

RESEARCH ARTICLE

Engineering Reports

Open Access

WILEY

Bayesian model calibration and damage detection for a digital twin of a bridge demonstrator

Thomas Titscher¹ | Thomas van Dijk² | Daniel Kadoke¹ |
Annika Robens-Radermacher¹ | Ralf Herrmann¹  | Jörg F. Unger¹ 

¹Safety of Structures, Bundesanstalt für Materialforschung und-prüfung, Berlin, Germany

²Structural Dynamics, TNO, Delft, Netherlands

Correspondence

Jörg F. Unger, Safety of Structures, Bundesanstalt für Materialforschung und-prüfung, Unter den Eichen 87, Berlin 12205, Germany.

Email: joerg.unger@bam.de

[Correction added on 17 May 2023, after first online publication: Ralf Herrmann's surname has been corrected from 'Ralf Herrman' to 'Ralf Herrmann' and ORCID has been added in this version.]

Abstract

Using digital twins for decision making is a very promising concept which combines simulation models with corresponding experimental sensor data in order to support maintenance decisions or to investigate the reliability. The quality of the prognosis strongly depends on both the data quality and the quality of the digital twin. The latter comprises both the modeling assumptions as well as the correct parameters of these models. This article discusses the challenges when applying this concept to real measurement data for a demonstrator bridge in the lab, including the data management, the iterative development of the simulation model as well as the identification/updating procedure using Bayesian inference with a potentially large number of parameters. The investigated scenarios include both the iterative identification of the structural model parameters as well as scenarios related to a damage identification. In addition, the article aims at providing all models and data in a reproducible way such that other researcher can use this setup to validate their methodologies.

KEYWORDS

damage detection, finite element analysis, load identification, model updating, parameter estimation, system identification, variational Bayesian statistics

1 | INTRODUCTION

Designing structures and components is often done using numerical methods due to reduced costs and time compared to experimental prototypes and better design optimization procedures. Once the design is complete and the structure or component has been manufactured, these complex numerical design models are often disregarded, even though there is often a relevant difference between the structure or component as-designed and as-built. Maintenance decisions and quality control that are driven by monitoring data such as the deflection of a bridge, the opening of a crack or wear due to abrasion are generally based on limit values calculated using the as-designed structure. The concept of Bayesian model updating allows to bridge the scale by combining monitoring data and simulation models to identify the as-built situation rather than the as-design numerical model, and then update the complex design models continuously during the life time of the structure. This has a huge potential for industrial applications.

A schematic overview of the general approach is shown in Figure 1 and illustrated in a video*. The real structure is equipped with a monitoring system that continuously measures the current state of the structure such as displacements,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Engineering Reports* published by John Wiley & Sons Ltd.

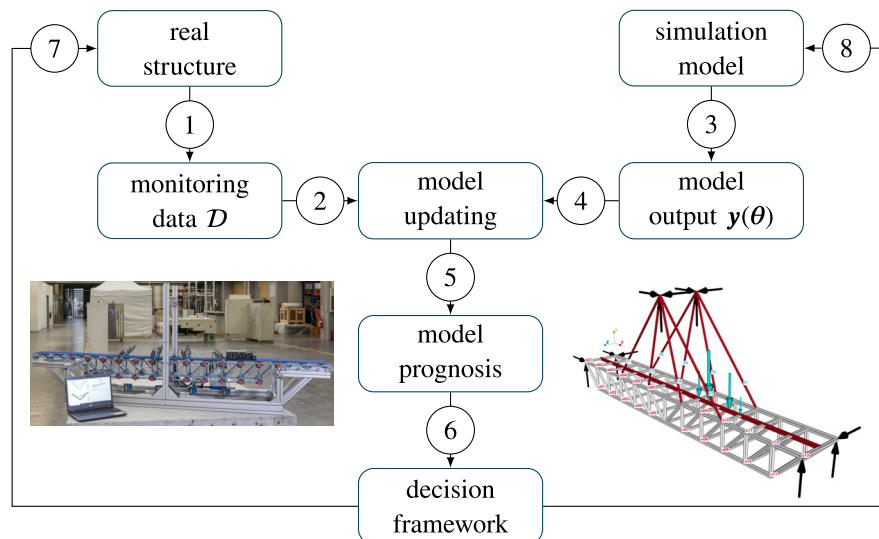


FIGURE 1 Concept of a digital twin.

strains or displacements at various positions (1). This data is fed into a data management system and used in the model updating (2) to improve the knowledge of the simulation model (3). The latter is often a parameterized finite element model. Its parameters θ are often constitutive parameters such as the effective Young's modulus of the different parts, but could also include the loading, boundary conditions or the geometry. Based on these parameters, the simulation model can predict the model response $y(\theta)$ that is corresponding to the measurement data D . In a deterministic setting, an objective function (e.g., the norm of the difference between measurement data D and the model prediction $y(\theta)$) can be used to apply an optimization algorithm that estimates the set of parameters θ that best explains the measurement data. Once the finite element model is updated, it can be used for predicting (5) key performance indicators *KPI*, for example, the stress at a relevant position, a deformation that has not been measured, or even a prediction of the remaining useful life. Based on the *KPIs*, a decision framework (6) is then installed which changes both the real structure (7) and the digital twin (8), for example, by reducing the maximum vehicle weight, or performing maintenance or repair. As such, this digital twin concept characterizes a closed cycle from the real structure to the simulation model (via model updating) and from the simulation model to the real structure via the decision framework.

The main advantage of the concept is that the measurement data is enriched with the simulation models to provide additional information such as the prediction of the reliability and the remaining useful life, inserting virtual sensors by extracting information from the numerical model at positions that are not directly measured (e.g., inaccessible or not directly measurable such as fatigue damage), identifying gradual changes (e.g., damage, creep) or using the numerical model to investigate different maintenance procedures.

In this approach, monitoring data is used to continuously update the numerical model. Furthermore, the incorporation of measurement uncertainties in the model updating using a Bayesian model updating procedure is demonstrated, with a specific variational approach that is computationally efficient and thus requires significantly less computational effort compared to traditional sampling based schemes. The implementation of such a digital twin poses many challenges and the goal of this work is to share the authors experiences when working on a lab-demonstrator bridge. The experimental data and the code are published as complementary material to encourage others to test their methods.

The first of those challenges is the necessity of an appropriate data management system. Based on the requirement (e.g., real-time access to the data, preprocessing of data to avoid transferring and storing large amounts of data), different options are possible. In particular when combining data sets acquired from different measurement systems, as demonstrated in this work, challenges related to the synchronization and a definition of metadata schema to correctly identify the different setups arise. Additionally, when working with real-world data, one has to deal with outliers and potentially malfunctioning sensors.¹ Another practical issue is the sensor calibration, including multiplication factors and offsets related to the initial value. Ideally, those are analyzed before the experiment, but in some cases (as illustrated in this article) this is not possible and those parameters might then be added as model parameters to be inferred. Another challenge is related to the development of the numerical model. In many cases, the design model can be used as an initial guess,

but an iterative procedure to identify the important physical phenomena that explain differences between the simulation model and the measurement data is often required.^{2,3} Finally, model updating procedures are required to calibrate these parametric models. In this context, it is of utmost importance to estimate both the quality of the models, the quality of the parameter estimates and in addition propagate all uncertainties from the data to the final model predictions that are used for decision making.^{4,5}

In addition to identifying the “initial” reference configuration, a digital twin can also be used to identify changes over the lifetime. This concept is common in structural health monitoring (e.g., References 6–8), where the evolution of latent parameters for example, related to a reduced stiffness is identified, often using dynamic response quantities such as eigenfrequencies or eigenmodes.^{9,10} Specific challenges are encountered, when the simulation model does not fully represent the physics of the real setup. This is often a relevant problem in civil structures, since the set of modeling assumptions (rigid support, no rotation, constant temperatures, constant cross sectional properties) is often only a coarse approximation of the reality. Some methods to cope with this problem are for example, the stochastic finite element method,¹¹ where an unknown source term is added to the right hand side of the partial differential equation (PDE), or the approach by Kennedy and O’Hagan,¹² where the model response is assumed to be a superposition of the original physics based model and an additional Gaussian process that describes the model bias. Specific applications of Bayesian model updating and damage identification for bridge monitoring problems are presented in References 13 and 14.

An appealing method to identify model parameters from experimental data are based on a Bayesian methodology¹⁵ that allows to select from a set of parameterized simulation models those that best represent the data. The ill-posedness of this problem (in particular related to different parameter combinations explaining all the measured data) is circumvented by providing a joint probability distribution of the parameter estimates.

For illustration purposes, the demonstrator has been modified for the Hannover fair 2021 in Germany as illustrated in a video[†]. The goal was to show, how a simulation model can be used to extract further information that have not been directly measured. In this setup, displacements were measured in vertical directions at six different positions as well as the forces in all the eight steel cables. Based on these measurements, the position and the weight of the car are continuously updated and displayed. In the video, it can be observed that the identified position of the virtual car is almost coinciding with the position of the real car, thus showing the applicability of the method to infer virtual sensor data (in this case the position of the car). Based on this information, maintenance measures, for example, automatically adjusting the weight restrictions can be introduced.

One of the challenges when building a digital twin is the model updating of the simulation model based on the measurement data. This is an extension of the digital twin concept where the model inputs (e.g., the loads) are measured and then used to compute additional outputs (e.g., for a fatigue analysis using the stresses based on the real loading scenario rather than the design loads). However, we have realized that in particular for civil structures the modeling assumptions from the design are often simplified for design purposes and conservative and thus do not accurately reflect the structural behavior in a real scenario. Thus, model updating using Bayesian inference techniques plays a central role in this digital twin concept and is the focus of the current article.

At first, the experimental setup is described, including procedures related to the data acquisition and preprocessing as well as the parameterized simulation model with the different scenarios (single beam, undamaged bridge, damaged bridge) that were investigated. Afterwards, a short overview about the variational inference procedure is given with a discussion on the validation of the modeling assumptions.

Finally, numerical results for the different scenarios are discussed. The focus is in particular on highlighting the challenges that are related to applying these methods to problems with real experimental data.

2 | EXPERIMENTAL SETUP

One of the goals in this research is the identification of material parameters that are the basis for the structural computation. In order to verify, if these parameters are really material properties and thus independent from the structure, different scenarios with varying complexity where sequentially investigated.

2.1 | Single beam in uniaxial loading

The main bridge structure is constructed using PASCO’s Structure System,¹⁶ which is developed for design based experiments. The demonstrator bridge is primarily built from the (blue) beams made of Acrylonitrile butadiene styrene

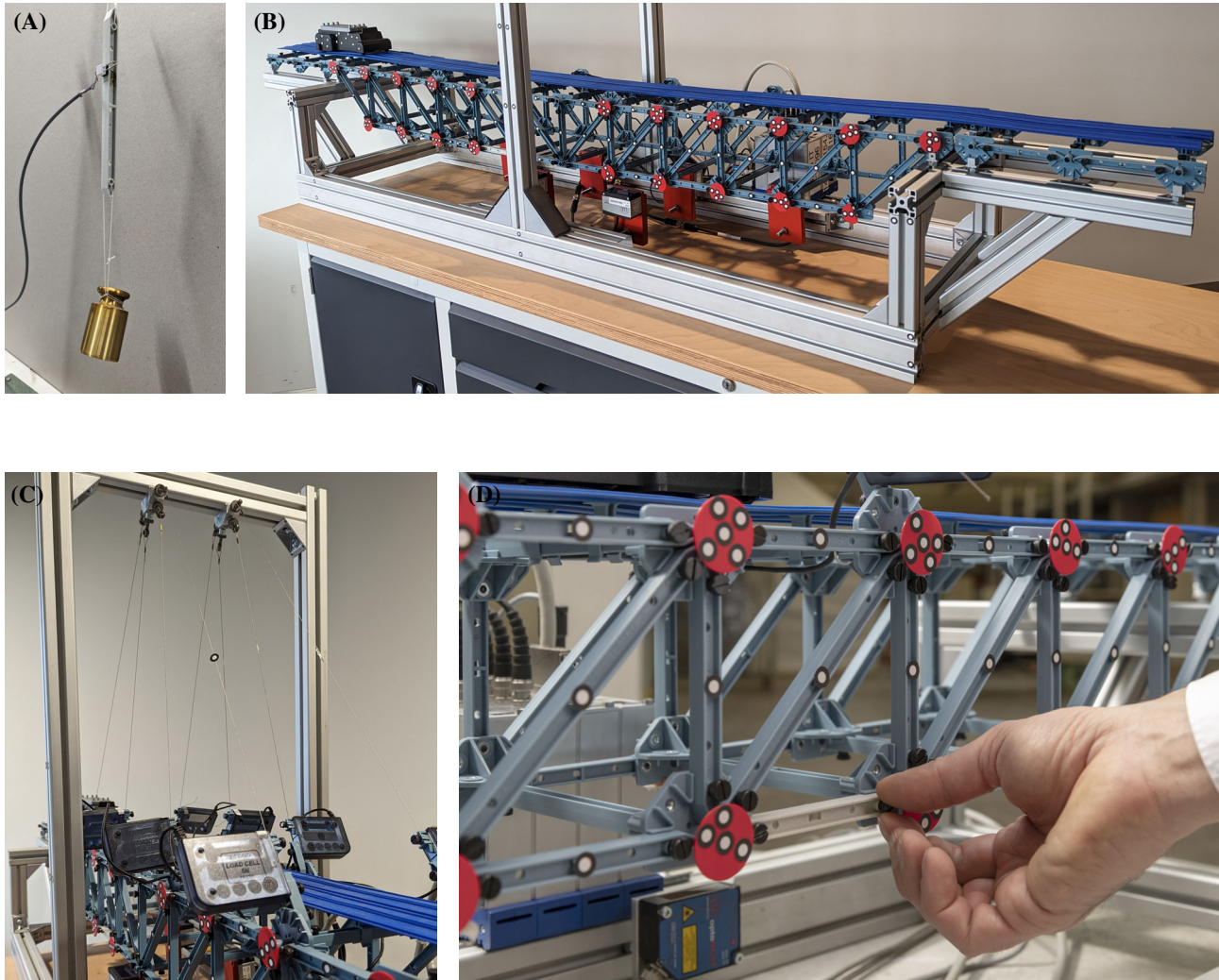


FIGURE 2 Overview of the different scenarios with increasing levels of complexity. (A) Single beam, (B) simply supported bridge, (C) cable-stayed bridge, (D) damage identification.

(ABS plastic) and a Young's modulus of $E_{\text{PASCO}} = 2.3 \text{ GPa}$.¹⁷ At first, a single beam element (essentially a truss) was loaded under tension as illustrated in Figure 2A. In this setup, the load was gradually increased from 224 to 1224 g in steps of 200 g. The resulting strains in the beam were measured, for each load case, using a strain gauge in a quarter bridge setup. The measurement time for each load case is approximately 30 s with a sampling frequency of 100 Hz.

Due to the simplicity of the setup, a finite element model is omitted and the linear, analytic model

$$\epsilon(n_w, E_{\text{beam}}^{\text{1D}}) = \frac{F(n_w)}{A} \frac{1}{E_{\text{beam}}^{\text{1D}}} \quad (1)$$

is used instead, where $F(n_w)$ is the gravitational force caused by the weights, $A = 51.74 \text{ mm}^2$ ¹⁸ is the beam cross section and $E_{\text{beam}}^{\text{1D}}$ is the Young's modulus.

