Structural Identification with Systematic Errors and Unknown Uncertainty Dependencies

James-A. Goulet[a,*], Ian F. C. Smith[a]

[a]*Applied Computing and Mechanics Laboratory (IMAC)*
School of Architecture, Civil and Environmental Engineering (ENAC)
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE (EPFL)
Lausanne, Switzerland

Abstract

When system identification methodologies are used to interpret measurement data taken from structures, uncertainty dependencies are in many cases unknown due to model simplifications and omissions. This paper presents how error-domain model falsification reveals properties of a structure when uncertainty dependencies are unknown and how incorrect assumptions regarding model-class adequacy are detected. An illustrative example is used to compare results with those from a residual minimization technique and Bayesian inference. Error-domain model falsification correctly identifies parameter values in situations where there are systematic errors, and can detect the presence of unrecognized systematic errors.

1. Introduction

It is now common to use in-situ measurements to determine properties of physics-based models in the desire to improve management decisions. During the last decade, measuring instruments evolved to a point where cheap and reliable sensing technologies are now commercially available [13, 25, 26, 37]. In spite of the large number of applications in civil engineering, our capacity to measure structures has outgrown our capacity to interpret data. Assumptions used by current system-identification methodologies are not always satisfied for practical applications [7, 36]. Parameters are usually inferred from models that are approximation of reality. When comparing several predictions with measurements, it is common to have dependencies between the prediction errors at various locations and for various types of predictions. Because of

*Corresponding author: james.a.goulet@gmail.com

the model simplifications and omissions, relationships between prediction errors are in many cases unknown. In this paper, this is referred as *uncertainty dependencies*. Most current methodologies either take uncertainties to be independent, or assume that they have known dependency patterns.

This paper presents how error-domain model falsification can identify properties of a structure without requiring to know the relationship between prediction errors and how wrong assumptions regarding model adequacy can be detected. This paper also underline some limitations of existing methodologies. These methods currently used in civil engineering are described Section 2. Section 3 presents the error-domain model falsification methodology and Section 4 provides an illustrative example comparing system identification methodologies for three scenarios; no systematic bias in model predictions, with recognized systematic bias and with unrecognized errors.

2. Common system identification methodologies in civil-engineering

System identification (SI) is the task of finding descriptions of systems that are compatible with observations. Note that the goal is not to update or to calibrate model parameters to improve the agreement between predicted and measured values. Model-based system identification (SI) uses physics-based models for inferring parameter values. Two approaches used in the context of large-scale structures are presented in this section.

2.1. Residual minimization

Several proposals involve minimizing the residual of the difference between predicted and measured values by adjusting the parameters $\boldsymbol{\theta}$ of a model. $\boldsymbol{\theta}$ is a vector defined in the parameter domain $\Theta \subseteq \mathbb{R}^{n_p}$, where n_p is the number of parameters to identify. One assumption used by residual minimization approaches, is that the difference between predicted and measured values is governed by the choice of parameter values [30]. Furthermore, most approaches are based on the minimization of the sum (weighed or not) of the squares of the differences between predicted ($\mathbf{g}_i(\boldsymbol{\theta})$) and measured ($y_i$) values at location $i \in \{1, \ldots, n_m\}$:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n_m} \mathbf{w}_i \left(\mathbf{g}_i(\boldsymbol{\theta}) - y_i\right)^2 \tag{1}$$

where $\hat{\boldsymbol{\theta}}$ contains the most likely parameter values, $\mathbf{w}_i$ are weighting factors and n_m is the number of measurements used. Mottershead et al. [30] suggested using $\mathbf{w}_i =$

y_i^{-2}. An alternative is to use $\mathbf{w}_i = \sigma^{-2}$ where σ is the standard deviation of the uncertainty for each comparison point i. With this formulation, the optimality of $\hat{\boldsymbol{\theta}}$ holds if the the residuals of the difference between predicted and measured values are distributed as a zero-mean independent Gaussian random distribution $\mathcal{N}(0, \sigma)$. For civil structures, this assumption is seldom met because a model $\mathbf{g}(\boldsymbol{\theta})$ is by definition an approximation of reality. Thus, as noted by Mahadevan et al. [22, 28, 33], the assumption of independence may not be fulfilled due to the systematic bias present in models. Several authors have also argued that while calibrated parameter values may be useful for interpolation, they are usually inappropriate for extrapolation and even less for use in other models [10, 27]. In the case where models are inexact representation of structures, their calibration may not reveal reliable information related to the condition and performance of structures.

2.2. Bayesian inference

Bayesian inference uses conditional probability to update the prior knowledge of model parameters using measurements. Equation 2 is the conditional probability relationship:

$$P(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{y})} \tag{2}$$

where the prior knowledge of physical parameters $P(\boldsymbol{\theta})$ is updated with a likelihood function $P(\boldsymbol{y}|\boldsymbol{\theta})$ using measured data $\boldsymbol{y}$. The posterior probability density function $P(\boldsymbol{\theta}|\boldsymbol{y})$ is obtained using the normalization constant $P(\boldsymbol{y})$. This method creates a mapping (e.g. a function, [40]) between the error domain, $\Xi \subseteq \mathbb{R}^{n_m}$ and the parameter domain Θ. The error domain corresponds to the residuals of the differences between predicted and measured values, $\boldsymbol{\epsilon}_o = [\epsilon_{o,1}, \ldots, \epsilon_{o,n_m}]^T$.

In most formulations reported in the literature, the likelihood function $P(\boldsymbol{y}|\boldsymbol{\theta})$ is based on a Gaussian distribution (see, Equation 3).

$$P(\boldsymbol{y}|\boldsymbol{\theta}) \propto const. \exp\left(-\frac{1}{2}(\mathbf{g}(\boldsymbol{\theta}) - \boldsymbol{y})^T \Sigma^{-1} (\mathbf{g}(\boldsymbol{\theta}) - \boldsymbol{y})\right) \tag{3}$$

In Equation 3, $\mathbf{g}(\boldsymbol{\theta})$ is a vector containing model predictions and $\boldsymbol{y}$ a vector containing measurements. Σ is a covariance matrix containing uncertainties (variances) and correlation coefficients for each comparison point where predicted and measured values are available. This map quantafies likelihood of parameter sets (n_p-dimensional space) based on the error values (n_m-dimensional space).

