

A Probabilistic Framework for Forward Model-Driven SHM

P. Gardner, C. Lord, R. J. Barthorpe

Sheffield University, UK, pagardner1@sheffield.ac.uk

Abstract

A challenge for many structural health monitoring (SHM) technologies is the lack of available damage state data. This problem arises due to cost implications of damaging a structure in addition to issues associated with the feasibility and safety of testing a structure in multiple damage scenarios. Many data-driven approaches to SHM are successful when the appropriate damage state data are available, however the problem of obtaining data for various damage states of interest restricts their use in industry. Forward model-driven approaches to SHM seek to aid this challenge. The methodology uses validated physical models to generate predictions of the system at different damage states, providing machine learning strategies with training data to infer decision bounds. In order to obtain statistically representative damage state data from physical models it is the authors' belief that a multi-level uncertainty integration approach is required. Component or sub-system level physical models, for which validation data is more easily obtained, may be calibrated over different damage states. These validated sub-system physical models may then be incorporated into the full-system model, providing probabilistic damage state predictions at a full-system level. This paper outlines such a framework using uncertainty quantification technologies and statistical methods for combining sub-system probabilistic models whilst accounting for model discrepancies. The key stages of forward model-driven SHM are presented, highlighting relevant technologies and application considerations. Additionally, a discussion of integration with current data-driven approaches and the appropriate machine learning tools is given for a forward model-driven SHM approach.

1. Introduction

Structural health monitoring strategies have typically been categorised into two types of approach: data-driven and model-driven (1, 2). A data-driven framework is one in which machine learning or pattern recognition algorithms are used to make health decisions based on features from in-service data. The approaches can be divided further into supervised and unsupervised categories, distinguished by whether labels for data (e.g. the damage state of the structure) are known or unknown respectively. As a consequence, supervised data-driven methods require in-service, labelled data from all damage states of interest in order to infer robust decision thresholds. This is often not economically viable or feasible at a full-system level, resulting in a significant challenge to their implementation. In addition, unsupervised techniques suffer from all the complexities of performing density estimation, as well as challenges in obtaining labels when in-service data appears outside the normal condition. In contrast, model-driven frameworks are often seen as the application of model



updating; whereby parameters of a physical model (herein defined as a *simulator*) are adjusted by reducing the residual between the simulator predictions and in-service data. These methods suffer from several issues when used for identifying health states. Among these are problems in parametrisation of the simulator when damage types and locations are unknown; this can often lead to updating a vast quantity of parameters. Another problem is that the presence of uncertainties contained within in-service data (e.g. environmental conditions) or the model form are often confounded in the parameter estimation problem. These issues often result in parameters losing physical meaning and becoming extremely difficult to interpret in an SHM context. The approaches are often not practicable in an online setting as multiple runs of expensive FE simulators are required. Subsequently, both data-driven and model-driven approaches have barriers to implementation for industrial contexts.

A forward model-driven framework seeks to resolve the issues posed by both data-driven and model-driven methodologies; namely the problems associated with the lack of available data and interpretable simulators that handle uncertainties in a rigorous manner. The distinguishing characteristics of a forward model-driven approach is the utilisation of simulators in a forward manner in order to predict statistically representative damage state features, that can then be input into machine learning or pattern recognition techniques. The contributions of this paper are as follows:

- The proposed framework outlined in this paper. A probabilistic approach using validated sub-system simulators to make full-system predictions of damage states, and using these to train classifiers for health state identification.
- A case study on a calibration approach, where both parameters and model discrepancies are inferred, resulting in statistically representative predictions of damage states.
- A multi-level uncertainty integration strategy whereby parameter uncertainties and model discrepancies are propagated through to a full-system level.
- The use of simulators in creating predictions of damage state distributions and using these in probabilistic health decision strategies such as a Bayes risk formulation.

The outline of this paper is as follow; a probabilistic framework for forward-model driven SHM is outlined, where each main component is discussed. These subsections are model selection, damage feature identification, calibration and validation, multi-level uncertainty integration and health decision strategies. Finally, conclusions and further work are presented.

