The University Of Sheffield.

Verification and validation of physics-based models for structural health monitoring

James Wilson

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Engineering

Department of Mechanical Engineering

Submission date

Wednesday 26th July, 2023

Acknowledgements

I would like to thank my supervisors Rob, Graeme and Paul for their guidance and support over the last few years.

I would also like to extend my thanks to the Dynamics Research Group as a whole (including many former members); you have created a welcoming and friendly environment for PhD study, with a healthy dose of distractions where necessary. Particular thanks go to Emiliano, Tina and Marcus for their help at various stages of my work, and of course to Furkan. The lab staff at the department also have my sincere thanks: Mathew, Robin and Michael possess formidable collective wisdom and expertise, without which I wouldn't have got very far.

My friends and family deserve thanks and credit for support and guidance wherever it was required. Max – I hope the mutual griping helped you as much as it did me.

Last but not least, thanks go to Bridget, my primary source for advice and encouragement.

Abstract

Structural health monitoring (SHM) refers to a group of methods through which engineers aim to infer the health state of a piece of engineering infrastructure given some measured data from that structure. These inferences are then used to inform ongoing maintenance strategies to ensure that the structure remains operating in its optimal condition, without compromising performance or safety. Given the age of much of the existing infrastructure around the world, life-extending technology such as SHM is clearly of significant potential benefit, and the use of SHM when developing new infrastructure offers the potential for maximising the return on the investment in that infrastructure.

SHM strategies can generally be divided into two types. Data-driven SHM uses data taken from the structure of interest to train a statistical model, such that the trained model will be able to label new data from the structure as being indicative of a particular health state. Model-driven SHM traditionally uses a physics-based model that can be adjusted to fit its predictions to live data taken from the structure, such that the adjustments to the model inputs give some indication of the structure's health state. Data-driven SHM methods are reliant on large sets of training data on which to develop statistical models; this is often difficult to acquire in practice, particularly when data are required from a structure in its damaged states. This issue can be mitigated by using a physics-based model to simulate the training data for the statistical model; however, in order for the physics-based model predictions to be considered trustworthy, the model must be validated against experimental data. This means that damage-state data are still required for the implementation of physics-based models in SHM in order to ensure the accuracy of their predictions. The acquisition of damage-state data from structures for this purpose therefore presents a significant set of problems for SHM methods.

It is possible to reduce the difficulties associated with model validation for SHM by carrying out validation on models of the subassemblies and components that make up a larger assembly structure. The data required for this *hierarchical* validation task should be easier to acquire cheaply compared to validation data drawn from the full structure. If the submodels representing these individual substructures can each be validated then it should be possible to recover an assembly-level model which, given the validation tasks carried out, could be used to make confident predictions regarding the structure in a range of health states.

A framework is presented in this thesis which summarises the activities required to carry out this hierarchical validation strategy, which would then enable a model to be developed with demonstrable accuracy and quantifiable uncertainty in its predictions. This framework is applied to a target structure – a truss bridge – and the model is used to carry out a series of SHM tasks on test data drawn from the structure. These tasks are carried out using the validated model in a forward manner to generate training data for statistical damage recognition models, which are then compared – in terms of performance – to traditional data-driven methods.

Based on the research presented in this thesis, it is shown that model uncertainty can be accurately quantified through the hierarchical validation process by comparing the model predictions to the experimental test data before and after the validation process. After this it is shown that it is possible to develop accurate damage classification algorithms using validated model predictions, with the SHM methods developed via the hierarchical validation framework performing favourably compared to traditional data-driven methods when exposed to the test data. Further research areas that would advance the methodologies presented in the thesis are then outlined following discussions of the results.