Early applications of Bayesian inference for identifying the properties of structures were made in the 1990's [2, 4, 23]. Since then, many applications and extensions have been proposed [11, 12, 15, 19, 20, 24, 29, 32, 39, 42–44]. In most applications to civil structures, authors have assumed that modelling uncertainties can be represented by independent Gaussian noise centered on zero.

In the traditional Bayesian inference scheme, there is no requirement to define the uncertainties as independent. When known, correlations can be included in the covariance matrix Σ. Outside of the scope of structural identification, authors such as Cheung et al. [12] included spatially correlated uncertainties during the Bayesian identification of turbulent flow models. In the field of geophysics, Arroyo and Ordaz [3] include spatial decencies in multivariate Bayesian regression analyzes. Also, McFarland and Mahadevan [28] included uncertainty dependencies in model validation and calibration. For measurement uncertainties, they estimated the covariance matrix Σ using experimental data. For model uncertainties, the covariance matrix was evaluated based on the model output covariance obtained from a Monte Carlo analysis. In these studies, uncertainty dependencies due to model simplifications and omissions were not included. Also, Simoen et al. [36] showed that using incorrect uncertainty correlations can lead to biased diagnosis.

3. Error-domain model-falsification

Current system-identification methodologies often assumes that uncertainties and dependencies are known for each comparison point where prediction and measurement are available. These assumptions can be fulfilled when working with models that exactly capture the physical behaviour of the system. In such cases, the discrepancy between model predictions and measurements is due to the choice of parameter values and Gaussian white noise associated with measurements. However, when identifying the parameter values and properties of civil-structures, most models are incomplete.

This section presents the *error-domain model falsification* methodology proposed by Goulet and Smith [17] and used in civil engineering applications [16, 18]. In this approach, model instances are either falsified or kept based on error values (e.g. the difference between predictions and measurements). The term *model instance* refers to a model class evaluated using a particular set of parameter values. The concept of falsification has been well-known in science for centuries. However, it was only in the

1930's that it was formalized by Karl Popper in *The logic of scientific discovery* [31]. Popper asserted that in science, models cannot be fully validated by data, they can only be falsified. Several authors, such as Tarantola [41], Beven [6, 7] and Beck [5], underlined the relevance of this philosophical perspective in fields related to model-based data interpretation. This principle is already used in other methodologies such as the Generalized likelihood uncertainty estimation approach (GLUE) developed by Beven and Binley [9] in the early 1990's in the field of environmental sciences. Like these methodologies, the approach presented here builds on the concept of model falsification mentioned above to provide tools suited for the identification of structures.

3.1. Model falsification

When describing the behaviour of a system, there may be several potentially adequate model classes ($\mathbf{g}(\ldots)$, $\mathbf{h}(\ldots)$, ..., etc.). Model classes take n_p physical parameters, $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_{n_p}]^T$, as arguments, which correspond to the system properties such as geometry, material characteristics, boundary conditions, loading, etc. Each combination of model class and parameter set (i.e. a model instance) leads to n_m predictions $\mathbf{g}_i(\boldsymbol{\theta})$ obtained at each location $i \in \{1, \ldots, n_m\}$. When taking a model class $\mathbf{g}(\ldots)$ and the right values for parameters $\boldsymbol{\theta}^*$, the value corresponding to the difference between a prediction $\mathbf{g}_i(\boldsymbol{\theta}^*)$ and its modeling error ($\epsilon^*_{model,i}$) is equal to the *"true"* value $\mathcal{Q}_i$ for the real system. The *"true"* value of $\mathcal{Q}_i$ is also equal to the difference between the measured value y_i and the measurement error ($\epsilon^*_{measure,i}$). This relation is presented in Equation 4.

$$\mathbf{g}_i(\boldsymbol{\theta}^*) - \epsilon^*_{model,i} = \mathcal{Q}_i = y_i - \epsilon^*_{measure,i} \tag{4}$$

In practice, neither the true value of $\mathcal{Q}_i$ nor error values ϵ^*_i are known precisely. Only a probability density function (*pdf*) describing possible errors ϵ_i can be estimated. $f_{U_i}(\epsilon_i)$ represents the *pdf* of a continuous random variable U_i. $f_{U_{c,i}}$ is the *pdf* describing the probability of the *observed residuals* $\epsilon_{o,i}$ of differences between predicted and measured values. Using Equation 4, $U_{c,i}$ is obtained by subtracting the modeling $U_{model,i}$ and measurement $U_{measure,i}$ uncertainties. The *pdf* of $U_{c,i}$ is presented in Figure 1. Note that random variables are used here to describe the outcome of either stochastic or deterministic processes.

A set of parameter values is falsified if the difference between its predicted and measured values is outside the interval defined by threshold bounds $[T_{low,i}, T_{high,i}]$

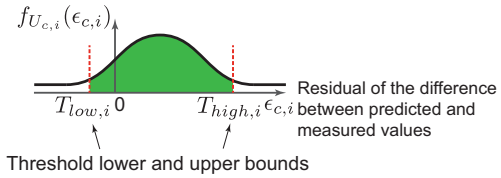


Figure 1: Model falsification - The combined probability density function describes the outcome of the random variable $U_{c,i}$. Threshold bounds are calculated to be the narrowest intervals $[T_{low,i}, T_{high,i}]$ that contains a target identification probability ϕ for all comparison points.

for any comparison point i. These bounds represent the shortest interval that satisfies the relation expressed in Equation 5 such that

$$\forall i \in \{1, \ldots, n_m\}, \left\{ T_{low,i}, T_{high,i} : \phi^{1/n_m} = \int_{T_{low,i}}^{T_{high,i}} f_{U_{c,i}}(\epsilon_{c,i}) d\epsilon_{c,i} \right\} \quad (5)$$

where $\phi \in]0, 1]$ is the target identification probability defined by the user. When $U_{c,i}$ is described by an unimodal symmetric *pdf*, threshold bounds can be computed using:

$$\begin{aligned} T_{low,i} &= F^{-1}_{U_{c,i}} \left(\frac{1}{2}(1 - \phi^{1/n_m}) \right) \\ T_{high,i} &= F^{-1}_{U_{c,i}} \left(1 - \frac{1}{2}(1 - \phi^{1/n_m}) \right) \end{aligned} \quad (6)$$

where $F^{-1}(x) : \; x \in [0, 1] \rightarrow \mathbb{R}$ represents the inverse cumulative distribution function. In these equations, the target probability ϕ is adjusted using the Šidák correction [34, 35] to account for the usage of multiple measurements to falsify model instances. Threshold bounds define the limit of a n_m-dimensions *hyper-rectangular* domain that has a probability larger or equal to ϕ of containing the correct residuals $\epsilon^*_{c,i}$ between predicted and measured values:

$$P(\cap_{i=1}^{n_m} T_{low,i} \leq U_{c,i} \leq T_{high,i}) \geq \phi \quad (7)$$