2. Forward Model-Driven Framework

Forward model-driven methods are comprised of two main components; generating representative damage state features from simulators and using those predictions to train machine learning or pattern recognition approaches. The second component has been well studied within the data-driven driven framework (3). Consequently, the main focus of research in establishing a framework for forward model-driven SHM is in developing

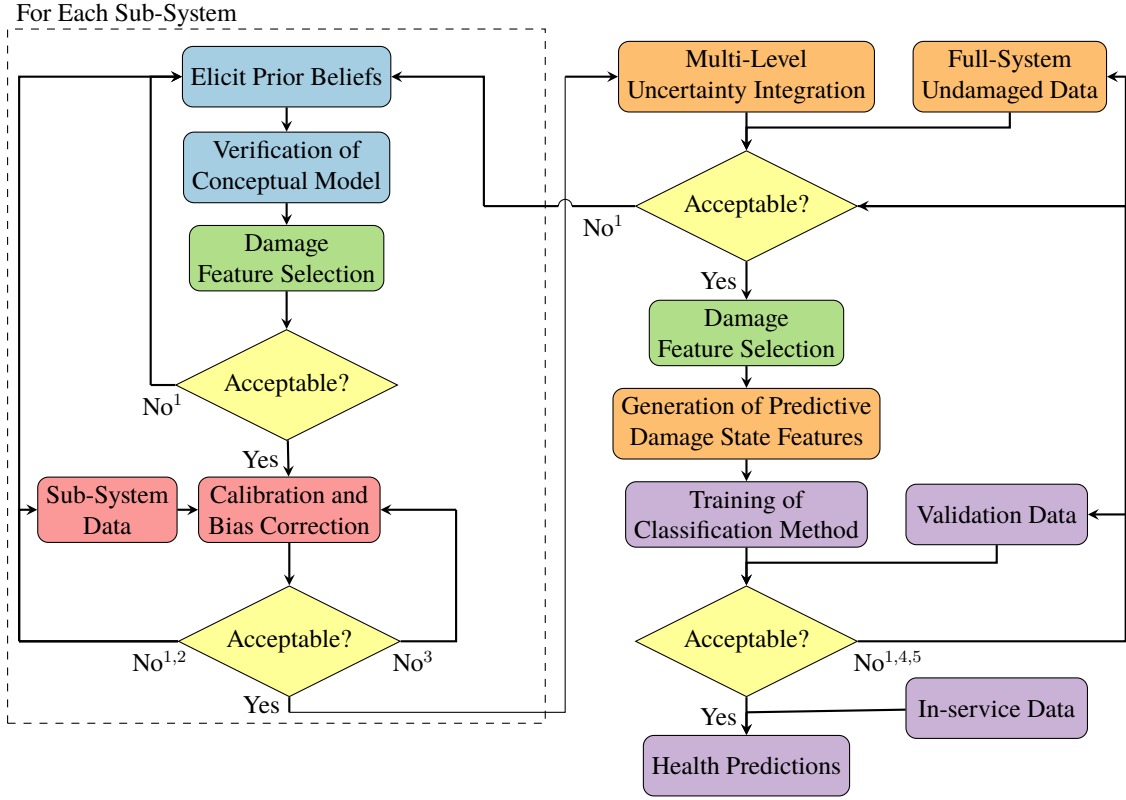


Figure 1. Flowchart for generating health predictions using the forward model-driven framework. Superscripts denote options for generating acceptable outputs. *1* denotes a review of model selection. *2* and *3* state that either more sub-system experimental data or simulator evaluations are required respectively. *4* results in the collection of more full-system undamaged validation data and *5*, that more validation data is required for the classification method.

methodologies and technologies for the first component. In the authors' opinion a non-deterministic philosophy is sensible if generating damage state features from simulators for several reasons. Firstly, a probabilistic approach provides an increased understanding in analysing the simulator for model form errors and parameter sensitivities. Secondly, real world data from structures are observations of uncertain processes. Moreover, the simulator can predict damage state distributions providing the option of using statistical hypothesis testing to validate the simulator, as well as the ability to use probabilistic health decision tools, such as a Bayes risk methodology (4). An overview of probabilistic forward model-driven SHM is outlined in the flowchart presented in Figure 1.

The proposed forward model-driven framework has five main elements:

- **Model Selection** (blue) - using prior beliefs about a structure and the processes to be modelled, in order to select an appropriate, verified model that captures the model form at the required level of fidelity.
- **Damage Feature Identification** (green) - the ability to use a simulator to investigate potential output quantities and mathematical transforms that are sensitive to the onset of damage. Having a simulator means that the measurement type and locations

can be explored before experimental or in-service data are acquired, aiding their selection.

- ***Calibration and Validation (red)*** - the ability to infer system parameters and model discrepancy of the simulator via inverse uncertainty quantification methods; where validation is performed probabilistically.
- ***Multi-Level Uncertainty Integration (orange)*** - the ability to use calibrated and validated component level simulators in order to propagate uncertainty to a full-system, removing the need for full-system damage state data.
- ***Health Decision Strategies (purple)*** - machine learning or pattern recognition methods used to infer decision bounds, as studied in data-driven research.