2.2 | Simply supported bridge

An overview of the experimental setup is illustrated in Figure 2B with the geometrical dimensions sketched in Figure 3. The setup comprises a small scale bridge (1.5 m span) featuring abutments, pillars, a deck, a moving load, displacement sensors, a stereo photogrammetry system and a data acquisition system. The features are described in the following sections.

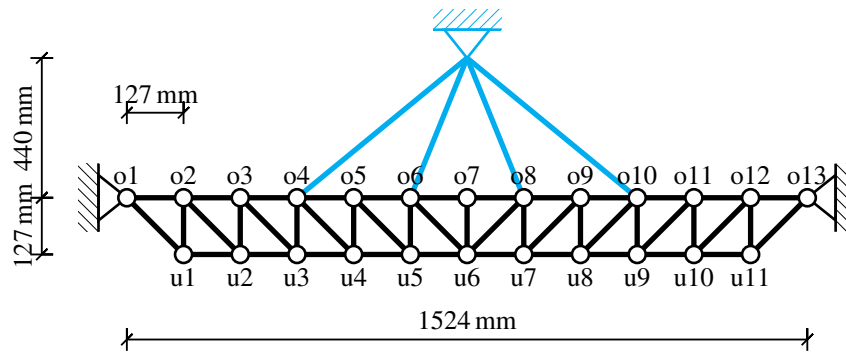


FIGURE 3 2D-Geometry of the cable-stayed demonstrator bridge. For the simply supported bridge, the cables and pillars (marked in blue) are omitted. The structure extends in 3D with a width of 194 mm.



FIGURE 4 PASCO connectors and connection. (A) Beam connection, (B) road connection clip, (C) self driving car for quasi-static loading.

The **support system** is made from extruded aluminum profiles. Besides abutments and pillars, it also provides support and space for the road to extend beyond the bridge for the miniature car to enter and leave the demonstrator bridge.

The main structure of the bridge is a Pratt deck truss bridge structure comprising of twelve segments in length direction, which are made from H-beams and assembled with connector pieces (see Figure 4A). No cross-braces in transverse direction are applied.

A PASCO road bed is connected to the bridge using clips, as shown in Figure 4B. The road bed is flexible compared with the structure, therefore it is not expected to add load bearing capabilities to the structure.

The bridge is **loaded** using a self-driving car with adjustable weight and velocity, see Figure 4C. The mass of the car is 400 g which can be increased by adding one or two masses of 250 g each. The velocity can be adjusted continuously between approximately 8 cm s^{-1} to 25 cm s^{-1} .

2.2.1 | Sensors

Laser displacement measurement

The laser displacement measurements are conducted at the six locations indicated in Figure 5 directly below the 3rd, 5th, 6th and 8th node in the front bottom and in the back only at the 5th and 6th node to analyze the symmetry. The sensors measure the vertical displacement. OptoNCDT ILD 1401-50 sensors from Micro Epsilon are used for this purpose, these sensors have a measuring range of 50 mm and a resolution of $5 \mu\text{m}$ in perfect static conditions and $25 \mu\text{m}$ for dynamic measurements at 1 kHz.

DIC displacement measurement

Displacements at numerous locations on the bridge and support structure are measured optically using a stereo camera setup with 5 or 10 Hz. Each beam at the front of the structure is equipped with high contrast markers to measure its deformation and curvature. The nodes on the front side of the structure are equipped with red discs with four markers,

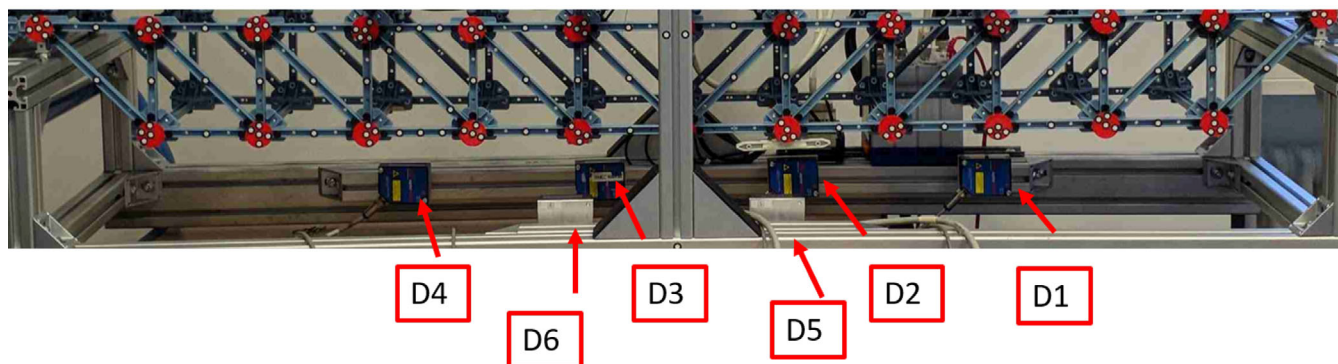


FIGURE 5 Laser sensor positions.

as shown in Figure 4C. This allows to compute the node rotation. In this work, however, only the vertical displacements at the node centers (all sensors with suffix _01 in Figure C1) are analyzed. Additional measurement points are attached to the support structure to validate the assumptions related to the boundary conditions of the model and to the car to measure its position while traversing the bridge.

2.2.2 | Data acquisition

The data of the laser sensors is acquired using a Gantner Q.bloxx voltage measurement module. The data acquisition system (DAQ) runs an Open Platform Communications Unified Architecture (OPC UA¹⁹) server from which the data is almost real-time extracted via the Ethernet using a laptop running an OPC UA client. The advantage of this kind of unified interface is that several user or services can access the measurement data independently and simultaneously directly from the DAQ. The small scale bridge can be monitored from any location provided there is an internet connection. The additional DIC displacement measurements are acquired via a separate stereo photogrammetry system ARAMIS 12M from Zeiss. Two 12 megapixel cameras placed in about 1.88 m distance from the bridge record the measurement points during the experiments. The displacements at each measurement point is then triangulated from those two camera images in a post-processing step using the GOM ARAMIS software version 2019 and 2020.

As illustrated in Figure 6, each measured data set is annotated with metadata including the experimental setup, time stamps, measurement frequencies and signal units. The first benefit is to formulate human-readable queries like “Load all displacement sensor data of the cable-stayed bridge.” or more specific for the other scenarios “Load data from laser sensor D4 and stereo sensors u07 and o04 where the bridge had no cables and damage in segment 7.” The second benefit is to automatically select the suitable finite element model related to the relevant scenario, which is in particular helpful when working exploratively with frequent changes in the used data and/or models.

2.2.3 | Data pre-processing

The data processing consists of four main steps that are illustrated in Figure 7. First, the signals of the two measurement systems are *synchronized*. The time shift Δt between both systems is determined by maximizing the signal correlation at points D5 and D6 (see Figures 5 and C1) which are measured by both the laser sensor and the stereo system. Therefore, the time stamps $t_{\text{stereo},i}$ and $t_{\text{laser},j} + \Delta t$ are joined into the overlapping interval t_k and each displacement signal $u_{\text{stereo},i}$ and $u_{\text{laser},j}$ is interpolated onto t_k . The Pearson correlation coefficient between the resulting signals $u_{\text{stereo},k}$ and $u_{\text{laser},k}$ is used to indicate the quality of the match. Powell’s conjugate direction method is then used to find the Δt that maximizes this correlation coefficient. For most car passes, the resulting coefficient of correlation is well above 0.99. Exceptional cases with < 0.9 were detected and analyzed. Establishing an automatic outlier analysis that removes points that exceed a distance of two standard deviations from the mean fixed those cases.

Next, a stereo measurement point on the car’s front axle (*car signal* in Figure 7) is used to determine the time period $[t_0, t_1]$, where the car is on the bridge. Data obtained outside of this interval is *cut off*.

The data is sampled with 5 Hz–100 Hz and clearly correlated. Treating it as individual, uncorrelated measurements would highly overestimate the posterior parameter precision. Thus, in the third step of the data processing, a 4th-order

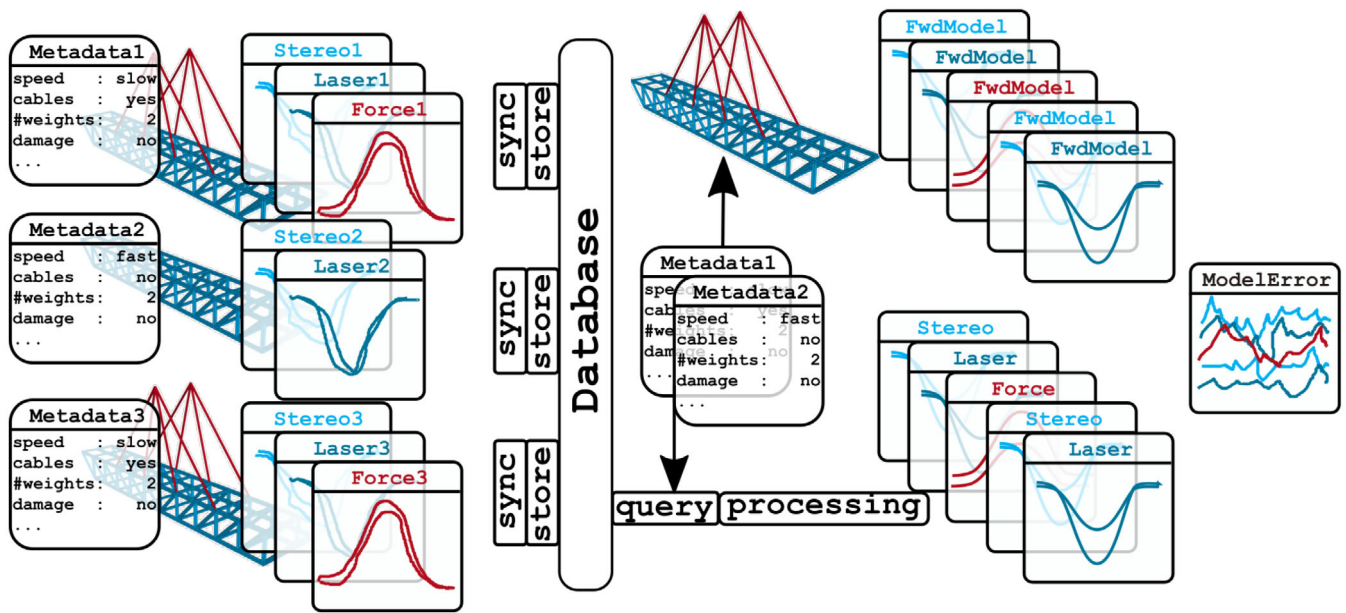


FIGURE 6 Annotation of the measured data with metadata information to easily query data sets and automatically derive the numerical model. The approach is generalized for all bridge scenarios and includes laser, DIC and force measurements.

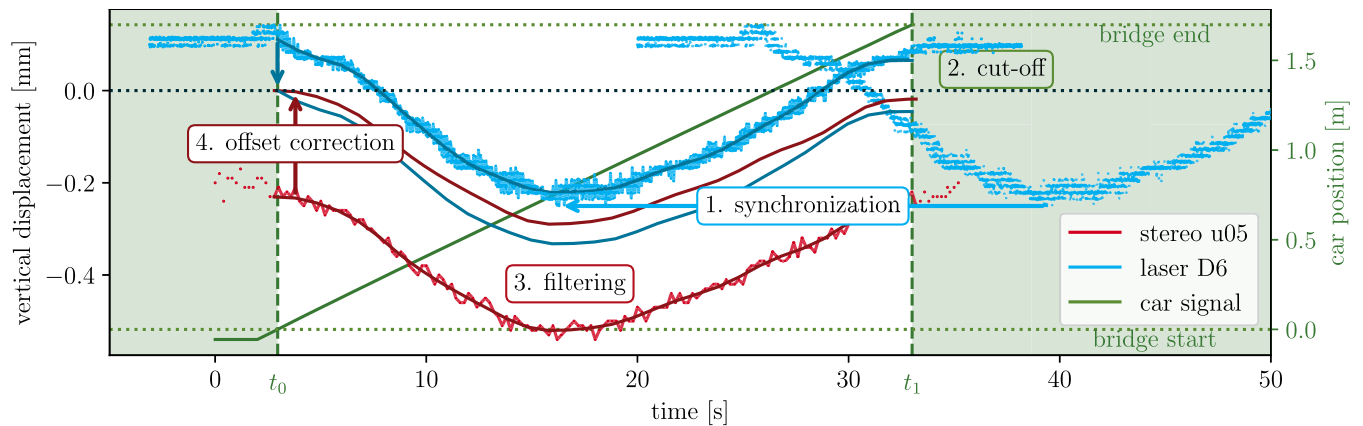


FIGURE 7 Illustration of the synchronization process between the laser, force measurements and the DIC measurements via the cross correlation.

Butterworth lowpass H is used to *filter* the raw data to a frequency of $f = 1$ Hz. Then, the filtered data is resampled such that only one data point every $1/f$ time points remains, which roughly corresponds to a temporal correlation length of 1 s.

The last processing step is an *offset correction*. Following the model assumption that the bridge is unloaded right before the front axle enters the bridge, the value of each sensor right before t_0 is used as its offset. However, the aluminum frame is not perfectly stiff and the car approaching the bridge before t_0 does indeed influence the sensor signal. Another indicator that this assumption is not fully correct is the remaining deviation from zero at the end time t_1 . To further analyze this uncertain zero-state of the bridge displacement measured by the laser sensors and the stereo system, additional offset parameters θ_0 can be introduced, which are inferred simultaneously with the other model parameters.

Finally, the processed data d_i for each sensor i is computed as

$$d_i(t, \theta_{i,0}) = \underbrace{H(d_{i,\text{raw}}(t \overset{\text{synchronization}}{+ \Delta t}), f)}_{\text{filter}} - \underbrace{d_{i,\text{raw}}(t_0 + \Delta t) + \theta_{i,0}}_{\text{zero-state offset}} \quad \text{with } \underbrace{t \in [t_0, t_1]}_{\text{cut-off}} \quad (2)$$

Note that in our approach only a single offset per sensor was used, which results in potentially nonzero and in particular not identical sensor values before the car enters the bridge and once it has left the bridge, which is obviously a violation of the physical modeling assumptions. However, other approaches, for example, adding an offset to both configurations and then interpolating via the car position requires additional assumptions and almost doubles the number of unknowns and thus might not give more accurate results.

2.2.4 | Finite element model

A two-dimensional quasi-static finite element (FE) model with small-strain-assumptions is used to model the demonstrator bridge with the geometry given in Figure 3. The main bridge structure is modeled either with Euler-Bernoulli beams with the local element stiffness given in Appendix B. Their cross section, section modulus and torsion modulus are computed using the manufacturers information.¹⁸ The length of the axis-aligned beams are computed as the mean from *all* horizontal and vertical node distances from *all* stereo data sets and the diagonal length is derived accordingly. This is due to the fact that the initial geometry is already deformed due to the dead load (thus smaller) and the length of a segment is a superposition of the length of a single beam plus the connection. The material's shear modulus is set to 1 kPa, which is numerically small compared to the manufacturer's Young's modulus estimate of $E_{\text{beam}} = 2.3$ GPa, to virtually eliminate any torsion resistance.