If this criteria is used to falsify model instances, there is a probability larger or equal to ϕ of not wrongly falsifying valid model instances, irrespective of the relationships between residuals $\epsilon_{c,i}$. Model instances are falsified if they do not satisfy the inequalities

$$\forall i \in [1, \ldots, n_m] : \; T_{low,i} \leq \mathbf{g}_i(\boldsymbol{\theta}) - y_i \leq T_{high,i} \quad (8)$$

Model instances that are not falsified are considered to be *candidate models* and are all considered as equal in the sense that they are all possible explanations of the observed behavior. A model class $\mathbf{g}(\ldots)$ is falsified if all possible sets of parameter values $\boldsymbol{\theta}$

are falsified by observations. This allows the falsification of a model class without requiring a comparison with others. When a model class is falsified, it is generally an indication that there are flaws in initial assumptions [8].

Figure 2 illustrates the concepts of threshold definition when using two measurements. In this figure, threshold bounds $T_{low,i}$ and $T_{high,i}$ are found separately for each comparison point to include a probability $\phi^{1/2}$ where $\phi = 0.95$. When these threshold bounds are projected on the bi-variate *pdf*, they define a rectangular boundary used to separate candidate and falsified models. This region include a probability ϕ. Thus, when used as criterion to falsify model instance, this rectangular region has a probability larger or equal to ϕ of correctly identifying the right model irrespectively of the relationship between the error at comparison point #1 and #2.

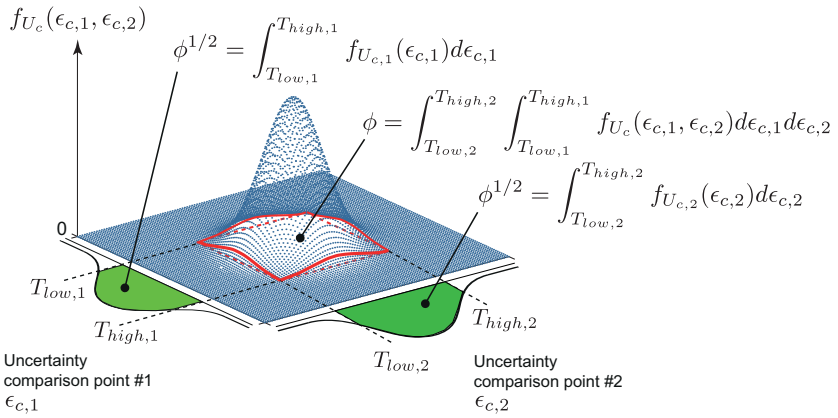


Figure 2: Threshold definition using two measurements to falsify model. The probability content included in threshold bounds is of $\phi^{1/2}$ for each comparison point taken individually and of ϕ for the bi-variate *pdf* of $U_{c,1}$ and $U_{c,2}$.

3.2. Comparison of rectangular and ellipsoidal threshold bounds

This example shows how error-domain model falsification can lead to probabilistically valid diagnostics irrespectively of the relationships between the outcome of random variables. Figure 3 compares the probability contained in rectangular and ellipsoidal threshold bounds commonly used in multivariate hypothesis testing [28, 33, 38]. In this example, realizations of a bi-variate random variable $X \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ are generated for a mean $\boldsymbol{\mu} = [0, 0]^T$ and a variance $\boldsymbol{\sigma} = [1, 1]^T$. In plots presented in Figure 3 a), b) and c), the correlation coefficient defining this random variable is set respectively to 0.9, 0.4 and -0.9. For a normal random variable, the smallest region possible

including a probability ϕ is bounded by the *Mahalanobis distance* $D_M(\boldsymbol{x})$:

$$D_M(\boldsymbol{x})^2 = (\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \tag{9}$$

where $\boldsymbol{x}$ is a vector containing a realization of X. Even if the size of the region defined by the Mahalanobis distance is minimal, its computation requires the definition of the correlation coefficients in Σ. In order to calculate the Mahalanobis distance, the correlation is set to 0.9 for all three scenarios a), b) and c), where only for scenario a) this choice is correct.

The probability that realizations of X are include in the ellipsoidal and rectangular region, P_M and P_T respectively, is expressed in Equations 10 and 11:

$$P_M = P(D_M(\boldsymbol{x})^2 \leq \chi^2_\alpha(n_m)) \tag{10}$$

$$P_T = P(\cap_{i=1}^{n_m} T_{low,i} \leq x_i - \mu_i \leq T_{high,i}) \tag{11}$$

In Equation 10, $\chi^2_\alpha(n_m)$ is the value of a chi-squared distribution having n_m degrees of freedom, found for a target cumulative probability $\alpha = 1 - \phi = 0.05$. Threshold bounds T_{low} and T_{high} are found using Equation 6 and a target reliability $\phi = 0.95$.

For each scenario a), b) and c), 1000 realizations $\boldsymbol{x} = [x_1, x_2]^T$ of X are generated. The ellipsoidal bound, represented by the solid line (e.g. the Mahalanobis distance), includes a proportion $P_M = 0.95$ of the realizations of X, only when the correlation is correctly evaluated (e.g. scenario a)). In Figure 3 b) and c), the ellipsoidal bound (supposed to include 95% of the realizations of X) only include 64% and 42% respectively. For all three scenarios, the rectangular threshold bounds represented by dashed lines, contains a proportion P_T of the realizations of X at least equal to the target ϕ. Defining rectangular threshold bounds using the approach presented above is conservative with respect to the target probability ϕ for any dependent variable [21, 34]. This property enables error-domain model falsification to reveal the condition and properties of civil structure, where uncertainty dependencies are unknown.