The flowchart (Fig. 1) is described as follows. The structure is divided into sub-systems for which damage state data is obtainable via experimental tests; this data is collected during certification or qualification stages of a product life-cycle. For each sub-system prior beliefs are obtained and used to create a verified simulator of that sub-system using model selection methodologies. Once generated, the sub-system simulator is interrogated to identify damage sensitive outputs that can be used to validate the sub-system level damage mechanisms (these may be different from the features later used at full-system level). At this stage the verified performance of the simulator is evaluated potentially leading to a repetition of the model selection process.

Once satisfactory, the sub-system simulator is calibrated and validated using experimental data via an inverse uncertainty quantification process that infers both parameter and model discrepancy uncertainties. Validation of the identified damage features using probabilistic metrics, such as hypothesis testing, are applied. If the simulator fails the validation process three outcomes are possible; repeat model selection; acquire more experimental data; or evaluate more simulator runs. These processes (inside the -- section of Fig. 1) are repeated for each sub-system, resulting in the generation of validated simulators that capture the appropriate damage mechanics.

Subsequently, multi-level uncertainty integration is performed, integrating the sub-system simulators into a full-system predictive simulator. A key assumption is that the damage mechanics of interest can be captured fully by a series of sub-system models and that their parameters, model discrepancies and uncertainties, once propagated to the full-system, have the ability to describe how the full-system behaves under these damage mechanics. Undamaged full-system observational data are acquired to validate the undamaged simulator's model form; a mismatch leads to model selection at a full or sub-system level. The valid full-system simulator is utilised to identify damage sensitive features for the employed health decision technique. As with the sub-system, the type and locations for measurements are inferred using the full-system simulator. Consequently, the selected damage state features are generated from the full-system simulator and used to train the health decision strategy. The technique is validated using held out simulator predictions, which can be supplemented with undamaged data from the structure. This stage provides the opportunity to either re-evaluate the model selection process, collect more full-system undamaged data and revalidate the simulator, or acquire more validation data for the health decision strategy. Finally, online health state predictions are made

using the health decision strategy. A predicted health state's type, location and extent are identified from the simulator, which can be subsequently used as a prognostic tool.

The proceeding subsections outline research into the five key elements of forward model driven SHM. Case studies are provided for calibration and validation as well as multi-level uncertainty integration, which are the key areas of development in this paper. Descriptions of the technologies and approaches that can be used are highlighted with examples demonstrating implementation.

2.1. Model Selection

Model selection is a challenging problem that, when applied properly, can significantly improve simulator predictive performance and hence help generate a more robust full-system simulator for forward model-driven SHM. The simplest form of model selection involves eliciting prior beliefs, generating a class of models that are possible, and evaluating their performance based on a threshold for a given metric, with a common method being the Bayesian Information Criterion (BIC) (5). Further research into how sub-system model selection may affect the full-system simulator as well as the appropriate procedures for integrating model selection into forward model-driven SHM are required to improve extrapolation abilities of damage predictions from simulators.

2.2. Damage Feature Identification

Damage feature identification is a positive by-product of a forward model-driven framework. The obtained simulator, at either a sub-system or full-system level, can be used to generate a variety of different output quantities and their transforms. By exploring the effect of damage mechanism parameters on these outputs, via sensitivity analysis (6), damage sensitive feature sets can be identified. In addition, current sensor placement optimisation methodologies could be applied, meaning that the type and location of measurements to collect on the structure can be determined *a priori* in a rigorous manner. Damage feature identification is therefore a significant area of further research.

2.3. Calibration and Validation

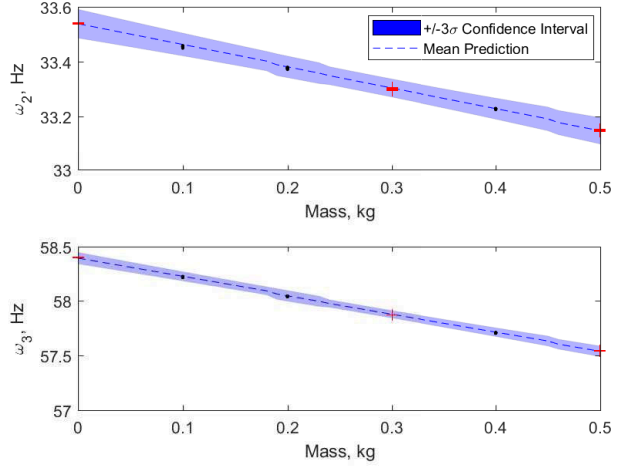
The success of a forward model-driven framework relies on the ability to generate a validated full-system simulator that is statistically representative of real world observations. As a consequence, calibration and validation are vital in producing robust health decisions.