One evaluation $\text{FE}(E_{\text{beam}}, \mathbf{x}_{\text{car}}, \mathbf{sensors})$ is performed as follows. A vector of external forces $\mathbf{F}_i$ is assembled for each car position in $\mathbf{x}_{\text{car},i}$. The sparse stiffness matrix $\mathbf{K}$ is build and homogeneous Dirichlet boundary conditions are applied. The system $\mathbf{K}\mathbf{u}_i = \mathbf{F}_i$ is then solved for each force vector resulting in a displacement solution field $\mathbf{u}_i$ for each car position $\mathbf{x}_{\text{car},i}$. These solution fields are passed to virtual sensors to extract their virtual measurements: In the case of the displacement measurements of laser and stereo sensors, they simply extract the vertical nodal value at their respective position.

It is important to note that a LU decomposition of $\mathbf{K}$ is used to efficiently solve the system and that this decomposition is only rebuilt if any of the Young's moduli $E_{\text{beam/cable}}$ changes. This makes subsequent evaluations of the same Young's moduli parameters for different $\mathbf{x}_{\text{car}}$ or different sensors very efficient.

2.3 | Cable-stayed bridge

In the third scenario, the eight steel cables connecting the pillars with the deck were added to support the structure as illustrated in Figure 2C. Apart from these cables and the corresponding force sensors in the cables, the setup was completely identical to Section 2.2 including the laser and DIC displacement measurements, the data acquisition system, the data pre-processing and the FEM model. As a consequence, only the additional features are discussed.

Eight steel cables (diameter 0.54 mm and stiffness 200 GPa) connected the pillar with the deck, making it a cable stayed Pratt deck truss bridge. The connection for the eight cables is adjustable and mounted between the two pillars as illustrated in Figure 8A. This connection can be used to fine-tune the tension of cables by fastening or loosening threaded eyes.

2.3.1 | Sensors

Cable load measurements

The cable tension is measured in the eight cables using 5 N load cells from PASCO, the PS-2201 and acquired using a Gantner Q.bloxx strain gauge measurement module. These load cells have a range of 5 N with a resolution of 0.001 N and an accuracy of 1%. These sensors are placed in line with the load and they measure the cable load using strain gauges connected to a shear beam load cell.

The sensor positions of the force sensors are illustrated in Figure 8B. As the laser sensor data and the force sensor data are both acquired with the Gantner module, the offset correction described in Section 2.2.3 is also used for the force data (using the maximum correlation in the displacement sensor data).

2.3.2 | Finite element model

In this scenario, the finite element model was extended to 3D as illustrated in Figure 9. In this scenario, only beam elements could be used, because in the 3D-scenario the system would otherwise be statically underdetermined (due to the missing diagonals in the depth-direction).

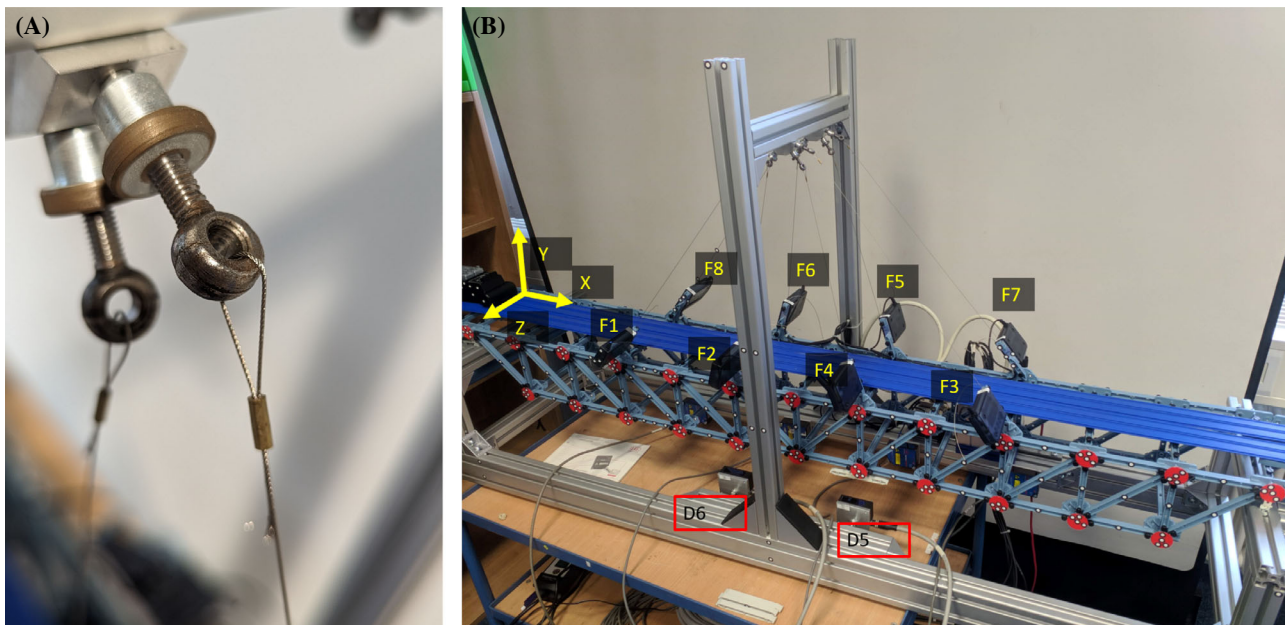


FIGURE 8 Cables and force sensors for the scenario cable-stayed bridge. (A) Adjustable cable connection at the top, (B) Force sensor positions.

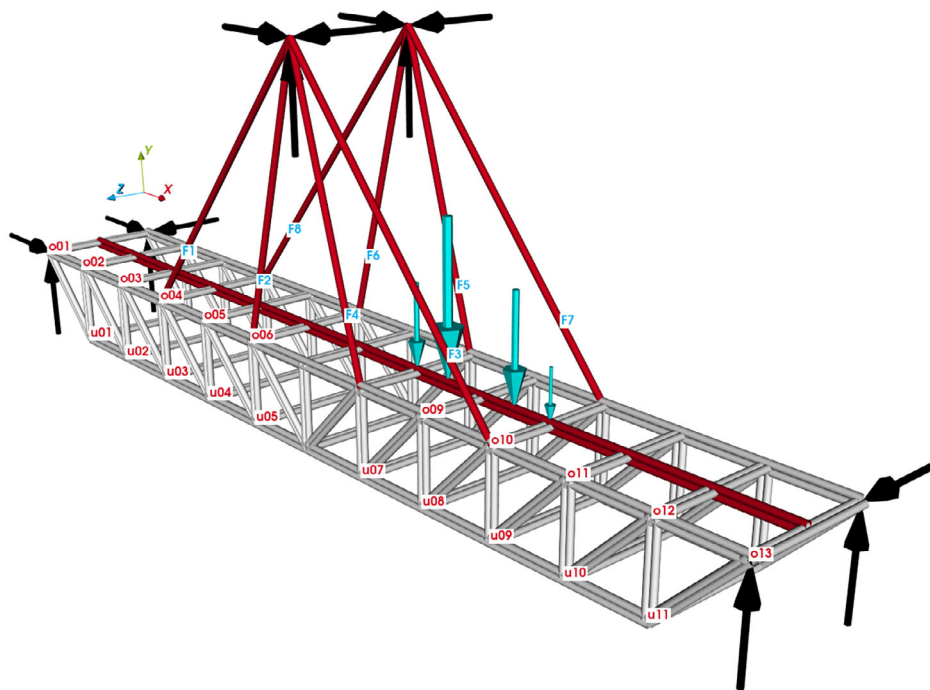


FIGURE 9 Finite element model of the bridge including the nodal forces related to the car and the geometrical dimensions.

As described in Section 2.2, the car drives on a very flexible road bed, also referred to as deck. It is also modeled by beam elements and highlighted red in Figure 9. They only serve as a point of load application and their low stiffness of $E_{\text{deck}} = 1 \text{ kPa}$ ensures that they do not add additional stiffness to the structure.

The teal arrows in Figure 9 represent the nodal forces generated by the car with its front axle at one specific position x_{car} . Its weight $w_{\text{car}} = 1.071 \text{ kg}$ is split 40 : 60 between the front and the rear axle, each $l_{\text{car}} = 170 \text{ mm}$ apart. Both forces are then, again, distributed to the two nodes of their respective finite element according to the shape functions.

As shown in Figure 8, the connection from the bridge to the tower consists of steel cables and the cable force sensors. This comes with two issues. First, the cable force sensors are not perfectly in line with the cables which causes a

multiplicative error in their measurements. This is taken into account by treating these factors – for each sensor separately – as additional unknown model parameters $F_1 \dots F_8$. Second, due to the internal measurement mechanism, the cable force sensors have a very high, unknown compliance. In analogy to two springs in series, where the individual compliances add up, the sensor compliance is assumed to dominate the overall connection. Thus, the finite element model treats the whole connection as a single truss element with the constitutive relationship

$$\sigma_{\text{cable}} = \frac{F_{\text{cable}}}{A_{\text{cable}}}, \quad F_{\text{cable}} = E_{\text{cable}} \frac{u_1 - u_0}{L_{\text{sensor}}}, \quad (3)$$

where the Young's modulus E_{cable} is a single (unknown) parameter for all cables and the (constant) sensor length $L_{\text{sensor}} = 50$ mm replaces the actual cable length.

The force sensor values F_{cable} in the simulation model are computed from the associated truss elements using Equation (3).

2.4 | Damage identification

In the final scenario, some of the beams in the bottom chord were replaced by a softer material representing damage scenarios. The goal was to identify, locate and quantify the amount of stiffness reduction purely based on the measurements. As shown in Figure 2D, there are two materials from which beams are made. The standard stiffer blue beams are made of Acrylonitrile butadiene styrene (ABS plastic) and a Young's modulus of $E_{\text{PASCO}} = 2.3$ GPa.¹⁷ The more compliant gray beams are made of 50D Durometer thermoplastic rubber with no further material properties provided by the manufacturer. For testing purposes, the beams 6, 7, and 8 on the front layer between (u6, u7), (u7, u8), and (u8, u9) according to Figure 3 are replaced with the softer beams.

3 | BAYESIAN MODEL UPDATING

In a Bayesian version of the model updating illustrated in Figure 10, the parameters θ of a simulation model (3) are no longer assumed to be deterministic, but characterized by a probability distribution. Prior engineering knowledge of these parameters (before looking at the data) can be incorporated (4), for example, that the Young's modulus is in the order of 40 MPa. The likelihood function (5) characterizes the conditional probability that the data D has been generated by the simulation model with parameters θ — similar to the objective function in a deterministic optimization — with additionally taking into account the measurement accuracy. Based on the Bayesian model updating with the monitoring data D from the real structure (1,2), a posterior parameter distribution $p(\theta|D)$ (6) is obtained, which characterizes

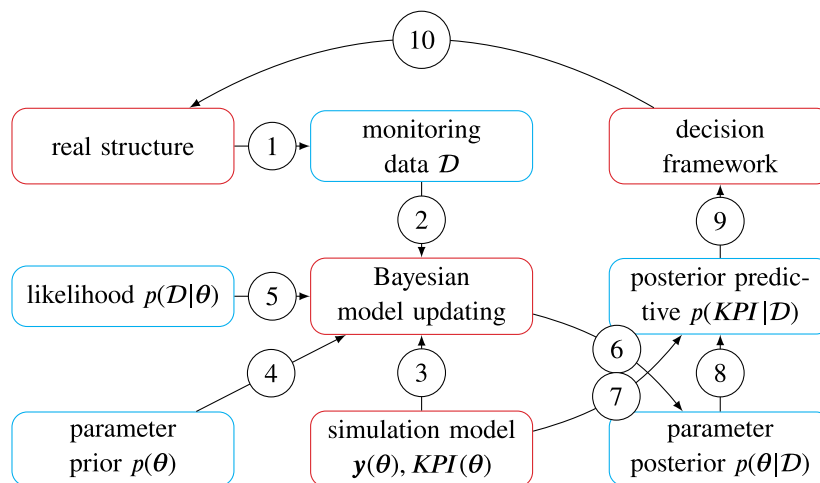


FIGURE 10 Concept of a Bayes digital twin.

the knowledge about these parameters after the model updating (taking into account the data). Based on this posterior parameter distribution (8), the simulation model can be used to predict a posterior predictive distribution of the key performance indicators (7). A decision framework is then installed to change the real structure (10). Note that the simulation model predicts probability distributions of these key performance indicators, such that the decision framework has to be adapted accordingly.

3.1 | Introduction

Inferring the unknown parameters θ of a model $\mathcal{M}$ is, in general, an inverse problem that involves finding a parameter set that minimizes the distance of the model response $\mathbf{g}(\theta)$ to some measured data $\mathbf{y}$. A common deterministic approach in solving these problems is an optimization algorithm that minimizes the least-square distance between model and data²⁰ which is successfully used to infer parameters of structural finite element (FE) models.²¹ Following the aphorism by George Box “*All models are wrong, but some are useful*”, a practical model can hardly fully explain the given data and some uncertainty remains. In addition, the results of deterministic analyses carry no information about the probability or reliability of the inferred parameter.

Probabilistic inference methods, on the other hand, aim to include and quantify uncertainties in their results, by modeling the uncertainties with additional variables Φ , extending the parameter vector to $\Theta = [\theta, \Phi]$. Given a prior probability density function (pdf) that reflects the prior knowledge on the parameters Θ and the likelihood that the measurements $\mathbf{y}$ are observed for given parameters Θ , Bayes’ theorem is used to compute a posterior pdf

$$\underbrace{P(\Theta|\mathbf{y}, \mathcal{M})}_{\text{posterior}} = \frac{\overbrace{P(\mathbf{y}|\Theta, \mathcal{M})}^{\text{likelihood}} \overbrace{P(\Theta|\mathcal{M})}^{\text{prior}}}{\underbrace{\int P(\mathbf{y}|\Theta, \mathcal{M})P(\Theta|\mathcal{M})d\Theta}_{\text{evidence } P(\mathbf{y}|\mathcal{M})}} \quad (4)$$

$$= \frac{P(\mathbf{y}, \Theta, \mathcal{M})}{P(\mathbf{y}|\mathcal{M})} \quad (5)$$

that reflects the updated knowledge about the parameters after observing the data.²² The normalization term in the denominator of Equation(4) describes the evidence for the data $\mathbf{y}$ considering the model $\mathcal{M}$.

Evaluating Equation(4) directly can only be done for a very limited number of parameters and thus solving practical problems requires approximation methods. Those methods fall into two main categories: sampling-based and variational Bayesian inference methods. The first category includes Markov Chain Monte Carlo (MCMC) methods that are reviewed in Reference 23. They are powerful tools for approximate Bayesian Inference, as they are asymptotically unbiased and require no assumptions on the form of the posterior. However, it is difficult for MCMC methods to estimate the evidence, as the samples are only proportional to the posterior. In contrast, (dynamic) nested sampling methods^{24,25} are used in this work, specifically the python implementation `dynesty`,²⁶ and aim at approximating the evidence, which is the basis of model comparison and model selection. The disadvantage of a very large ($\gg 1000$) number of required samples (thus model evaluations) that also scales with the number of parameters, is shared by both types of methods. In the context of computational mechanics, this is often prohibitively expensive due to the numerically expensive model evaluations.²⁷ In particular for real-time applications and digital twins, numerically efficient surrogates models like Gaussian processes²⁸ or reduced order models²⁹ are often required to employ sampling-based methods.