Table of contents

List of figures

List of tables

List of initialisms

ABC approximate Bayesian computation

ANN artificial neural network

DoF degree-of-freedom

EOVs environmental or operational variables

FEM finite element model(ling)

FMD-SHM forward model-driven structural health monitoring

FPR false positive rate

FRF frequency response function

GCI grid convergence index

GP Gaussian process

IMD-SHM inverse model-driven structural health monitoring

MAC modal assurance criterion

NDT non-destructive testing

RBF radial basis function

ROC receiver operating charactistic

SA sensitivity analysis

SHM structural health monitoring

SVM support vector machine

TPR true positive rate

UP uncertainty propagation

UQ uncertainty quantification

V&V verification and validation

Chapter 1

Introduction

1.1 Project background and context

Structural health monitoring refers to the field of civil and mechanical engineering that seeks to ensure that built infrastructure – both planned and historical – is evaluated in terms of its health state according to a specified strategy. The 'health state' of a structure refers to some condition that the structure is operating in based on the presence or absence of damage. Health states can be classified by the presence of damage alone (a binary classification space) or by damage location and extent. Damage progression is unavoidable in structures that are subjected to repeated or cyclic load cycles, and generally manifests in the form of crack propagation from some initial fault that occurs in manufacture to eventual failure. Other types of damage include abrasion and corrosion. Engineering structures can often operate in a range of health states, but the presence of damage may impact on operational performance. Critical damage refers to the set of health states within which the operational performance or safety of the structure is significantly compromised, such that the structure must be taken out of service for maintenance or end-of-life management. As such, it is clear that the ability to differentiate between the health states of a structure offers significant benefits to the life-cycle management of the structure by offering the ability to schedule maintenance in such a way as to minimise down-time and maximise productivity.

The scope of an SHM strategy is dependent on the requirements for the structure and the available investment. Rytter's Hierarchy presents a full range of SHM tasks, where each task is dependent on the completion of the previous tasks, and provides

further information on the health state of the structure (but requires more refined analysis) [1]:

1. Damage detection
2. Damage localisation
3. Damage classification
4. Damage assessment
5. Remaining life prognosis

Health monitoring strategies have been categorised into the following three options: run-to-failure, time-based monitoring and condition-based monitoring [2]. Run-to-failure refers to the use of a structure or system without monitoring until it eventually fails, while time-based monitoring uses regular maintenance checks that are pre-specified for the structure. Condition-based monitoring is enabled by the use of intelligent systems to infer the live health-state of the structure in order to provide maintenance when it is required to prevent failure. This strategy aims to reduce the level of risk in the use of the structure by preventing catastrophic failure, and also to minimise the maintenance costs by reducing unnecessary callouts. However, implementation of condition-based monitoring tends to have a high up-front cost due to the technology required.

Health monitoring methods can be either *offline* or *online*. Offline methods are carried out at discrete points in time to assess the health state of a structure at that time. Non-destructive testing (NDT) methods such as visual inspection and dye penetrant testing form a large part of offline monitoring techniques, and they are employed in many scenarios. Offline monitoring is a very mature field in both research and industry, and is employed in time-based monitoring strategies as discussed above. Online monitoring techniques make ongoing assessments of the health state of a structure in real time based on live data that are drawn from the structure, and are therefore crucial to condition-based monitoring. The majority of current SHM research concerns online monitoring due to the significant opportunities it avails in increased safety and reduced operation and maintenance costs. However, while offline methods have a history of successful applications in industry, online methods have, as yet, struggled to make this transition from academia in terms of widespread industry take-up [3].

Two streams of SHM have been theorised and investigated: data-driven SHM and model-driven SHM. Data-driven methods use datasets drawn from the structure to

train machine learning algorithms (known as 'black boxes') which can then monitor future data for signs of damage. Model-driven methods use physical representations (often known as 'white boxes') of structures to make inferences of the health of their real-life cousins. More recently, academic research has looked into combining the ideas of data- and model-driven SHM, resulting in ideas such as grey-box modelling for SHM. This method uses physical knowledge of a structure to constrain the learning process for a data-driven SHM strategy. Another bisection of ideas is forward model-driven structural health monitoring (FMD-SHM), which uses data from a white-box model to train a black-box machine learning-based system, which is then used to monitor a structure for damage.