Simulators contain simplifications or the absence of certain physics and therefore even with 'true' parameters will result in a mismatch with observational data. This difference between the assumed model form and observational data is known as model discrepancy and must be accounted for in the calibration process (7, 8). Bayesian history matching is one such method with a mechanism for handling model discrepancy.

Case Study Bayesian history matching is a calibration methodology that reduces the simulator's parameter input space whilst accounting for model discrepancy. The method achieves this using an implausibility metric: a measure of how likely it is that a given parameter combination will produce a given output. Implausibility is the distance between the experimental data and the simulator output weighted by the process's uncertainties. This allows a statistical model of the form shown in Eq. (1) to be calibrated.



(a)



(b)

Figure 2. Five storey building structure experimental setup (a). Predictive distributions from the combined Bayesian history matching and GP approach (b), + are experimental training points, · are experimental test points.

$$z_j(\mathbf{x}) = \eta_j(\mathbf{x}, \boldsymbol{\theta}) + \delta_j + e_j \quad (1)$$

Where $z_j(\mathbf{x})$ is the j th experimental output given inputs \mathbf{x} , $\eta_j(\mathbf{x}, \boldsymbol{\theta})$ is the j th simulator given inputs \mathbf{x} and parameters $\boldsymbol{\theta}$. The model discrepancy is δ and observational uncertainty e . In order to overcome computational cost the simulator is often replaced with an emulator - a statistical model of the simulator. Specifically, a Gaussian Process (GP) emulator is used, as the Bayesian formulation provides an understanding of the code uncertainty introduced by using the emulator, rather than the original simulator. Additionally, the GP will fit known simulator runs exactly, and as a consequence any information from simulator runs is preserved whilst allowing the emulator to interpolate effectively across the space. The implausibility metric can be defined as demonstrated in Eq. (2).

$$I_j(\mathbf{x}, \boldsymbol{\theta}) = \frac{|z_j(\mathbf{x}) - \mathbb{E}^*[\mathcal{GP}_j(\mathbf{x}, \boldsymbol{\theta})]|}{[V_{o,j} + V_{m,j} + V_{c,j}(\mathbf{x}, \boldsymbol{\theta})]^{\frac{1}{2}}} \quad (2)$$

Where V_o , V_m and V_c are variances associated with observational, model discrepancy and code uncertainties. The predictive mean of the GP emulator $\mathbb{E}^*[\mathcal{GP}_j(\mathbf{x}, \boldsymbol{\theta})]$ and the associated variance $V_{c,j}(\mathbf{x}, \boldsymbol{\theta})$ are integral to exploring a large part of the parameter space efficiently. A large implausibility value indicates that the parameter set was very unlikely to have produced an output that matches the experimental data given the uncertainties in the process. For multiple outputs the maximum implausibility is used. By comparing the implausibility metric to a threshold, parts of the input space can be removed. Pukelsheim's 3σ rule is a sensible cut-off as it states that any continuous unimodal distribution contains 99.73% of probability mass within three standard deviations from its mean (9). In order to preserve conciseness of this paper, the reader is referred to (10) for more information on Bayesian history matching.

An example of how Bayesian history matching can be used in calibrating sub-system models within the forward model driven framework is presented below. Modal testing was performed on a five storey building structure, shown in Fig. 2a, under different extents of pseudo-damage (added masses). The parameters, θ , were the material properties of the structure, for which $\pm 10\%$ of the nominal values were used as prior bounds. The inputs were, $x = \{0, 0.1, \dots, 0.5\}$ kg masses. The experimental outputs were the natural frequencies of the structure, specifically the second and third natural frequencies, as these were most sensitive to damage. Bayesian history matching was performed using experimental natural frequencies when $x_z = \{0, 0.3, 0.5\}$ kg with the remaining data used as a validation set. The identified plausible parameters were propagated through the emulators in order to obtain the distribution of the outputs for the calibrated parameter set.