3.3. Summary

Figure 4 summarizes the steps leading to the falsification of model instances or of a model class. The first step is to define the goal of the identification and to convert it into parameters to identify ($\boldsymbol{\theta}$) using a model class ($\mathbf{g}(\boldsymbol{\theta})$) and measurements ($\boldsymbol{y}$). Uncertainties associated with both the model and measurements are combined and used to define threshold bounds (T_{low}, T_{high}) including a target probability ϕ^{1/n_m}. Model

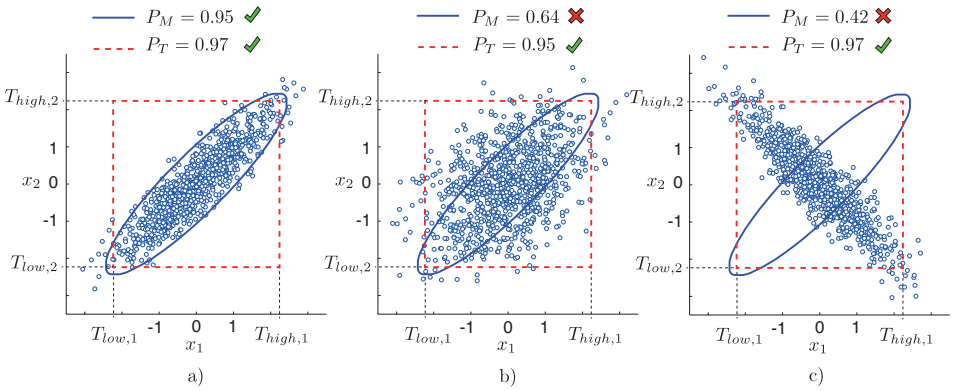


Figure 3: Comparison of the probability contained in rectangular and ellipsoidal threshold bounds when varying the correlation between random variables. In all three cases, the correlation assumed to compute the the ellipsoidal bound is set to 0.9. However, the correlation used to generate realizations of X is a) 0.9, b) 0.4 and c) -0.9. For all three scenarios, the rectangular region includes a proportion of the realisation of X $P_T \geq \phi$.

instances are generated and each predicted value is compared with measured values. If for any comparison point, the difference $\mathbf{g}_i(\boldsymbol{\theta}) - y_i$ is outside threshold bounds, the model instance is falsified. This process is repeated for all model instances and those that are not falsified are part of the candidate model set. The candidate model set contains plausible explanations of the observed behavior given uncertainties in models and measurements. If all model instances are falsified, it indicates that the model class $\mathbf{g}(\ldots)$ is also falsified by measured data. In these circumstances, it is possible to review the model class and assumptions initially chosen. When candidate models are obtained, it is possible to take additional measurements to falsify more instances or to review initial objectives.

4. Illustrative example

An illustrative example is presented to compare identification methodologies presented previously. This example is tailored to demonstrate the effect on diagnosis reliability, of using models that are idealized representations of reality for identifying the properties of structures. The system studied here is a cantilever beam. The "true" representation of the structure is presented in Figure 5(a), where the semi-rigid cantilever connection is modeled using a rotational spring having a stiffness parameter K. This beam has a Young's modulus $E^* = 70 \times 10^3$ MPa and the vertical force applied on its end is $F = 5 \times 10^3$ N.

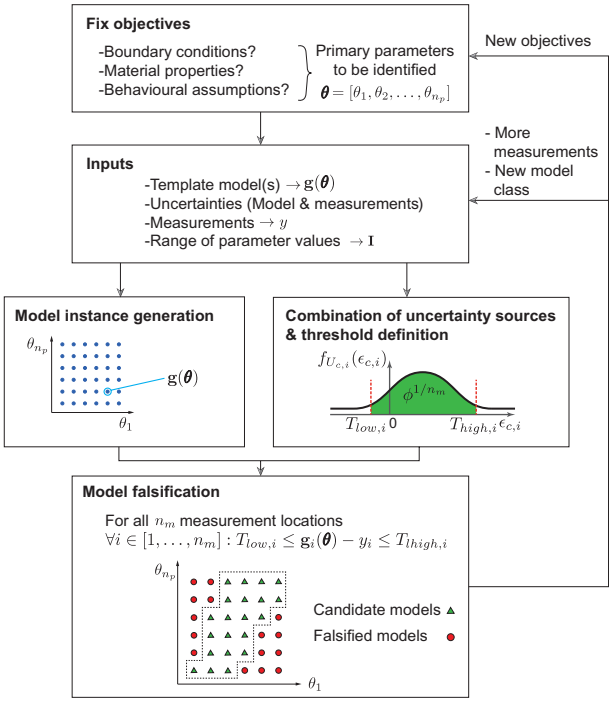
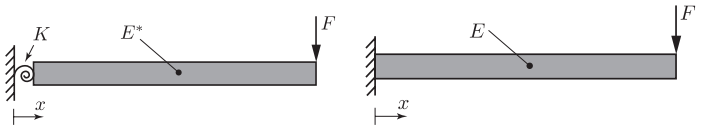


Figure 4: Flowchart describing error-domain model falsification

In order to be representative of full-scale situations where it is not possible to capture reality in a model, an idealized beam is used as model of the "true" system (see Figure 5(b)). This idealized structure does not include the partial rigidity of the cantilever connection. For this structure, the parameter to be identified is the Young's modulus E which has a possible range of values of [20,100]$\times 10^3$ MPa. To be consistent with the notation used previously, the parameter to be identified is denoted $\boldsymbol{\theta} = E$. The beam is 3000 mm long and has a square cross-section of 300 mm $\times$ 300 mm. Its inertia I is 6.75$\times 10^8$ mm^4.



(a) True cantilever beam (b) Idealized cantilever beam

Figure 5: True and idealized cantilever beams. The parameter to be identified using the idealized beam is the Young's modulus (E).

The vertical displacement $v(x)$ of the beam at any location $x \in [0, l]$ (l=3000 mm)

is described by

$$v(x) = \frac{Fx^2(3l - x)}{6EI} \tag{12}$$

For any location x the error introduced by the idealized model is $\epsilon(x) = -Flx/K$. Simulated measured values, $y(x)$, are obtained according to

$$y(x) = v^*(x) - \epsilon(x) + u_{meas} \tag{13}$$

where $v^*(x)$ is the displacement computed with correct parameter value E^* and where u_{meas} is a realization of $U_{meas} \sim \mathcal{N}(\mu_{meas}, \sigma_{meas})$, a Gaussian random variable describing sensor resolution uncertainty. The mean of this random variable is 0 and its standard deviation is 0.02 mm. Sensor resolution errors are independent of the measured locations. The combined uncertainty variance σ_c^2 is obtained by summing the variance of the model simplification and sensor resolution uncertainties ($\sigma_c^2 = \sigma_{model}^2 + \sigma_{meas}^2$). The combined uncertainty is represented by a random variable denoted $U_c \sim \mathcal{N}(\mu_c, \sigma_c)$.

