MCMC sampler and at thee end of the full chain, we performed a postprocessing step that clustered walkers according to their average log-posterior values and pruned walkers belonging to low-probability clusters. The objective was to suppress walkers stuck in local minima in probability wells while preserving the largest coherent mode of the posterior. Our implementation followed Hou et al. (2012), with some modifications to consider potential multiple low-probability clusters and to increase robustness against outliers.

Let the ensemble contain N walkers and let $\{\ln \pi_{i,t}\}_{t=1}^{T}$ denote the log-posterior distribution of walker i during the last $T = \lfloor \alpha T_{\text{burn}} \rfloor$ steps of burn-in, where $\alpha \in (0,1)$ is a user-specified fraction, which in this paper will be 0.2. Define the mean terminal log-posterior

$$\ell_i = \frac{1}{T}\sum_{t=1}^{T} \ln \pi_{i,t}, \quad i = 1,\ldots,N. \tag{59}$$

Sort these values in ascending order $\ell_{(1)} \le \ell_{(2)} \le \ldots \le \ell_{(N)}$, and denote the associated permutation of walker indices by σ. Define the first-order differences

$$d_k = \ell_{(k+1)} - \ell_{(k)}, \quad k = 1,\ldots,N-1, \tag{60}$$

and the median difference $\tilde{d} = \text{Median}\{d_k\}_{k=1}^{N-1}$. A jump was declared wherever

$$d_k > \gamma \tilde{d}, \tag{61}$$

where $\gamma > 0$ is a user-specified jump factor that we will choose as 5.0.

Let $J = \{k_1, k_2, \ldots, k_M\}$ be the ordered list of detected jump locations. These induce $M+1$ disjoint clusters

$$C_1 = \{1,\ldots,k_1\}, C_2 = \{k_1+1,\ldots,k_2\}, \ldots, C_{M+1} = \{k_M+1,\ldots,N\}. \tag{62}$$

For each cluster C_m with $m = 1,\ldots,M+1$, define its size $S_m = |C_m|$. We select the *largest* cluster index as

$$m^\star = \arg\max_m S_m, \tag{63}$$

and define the acceptance threshold $\ell_{\text{thr}} = \ell_{(\min C_{m^\star})}$. All walkers satisfying $\ell_i \ge \ell_{\text{thr}}$ were retained. Let K be the set of retained indices and P its complement.

For the next MCMC phase, let $\mathbf{x}_i^{\text{last}}$ denote the last state of walker i during burn-in. Construct an array $\mathbf{X}^{\text{new}} \in \mathbb{R}^{N \times d}$, where d is the parameter dimension, by

$$\mathbf{X}_i^{\text{new}} = \begin{cases} \mathbf{x}_i^{\text{last}}, & i \in K, \\ \text{resampled point}, & i \in P. \end{cases} \tag{64}$$

Resampling of pruned walkers was performed by convex mixing of two randomly drawn kept walkers

$$\mathbf{X}_i^{\text{new}} = w\mathbf{X}_a^{\text{new}} + (1-w)\mathbf{X}_b^{\text{new}}, \quad i \in P, \tag{65}$$

where $a, b \in K$ are drawn without replacement and $w \sim \text{Uniform}(0,1)$. This preserves the support of the dominant cluster while replenishing the ensemble with diversified points lying within the high-probability region.

At the end of the full chain, the same clustering formulation was reapplied to the final ensemble. In this second application, pruning was performed without resampling; we simply returned the subset of walkers satisfying $\ell_i \ge \ell_{\text{thr}}$ as the final reduced posterior ensemble.

B. Stable estimation of $D_\phi(S)$ measures from posterior samples

B.1. Estimation of global influence

Given a posterior sample $\{\boldsymbol{\theta}^{(i)}\}_{i=1}^{N} \sim \pi(\boldsymbol{\theta} \mid Y)$, we can estimate $D_\phi(S)$ from Equation 25 empirically as

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.

$$\widehat{D}_{\phi}(S) = \log\left(\frac{1}{N}\sum_{i=1}^{N}\pi_S\left(\boldsymbol{\theta}^{(i)}\right)^{-1}\right) + \frac{1}{N}\sum_{i=1}^{N}\log\pi_S\left(\boldsymbol{\theta}^{(i)}\right). \tag{66}$$

To prevent numerical underflow or overflow in evaluating likelihoods and ratios, all computations were performed in log-space. Let

$$\ell_i = \log\pi_S\left(\boldsymbol{\theta}^{(i)}\right) = \log\pi\left(Y\,|\,\boldsymbol{\theta}^{(i)}\right) - \log\pi\left(Y_{S^c}\,|\,\boldsymbol{\theta}^{(i)}\right). \tag{67}$$