3.2 | Variational inference

In the category of variational inference methods, the true posterior $P(\Theta|\mathbf{y})$ is approximated by an analytic function $q(\Theta)$ from a selected family of stochastic distributions. Note that the model $\mathcal{M}$ from Equation(4) is dropped for convenience assuming all pdfs are conditioned on the same model. The parameters of those distributions are found by minimizing the Kullback-Leibler divergence as a similarity measure:

$$KL(q||P) = \int q(\Theta) \log \frac{q(\Theta)}{P(\Theta|\mathbf{y})} d\Theta. \quad (6)$$

Several of such algorithms are reviewed in Reference 30 and the one used in this work closely follows the variational Bayes (VB) method presented in Reference 31, which uses the linearization of the model, the mean field approximation and the conjugate-exponential restriction. This transforms the probabilistic parameter estimation into an optimization problem that requires very few iterations (= model evaluations) to converge. A short overview is given here, and a complete derivation is provided in Reference 32, since slight modifications related to the original article where necessary.

Bayes' theorem in Equation(5) is used to replace the unknown, true posterior in Equation(6) to obtain

$$KL(q||P) + \underbrace{\int q(\Theta) \log \frac{P(\mathbf{y}, \Theta)}{q(\Theta)} d\Theta}_{\text{variational Bayes free energy } F} = \underbrace{\log P(\mathbf{y})}_{\text{log evidence}} \quad (7)$$

Since the log evidence is independent of the parameters Θ , minimizing KL is equivalent to maximizing the free energy F , which is then the objective of the optimization.

The deviation of a model response $\mathbf{g}(\theta)$ to measured data $\mathbf{y}$ is defined as the model error $\mathbf{k}$

$$\mathbf{k}(\theta) = \mathbf{y} - \mathbf{g}(\theta) = \mathbf{e}(\Phi) \text{ with } \mathbf{e} \sim \mathcal{N}(0, \Phi^{-1}\mathbf{I}), \quad (8)$$

where this deviation is specifically modeled as an uncorrelated, zero-mean Gaussian noise $\mathbf{e}$ with the unknown precision Φ . To make the free energy in Equation(7) tractable for the optimization, additional assumptions and simplifications are made that are briefly discussed here.

First, the possibly nonlinear model is approximated by its first-order Taylor expansion

$$\mathbf{g}(\theta) \approx \mathbf{g}(\mathbf{m}) + \mathbf{J}(\mathbf{m})(\theta - \mathbf{m}) \quad (9)$$

with the Jacobian matrix $\mathbf{J} = \partial \mathbf{g} / \partial \theta$ evaluated at the expansion point $\mathbf{m}$, which is the current estimate for the mean of the posterior distribution of θ .

Second, the mean field approximation for $q(\Theta)$ – recall that $\Theta = [\theta, \Phi]$ contains both model and noise parameters – states that the joint posterior distribution can be factorized into independent groups

$$q(\Theta) = \prod_j q_j(\Theta_j), \quad (10)$$

where each group has its own approximate posterior distribution q_j . Parameters of separate groups are independent, but can be correlated within their group. A logical grouping in this work is the separation of model parameters θ and the independent noise parameters Φ_i . A more general definition of the model error in Equation(8) reads

$$\mathbf{k}(\theta) = \begin{bmatrix} \mathbf{k}_0(\theta) \\ \mathbf{k}_1(\theta) \\ \vdots \\ \mathbf{k}_n(\theta) \end{bmatrix} \sim \begin{bmatrix} \mathcal{N}(0, \Phi_0^{-1}\mathbf{I}) \\ \mathcal{N}(0, \Phi_1^{-1}\mathbf{I}) \\ \vdots \\ \mathcal{N}(0, \Phi_n^{-1}\mathbf{I}) \end{bmatrix} \quad (11)$$

that is able to model n noise groups, potentially coming from n independent sensors or measurement systems. Thus, instead of a single noise parameter Φ , now the parameter Φ_i describes the noise of the i -th noise group with its model error contribution $\mathbf{k}_i$.

Last, the conjugate-exponential restriction imposes the use of priors that are conjugate to the likelihood and come from the exponential family. Here, the posterior for the model parameters is defined as a multivariate normal $q(\theta) \sim \mathcal{N}(\mathbf{m}, \Lambda^{-1})$ with the means $\mathbf{m}$ and the precision matrix Λ . The posterior for the i -th noise parameter is defined as a Gamma distribution $q(\Phi_i) \sim \text{Ga}(s_i, c_i)$ with the shape parameter c_i and the scale parameter s_i . For conjugacy, the corresponding prior distributions with subscript $_0$ read $P(\theta) \sim \mathcal{N}(\mathbf{m}_0, \Lambda_0^{-1})$ and $P(\Phi_i) \sim \text{Ga}(s_{i,0}, c_{i,0})$.

These assumptions and restrictions of the distributions lead to the log likelihood (used to compare the VB results with sampling algorithms)

$$\log P(\mathbf{y}|\Theta) = \sum_i \left[-\frac{N_i}{2} \log \left(\frac{2\pi}{s_i c_i} \right) - \frac{1}{2} s_i c_i \mathbf{k}_i(\mathbf{m}) \right] \quad (12)$$

and also to a closed form for the VB free energy F shown in Appendix A. The latter one has its maximum at the posterior distribution and the update equations for the parameters can be derived by using the necessary maximum condition that the partial derivatives of F w.r.t. to the parameters of the distributions need to be zero. Specifically, $\partial F / \partial s_i = 0$ leads to the update equations of the noise parameters

$$c_i = \frac{N_i}{2} + c_{i,0} \quad (13)$$

$$\frac{1}{s_i} = \frac{1}{s_{i,0}} + \frac{1}{2} \mathbf{k}_i^T \mathbf{k}_i + \frac{1}{2} \text{tr}(\Lambda^{-1} \mathbf{J}_i^T \mathbf{J}_i). \quad (14)$$

The model update equations for the precision Λ and the means $\mathbf{m}$ come from $\partial F / \partial \Lambda = 0$ and $\partial F / \partial \mathbf{m} = 0$, respectively, and read

$$\Lambda = \Lambda_0 + \sum_i s_i c_i \mathbf{J}_i^T \mathbf{J}_i \quad (15)$$

$$\Lambda \mathbf{m} = \Lambda_0 \mathbf{m}_0 + \sum_i s_i c_i \mathbf{J}_i^T (-\mathbf{k}_i + \mathbf{J}_i \mathbf{m}_{\text{old}}), \quad (16)$$

where $\mathbf{m}_{\text{old}}$ refers to the model parameter means of a previous iteration. These update equations are evaluated repeatedly until the free energy from one iteration to another remains below a given tolerance, here $\Delta F = 0.1$.

This work provides an additional post-processing check that indicates the validity of those assumptions (Section 3.3). This check is also performed in the examples where the results from VB and nested sampling for the posteriors and the evidences are compared.

3.3 | Nonlinearity measure

The third assumption of the presented variational Bayes method is the fact that the prior and posterior distributions are conjugate multivariate normal distributions. If the model parameters θ are normally distributed, so is their linear transformation $\mathbf{A}\theta + \mathbf{b}$ or, in our case, $\mathbf{J}\theta + \mathbf{k}$. Another interpretation is a direct conclusion of the variational Bayes derivation that replaces the actual model with its linear Taylor expansion in Equation (9). If the model is indeed linear in its parameters, this step is exact and involves no approximations. As a result, the true posterior is indeed a multivariate normal distribution, the KL divergence to the parameterized VB posterior vanishes and, as a consequence of Equation (7), the VB free energy corresponds to the log-evidence. Obviously, if the Taylor expansion includes higher order nonzero terms, the true posterior deviates from a multivariate normal distribution and the aforementioned properties are only approximations. Generally, the higher the nonlinearity of the model, the worse this approximation usually is.

Practically, the nonlinearity of a model may not be known a priori. Within this work, the error in the posterior distributions (when this linearity assumption is violated) is illustrated by comparing the VB results with a sampling method for different examples, see Figure 15A, and further explanations in the corresponding Section 4.2.

Consequently, the following post-processing measure to indicate the degree of nonlinearity within the posterior distribution $q(\theta) \sim \mathcal{N}(\mathbf{m}, \Lambda^{-1})$ is proposed. For each parameter θ_i with its mean m_i and its standard deviation $\sigma_i = \Lambda_{ii}^{1/2}$, a modified parameter vector $\mathbf{m}' = \mathbf{m}$ is defined, where the i -th component is changed according to

$$m'_i \leftarrow m_i + j\sigma_i. \quad (17)$$

The values of j define the relevant parameter range and $\mathbf{j} = (-3, -2, \dots, 3)$ is used in this work. This results in a total of $\text{dim}(\mathbf{j})$ parameter sets which are used to evaluate both the true model error $\mathbf{k}(\mathbf{m}')$ and its linearization $\mathbf{k}(\mathbf{m}) + \frac{\partial \mathbf{k}}{\partial \theta} \big|_{\mathbf{m}} (\mathbf{m} - \mathbf{m}')^{\ddagger}$.

The results are arranged in a matrix $\mathbf{K}_i$ for the true model error and $\mathbf{K}_{i,\text{lin}}$ for its approximation, both of dimension $(\text{dim}(\mathbf{k}) \times \text{dim}(\mathbf{j}))$. The measure of nonlinearity L_i for the i -th parameter is then defined as

$$L_i = \frac{\|\mathbf{K}_i - \mathbf{K}_{i,\text{lin}}\|}{\|\mathbf{K}_{i,\text{lin}}\|} \quad (18)$$

where $\|\cdot\|$ stands for the Frobenius norm. Evaluating a perfect model at the mean can result in $\mathbf{k}(\mathbf{m}) \equiv \mathbf{k}_{\text{lin}}(\mathbf{m}) \equiv \mathbf{0}$. As long as the model actually depends on the parameter, the deviation $\mathbf{m}'$ will cause a nonzero model error and the denominator of Equation (18) will not become zero. An example to illustrate the application is illustrated in Appendix D.

4 | RESULTS

In the following section, the method described above is applied to infer the parameters of the demonstrator bridge with step-wise increasing model complexity. In a first scenario according to Section 2.1, a single beam (Figure 11A) is analyzed to identify its Young's modulus. Afterwards, a two-dimensional slice of the bridge according to Figures 2B and 11B is used to illustrate the methods and compare the results using the variational methods with the ones obtained by the established dynamic nested sampling.²⁶ This example is also used to validate the modeling assumptions related to connections.

In a third scenario, the cable-stayed bridge (Section 2.3) is analyzed and a complete system identification is performed (Figure 11C). This information is then used as prior information in the final scenario according to Section 2.4 for a damage identification procedure (Figure 11D).

4.1 | Single beam in uniaxial loading

The main building block of the demonstrator bridge is the blue beam element and its Young's modulus significantly influences the structural behavior. The manufacturer provides the Young's modulus of the blue beam element as $E_{\text{PASCO},0} = 2.3 \text{ GPa}$ ¹⁷ and the purpose of the first experiment is to update this parameter using strain measurements $y_\epsilon(n_w)$ for different load amplitudes.

The measured strains are illustrated in Figure 12A. The mean of each signal is then used as the data point y_ϵ of the measurement.

The Young's modulus parameter is inferred based on the strain measurements for each load step. Thus, the corresponding model error has six entries, one for each number of weights, and reads

$$\mathbf{k}(E_{\text{beam}}^{\text{1D}}) = \boldsymbol{\epsilon}(E_{\text{beam}}^{\text{1D}}) - \mathbf{y}_\epsilon. \quad (19)$$

The prior distribution is given by $P(E_{\text{beam},0}^{\text{1D}}) \sim \mathcal{N}(\mu = E_{\text{PASCO}}, \sigma = 0.2E_{\text{PASCO}})$. The prior distribution for the sensor precision $P(\Phi_{\text{strain},0})$ derived from separated measurements in an unloaded state with an experimental standard deviation of about $5 \mu\text{m/m}$. The parameters of the Gamma distribution are chosen such that its 5% percentile is $1/(1 \mu\text{m/m})^2$ and the 95% percentile is $1/(10 \mu\text{m/m})^2$.

The results of the inference using the variational Bayes algorithm is visualized in Figure 12A. The solid line represents the model response with the posterior mean of $E_{\text{beam}}^{\text{1D}}$. The shaded regions around this response represent the propagated uncertainty with distance σ_E , $2\sigma_E$ and $3\sigma_E$ with

$$\sigma_E(n_w) = \sqrt{\left(\frac{dk(E_{\text{beam}}^{\text{1D}}, n_w)}{dE_{\text{beam}}^{\text{1D}}}\right)^2 \text{COV}(E_{\text{beam}}^{\text{1D}}) + \frac{1}{\Phi_{\text{strain}}}}, \quad (20)$$

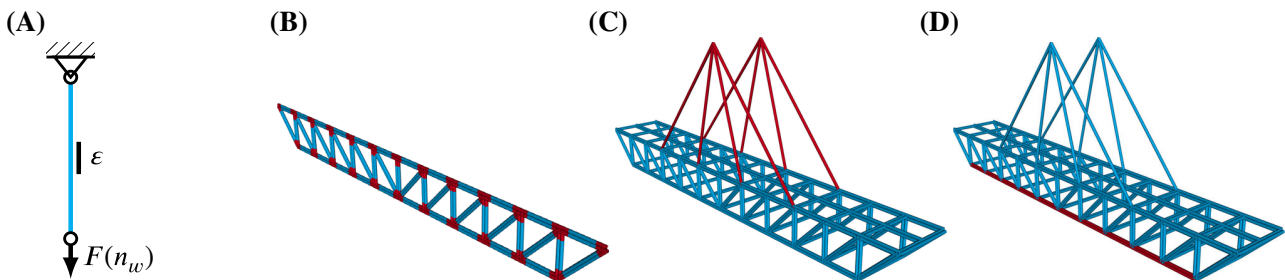


FIGURE 11 Overview over the following numerical studies. (A) Single beam experiment, (B) model selection without cables, (C) system identification, (D) damage identification.