4.1. Comparison of system-identification approaches

The effect of using an idealized model is studied for three identification methodologies; residual minimization using weighed least-square regression, Bayesian inference and error-domain model falsification. These three approaches are compared for several cases using different number of measurements: $n_m = \{1, 2, 10, 50\}$. For each value of n_m, the displacement is evaluated for $x = x_{start} + i \cdot (l - x_{start})/n_m \; \forall i \in \{1, \ldots, n_m\}$, using Equation 13. x_{start} is the minimal distance from the cantilever support where measurements are taken. In this example, $x_{start} = 500$ mm. Also, these approaches are compared for three scenarios, firstly for a case where there is no systematic error, secondly for a case where there are systematic errors and these are recognized and thirdly for a case where there are unrecognized errors. For the two first scenarios, a diagnostic is deemed correct if it accurately identifies the value E^*. For the third scenario, the identification is correct if either it accurately identifies the value for E^* or if it returns an empty set indicating that the entire model class is falsified.

For each approach, the domain of possible solutions Θ is explored by solving the model class $\mathbf{g}(\ldots)$ for each parameter set $\boldsymbol{\theta}$ corresponding to the possible range for this parameter subdivided in 100. For each of these model instances, a vector $\boldsymbol{\epsilon_o}(\boldsymbol{\theta})$ contains the observed residual of the difference between predicted and measured values for all measurement locations.

$$\boldsymbol{\epsilon_o}(\boldsymbol{\theta}) = \mathbf{g}(\boldsymbol{\theta}) - \boldsymbol{y} \tag{14}$$

Residual minimization

The first identification approach finds parameters $\hat{\boldsymbol{\theta}}$, that are optimal in a least-squares sense:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \boldsymbol{\epsilon_o}(\boldsymbol{\theta})^T \boldsymbol{W} \boldsymbol{\epsilon_o}(\boldsymbol{\theta}) \tag{15}$$

The weighing matrix $\boldsymbol{W}$ is set to $[\text{diag}(\boldsymbol{y})]^{-2}$. In this approach, the goal is to calibrate model parameters to obtain the smallest weighted sum of the square of the residuals. Uncertainties are assumed to be centred on zero, Gaussian and independent. The result is a single optimal parameter value $\hat{\boldsymbol{\theta}}$.

Bayesian inference

The second identification approach is used to obtain the posterior *pdf* describing what the correct parameter values should be. Here, it is assumed that the only prior knowledge available is the minimal and maximal bound for parameters to be identified. Thus, prior knowledge is represented by a uniform distribution as described in Equation 16.

$$P(\boldsymbol{\theta}) = \begin{cases} \text{constant}, & \text{if } \boldsymbol{\theta} \in \Theta \\ 0, & \text{if } \boldsymbol{\theta} \notin \Theta \end{cases} \tag{16}$$

The function mapping the residual values to likelihood of parameter values is

$$P(\boldsymbol{y}|\boldsymbol{\theta}) = (2\pi)^{-n_m/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (\boldsymbol{\epsilon_o}(\boldsymbol{\theta}) - \overline{\boldsymbol{U}}_c)^T \Sigma^{-1} (\boldsymbol{\epsilon_o}(\boldsymbol{\theta}) - \overline{\boldsymbol{U}}_c) \right) \tag{17}$$

where $\overline{\boldsymbol{U}}_c$ is a vector containing the mean of the combined uncertainty *pdf* for each location i. Equation 17 is the multivariate Gaussian distribution. The normalization constant $P(\boldsymbol{y})$ (see Equation 2) is computed by integrating $\int_\Theta P(\boldsymbol{y}|\boldsymbol{\theta}) P(\boldsymbol{\theta}) d\Theta$. For this method, uncertainties are assumed to be independent. Thus, the covariance matrix Σ is a diagonal matrix containing the variance $\sigma^2_{c,i}$ for each comparison point i where measured and predicted values are available. The result is a posterior *pdf* describing the probability of each parameter value E.

Error-domain model falsification

The third approach compared is error-domain model falsification. Here, the target reliability is set to $\phi = 0.95$ and threshold bounds are obtained using Equation 6. Model instances are either accepted as candidate or falsified based on Equation 8. The

result is an interval describing what the true parameter value E^* can be, given the modelling and measurement uncertainties.

4.1.1. First scenario: No systematic errors

In the first scenario, the true value for the rotational stiffness $K \to \infty$, so there is no systematic error between the simulated measurements and model predictions obtained using E^* as parameter value. The only uncertainty is due to sensor resolution. Figure 6 compares the optimal parameter values $\hat{\boldsymbol{\theta}}$, the Bayesian posterior *pdf* and the candidate model set, with true parameter value E^*. Each graph represents the results obtain for a number of measurements n_m. The horizontal axis corresponds to the values for E and the vertical axis is the relative frequency of the posterior *pdf* histogram.

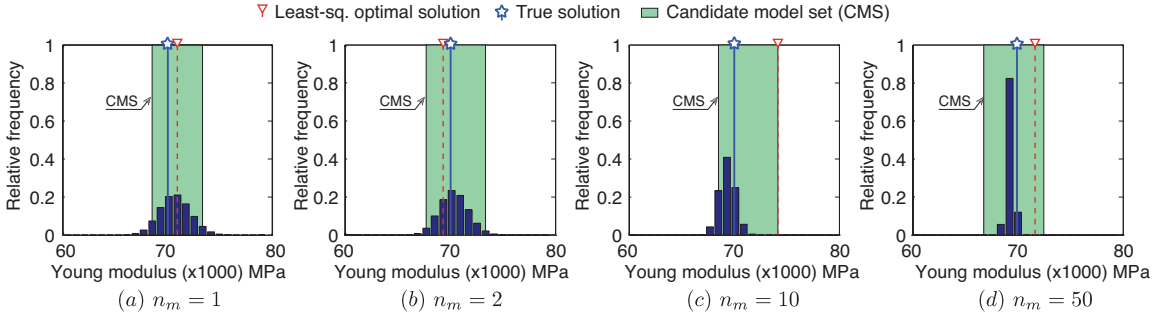


Figure 6: Comparison of parameter values identified using least-squares parameter identification, Bayesian inference and error-domain model falsification with the correct parameter value for Young's modulus (E^*). The number of measurements varies from 1 (a) to 50 (d). In this scenario, there are no systematic errors and uncertainties are rightfully assumed to be independent.