Data-driven SHM methods have the advantage of being a flexible way of learning the trends behind structural data, such that confident predictions can be made on the health state of the structures when new data are presented. Several successful demonstrations of the method have been made with applications across Rytter's Hierarchy [1], and efforts to incorporate probability in the methods have also shown promise. However, key issues remain that limit the applicability of the methods. The first among these is the lack of labelled training data that is required for machine learning algorithms; these data are generally very difficult or expensive to collect in SHM applications, as data must be drawn from the structure in its damaged states in order to provide informative training sets. In the context of machine learning theory, labels refer to the outputs of a model such that the data associated with a certain label are representative of data drawn from a system in the condition described by that label. In SHM applications, labels are generally associated with certain health states of the structure, although the label space may also describe a range of operating conditions.

Model-updating SHM methods have also been successfully demonstrated in a number of SHM applications. The use of the physics-based model often allows for a more informative set of predictions to be made on the state of the structure, with the law-based approach enabling remaining life prognosis where required. However, the model-updating approach presents serious computational difficulties (often related to the lack of solution uniqueness), as many model runs are often required in order to arrive at a solution for new data. The model-based method is also dependent on labelled damage-state data from the structure in order for the physics-based models to be validated, and the effects of model bias are also an outstanding issue.

One method of circumventing the problems involved in developing a stable model-updating SHM strategy is the use of model-based methods in a *forward* direction – as in FMD-SHM [4]. In this technique, the models are used to simulate training data for a statistical model which, when presented with test data, would then be tasked with identifying damage in the structure. This could reduce the computational burden associated with traditional model-based SHM methods (another way of looking at it is that it removes the need for real-life training data from data-driven methods, as these data are simulated instead), but does not necessarily alleviate the requirement for informative validation datasets.

It is hypothesised in this thesis that it may be possible to avoid the dependency of model-based methods on high-cost system-level validation data by approaching the validation problem in a hierarchical sense. If the model of a complex assembly can be broken down into a set of simpler submodels, these could then be validated separately, using validation data drawn from the components or subassemblies they represent. This could potentially be a much less expensive task than carrying out validation at the system-level, as only the components where damage would be expected to manifest would need to be tested in their damaged states. Once the submodels are validated across the full range of operational and damage conditions expected in the assembly structure, they could then be reassembled in order to recover the system-level predictive model, but with quantifiable confidence now attached to any further predictions.

Uncertainty is virtually unavoidable in engineering because lack of knowledge is present in any prediction or measurement of a system or structure's behaviour. Overall uncertainty can be divided into two contributing types of uncertainty: epistemic and aleatoric uncertainty. Aleatoric uncertainty refers to the inherent randomness in a system and causes a scattering of measurements or predictions around the true value – this is generally estimable through verification and validation (V&V) [4]. Epistemic uncertainty refers to the reducible lack of knowledge of a system, and it generally leads to a biased error in measurements or predictions [4]. Epistemic error can generally be reduced by adding complexity to a model to more accurately represent the underlying physical processes, or by improving experimental practice in the case of recording measurements. However, the biases it causes are often of unknown magnitude and the problem of epistemic uncertainty in engineering modelling is an outstanding research area.

A key issue in modelling is the quantification and handling of model discrepancy, which is an example of epistemic uncertainty in a problem. Model discrepancy is defined as the difference between the predictions of a model and the observed data when the true parameters of the model are known, due to estimations and discretisations of physical processes represented in the model [5]. Numerical error can also contribute significantly to overall model bias [6]. Error is caused by uncertainty, and it is a measure of the difference between an estimation and the true solution. Model discrepancy often results in systematic errors; a model discrepancy function can be defined during the V&V process to account for this.

Verification is defined as 'the process of determining that a computational model accurately represents the underlying mathematical model and its solution', and is divided into code verification and calculation verification [7]. Complete verification accounts for both mesh convergence and discretisation error (although the two are closely linked) [8]. Validation is defined as 'the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model' [7]. Put more succinctly, verification asks 'does the model solve the equations right?' while the validation process asks 'is the model solving the right equations' [9].

V&V provides the means by which uncertainty and error can be accounted for in the application of predictive models. It allows the modeller to determine a model that is demonstrably accurate to a level that is pre-determined according to the tasks at hand, and to quantify the uncertainty that remains in the model predictions. Problems exist in the implementation of rigorous V&V strategies when applied to SHM, which will be investigated in this thesis through techniques such as dynamic substructuring and hierarchical modelling.