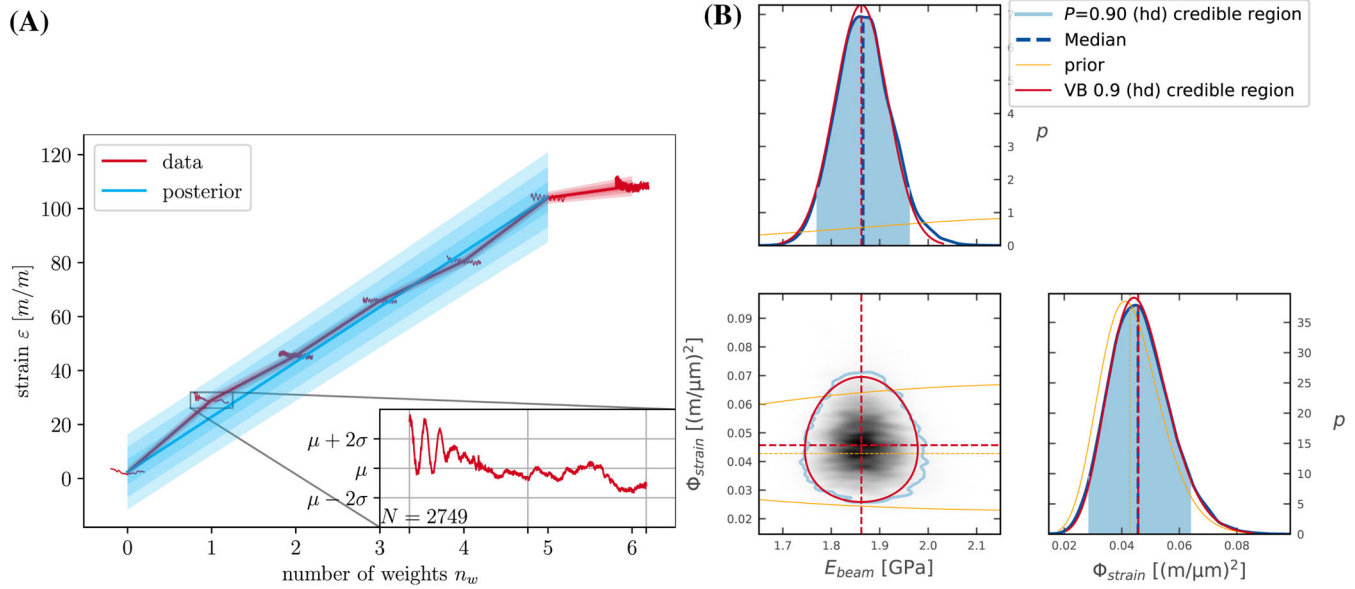


FIGURE 12 Result of the Young's modulus inference of the single beam in uniaxial loading scenario (see Figure 11A). (A) Measurement data and uncertainty of model prediction. (B) Posterior distribution computed from sampling (blue) and VB (red).

where the first summand is related to the variance of the model parameters propagated through the model response by linearizing the model response at the posterior mean, and the second one to the variance of the noise computed from the mean of the posterior noise precision Φ_{strain} .

Even though the model response is inversely proportional to the parameter, the nonlinearity measure $L = 0.060$ is almost zero – mainly due to a very narrow posterior. As a consequence, the variational Bayes free energy of $F = -18.835$ matches the evidence $\log P(\mathbf{y}) = -18.836$ from the nested sampling and the corresponding posterior distributions in Figure 12B coincide.

Repeating the experiment with the more compliant, gray beams results in a beam stiffness of $q(E_{\text{beam,gray}}^{1D}) \sim \mathcal{N}(\mu = 0.358, \sigma = 0.0028)[\text{GPa}]$ which is about $0.19E_{\text{beam}}^{1D}$.

4.2 | Simply supported bridge

The geometry of this scenario presented in Section 2.2 is based on a cable-less, two-dimensional slice of the bridge shown in Figure 11B to reduce both the number of parameters and the numerical cost. The stereo measurements at 21 locations of a single car pass is used as data.

First, the inference is run with the beam element's stiffness E_{beam}^{2D} as the only model parameter, the noise precision Φ_{stereo} as the only noise parameter and the priors are assumed to be

$$P(E_{\text{beam},0}^{2D}) \sim \mathcal{N}(\mu = E_{\text{PASCO}}, \sigma = 0.2E_{\text{PASCO}}) \text{ and} \quad (21)$$

$$P(\Phi_{\text{stereo},0}) \sim \text{Ga}(\text{scale} = 625, \text{shape} = 1)[\text{mm}^{-2}], \quad (22)$$

such that the mean of the noise *standard deviation* with $(625 \text{ mm}^{-2})^{-\frac{1}{2}} = 0.04 \text{ mm}$ is roughly twice as high as the standard deviation of a noisy zero signal.

Figure 13A shows the data of this experiment together with the model response and the propagated uncertainties according to Equation(20) (with trivial adjustments).

The comparison of the variational Bayes results to the ones obtained by nested sampling in Figure 13B show a perfect match in both the individual distributions as well as their correlation. This is also indicated by the nonlinearity measure of $L = 0.0007$.

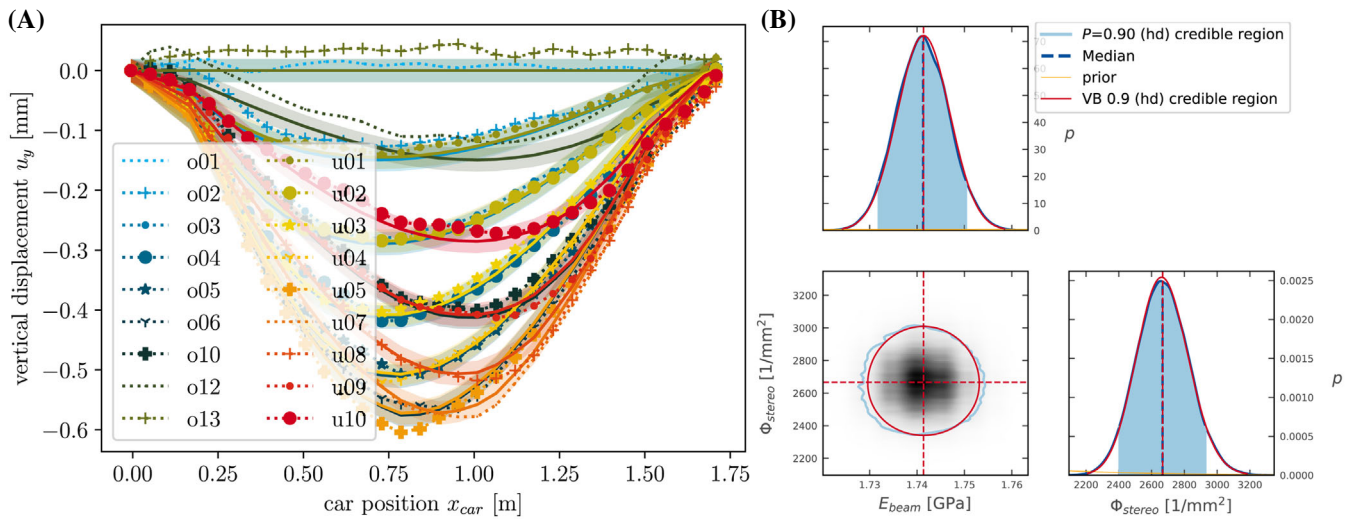


FIGURE 13 Results of the Young's modulus inference for the simply supported bridge scenario (see Figure 11B). (A) Model response (solid), $\pm\sigma$ (shaded) compared to the data (dotted). (B) Posterior distribution from sampling (blue) and VB (red).

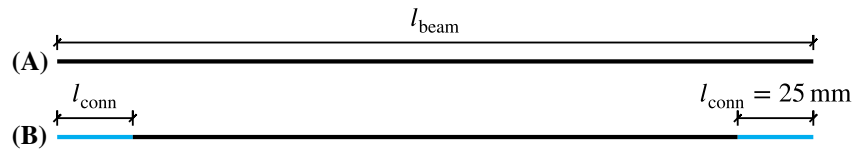


FIGURE 14 Two modeling approaches with (A) a single beam element and (B) using two additional beam elements representing the connections.

There is, however, a mismatch between the model parameter posterior of this experiment compared to the one of the uniaxial test. This is visualized in Figure 18. A possible reason is the simplified modeling of the real bridge consisting of beams *and* connections with a single beam element. Thus, in an adjusted numerical model, each beam is subdivided into three beams according to Figure 14. The stiffness of the highlighted beams associated with the connections is described by an additional unknown parameter E_{conn}^{2D} .

Figure 15A shows the posterior plots of the inference using now two parameters to describe the stiffness in the beams—one for the actual beam, and another one for the two additional beams representing the connection elements at each end. The posterior means obtained by both inference types coincide, but the shapes of the distributions differ. Especially in the pair plot between E_{beam}^{2D} and E_{conn}^{2D} , both deviate. The nested sampling shows a narrow, but curved distribution that only matches the VB distribution around the posterior mean. This deviation indicates that the true posterior is not a normal distribution and that the VB assumptions are violated. One result is a significant difference of the variational Bayes free energy to the log-evidence of ≈ 0.5 , shown in Table 1.

This problem, however, would go unnoticed when only considering the VB results and actually motivated the nonlinearity measure L in Section 3.3: The values of $L_{E_{beam}^{2D}} = 0.509$ and $L_{E_{conn}^{2D}} = 1.001$ immediately indicate a strong nonlinearity ($L \gg 0$) around the posterior mean. The finite element beam model solves for the displacement field $\mathbf{Ku} = \mathbf{F}$ with $\mathbf{K} \propto E_{beam}^{2D}, E_{conn}^{2D}$. Thus, the model parameters are (qualitatively) inverse proportional to the model response $\mathbf{u}$. In contrast to the two analyses above that also had the inverse relationship between parameters and model responses, the posterior distributions in this study are much wider.

A simple remedy is a parameter inversion, that is, the inference of compliances $C_{beam/conn} = 1/E_{beam/conn}$ instead of stiffnesses E . The resulting posterior distributions are shown in Figure 15B and are in almost perfect agreement. The corresponding nonlinearity measures $L_{C_{beam}^{2D}} = 0.0002$ and $L_{C_{conn}^{2D}} = 0.0005$ are almost zero and Table 1 shows that the VB free energy matches the log-evidence of the nested sampling.

Both studies with the additional connection parameter show a strong correlation to the beam stiffness parameter that deforms the joint distribution almost to a line, as seen in the pair plot. The projection into the individual model parameter

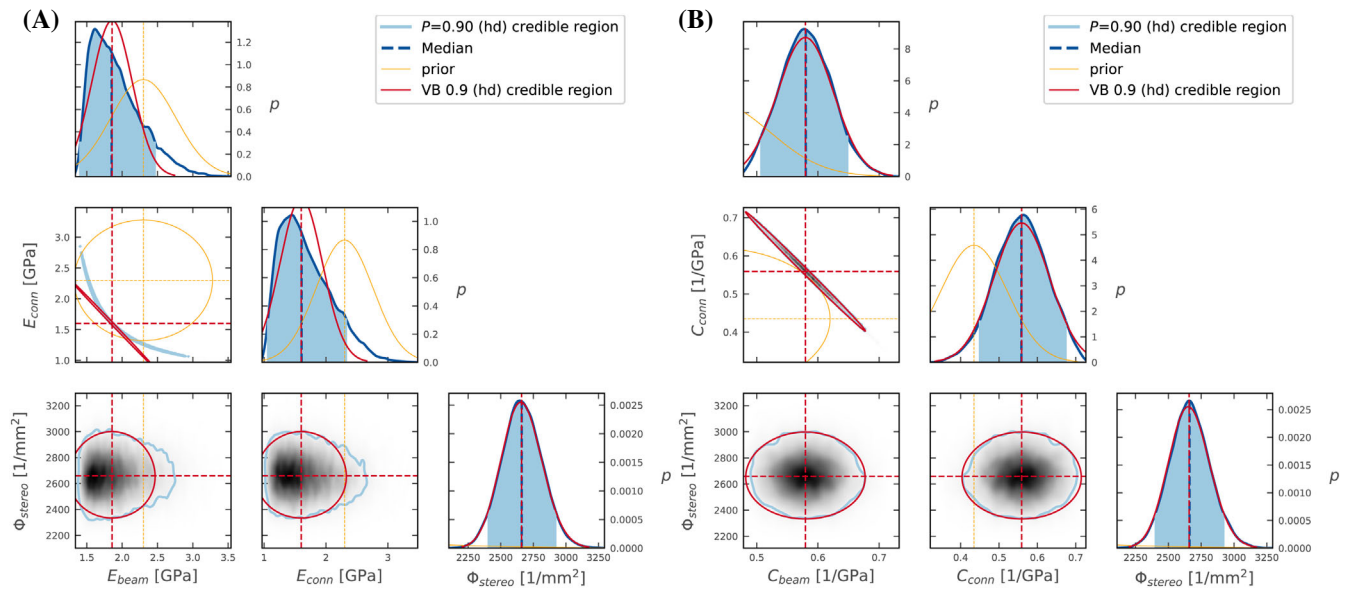


FIGURE 15 Results of stiffness vs. compliance inference of the 2D simply supported bridge using additional connection beam elements with either nested sampling (P) or the variational inference (VB). (A) Inferring stiffness, (B) Inferring compliances

TABLE 1 Comparison of the variational Bayes free energy F_{VB} with the log evidence $\log(z)$ obtained from nested sampling using dynesty. The columns with $\#_0$ show the number of iterations and t_0 refers to the total run time.

Case	F_{VB}	Dynesty $\log(z)$	$\#_{VB}$	$\#_{dynesty}$	$t_{VB}[s]$	$t_{dynesty}[s]$
$\theta = (E_{beam}, E_{conn})$	1447.511	1448.088	10	81,864	0.211	258.721
$\theta = (C_{beam}, C_{conn})$	1447.040	1447.115	3	64,362	0.098	200.717
$\theta = E_{beam}$	1448.880	1448.873	4	59,764	0.035	90.579
$\theta = C_{beam}$	1448.887	1448.957	3	52,772	0.028	7.436

spaces then results in posterior distributions that are not significantly narrower than their priors. The physical reason for the correlation is the fact that the subdivided beams can be interpreted as a series of springs where the compliances in the beam axis add up. With the total length $l \approx 126.7$ mm and the posterior mean of E_{beam} from Figure 13B, the corresponding spring compliance for the single model parameter yields $c_{total} = l/(AE_{beam}^{2D}) = 1.406$ mm N⁻¹. Summing up the individual contributions from Figure 15A yields $c_{beam} + c_{conn} = (0.799 + 0.603)$ mm N⁻¹ = 1.402 mm N⁻¹ which closely matches c_{total} and proves this theory.

Combining this interpretation with the evidences in Table 1 that show higher values for the single parameter models, subsequent analyses are performed with a single E_{beam} model parameter.

Table 1 also shows performance measurements for each method, measured on a single core of a AMD Ryzen 5 2600 (as all subsequent performance measurements). Performing one VB iteration includes evaluating the derivative of the model error w.r.t. the parameters. Here, this is done via central differences which itself requires additional model evaluations and results in a higher time per iteration compared to the nested sampling. As VB requires only a very small number of iterations (see $\#_{VB}$), its overall run time is *much* smaller.

4.3 | System identification of the cable-stayed bridge

In this numerical example, the complete, three-dimensional bridge including cables as shown in Figure 11C is analyzed using the data of five separate bridge crossings. As described in Section 2.3, this introduces an additional model parameter E_{cable}^{3D} for the cable stiffness and eight additional calibration parameters $F_1 \dots F_8$ that account for the unknown force sensor properties. This example also demonstrates the performance of the variational Bayes method by additionally

TABLE 2 System identification parameters and their prior distributions. With $\epsilon = 10^{-6}$, the noise parameter prior is an noninformative Gamma distribution. The # column refers to the number of parameters for the corresponding parameter type.