For any value of $n_m \in \{1, 2, 10, 50\}$, the least-squares optimal solution provides estimates of $\hat{E}$ are close to the true value E^*. However, in all cases $\hat{E} \neq E^*$. For Bayesian inference, the true parameter values is correctly identified by the posterior *pdf* for all values of n_m. Note that the posterior *pdf* becomes narrower as n_m increases. This indicates that the more measurements are used, the more precise the identification is, because each new measurement brings additional information.

The shaded region represents the candidate models identified by error-domain model falsification and the white region represents falsified model instances. Analogously to Bayesian inference, the identification is correct for any value of n_m. For the first scenario where no systematic errors are present, the three approaches lead to correct identifications that are sufficiently accurate for engineering purposes, considering that the initial parameter range was from 20 to 100 GPA.

4.1.2. Second scenario: With recognized systematic errors

In the second scenario, systematic errors are introduced by setting the true value K for the rotational stiffness to 12×10^{10} N·mm/rad. The magnitude of this systematic bias is chosen to represent situations governed by modelling errors rather than by measurement errors. For the purpose of this illustrative example, the effect of model simplifications is described by a Gaussian distribution having a mean $\overline{U}_{model}$ corresponding to -10% of the measured values and a coefficient of variation of 15%. This estimation of modelling uncertainties is intended to be a conservative assumption where the correct residual values lie within the interval $\overline{U}_c \pm 2\sigma_c$. Here, the model simplification uncertainty is not centered on zero because it is known that the simplification related to the beam connection is likely to cause an underestimation of predicted displacement (in absolute terms). In order to represent the situations that civil engineers face during the identification of full-scale structures, it is assumed that there is no information available to quantify the effect of spatial dependencies of prediction errors.

Figure 7 compares the identification results with true parameter value E^* for a number of measurement $n_m = \{1, 2, 10, 50\}$. Least-squares residual minimization

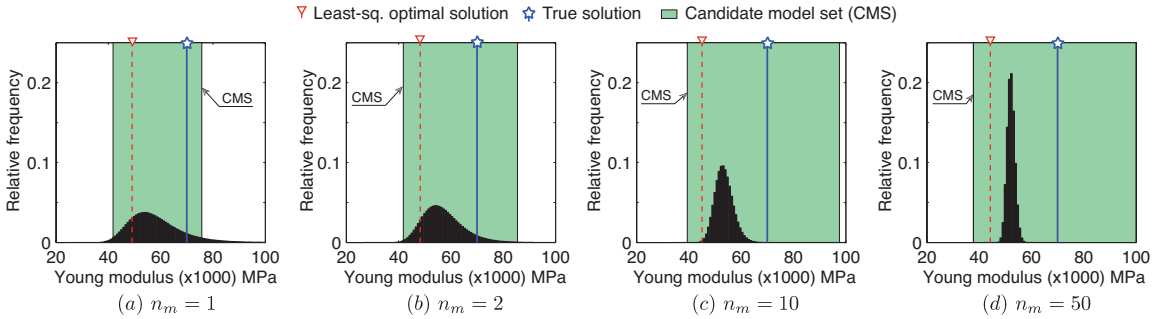


Figure 7: Comparison of parameter values identified using least-squares parameter identification, Bayesian inference and error-domain model falsification with the correct parameter value for Young's modulus (E^*). The number of measurements varies from 1 (a) to 50 (d).

fails to provide correct identifications for any value of n_m. Even if the posterior *pdf* is strongly biased, an acceptable identification is obtained for $n_m = \{1, 2\}$. For larger number of measurements, the correct parameter value E^* is probabilistically excluded. Furthermore, the more measurements are used, the narrower the posterior *pdf* becomes, such that the importance of the identification error increases with the number of measurements used. This could have lead to the belief that the identification is correct because of its high precision. Thus, even if the magnitude of uncertainties are ade-

quately estimated, if wrong assumptions are made regarding uncertainty dependencies, the posterior *pdf* obtained using Bayesian inference can be biased.

For error-domain model falsification, correct identifications are achieved for all values of n_m. Note that the size of the candidate model set increases with the number of measurement used because errors are strongly dependent. Therefore, each new measurement brings almost no new information to further discard model instances. When more measurements are included, threshold bounds are widened according to Equation 5, to include the effect of unknown uncertainty dependencies.

This example illustrates that using wrong values of uncertainty correlation with Bayesian inference may lead to biased identifications. It also shows that error-domain model falsification can achieve correct identifications without having to define uncertainty dependencies.

4.1.3. Third scenario: With unrecognized errors

The last scenario compares the identification performance when model simplification errors are present and unrecognized. Here, the true value K for the rotational stiffness is set to 12×10^{10} N·mm/rad as in the previous scenario. However, in this case, modelling uncertainties are set to 0 to represent unrecognized systematic errors. Only measurement errors have a standard deviation of 0.02 mm.

Figure 8 compares identification results for Bayesian inference, weighed least-squares residual minimization and error-domain model falsification with true parameter value E^*, for a number of measurements $n_m = \{1, 2, 10, 50\}$. It shows that for

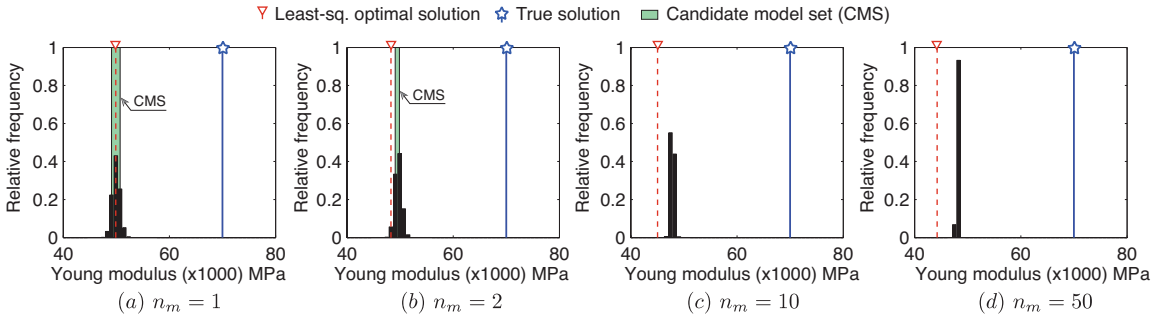


Figure 8: Comparison of parameter values identified using least-squares parameter identification, Bayesian inference and error-domain model falsification with the correct parameter value for Young's modulus (E^*). The number of measurements varies from 1 (a) to 50 (d). The systematic bias in model predictions are not recognized.

any number of measurements n_m, Bayesian inference and weighed least-squares residual minimization lead to biased identifications. As in the previous case, the posterior

pdf gets narrower when the number of measurements increases. Again, it can lead to the belief that the identification is correct because the identification results are highly precise. These approaches are unable to signal that initial assumptions regarding the model adequacy were wrong.