1.2 Project motivation

The necessity to validate models to ensure confidence in their use has long been a stumbling block to their widespread uptake in an SHM context, due to the problem of lacking sufficient data. This is because V&V to support model-based SHM strategies would usually be carried out using assembly-level data from a potentially unique, high-value, large-scale structure in both its undamaged and damaged states. Each of these aspects of common structures suitable for SHM makes the acquisition of validation data very difficult, and in many cases not viable. This therefore motivates

further research into the field of V&V for SHM. The requirements of the V&V process currently hold back industry implementation of model-based SHM, and therefore development of the practice offers potentially very significant benefits by providing solutions to the issues discussed here.

The lack of comprehensive assembly-level data also affects the implementation of data-driven SHM strategies for similar reasons to those outlined above: in data-driven methods, training data are required for the machine learning algorithms. In the absence of assembly-level damage-state data, data-driven methods are generally limited to novelty detection to infer damage (although some unsupervised methods such as clustering are able to provide advancements on this). It is therefore possible that the development of methods to mitigate this could expand the potential effectiveness of SHM strategies in a wide range of scenarios.

1.3 Aims and objectives

It is hypothesised that the issues associated with the acquisition of validation data for physics-based models can be circumvented by approaching V&V in a hierarchical manner. This would entail breaking a complex model of a fully assembled structure down to a set of subassembly- or component-level structures. These smaller substructure models would then be validated separately, a process that could offer much better feasibility than assembly-level validation. This is because the V&V strategy could utilise symmetries and modular aspects of the assembly; the ease of testing could be improved, allowing laboratory data acquisition at lower cost; and damage-state testing would be at lower cost for simpler structures.

The central aim of the project is to develop and evaluate a novel framework for V&V of models in an SHM context, focusing on the application of predictive models to FMD-SHM. This framework will set out best practices for the hierarchical V&V of models, and will seek to provide a step forward for the use of model-driven SHM in industry. The tasks required to achieve this aim can therefore be categorised as follows:

1. Develop the theoretical background required to execute the framework.
2. Demonstrate that the framework is appropriate for use in an SHM context.

The aims above required research to be carried out into the fields of dynamic substructuring, traditional V&V, and uncertainty propagation. These will, respectively, pro-

vide the capabilities to assemble validated models in a hierarchy, determine best practices for V&V of substructures and allow the uncertainty quantified at substructure-level to be carried through to the assembly model. Based on the above, the following objectives are defined for this project:

1. Create a hypothesis framework for hierarchical V&V.
2. Identify target structure(s) and acquire data with which to demonstrate elements of the framework.
3. Evaluate and demonstrate the use of suitable metrics and criteria for substructure model validation and uncertainty quantification (UQ) for SHM.
4. Develop a robust methodology for uncertainty propagation (UP) upwards through model hierarchy using numerical case studies.
5. Identify the impact on the handling of features when employing hierarchical models in an FMD-SHM context.
6. Demonstrate using experimental data how appropriate validation, UQ and UP techniques can be applied in practice for hierarchical V&V.
7. Use the hierarchical V&V framework to demonstrate and evaluate the application of FMD-SHM for the target structure.
8. Evaluate and refine the framework to identify how it would interface with existing practices in industrial health monitoring and to outline requirements for future research.

The main contribution of this research will be the framework for, and demonstration of, the application of hierarchical validation to model-based methods of SHM. If successful, this will provide a benchmark from which further studies on model-based SHM studies can proceed, with demonstrations that the data-acquisition problem can be reduced in the development of predictive models. The framework will be presented in basic form early in the thesis, and investigations into its application will inform an evaluation of the original ideas, from which an updated framework can be formulated.

Further contributions in this thesis will be based on the evaluation of validation methods for FMD-SHM methods on a realistic, novel, dataset. The data used in this thesis are drawn mostly from a laboratory-scale truss bridge, which is very representative of structures seen in built infrastructure worldwide.

1.4 Thesis layout

The chapters of this thesis proceed as follows:

2. Literature review: the existing literature concerning methods and applications of SHM and V&V is summarised, indicating the current state of the art and potential for further research.