Bayesian history matching defines the description of model discrepancy as a variance, however it will have a functional form. In order to infer this functional form a GP regression model (with noise) was used to infer the model discrepancy. A maximum *a posteriori* (MAP) estimate was taken of the output distributions from Bayesian history matching and used as inputs to a GP regression model for inferring the experimental outputs. This was performed using the same training and validation data split as in the Bayesian history matching step. The predicted outputs of the combined Bayesian history matching and GP regression approach are presented in Fig. 2b. The normalised mean squared error (NMSE) for each output were 0.04 and 0.02 respectively and the KL-divergences were all below 3, with most being close to 2. In keeping with the brevity of this paper the reader is referred to (10) for more details on the analysis.

2.4. Multi-Level Uncertainty Integration

An important aspect of a forward model-driven approach is the ability to create validated simulators of full-systems without conducting damage experiments at the full-system level. In order for this to be viable a multi-level uncertainty integration methodology must be developed. This process takes sub-system level simulators where key model forms can be validated, such as the functional relationship when damages types, extents and locations are applied, and scales the uncertainties and model discrepancies through to a full-system prediction. This means that the damage mechanisms are validated at a sub-system level, reducing the need for validation at a full-system level. This is possible if the damage mechanics can be captured at a sub-system level and appropriately scaled up. In order to demonstrate multi-level uncertainty integration a numerical case study on a four degree of freedom sway frame is presented.

Case Study. A four degree of freedom sway frame is constructed as displayed in Fig. 3a where the frame is divided into two component types - beams and plates. The key modelling assumptions in this case study are that the beams and plates do not contribute to the mass and stiffness of the full-system respectively. The joints are perfect between each beam and its respective plate and that damping is ignored. Each beam is modelled as a cantilever beam, where the tip stiffness K contributes to the full-system stiffness. Table 1 presents properties for each component with the structure comprised of eight beams and four plates. For simplicity it is assumed that the parameter distributions of Young's modulus and density are already known. In practice these would be inferred in the calibration stage, as mentioned in Section 2.3. Damage is introduced into the structure by applying a single

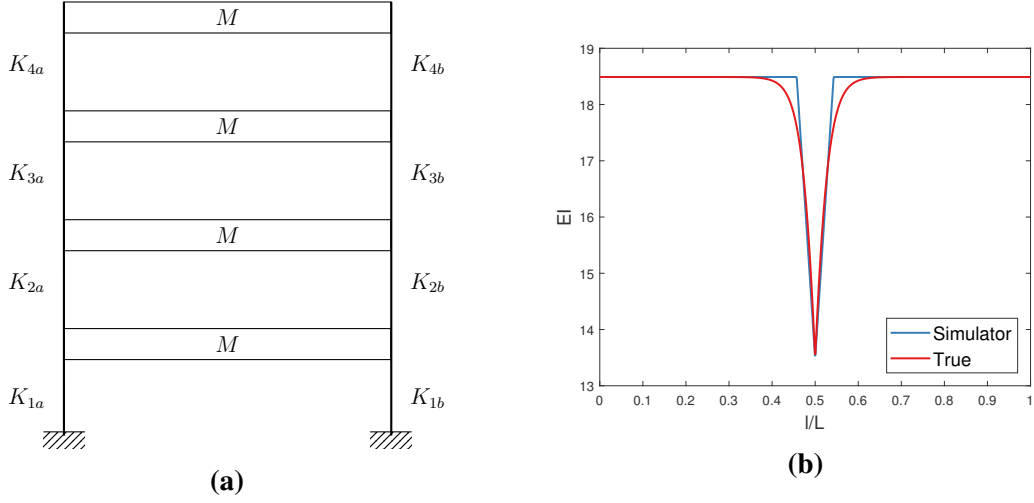


Figure 3. Four degree of freedom sway frame (a). EI curves for simulator and ‘true’ crack models. A crack length is 0.1% of the beam thickness, located at the midpoint, is presented (b).

Property	Value
Length, l_b	17.5cm
Width, w_b	2.5cm
Thickness, t_b	0.5cm
Young’s Modulus, E	$\mathcal{N}(7.1, 0.5^2)$ GPa

Property	Value
Length, l_p	30cm
Width, w_p	25cm
Thickness, t_p	2.5cm
Density, ρ	$\mathcal{N}(2700, 100^2)$ kg/m ³

Table 1. Properties of the sway frame components: beam (left) and plate (right).

open crack to the midpoint of the beam. In this paper the ‘true’ functional relationship between stiffness and crack size for a cantilever beam is formulated using the crack model for a continuous beam proposed by Christides and Barr (11), involving an exponential function of EI (where I is the second moment of area). The simulator crack model is an idealised bilinear function of EI proposed by Sinha et al (12). The two open crack models in Fig. 3b demonstrate that model discrepancy will exist even for simple structures and damage types. The full-system is damaged by including the reduction in tip stiffness of a cantilever beam under different crack lengths, into the four degree of freedom system’s equation of motion. This allows any beam to be damaged at any of the four locations. The only difference between the sway frame simulator and the ‘true’ physics therefore is the stiffness model that governs an open crack in a cantilever beam.