Parameter type		#	Prior
Beam stiffness	E_{beam}^{3D}	1	$\mathcal{N}(E_{\text{PASCO}}, 0.2E_{\text{PASCO}})$
Cable stiffness	E_{cable}^{3D}	1	$\mathcal{N}(E_{\text{PASCO}}, 0.2E_{\text{PASCO}})$
Calibration factors	F_i	8	$\mathcal{N}(-0.5, 0.5)[V\ N^{-1}]$
Stereo offsets	$\theta_{o,\text{stereo}}$	105	$\mathcal{N}(0, 1)[\text{mm}]$
Laser offsets	$\theta_{o,\text{laser}}$	30	$\mathcal{N}(0, 1)[\text{mm}]$
Force offsets	$\theta_{o,\text{force}}$	40	$\mathcal{N}(0, 1)[V]$
Stereo noise	Φ_{stereo}	1	$\text{Ga}(\text{scale} = \epsilon, \text{shape} = 1/\epsilon)[\text{mm}^{-2}]$
Laser noise	Φ_{laser}	1	$\text{Ga}(\text{scale} = \epsilon, \text{shape} = 1/\epsilon)[\text{mm}^{-2}]$
Force noise	Φ_{force}	1	$\text{Ga}(\text{scale} = \epsilon, \text{shape} = 1/\epsilon)[V^{-2}]$
	Σ	188	

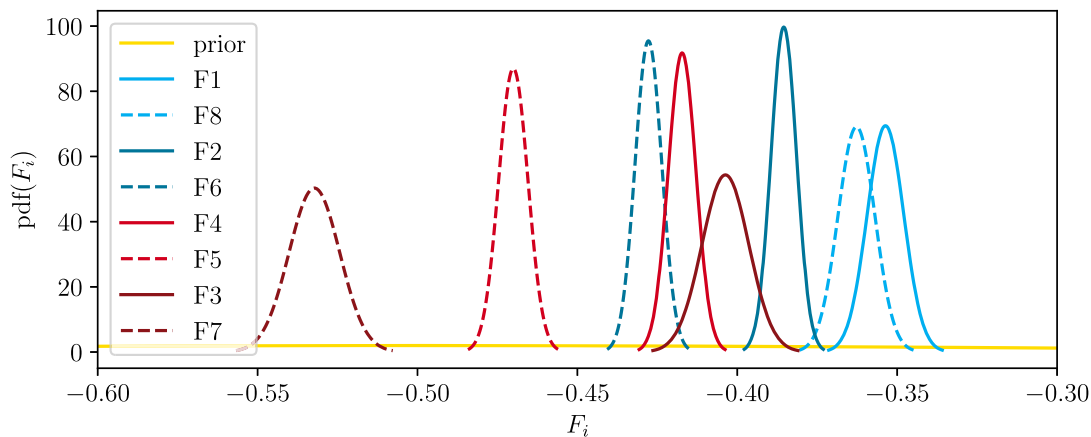


FIGURE 16 Posterior distributions of the force calibration factors.

inferring all offset parameters θ_o described in Section 2.2.3 and summarized in Table 2, that is, compensating for the wrong modeling assumption of the structure not being influenced by the car before it enters the bridge. As pointed out in Section 2.3.2, the cable stiffness is related to a superposition of the relatively stiff steel cable and the more compliant force sensor. Since the sensor itself is very compliant with a stiffness similar to the beam and due to a lack of additional knowledge, the cable stiffness prior was assumed similar to the beam stiffness prior.

Based on Equation (2), an offset correction is introduced and the corresponding parameters are identified in parallel to the model parameters. Note that only a single offset per sensor is used, which means the reference state (corresponding to the unloaded bridge before the car enters/after it leaves the bridge) is identified. A similar prior for these offset parameters (for stereo, laser and force) has been used due to the absolute values (irrespective of their different physical quantities and units) being in the same order. For their noise standard deviation, a noninformative prior was chosen to reflect our lack of knowledge, since this term comprises both the actual measurement uncertainty but in addition the model bias (the inability of the model to represent the actual physics).

The whole analysis converged after 6 iterations with a total run time of 2.33 s. The majority of the time is spent in computing the Jacobian. The derivative w.r.t the two stiffness parameters is obtained by central differences and requires additional finite element model evaluations. With the highest nonlinearity measures $L_{E_{\text{beam}}^{3D}} = 0.0012$ and $L_{E_{\text{cable}}^{3D}} = 0.0011$, this is a good approximation. As the other model parameters are either linear factors or offsets, the derivative w.r.t them is computed analytically with negligible computational costs. Obviously, their nonlinearity measure is numerically zero.

The inferred calibration factors for the force sensors are exemplarily shown in Figure 16. With Figure 8B as a reference, the posteriors for the force sensors at the front side of the bridge are shown in solid lines, the dashed ones are related

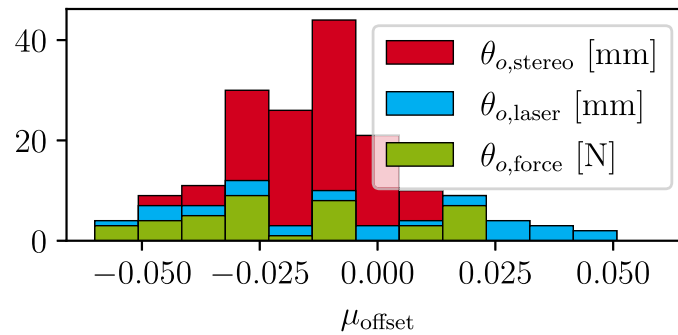


FIGURE 17 Histogram of the posterior means for the offsets. The contributions of each sensor type is added up in the histogram bins.

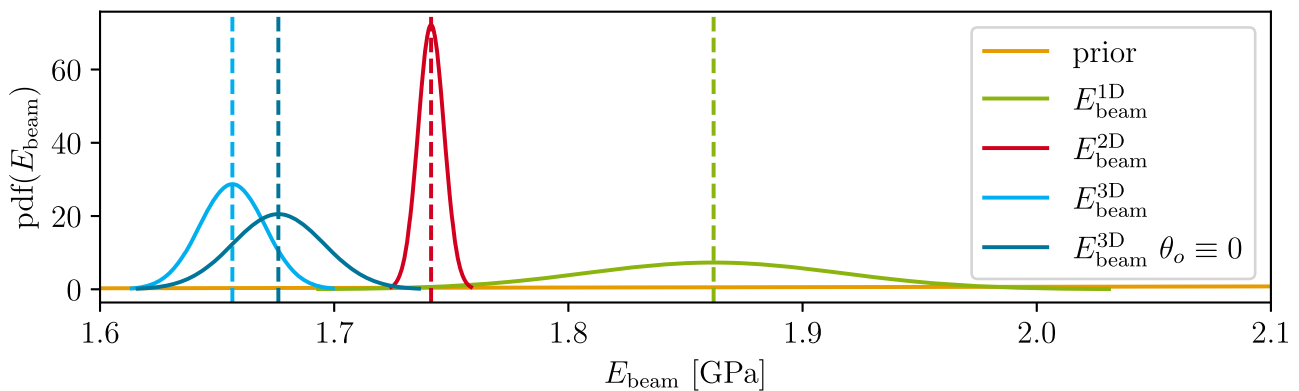


FIGURE 18 Summary of the E_{beam} posterior from the uniaxial loading E_{beam}^{1D} , for the 2D simply supported bridge E_{beam}^{2D} for the cable-stayed bridge E_{beam}^{3D} where in the last case all displacement offsets θ_o are assumed to vanish.

to the sensors on the back side. Ideally, this symmetry would cause identical posteriors for each front-back-pair. Due to symmetry in the axis of the tower, one would furthermore expect the posteriors of outer cable factors (colored blue), as well as the group of inner cable factors (colored red). Reasons for the visible deviation from this symmetry include variations in the material parameters of each component and a slight misalignment of the force sensors w.r.t. the cable direction. Additionally, an explicit calibration of each force sensor is not performed, so the conversion factor from the measured force to the sensor signal in m V can indeed vary from sensor to sensor. However, the inferred calibration factors are in a reasonable range and show a significant increase in their precision compared to the prior.

The properties of the offset parameter distributions are analyzed in Figure 17. The distribution of the posterior means is almost symmetrical around zero. The highest means for the offsets of the displacement measurements reach a value of ± 0.05 mm that is significant compared to maximum displacement signals of $-0.1 \dots -0.6$ mm (from Figure 13A).

Last, the inferred beam stiffnesses of all previous experiments are collected in Figure 18, including the results of the three-dimensional model of this section with inferred and fixed offset parameters $\theta_o \equiv 0$, the two dimensional problem assuming symmetry as discussed in the previous section and the data from the uniaxial tests in Section 4.1. In theory, the beam stiffness E_{beam}^{xD} is supposed to describe the actual material properties of the ABS plastic used to build the beam elements which would result in matching posterior means regardless of the experimental setup. The actual posterior distributions however show different results depending on the experimental data used. There are multiple reasons that explain these differences. Certain modeling choices like modeling the connections and the beam (to avoid the very high correlation of both) with a single parameter merges both properties. Another reason is the assumption of a perfectly stiff support at the beginning of the bridge and a floating support at its end (allowing only horizontal displacements). As can be observed from the data, the supports are not rigid thus resulting in a different offset before the car enters the bridge (being on the left-hand side support) and once it has left the bridge (being on the right-hand side support), which was not included in our model. In theory, this could be added by identifying two independent offsets, but we chose not to double the number of offsets. The difference between the identified mean of E_{beam} for the 3D case with

($E(E_{\text{beam}}) = 1.66$ GPa) and without offsets ($E(E_{\text{beam}}) = 1.68$ GPa) is rather small, but the error compared to the uniaxial experiment ($E(E_{\text{beam}}) = 1.86$ GPa) is much more prominent. Last but not least, the material parameter is assumed to be fixed for all beams which neglects the effect of manufacturing tolerances.

Another modeling assumption is related to the statistical model. In the current implementation, the measurements are assumed to be uncorrelated. This is a common choice (since it simplifies the mathematical treatment) and often there is a lack of a better representation due to the choice of a specific correlation model (with additional hyperparameters) not being objective either. Finally, the prior and posterior distributions are assumed to be normal and Gamma distributions, which is only exact for a linear forward model, otherwise a linearization of the forward model at the posterior mean might introduce additional errors. For simulations with a large number of parameters, where a reference result with sampling cannot be obtained, the nonlinearity measure is the only indication of fulfilling this property.

4.4 | Damage identification

The last numerical example demonstrates the identification of damaged beams within the structure discussed in Section 2.4. Therefore, four car passes are analyzed: One corresponding to the previous experiment with no damage as a reference and three experiments where a beam is replaced by a more compliant one in segments 6, 7, and 8 of the bridge. The model of the previous Section 4.3 is extended by 10 additional damage parameters that belong to the 10 possibly damaged bridge segments highlighted in Figure 11D (the beam elements in the bottom front).

The beams in each segment $i = 1 \dots 10$ have their Young's modulus modified with a damage parameter ω_i

$$E_i = (1 - \omega_i)E_{\text{beam}}^{3D} \quad (23)$$

ranging from zero for undamaged material to one for a fully damaged state. Note that a value below zero sounds unreasonable, but effectively leads to a higher effective stiffness. A value above one, however, results in a negative Young's modulus and must be avoided. Hard limits like $\min(\max(\omega_i, 0), 1)$ would lead to a zero Jacobian once ω_i leaves the range $[0, 1]$, falsely indicating that the parameter has no influence on the model. Instead, an unrestricted damage parameter W_i is used as the statistical parameter which is then transformed to the actual damage parameter ω_i

$$\omega_i = S(W_i, \epsilon) \quad (24)$$

$$S(x, \epsilon) = \frac{\epsilon(e^x - 1)}{\epsilon e^x + 1} \text{ such that } S(x = 0, \epsilon) = 0 \quad (25)$$

using a modified sigmoid function S with $\epsilon = 0.1$, as shown in Figure 19 that limits ω_i to $[-0.1, 1]$.

The prior distributions for E_{beam}^{3D} , E_{cable}^{3D} and the noises are defined as the rather narrow posterior distributions of the previous Section 4.3. The calibration factors F_i are treated as deterministic parameters here with the value taken as the posterior mean of the previous system identification. As new data sets have potentially different offset parameters, the prior $P(\theta_{0,0}) \sim \mathcal{N}(0, 1)$ is used. The priors of the newly introduced damage parameters are assumed to be rather wide with $P(W_{i,0}) \sim \mathcal{N}(0, 1/2)$ and the transformation to $\omega_{i,0}$ is shown as *prior* in Figure 20.

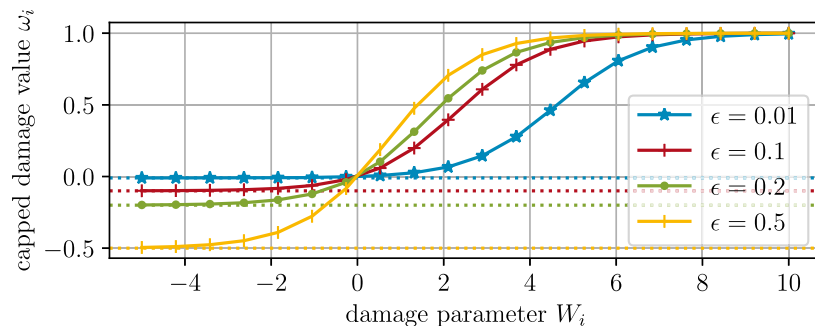


FIGURE 19 Sigmoid function used to limit the damage to $(-\epsilon, 1)$.

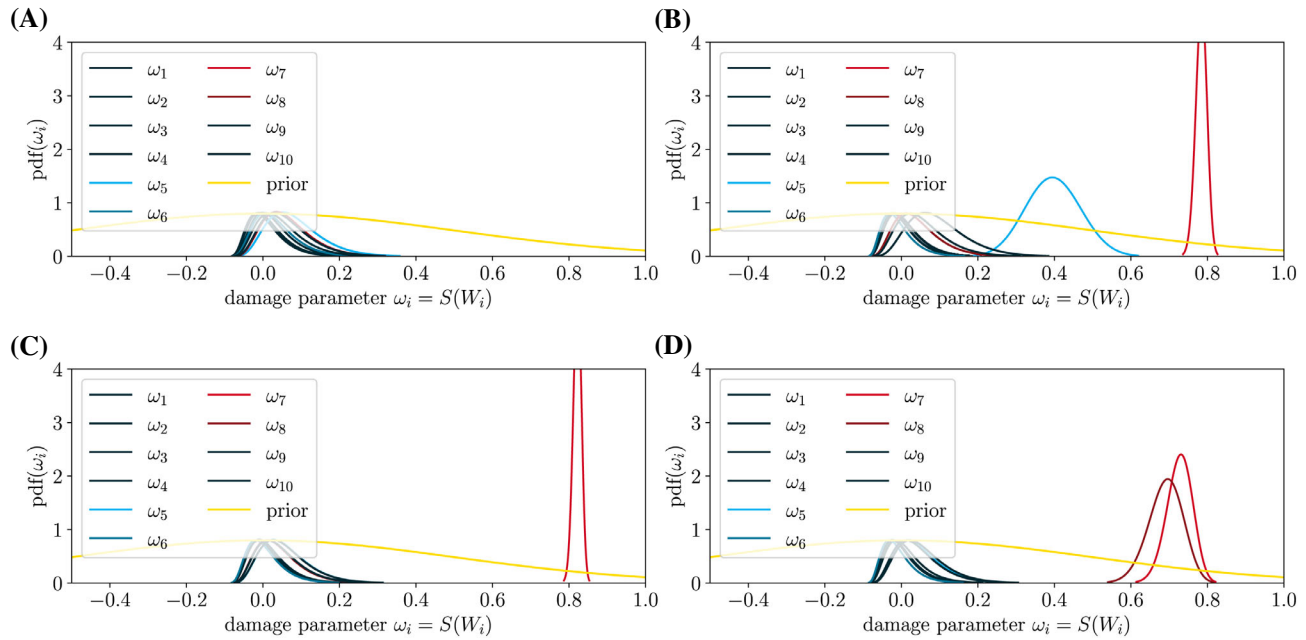


FIGURE 20 Posterior distributions of each damage variable W_i . (A) No damage for reference, (B) damage in segment 6, (C) damage in segment 7, (D) damage in segment 8.