Define $M = \max_i -\ell_i$ to stabilize exponentiation, and the scaled weights

$$\tilde{w}_i = \exp(-\ell_i - M) \quad \text{so that} \quad 0 < \tilde{w}_i \leq 1. \tag{68}$$

Then, the log-space equivalent of the first term in $D_{\phi}(S)$ was written using the log-sum-exp operation as

$$\log\mathbb{E}_{\boldsymbol{\theta}|Y}\left[\pi_S(\boldsymbol{\theta})^{-1}\right] \approx -\log N + M + \log\left(\sum_{i=1}^{N}\tilde{w}_i\right) = -\log N + \text{logsumexp}(-\ell_i). \tag{69}$$

Similarly, the second term was obtained as the sample mean of ℓ_i as

$$\mathbb{E}_{\boldsymbol{\theta}|Y}[\log\pi_S(\boldsymbol{\theta})] \approx \frac{1}{N}\sum_{i=1}^{N}\ell_i. \tag{70}$$

Finally, combining both terms gives the numerically stable estimator

$$\widehat{D}_{\phi}(S) = [-\log N + \text{logsumexp}(-\ell_i)] + \frac{1}{N}\sum_{i=1}^{N}\ell_i. \tag{71}$$

This log-space formulation ensured stable and accurate computation of $D_{\phi}(S)$ even when the likelihood ratios $\pi_S(\boldsymbol{\theta})$ span several orders of magnitude, a situation typical in high-dimensional or data-intensive Bayesian models.

B.2. Estimation of marginal influence with KDE approach

Analogously, we approximate the two expectations of Equation 28 by empirical averages using posterior samples $\left\{\theta_j^{(i)}\right\}_{i=1}^{N}$ as

$$\widehat{D}_{\phi}^{(j)}(S) = \log\left(\frac{1}{N}\sum_{k=1}^{N}\widehat{\pi}_S^{(j)}\left(\theta_j^{(k)}\right)^{-1}\right) + \frac{1}{N}\sum_{k=1}^{N}\log\left(\widehat{\pi}_S^{(j)}\left(\theta_j^{(k)}\right)\right). \tag{72}$$

Because $w_i = \pi_S\left(\boldsymbol{\theta}^{(i)}\right)$ may be very small or large, we compute the kernel weights in log-space as well for numerical stability. Let $\ell_i = \log\pi_S\left(\boldsymbol{\theta}^{(i)}\right)$, $M = \max_i -\ell_i$, and set scaled weights $\tilde{w}_i = \exp(-\ell_i - M)$. Then

$$\widehat{\pi}_S^{(j)}\left(\theta_j^{(k)}\right) = \frac{\sum_{i=1}^{N}\tilde{w}_i K_h\left(\theta_j^{(k)} - \theta_j^{(i)}\right)}{\sum_{i=1}^{N}K_h\left(\theta_j^{(k)} - \theta_j^{(i)}\right)} \tag{73}$$

and the corresponding log-smoothed value is

$$\log\widehat{\pi}_S^{(j)}\left(\theta_j^{(k)}\right) = \log\left(\sum_{i=1}^{N}\tilde{w}_i K_h\left(\theta_j^{(k)} - \theta_j^{(i)}\right)\right) - \log\left(\sum_{i=1}^{N}K_h\left(\theta_j^{(k)} - \theta_j^{(i)}\right)\right) + M. \tag{74}$$

B.3. Estimation of marginal influence with Fixed Mean approach

The marginal influence of Equation 31 was analogously approximated by samples $\theta_j^{(i)}$ as

$$\widehat{D}_{\phi}^{(j),(\text{fix})}(S) = \log\left(\frac{1}{N}\sum_{i=1}^{N}\pi_S\left(\theta_j^{(i)}, \bar{\theta}_{-j}\right)^{-1}\right) + \frac{1}{N}\sum_{i=1}^{N}\log\left(\pi_S\left(\theta_j^{(i)}, \bar{\theta}_{-j}\right)\right). \tag{75}$$

Cite this article: Andrés Arcones D, Paul A, Weiser M, Sanio D, Mark P and Unger JF (2026). Bayesian tendon breakage localization under model uncertainty using distributed fiber optic sensors. *Data-Centric Engineering*, 7, e46. doi:10.1017/dce.2026.10077

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 12:21:44, subject to the Cambridge Core terms of use.