3. Hierarchical validation hypothesis: an initial framework for hierarchical validation in a model-based SHM context is presented. Motivations for the framework and foreseen advantages and disadvantages are discussed.

4. Hierarchical validation theory and methods: the technical processes for carrying out hierarchical validation are presented and tested using numerical examples.

5. Feature selection in a hierarchy: the selection and handling of damage-sensitive features within an FMD-SHM context based on hierarchical validation are discussed. Potential benefits are highlighted alongside potential pitfalls, and ideas are tested based on experimental data.

6. Quantification and propagation of uncertainty in hierarchical validation: the quantification of uncertainty in an assembly via validation of submodels of that assembly is investigated through an experimental case study.

7. SHM case studies: the validated models developed in the previous chapter are tested on a range of SHM tasks and compared against traditional methods.

8. Conclusions: the performance of the previously presented framework is evaluated based on the results of the tests carried out in previous chapters. A new framework is presented based on the findings of the research. Future areas of research which follow on from this thesis are identified.

Chapter 2

Literature review

This chapter provides a review of the current state of the art in SHM, with a particular focus on FMD-SHM methods. FMD-SHM is the strategy by which a sufficiently validated model is used to generate training data for data-driven damage identification methods [4, 10]. The chapter will cover the existing literature in this field, and it aims to contrast FMD-SHM techniques with other, more established SHM techniques; these are broadly divided into inverse model-driven methods and data-driven methods [2, 4]. Other areas of interest covered in this chapter include V&V and uncertainty analysis in physics-based and data-based models. The review will conclude by highlighting the areas which are still open for further research. Detailed methodological literature is not presented in this chapter; instead, techniques and methods will be referenced where they are applied throughout the thesis. This is intended to allow this chapter to deliver an overall review of research in SHM and to frame the more detailed discussions that follow in the later chapters.

2.1 Structural health monitoring: An overview

Structural health monitoring has been a fast-growing area of research in the fields of civil and mechanical engineering since its inception in the mid 20^{th} century [3]. The SHM field is often split into four categories [2, 3]:

1. Machine condition monitoring
2. Global monitoring
3. Large area monitoring
4. Local monitoring

SHM tasks are generally aligned with a particular maintenance strategy, including 'run-to-break' maintenance, time-based preventative maintenance and condition-based maintenance [2]. These strategies represent progressively greater investment, but offer greater long-term benefit in terms of reduced structural risk and reduced maintenance or repair costs. Condition-based maintenance is reliant on continuous analysis of the health state of a structure, and is therefore dependent on some form of online SHM.

Cases favourable for implementation of SHM strategies include structures with low safety factors, structures which have catastrophic consequences of failure (either because of cost or safety), structures for which damage progresses from incidence to a critical state on a relatively short time scale, and structures for which NDT is difficult or impractical [3]. The extent of a prospective SHM strategy is usually based on a trade-off between the diagnostic capability of the strategy and its cost and complexity. The scope of SHM systems is clarified neatly in Rytter's Hierarchy [1], where the SHM tasks are ordered from damage detection, through damage location, assessment and remaining life prognosis. The hierarchy is organised such that the tasks progressively increase in complexity, with each task requiring completion of those before it [1]. Damage assessment is often split into classification of damage type and estimation of the severity of the damage. Future SHM research areas highlighted in [3] include the following:

1. Methods to increase area coverage per sensor
2. Techniques for performance validation
3. Development of technological capabilities and business applicability
4. Efficient data handling to give operators information for decision making as opposed to raw data
5. Fusion of data from multiple sources to improve prognostic capabilities

This thesis concerns the field of model validation for SHM, and it is therefore situated in the research following point 2 of the above. Point 3, concerning general technological development and business application of SHM research, is critical to most engineering research. Issues relating to the development of the FMD-SHM framework would also fall within the research described by point 4.

A further series of significant works reviewing the field of SHM are recognised in this chapter. The first two were produced by the Los Alamos National Laboratory.