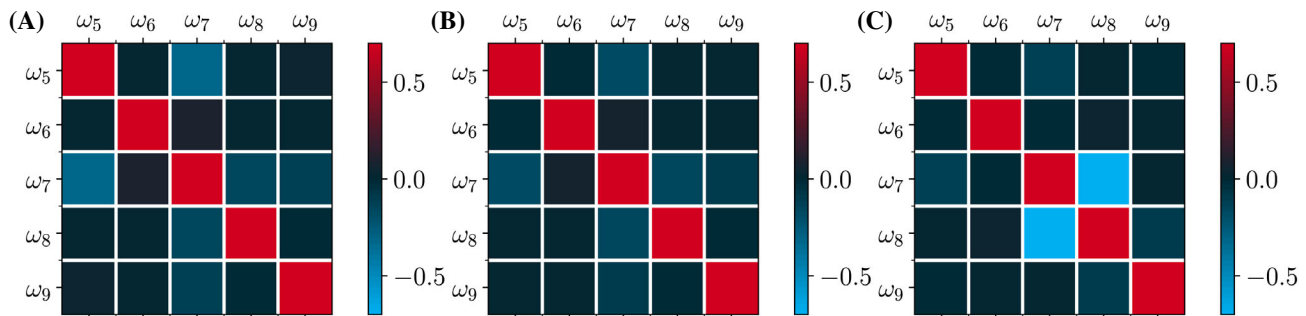


FIGURE 21 Correlations in the posterior distributions for the damage variables. (A) Damage in segment 6, (B) damage in segment 7, (C) damage in segment 8.

The inference using undamaged reference data finished after just 2 iterations, because the prior distributions were already close to the posteriors. The other data sets took 5 – 6 iterations with the highest nonlinearity of a stiffness parameter $\max(L_E) < 0.001$ and the highest nonlinearity of a damage parameter $\max(L_{W_i}) = 0.038$, indicating a correct VB run.

Figure 20 summarizes the (transformed) posteriors of the damage variables ω_i . As expected, using the undamaged data set shows that all damage variables are significantly narrower than the prior distribution with a mean very close to zero, indicating no damage. The data set with damage in segment 6 results in high damage values in the surrounding segments 5 and 7. An almost ideal result is obtained for the data set with damage in segment 7 with the matching posterior of ω_7 as the only one deviating from a zero mean. As mentioned at the end of Section 4.1, the ratio between the blue and gray beams is approximately $E_{\text{beam,gray}}^{1D} \approx 0.19 E_{\text{beam}}^{1D}$, which corresponds to a damage value of $\omega \approx 0.81$ that is close to the posterior mean of ω_7 . Finally, the data set with damage in segment 8 results in high values for only ω_8 and ω_7 . When comparing the corresponding correlations illustrated in Figure 21, it is seen that in particular for damage in segment 6 the damage variables in segments 5 and 7 are negatively correlated, and even more pronounced for the damage identification with damage in segment 8, where segments 7 and 8 are negatively correlated. This indicates the ill-posedness of the inverse problem resulting in a set of solutions that all produce similar results. If a model developer requires to have a well posed problem either more data is required (that in particular allows to distinguish between these different solutions) or the set of parameters could be modified.

5 | SUMMARY AND CONCLUSION

Representing real-world structures in a computational model usually involves challenges related to data generation, data processing as well as the development of a simulation model implying assumptions on the physics as well as the calibration of the corresponding model parameters.

In this context, we outlined the requirement to establish a metadata schema to allow a simple access to the data—both from a article developer point of view as well as a for a researcher that intends to reuse the published data. It was also shown how different asynchronous data sets (due to different data acquisition systems) can be jointly analyzed by identifying the time lag as an additional parameter.

The main challenge in building digital twins is related to the modeling assumptions—both for the physical model as well as for the inference problem. It was shown that—using Bayesian inference procedures—the ill-posedness of the inverse problem can be identified, resulting in our case in the decision to model the connections and the beams not as separate entities. Furthermore, posterior distributions of the identified parameters are obtained that provide additional accuracy information compared to a deterministic optimization approach. The identification process was performed for different scenarios (single beam in uniaxial loading, simply supported bridge, cable-stayed bridge) showing that the identified model parameters (Young's modulus of the beams) depend on the scenario with corresponding modeling assumption and might differ between the different scenarios. In the example, the mean varied in the order of about 10%. This is due to the fact that in structural engineering many model simplifications are made. This includes the initial conditions (see e.g., the initial geometry of the bridge and the simplification of using 1d-elements), constitutive equations (e.g., all beams are elastic, but friction in the joints is present), symmetry effects (the undamaged model is supposed to be symmetric, but could also be parameterized to allow for unsymmetrical solutions), measurement uncertainty (e.g., due to the force sensors with a finite stiffness and a not completely accurate alignment with the cable) or boundary conditions (e.g., how the car was modeled). It is to be noted that even when analyzing all configurations simultaneously, this is usually not captured by a standard Bayesian inference procedure, where the most probable parameter set is computed. In particular when combining data sets with a different amount of measurement data (e.g., single scalar measurements with time series data) the probabilistic model used is of mayor importance.

This relates to the second challenge when using Bayesian inference procedures, which is linked to the postulated data generation process. A main challenge is the correlation model used for the measurement data. For structural problems with time series data, the measurements are certainly correlated in both space and time, but the definition of an appropriate parameterized kernel function with potential hyperparameters such as the correlation length is not objective and might significantly influence the results. The problem is even more complicated due to the fact that the noise model comprises both the measurement noise as well as a model bias (the inability of the model to actually represent the data), the latter leads to strong correlations not only related to space and time but caused by the model. In this article, the measurement data was resampled to 1 Hz, future extension might directly use a correlation model in the inference procedure. However, due to the large number of time steps (up to 30.000) this requires some further development coping with sparse correlation matrices that is outside the scope of the current article (see for example, Reference 33).

Another challenge is the definition of appropriate priors, in particular for the noise models (which is due to model bias not only related to the measurement accuracy of the sensor itself). And finally, the parameterization of the physics based model (e.g., identifying a single Young's modulus for all scenarios or separating between single beam, simply supported bridge and cable-stayed bridge and adding some additional likelihood terms penalizing a difference). Even though different models could in theory be compared using the model evidence, in practice this is sometimes difficult due to a different parameterization (e.g., requiring offsets not present in another model) where the choice of the prior is often not fully objective.

Using variational inference procedures allowed to work with a large number of parameters (up to 188 in seconds) and also provided without additional computational effort an approximation of the free energy to approximate the model evidence. The large number of parameters was required to compensate for model assumptions (deformation state is the same when the car is on the left or right support) that were not supported by the experimental data requiring to add offsets as additional latent variables in the identification procedure or due to unknown calibration factors that resulted from the not fully cable-aligned orientation of the force sensors. A procedure for validating the linearity assumption in the variational inference procedure was outlined and it was shown that the number of forward model evaluations is significantly reduced compared to sampling based approaches (a reduction factor in the order of $1e-3$) making this approach applicable to more complex FEM-problems. In particular, the authors realized that setting up the inference

problem for a real setup requires an iterative process making sampling based approaches at least in the development phase prohibitively expensive.

Finally, the authors want to point out that the estimation of model parameters presented in scientific journals for the validation of a new model is often poorly documented and not reproducible (e.g., the actual measurement data, prior/starting values, objective/likelihood are often not provided). As illustrated in this article, this might lead to significant deviations and can only be solved when the complete parameter estimation process is included in the provided software repositories.

AUTHOR CONTRIBUTIONS

Thomas Titscher: data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); software (lead); visualization (lead); writing – original draft (lead). **Thomas van Dijk:** formal analysis (equal); software (equal); writing – review and editing (supporting). **Daniel Kadoke:** formal analysis (equal). **Annika Robens-Radermacher:** formal analysis (supporting); investigation (supporting); methodology (supporting); writing – review and editing (equal). **Ralf Herrman:** formal analysis (supporting). **Jörg F. Unger:** conceptualization (lead); funding acquisition (lead); methodology (equal); project administration (lead); supervision (equal); writing – original draft (supporting); writing – review and editing (equal).

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/eng2.12669>.

DATA AVAILABILITY STATEMENT

For this article, all measured data sets as well as the source code to reproduce the results and this article are provided in <https://zenodo.org/record/7874659#.ZEtXlc5BxaQ>

ENDNOTES

*<https://www.youtube.com/watch?v=ReWmrbQgQLU>.

†<https://youtu.be/VVWqJxiMzdM>

‡Note that an alternative formulation could be based on the model response $\mathbf{g}$ instead of the model error $\mathbf{k}$. However, the update Equations (13) to (16), the free energy in Equation (A1) and the log likelihood in Equation (12) directly contain the model error $\mathbf{k}$, so it is simply more convenient to use $\mathbf{k}$ directly.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

ORCID

Ralf Herrmann  <https://orcid.org/0000-0001-6941-9983>

Jörg F. Unger  <https://orcid.org/0000-0003-0035-0951>

REFERENCES

- Behmanesh I, Moaveni B. Probabilistic identification of simulated damage on the Dowling Hall footbridge through Bayesian finite element model updating. *Struct Control Health Monit*. 2014;22(3):463–483. doi:[10.1002/stc.1684](https://doi.org/10.1002/stc.1684)
- Oberkampf WL, DeLand SM, Rutherford BM, Diegert KV, Alvin KF. Error and uncertainty in modeling and simulation. *Reliab Eng Syst Saf*. 2002;75(3):333–357. doi:[10.1016/S0951-8320\(01\)00120-X](https://doi.org/10.1016/S0951-8320(01)00120-X)
- Unger JF, Teughels A, Roeck GD. System Identification and Damage Detection of a Prestressed Concrete Beam. *J Struct Eng*. 2006;132(11):1691–1698. doi:[10.1061/\(ASCE\)0733-9445\(2006\)132:11\(1691\)](https://doi.org/10.1061/(ASCE)0733-9445(2006)132:11(1691))
- Oakley J, O'Hagan A. Bayesian inference for the uncertainty distribution of computer model outputs. *Biometrika*. 2002;89(4):769–784. doi:[10.1093/biomet/89.4.769](https://doi.org/10.1093/biomet/89.4.769)
- Volodina V, Challenor P. The importance of uncertainty quantification in model reproducibility. *Philos Trans R Soc A Math Phys Eng Sci*. 2021;379(2197). doi:[10.1098/rsta.2020.0071](https://doi.org/10.1098/rsta.2020.0071)

6. Vanik MW, Beck JL, Au SK. Bayesian Probabilistic Approach to Structural Health Monitoring. *J Eng Mech*. 2000;126(7):738-745. doi:10.1061/(ASCE)0733-9399(2000)126:7(738)
7. Cheung S, Beck J. Bayesian Model Updating Using Hybrid Monte Carlo Simulation with Application to Structural Dynamic Models with Many Uncertain Parameters. *J Eng Mech*. 2009;135(4):243-255. doi:10.1061/(ASCE)0733-9399(2009)135:4(243)
8. Behmanesh I, Moaveni B, Lombaert G, Papadimitriou C. Hierarchical Bayesian model updating for structural identification. *Mech Syst Signal Process*. 2015;64-65:360-376. doi:10.1016/j.ymssp.2015.03.026
9. Ebrahimian H, Astroza R, Conte JP, Callafon dRA Nonlinear finite element model updating for damage identification of civil structures using batch Bayesian estimation. *Mech Syst Signal Pr*. 2017;84:194-222. doi:10.1016/j.ymssp.2016.02.002
10. Sohn H, Law KH. A Bayesian probabilistic approach for structure damage detection. *Earthq Eng Struct Dyn*. 1997;26(12):1259-1281. doi:10.1002/(sici)1096-9845(199712)26:12%3C1259::aid-eqe709%3E3.0.co;2-3
11. Duffin C, Cripps E, Stemler T, Girolami M. Statistical finite elements for misspecified models. *Proc Natl Acad Sci*. 2021;118(2):e2015006118. doi:10.1073/pnas.2015006118
12. Kennedy MC, O'Hagan A. Bayesian calibration of computer models. *J Royal Stat Soc B*. 2001;63(3):425-464. doi:10.1111/1467-9868.00294
13. Sousa H, Rozsas A, Slobbe A, Courage W. A novel pro-active approach towards SHM-based bridge management supported by FE analysis and Bayesian methods. *Struct Infrastruct Eng*. 2020;16(2):233-246. doi:10.1080/15732479.2019.1649287
14. Febrianto E, Butler L, Girolami M, Cirak F. Digital twinning of self-sensing structures using the statistical finite element method, <https://arxiv.org/abs/2103.13729>; 2021.
15. Huang Y, Shao C, Wu B, Beck JL, Li H. State-of-the-art review on Bayesian inference in structural system identification and damage assessment. *Adv Struct Eng*. 2018;22(6):1329-1351. doi:10.1177/1369433218811540
16. PASCO. Structures System. <https://www.pasco.com/products/guides/structures-system>
17. Material properties of polymer used for blue beams—Knowledge Base. <https://www.pasco.com/support/knowledge-base/49>
18. Cross section of beams for structures sets—Knowledge Base. https://www.pasco.com/media/files/static/kb/cross_section.png
19. Leitner SH, Mahnke W. OPC UA—service-oriented architecture for industrial applications. *ABB Corporate Res Center*. 2006;48(61-66):22.
20. Tarantola A, Valette B. Generalized nonlinear inverse problems solved using the least squares criterion. *Rev Geophys*. 1982;20(2):219. doi:10.1029/rg020i002p00219
21. Mottershead JE, Link M, Friswell MI. The sensitivity method in finite element model updating: A tutorial. *Mech Syst Signal Pr*. 2011;25(7):2275-2296. doi:10.1016/j.ymssp.2010.10.012
22. Tarantola A. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for industrial and applied mathematics. 2005.
23. Schoot V dR, Depaoli S, King R, et al. Bayesian statistics and modelling. *Nat Rev Methods Primers*. 2021;1(1):1-26. doi:10.1038/s43586-020-00001-2
24. Skilling J. Nested sampling for general Bayesian computation. *Bayesian Anal*. 2006;1(4):833-859. doi:10.1214/06-ba127
25. Higson E, Handley W, Hobson M, Lasenby A. Dynamic nested sampling: An improved algorithm for parameter estimation and evidence calculation. *Stat Comput*. 2018;29(5):891-913. doi:10.1007/s11222-018-9844-0
26. Speagle JS. dynesty: A dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Mon Not R Astron Soc*. 2020;493(3):3132-3158. doi:10.1093/mnras/staa278
27. Khodadadian A, Noii N, Parvizi M, Abbaszadeh M, Wick T, Heitzinger C. A Bayesian estimation method for variational phase-field fracture problems. *Comput Mech*. 2020;66(4):827-849. doi:10.1007/s00466-020-01876-4
28. Bogdanor MJ, Oskay C, Clay SB. Multiscale modeling of failure in composites under model parameter uncertainty. *Comput Mech*. 2015;56(3):389-404. doi:10.1007/s00466-015-1177-7
29. Rubio PB, Louf F, Chamoin L. Fast model updating coupling Bayesian inference and PGD model reduction. *Comput Mech*. 2018;62(6):1485-1509. doi:10.1007/s00466-018-1575-8
30. Blei DM, Kucukelbir A, McAuliffe JD. Variational Inference: A Review for Statisticians. *J Am Stat Assoc*. 2017;112(518):859-877. doi:10.1080/01621459.2017.1285773
31. Chappell M, Groves A, Whitcher B, Woolrich M. Variational Bayesian Inference for a Nonlinear Forward Model. *IEEE Trans Signal Process*. 2009;57(1):223-236. doi:10.1109/tsp.2008.2005752
32. Titscher T, Robens-Radermacher A, Jafari A, Coelho-Lima I, Klawonn A, Unger J. Bayesian Inference. <https://github.com/BAMresearch/bayem>
33. Koune J. *Bayesian system identification for structures considering spatial and temporal dependencies*, Master's thesis. TU Delft Mechanical, Maritime and Materials Engineering; 2021.
34. Craig M vaby-avb. https://github.com/physimals/vaby_avb; 2021.
35. Torstenfelt B. *Finite elements—from the early beginning to the very end*. LiU-Tryck, Linköping; 2007.