For error-domain model falsification, when one or two measurement locations are used, the approach leads to biased identifications because it finds candidate model sets that do not include the correct solution. For any values $n_m > 2$, the approach leads to a correct identification by returning no candidate model. It identifies that the initial assumptions made regarding the model adequacy were wrong. Therefore, given a sufficient number of measurements, error-domain model falsification can identify when initial assumptions are flawed.

4.2. Summary of results

Table 1 summarizes the comparison of identification methodologies. Only error-domain model falsification leads to a correct identification for all scenarios tested, provided that a sufficient number of measurements are used. Bayesian inference using a

Table 1: Summary of the identification methodology comparison on the basis of their capacity to provide correct identification for the cantilever beam example.

	Residual minimization	Bayesian inference	Error-domain model falsification
No systematic errors	✓	✓	✓
recognized systematic errors	✖	✖[1]	✓
Unrecognized errors	✖	✖	✓[2]

[1] When uncertainty dependencies assumed to be independent
[2] Provided that a sufficient number of measurements are used.

likelihood function based on Gaussian distribution, lead to correct identifications when systematic effects can be either fully described or removed. When Bayesian inference is used with incorrect uncertainty dependency values, the approach may return biased posterior *pdf*s. For all scenarios involving model simplifications, considering more than one solution as a candidate model is mandatory to obtain an unbiased identification.

In traditional Bayesian inference, the diagnosis becomes more precise as measurements are added. However, as illustrated in this paper, if wrong choices of uncertainty correlation are made, the diagnosis will be precise but biased. The falsification approach presented in the paper is not affected by this drawback because it does not require to know the relationship between prediction errors.

5. Discussion & limitations

For the previous example, it is trivial to quantify the dependencies introduced by model simplifications and to parametrize the boundary conditions as it has already been proposed [1, 14]. However, for full-scale civil-structures, model simplifications are inevitable and the relationship between errors are in most cases unquantifiable. Thus, most full-scale identification tasks correspond to either the second or third scenario where systematic bias is introduced by model simplifications and it may or may not be recognized.

For real structures, model simplifications are usually related to the omission of load-carrying elements such as pedestrian barriers on bridges, connection details and secondary load-carrying elements such as stiffeners and gusset plates. Additional simplifications commonly come from the representation of boundary conditions and loads at discrete locations rather than over distributed surfaces, from simplifications related to the soil-structure interaction and from finite-element approximations. All of these simplifications may systematically affect the static and dynamic predictions at several locations and to varying degrees. Figure 9 presents an example of bridge model where even state-of-the-art practices lead to omissions and simplifications in the representation of reality. Quantifying dependencies between prediction errors is a difficult task and there is no conservative value for these dependencies.

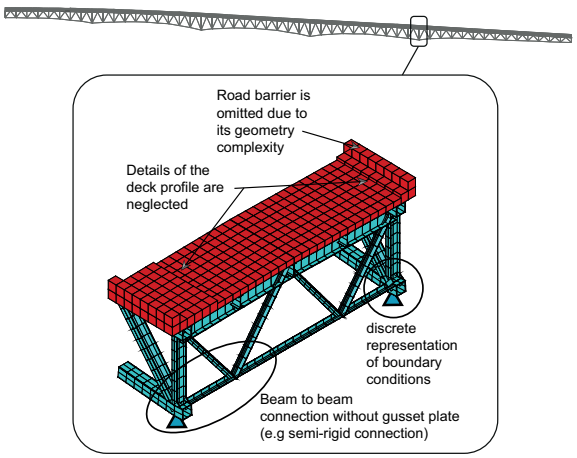


Figure 9: Example of model simplifications on a full-scale civil-structure. This bridge is a composite steel truss structure. The level of detail involved in the system is too important to be fully captured by a model.

In its current form, the falsification methodology presented may be computationally

demanding when there are a large number of model parameters to identify. The use of feature selection methods and more efficient sampling techniques is currently under study.

6. Conclusions

The following conclusions are drawn:

- Error-domain model falsification provides correct identifications in situations where there are aleatory and systematic errors, without requiring to know the relationships between prediction errors. Furthermore, incorrect assumptions related to model relevance may be detected through model-class falsification.
- With Bayesian inference, when there are aleatory and systematic errors, assuming that uncertainties are independent may bias the posterior *pdf*.
- When taken out of the scope of model calibration, residual minimization might lead to biased identifications, especially when multiple possible solutions are not included.

Acknowledgements

Parts of this research is funded by the Swiss National Science Foundation under contract no. 200020-117670/1. Authors would like to thank anonymous reviewers, Prof. J.E. Mottershead and Prof. K.J. Beven for their comments on the manuscript.

[1] H. Ahmadian, J.E. Mottershead, and MI Friswell. Physical realization of generic-element parameters in model updating. *Journal of vibration and acoustics*, 124:628, 2002.

[2] K.F. Alvin. Finite element model update via Bayesian estimation and minimization of dynamic residuals. Technical report, Sandia National Labs, Albuquerque, NM, 1996.

[3] D. Arroyo and M. Ordaz. Multivariate Bayesian regression analysis applied to ground-motion prediction equations, part 1: Theory and synthetic example. *Bulletin of the Seismological Society of America*, 100(4):1551–1567, 2010.

[4] J.L. Beck and L.S. Katafygiotis. Updating models and their uncertainties. i: Bayesian statistical framework. *Journal of Engineering Mechanics*, 124(4):455–461, 1998.

[5] M. B. Beck. Water quality modeling: A review of the analysis of uncertainty. *Water Resour. Res.*, 23(8):1393–1442, 1987.