The multi-level uncertainty integration strategy proposed in this paper, for this motivating case study, is as follows:

1. A simulator of a continuous cantilever beam with a bilinear crack model is generated $\eta_b(\mathbf{x}_b\boldsymbol{\theta}_b)$, where the inputs are crack length $\mathbf{x}_b = \{0, 0.1, \dots, 0.9\} \times t_b$ and the parameters $\boldsymbol{\theta}_b$ are the properties shown in Table 1. The output of the model \mathbf{y}_b is the tip stiffness.
2. Experimental data is collected for the beam component, namely force-deflection curves for different crack lengths. Experimental tip deflections are measured with an

observational uncertainty distributed $\mathcal{N}(0, 1^2)$ mm - here are calculated using the ‘true’ crack model. The experimental tip stiffness $z_b(\mathbf{x}_b)$ are calculated using linear least squares regression. 50 repeats are used to obtain the experimental uncertainties.

3. Using the statistical model $z_b(\mathbf{x}_b) = \eta_b(\mathbf{x}_b, \boldsymbol{\theta}_b) + \delta(\mathbf{x}_b) + e_b$ (where $z_b(\mathbf{x}_b) = K(\mathbf{x}_b)$ - the tip stiffness) the model discrepancy $\delta_b(\mathbf{x}_b)$ and observational uncertainty e_b are estimated using GP regression. A MAP estimates of the material properties are used to generate simulator predictions, and the GP regression model is fitted between the crack length and tip stiffness residual, $\Delta K(\mathbf{x}_b) = z_b(\mathbf{x}_b) - \mathbf{y}_b(\mathbf{x}_b)$ (where $z_b(\mathbf{x}_b)$ contains 50 repeats). A comparison of the bias and non-bias correction simulators and the experimental data are presented in Fig. 4 (where 1000 Monte Carlo realisations have been used for the non-bias correct simulator).
4. A simulator of the sway frame $\eta_s(\mathbf{x}_s, \mathbf{x}_b, \boldsymbol{\theta}_p, \boldsymbol{\theta}_b)$, is generated from a four degree of freedom system. The inputs $\mathbf{x}_s = \{1, 2, 3, 4\}$ are the floor in which the crack is applied to beam a , and the parameters $\boldsymbol{\theta}_p$ are those associated with the plates in Table 1a. The outputs of the simulator are the percentage difference of the four natural frequencies of the system (the damage state feature used in this case study) i.e., $\mathbf{y}_s = \{\Delta\omega_1, \Delta\omega_2, \Delta\omega_3, \Delta\omega_4\}$. 1000 Monte Carlo realisations of the material properties and the bias corrected stiffness are used to obtain the damage state distributions shown in Fig. 5a.

Due to the numerical nature of the study, validation data at all damage states are obtained to show the effectiveness of the multi-level uncertainty integration strategy. In practice only the undamaged state distributions would be obtained experimentally. The ‘true’ percentage differences for the four natural frequencies $z_s(\mathbf{x}_s, \mathbf{x}_b)$, obtained for 1000 Monte Carlo realisations of the material properties using the ‘true’ crack model, are presented in Fig. 5b. The NMSEs between the mean predictions and ‘true’ outputs at a full-system level for the bias and non-bias corrected simulator were 0.003 and 101.151 respectively. At a sub-system level the NMSEs were 0.006 and 91.33 respectively, demonstrating how errors scale through to the full-system and the need for a model discrepancy approach. A more detailed comparison of the distributions, when $\mathbf{x}_s = 1$ for $\Delta\omega_1$, are displayed in Fig. 6.

The case study motivates the need for multi-level uncertainty integration. The study can easily be extended to inferring sub-system simulators that model, for example, the joint interactions, if bolted joints were included. This would be an additional damage mechanism, though loosening of bolts. At a full-system level, calibration may also be required. This means parameter uncertainties and model discrepancies that only apply at full-system level could be inferred. A challenge with a multi-level uncertainty integration approach is understanding where key damage mechanisms apply, and therefore what sub-systems to simulate. This is left as an area of additional research.