How to cite this article: Titscher T, van Dijk T, Kadoke D, Robens-Radermacher A, Herrmann R, Unger JF. Bayesian model calibration and damage detection for a digital twin of a bridge demonstrator. *Engineering Reports*. 2023;5(11):e12669. doi: 10.1002/eng2.12669

APPENDIX A. VARIATIONAL FREE ENERGY

The free energy is given by Equation (A1), which results from substituting the log likelihood and the posteriors expressions into the free energy definition in Equation (7). The expression in Equation (A1) differs from the one used in the Reference 31, but it is inline with the latest implementations from the same authors of Reference 31, available at Reference 34. A detailed derivation can be found in Reference 32.

$$\begin{aligned}
 F_{vb} = & \left(\frac{N}{2} + c_0 - 1 \right) (\log[s] + \psi(c)) - \frac{1}{2} sc(\mathbf{k}_m^T \mathbf{k}_m + \text{tr}(\Lambda^{-1} \mathbf{J}_k^T \mathbf{J}_k)) \\
 & - \frac{1}{2} ((\mathbf{m} - \mathbf{m}_0)^T \Lambda_0 (\mathbf{m} - \mathbf{m}_0) + \text{tr}(\Lambda^{-1} \Lambda_0)) - \frac{sc}{s_0} \\
 & - \frac{1}{2} \log[\det \Lambda] + \log[\Gamma_c] + c \log[s] + c - (c - 1)(\log[s] + \psi(c)) \\
 & - \frac{N}{2} \log[2\pi] + \frac{1}{2} \log[\det \lambda_0] - \log[\Gamma_{c_0}] - c_0 \log[s_0] + \frac{1}{2} \text{tr}[\Lambda^{-1} \Lambda]
 \end{aligned} \tag{A1}$$

APPENDIX B. BEAM FORMULATION

The full beam element stiffness matrix in 3D

$$\mathbf{K}^e = \begin{bmatrix}
 a_1 & 0 & 0 & 0 & 0 & 0 & -a_1 & 0 & 0 & 0 & 0 & 0 \\
 0 & b_1 & 0 & 0 & 0 & b_2 & 0 & -b_1 & 0 & 0 & 0 & b_2 \\
 0 & 0 & c_1 & 0 & -c_2 & 0 & 0 & 0 & -c_1 & 0 & -c_2 & 0 \\
 0 & 0 & 0 & d_1 & 0 & 0 & 0 & 0 & 0 & -d_1 & 0 & 0 \\
 0 & 0 & -c_2 & 0 & c_3 & 0 & 0 & 0 & c_2 & 0 & c_4 & 0 \\
 0 & b_2 & 0 & 0 & 0 & b_3 & 0 & -b_2 & 0 & 0 & 0 & b_4 \\
 -a_1 & 0 & 0 & 0 & 0 & 0 & a_1 & 0 & 0 & 0 & 0 & 0 \\
 0 & -b_1 & 0 & 0 & 0 & -b_2 & 0 & b_1 & 0 & 0 & 0 & -b_2 \\
 0 & 0 & -c_1 & 0 & c_2 & 0 & 0 & 0 & c_1 & 0 & c_2 & 0 \\
 0 & 0 & 0 & -d_1 & 0 & 0 & 0 & 0 & 0 & d_1 & 0 & 0 \\
 0 & 0 & -c_2 & 0 & c_4 & 0 & 0 & 0 & c_2 & 0 & c_3 & 0 \\
 0 & b_2 & 0 & 0 & 0 & b_4 & 0 & -b_2 & 0 & 0 & 0 & b_3
 \end{bmatrix} \tag{B1}$$

with

$$\begin{aligned}
 a_1 &= EA/L, & d_1 &= GK/L, \\
 b_1 &= 12EI_z/L^3, & b_2 &= 6EI_z/L^2, & b_3 &= 4EI_z/L, & b_4 &= 2EI_z/L \\
 c_1 &= 12EI_y/L^3, & c_2 &= 6EI_y/L^2, & c_3 &= 4EI_y/L, & c_4 &= 2EI_y/L
 \end{aligned} \tag{B2}$$

is taken from Reference 35. In case of a 2D formulation, the corresponding rows and columns related to the 3 dimension are omitted.

APPENDIX C. STEREO SENSOR POSITIONS

Figure C1 shows the location and naming of the stereo measurement nodes, also referred to as stereo sensors. Only the center point of each connection—suffix _01—was analyzed in this work.

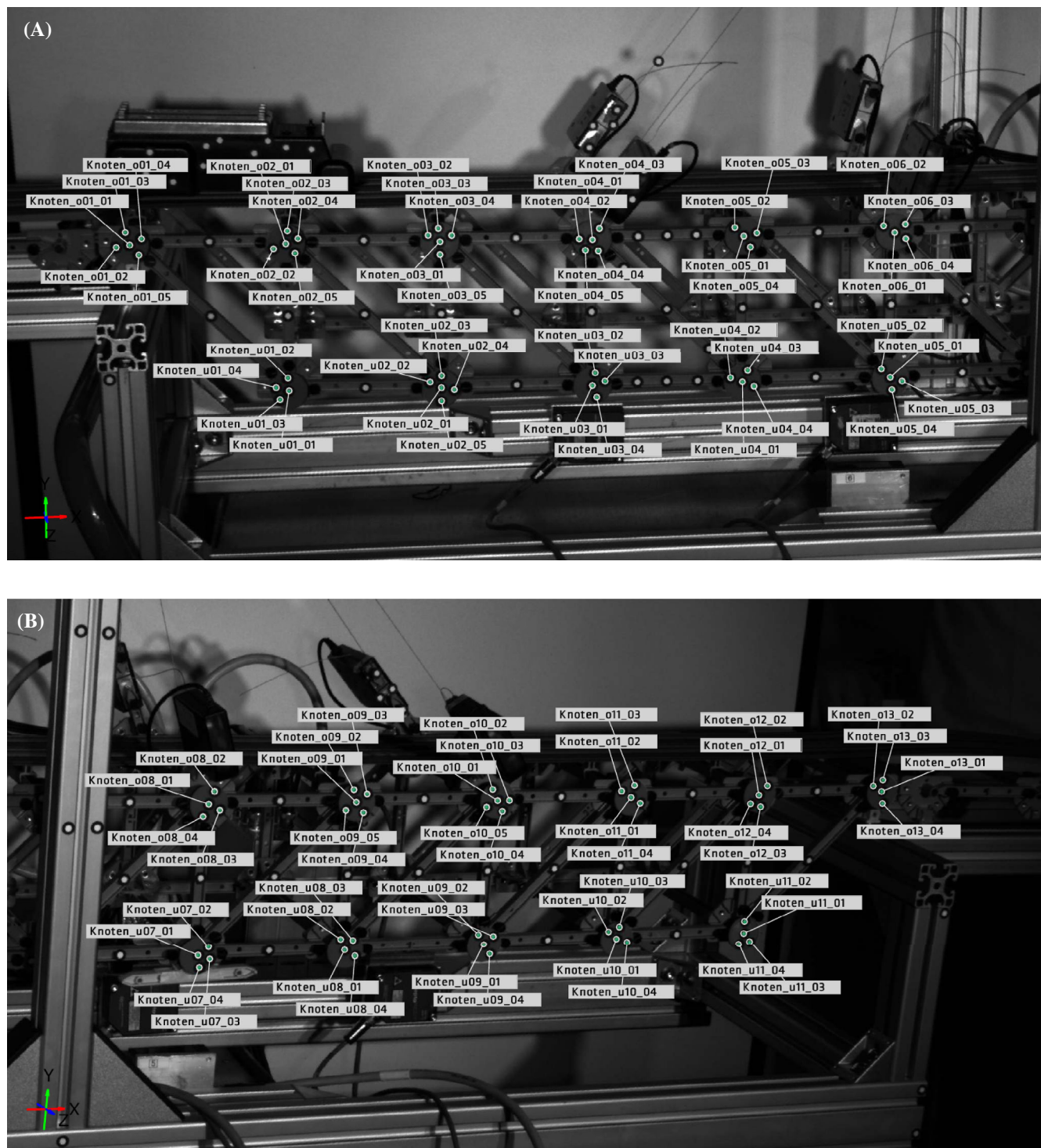


FIGURE C1 Stereo sensor locations. (A) Left side, (B) right side.

APPENDIX D. EXAMPLE NONLINEARITY TEST

To illustrate the linearity measure introduced in Section 3.3, a simple test case is crafted based on the data $\mathbf{y}$ generating process

$$\mathbf{y} = \mathbf{x} (\theta_0 + \theta_1^2 + \theta_2^2 + \theta_3^2) + \mathcal{N}_{21}(0, \sigma^2 \mathbb{I})$$

with $\mathbf{x} = (0, 0.05, 0.1, \dots, 1)$ and $\sigma = 1$,

(D1)

where $\mathbf{x}$ are the 21 measurement points where sensor data is available. The data is generated from the true values $\theta_{\text{true}} = (-4, -1, 1, 4)$ and the nonlinearity is analyzed w.r.t. the distribution $P(\theta) \sim \mathcal{N}_{21}(\theta_{\text{true}}, \text{diag}(1, 2, 1, 1))$ shown in the lower part of Figure D1. Note that this distribution would correspond to the estimated posterior (e.g., after the variational inference has converged) at which the linearity of the model response is to be verified.

The top part of Figure D1 illustrates the linearity analysis, where the individual marks correspond to a column-wise application of the norm. As expected, the linear parameter θ_0 causes both the true and linearized model errors to coincide. This perfect match results in the nonlinearity measure $L_0 = 0$.

For the quadratic parameter θ_1 around the point $\theta_{\text{true},1} = -1$, the true model error remains positive while its linearization changes signs. This significant deviation results in a high linearity measure $L_1 > 1$. Even though θ_2 and θ_3 are also quadratic, their nonlinearity is significantly lower. The mean of the θ_2 distribution is mirrored around the origin, which has no influence on the measure, but its distribution is narrower. The nonlinearity decreases further for θ_3 with its mean further away from 0.

Note that the parameterized multivariate normal distribution can only capture one peak in the posterior distribution. A crafted counter example would be a model error like $k(\theta) = |\theta| - 1$. Given a rather noninformative prior, the posterior would be a bimodal distribution with peaks at -1 and 1 , while VB would—depending on the prior mean—converge to one of them. Then, if the linearity checking range in Equation (17) does not cross the origin, the resulting nonlinearity measure of $L = 0$ would not notice that issue. Thus, a linearity measure of $L \approx 0$ is a good indicator that the VB assumptions are fulfilled, the modelers knowledge about the problem is still required to identify edge cases.

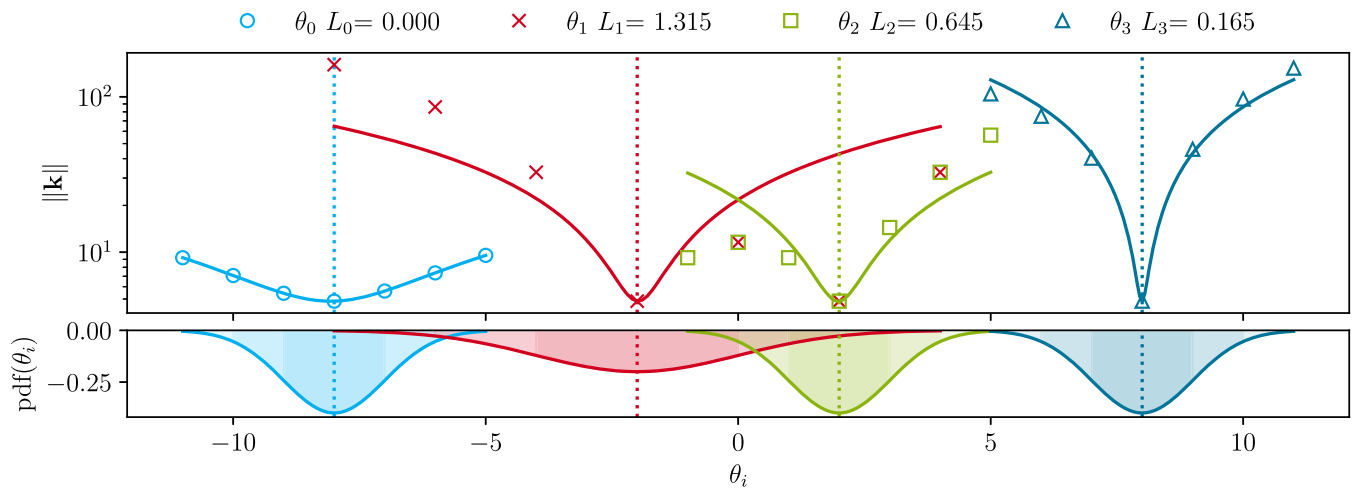


FIGURE D1 Top: comparison of the (norm of the) linearized (solid lines) model responses to the actual ones (marks). Bottom: pdf of the individual parameters.