[6] K. J. Beven. Towards a coherent philosophy for modelling the environment. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 458:1–20, 2002.

[7] K. J. Beven. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1-2):18–36, 2006.

[8] K. J. Beven. *Environmental modelling: an uncertain future?* Routledge, New-York, 2009.

[9] K. J. Beven and A. Binley. Future of distributed models: Model calibration and uncertainty prediction. *Hydrological processes*, 6(3):279–298, 1992.

[10] K.J. Beven. Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, 4(2):203–213, 2000.

[11] S.H. Cheung and J.L. Beck. Calculation of posterior probabilities for Bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Computer-Aided Civil and Infrastructure Engineering*, 25(5):304–321, 2010.

[12] S.H. Cheung, T.A. Oliver, E.E. Prudencio, S. Prudhomme, and R.D. Moser. Bayesian uncertainty analysis with applications to turbulence modeling. *Reliability Engineering & System Safety*, 96(9):1137–1149, 2011.

[13] C.R. Farrar, G. Park, D.W. Allen, and M.D. Todd. Sensor network paradigms for structural health monitoring. *Structural control and health monitoring*, 13(1):210–225, 2006.

[14] GML Gladwell and H. Ahmadian. Generic element matrices suitable for finite element model updating. *Mechanical Systems and Signal Processing*, 9(6):601–614, 1995.

[15] B. Goller, J.L. Beck, and G.I. Schuëller. Evidence-based identification of weighting factors in Bayesian model updating using modal data. *Journal of Engineering Mechanics*, in press, 2012.

[16] J.A. Goulet, C. Michel, and I.F.C. Smith. Hybrid probabilities and error-domain structural identification using ambient vibration monitoring. *Mechanical Systems and Signal Processing*, 37(1–2):199–212, 2013.

[17] J.A. Goulet and I.F.C. Smith. Uncertainty correlation in structural performance assessment. In *Proceedings of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering*, pages 2801–2806, Zürich, Switzerland, 2011.

[18] J.A. Goulet and I.F.C. Smith. Performance-driven measurement-system design for structural identification. *Journal of Computing In Civil Engineering*, 27(4):427–436, 2013.

[19] R. Hadidi and N. Gucunski. Probabilistic approach to the solution of inverse problems in civil engineering. *Journal of Computing In Civil Engineering*, 22(6):338–347, 2008.

[20] T. Haukaas and P. Gardoni. Model uncertainty in finite-element analysis: Bayesian finite elements. *Journal of Engineering Mechanics*, 137(8):519–526, 2011.

[21] JCGM. *Evaluation of measurement data – Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities*, volume JCGM 102:2011. JCGM Working Group of the Expression of Uncertainty in Measurement, 2011.

[22] X. Jiang and S. Mahadevan. Bayesian validation assessment of multivariate computational models. *Journal of Applied Statistics*, 35(1):49–65, 2008.

[23] L.S. Katafygiotis and J.L. Beck. Updating models and their uncertainties. II: Model identifiability. *Journal of Engineering Mechanics*, 124(4):463–467, 1998.

[24] M.C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[25] J.M. Ko and Y.Q. Ni. Technology developments in structural health monitoring of large-scale bridges. *Engineering structures*, 27(12):1715–1725, 2005.

[26] J.P. Lynch and K.J. Loh. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest*, 38(2):91–130, 2006.

[27] J. McFarland. *Uncertainty analysis for computer simulations through validation and calibration*. PhD thesis, Vanderbilt University, Nashville, Te, 2008.

[28] J. McFarland and S. Mahadevan. Multivariate significance testing and model calibration under uncertainty. *Computer Methods in Applied Mechanics and Engineering*, 197(29-32):2467–2479, 2008.

[29] T. Most. Assessment of structural simulation models by estimating uncertainties due to model selection and model simplification. *Computers & Structures*, 89(17-18):1664—1672, 2011.

[30] J.E. Mottershead, M. Link, and M.I. Friswell. The sensitivity method in finite element model updating: A tutorial. *Mechanical Systems and Signal Processing*, 25(7):2275–2296, 2011.

[31] K.R. Popper. *The logic of scientific discovery*. Routledge, New-York, third edition, 2002.

[32] R. Rebba and S. Mahadevan. Model predictive capability assessment under uncertainty. *AIAA Journal*, 44(10):2376–2384, 2006.

[33] R Rebba and S. Mahadevan. Validation of models with multivariate output. *Reliability Engineering & System Safety*, 91(8):861–871, 2006.

[34] Z. Šidák. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62:626–633, 1967.

[35] Z. Šidák. On probabilities of rectangles in multivariate student distributions: their dependence on correlations. *The Annals of Mathematical Statistics*, 42:169–175, 1971.

[36] E. Simoen, C. Papadimitriou, G. De Roeck, and G. Lombaert. Influence of the prediction error correlation model on bayesian FE model updating results. In *Life-Cycle and Sustainability of Civil Infrastructure Systems*, pages 192–199, 2013.

[37] B.F. Spencer, M.E. Ruiz-Sandoval, and N. Kurata. Smart sensing technology: opportunities and challenges. *Structural Control and Health Monitoring*, 11(4):349–368, 2004.

[38] M.S. Srivastava. *Methods of multivariate statistics*. Wiley, New-York, 2002.

[39] A. Strauss, D.M. Frangopol, and S. Kim. Use of monitoring extreme data for the performance prediction of structures: Bayesian updating. *Engineering Structures*, 30(12):3654–3666, 2008.

[40] A. Tarantola. *Inverse problem theory: Methods for data fitting and model parameter estimation*. Siam, Philadelphia, PA, USA, 2005.

[41] A. Tarantola. Popper, Bayes and the inverse problem. *Nature Physics*, 2(8):492–494, 2006.

[42] K.-V. Yuen, J.L. Beck, and L.S. Katafygiotis. Efficient model updating and health monitoring methodology using incomplete modal data without mode matching. *Structural Control & Health Monitoring*, 13(1):91–107, 2006.

[43] K.V. Yuen, S.K. Au, and J.L. Beck. Two-stage structural health monitoring approach for phase I benchmark studies. *Journal of Engineering Mechanics*, 130(1):16–33, 2004.

[44] E.L. Zhang, P. Feissel, and J. Antoni. A comprehensive Bayesian approach for model updating and quantification of modeling errors. *Probabilistic Engineering Mechanics*, 26(4):550–560, 2011.