2.5. Health Decision Strategies

Health decision strategies have been well studied in the data-driven framework, where a variety of classification methods have been successfully implemented when labelled damage state data are available (3). All of these techniques are applicable to a forward model-driven framework, with the key difference being that the labelled data is generated from a simulator

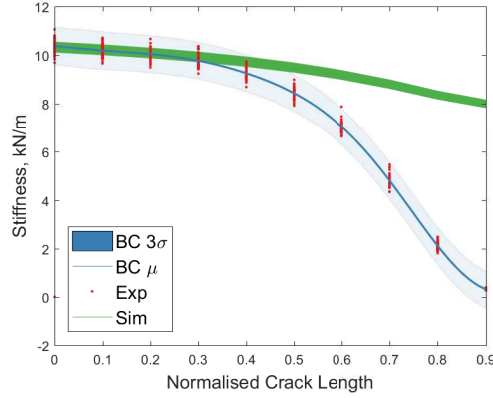


Figure 4. Beam stiffness curves where BC is the bias corrected simulator, Exp is the experimental data and Sim is the non-biased corrected simulator with 1000 Monte Carlo realisations.

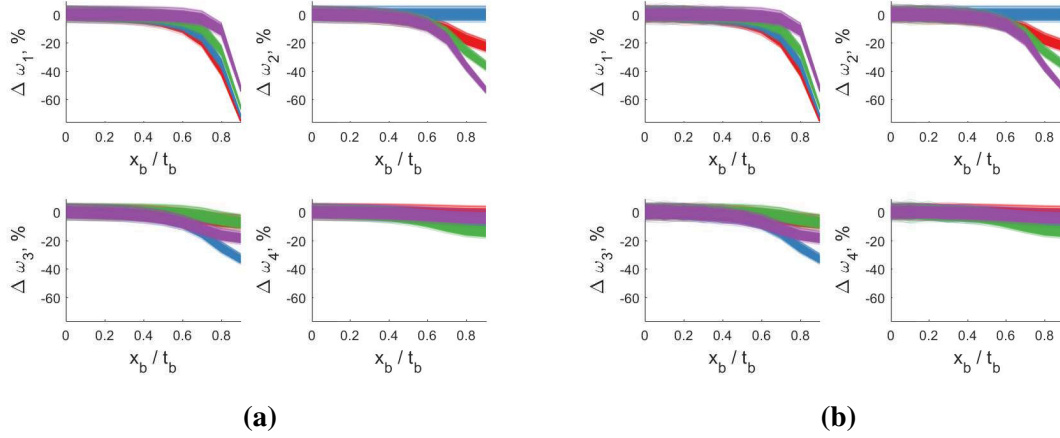


Figure 5. The percentage difference for the four natural frequencies presented for different crack length at locations: 1 (red), 2 (blue), 3 (green) and 4 (purple) - for the bias-corrected simulator (a) and ‘true’ (b) outputs.

rather than from observational data. A health decision strategy trained using a full-system simulator provides additional insight. Firstly, any classified observational data will relate to a damage state in the simulator, aiding the interpretation of the type, location and extent of damage in the structure. This means that once identified the simulator can be used as a model for prognosis. Additionally, any new damage state data can help to recalibrate and validate the simulator, improving the simulator’s predictive performance. Furthermore, the observational data can be added to the health decision strategy’s training set, aiding the calculation of decision bounds.

An alternative health decision strategy that is available to a forward model-driven approach is that of a Bayes risk methodology (4). The technique requires distributions of the damage feature for each damage state of interest. These are obtainable from the validated full-system simulator, generated by a forward model-driven framework. This approach has strengths in that it is both probabilistic and results in outcomes based on cost from asset management, aiding the decisions strategies interpretability. The implementation of Bayes risk in a

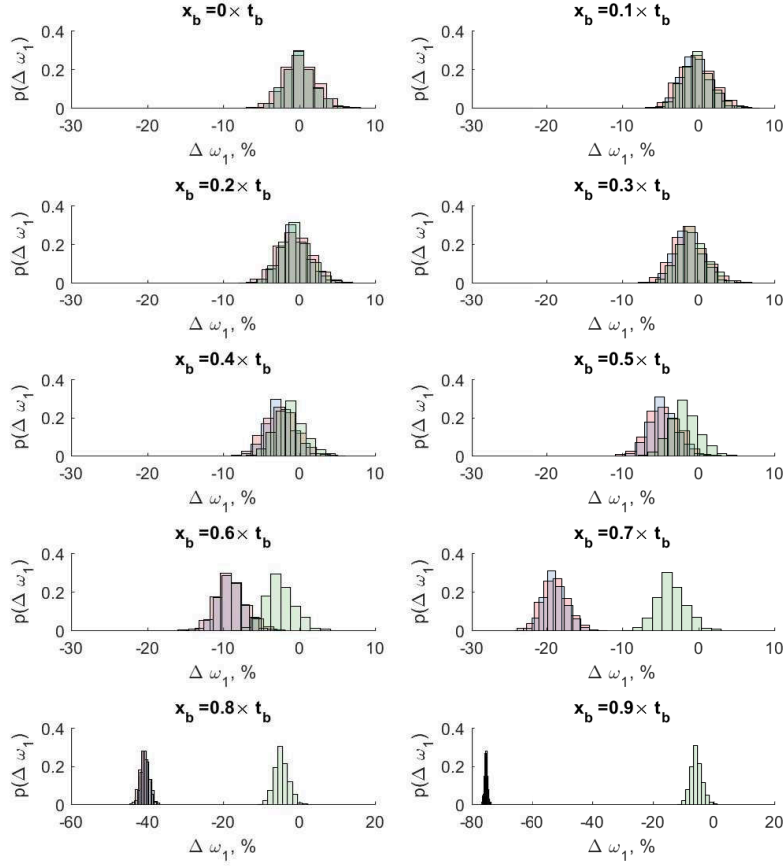


Figure 6. Comparison of the bias corrected simulator (blue), experimental (red), and non-bias corrected simulator (green), damage state distributions when $x_s = 1$ for $\Delta\omega_1$. The x-axis limits are different for $\{0.8, 0.9\} \times t_b$ due to the large percentage differences at these crack lengths.

forward model-driven setting is an area of further research.

3. Discussion

Forward model-driven SHM has been demonstrated to be an alternative to both data-driven and model-driven approaches. This is because it provides potential solutions to the lack of available damage state data in industrial applications, among other problems. The framework relies on the ability to generate valid, statistically representative predictions of damage state features, replacing the need for observational data of damage states. The method proposed in this paper is one where multiple sub-system level models are integrated, whilst inferring parameters and model discrepancies associated with damage mechanisms of interest. These are propagated through the sub-systems to a full-system level. Hence, the two key elements that have been investigated in this paper are the calibration and validation, as well as the multi-level uncertainty integration processes.

Bayesian history matching has been proposed as a method for performing calibration of sub-system models, allowing both inferences of the parameters and model discrepancy to occur (when combined with a GP regression model). In addition, multi-level uncertainty integration using a model discrepancy approach has also been demonstrated to be effective on a simple case study. As a consequence further work will be conducted in applying the two approaches on a more complex system.

Three areas of further research have been highlighted, namely integrating model selection methodologies, using simulators for damage feature selection, and using the validated full-system models in a Bayes risk health decision strategy. The inclusion of these methods into a full demonstration of forward model-driven SHM is left for further research.

In conclusion, forward model-driven SHM is a promising framework in which current industrial challenges could be overcome. These approaches have the potential to provide further benefits and insight to the current state-of-the-art data-driven methods helping SHM to become more universally adopted.

References

1. K. Worden and G. Manson. The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):515–537, 2007.
2. M. I. Friswell. Damage identification using inverse methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):393–410, feb 2007.
3. J. Ramon and F. Esquivel. *On Bayesian Networks for Structural Health and Condition Monitoring*. PhD thesis, 2017.
4. E. B. Flynn and M. D. Todd. A Bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mechanical Systems and Signal Processing*, 24(4):891–903, 2010.
5. D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 64(4):583–616, 2002.
6. J. E. Oakley and A. O’Hagan. Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 66(3):751–769, 2004.
7. 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.
8. J. Brynjarsdóttir and A. O’Hagan. Learning about physical parameters: the importance of model discrepancy. *Inverse Problems*, 30(11):114007, 2014.
9. Pukelsheim. The three sigma rule. *American Statistician*, 48(2):88–91, 1994.

10. P. Gardner, C. Lord, and R. J. Barthorpe. Bayesian History Matching for Forward Model-Driven Structural Health Monitoring. In *Proceedings of IMAC XXXVI*, 2018.
11. S. Christides and A. D. Barr. One-dimensional theory of cracked Bernoulli-Euler beams. *International Journal of Mechanical Sciences*, 26(11-12):639–648, 1984.
12. J. Sinha, M. Friswell, and S. Edwards. Simplified Models for the Location of Cracks in Beam Structures Using Measured Vibration Data. *Journal of Sound and Vibration*, 251(1):13–38, 2002.