Doebling et al. compiled an extensive review summarising an extensive range of SHM research up until 1996 [11]. The summaries are presented chronologically and are organised by method and application. Sohn et al. followed up on this work in 2001, providing an update on the state of SHM research in that period [12]. Finally, key insight and detail into the state of research in a range of fields of SHM up to 2013 can be found in Farrar and Worden's book [2].

2.2 Data-driven structural health monitoring

Data-driven methods of SHM refer to processes that make use of structural data only to infer the health state of a structure, without any dependence on physical knowledge. The methods usually rely on some statistical machine learning model, the opaque nature of which means the models are often referred to as black boxes; investigation of the internal mathematics of the models generally yields no physically interpretable meaning. The capability of the method selected is dependent on the training data available, and the requirements of the SHM strategy it serves.

It is possible for data-driven methods to serve the first three stages of Rytter's Hierarchy, provided adequate training data can be acquired, with labels where necessary [1]. However, the final level of the hierarchy, damage prognosis, requires knowledge of the physical laws governing damage propagation, and it is therefore unlikely to be feasible to carry out using purely data-driven techniques.

Data-driven methods can be separated into supervised and unsupervised learning methods. *Supervised* learning refers to the case in which the training dataset contains target labels, describing certain damage states and operating conditions in the case of SHM applications. These methods can be used for damage detection, damage location and damage assessment. *Unsupervised* learning techniques are based on datasets which have no labels or targets, which reduces the problem to a novelty detection exercise; these techniques can only be used for damage detection in an SHM context [2, 4].

Successful applications of data-driven SHM methods include, among others, damage detection in a bridge and in pipes [13, 14]. A recent review of data-driven SHM focused on deep learning techniques can be found in [15].

The dependence of data-driven methods on labelled training data is a significant issue in SHM, as damage-state data from the target structure are often difficult to

acquire. Efforts to alleviate this issue in a data-driven framework include ideas such as grey-box modelling and population-based SHM. Grey-box modelling refers to a set of techniques for data-driven SHM that incorporate physical laws into the machine learning algorithms [16]. Population-based SHM techniques use data from structures across a fleet that exhibit similarity to each other to enhance knowledge of the label space for those structures, generally utilising some form of transfer learning [17].

2.3 Inverse model-driven structural health monitoring

Physics-based models are used for a range of tasks in SHM strategies. Model-driven SHM, as opposed to data-driven SHM, is sometimes referred to as a white box method; this is because the model is usually based on physics-based laws and parameters which the modeller can tune, or at least understand.

The inverse model-driven, or model updating, approach to SHM – referred to in this thesis as IMD-SHM – is a method for damage identification based on knowledge of the physics of a structure. Using a set of physics-derived laws, a representative model is constructed to represent the target structure, often utilising finite element model(ling) (FEM) techniques. The model outputs will usually be dependent on a set of inputs and system parameters (where the parameters define the physical properties of the system and the inputs describe its operating condition), such that as new response data are acquired from the monitored system, the input vector can be updated to fit the model predictions to the data. From the updating of the inputs, which have known physical meaning, the health state of the system can be inferred; this process is known as parametric or sensitivity-based model updating [2, 4, 18].

An alternative updating method, direct model updating, applies the updating process to the parameter matrices of the model [19]. These techniques are generally less physically interpretable and can often suffer from issues of solution uniqueness. Direct model updating methods appear to have been largely left behind in recent research, with modellers preferring parametric methods.

Due to the large scale and complexity of many models used for SHM, the number of inputs can be very large, which can lead to significant ill-conditioning and ill-posedness in the updating problem [19]. Regularisation is introduced as a means to ensure that a unique solution for the updating problem can be found [19]. Several techniques of regularisation are presented in the literature [20, 21], many of which are

based on an addition of constraints by using knowledge of the physics of the system to eliminate certain parameters from the updating process. Methods for regularisation include use of the Moore-Penrose Pseudo-Inverse and statistical parameter subset selection; other methods are given in [19].

IMD-SHM has been shown to be successful in a number of applications, including a wall structure [22], a bridge [23] and historical structures [24]. However, significant issues remain with the computational cost of model updating problems [4].

2.4 Forward model-driven structural health monitoring

A novel method for model use in SHM is forward model-driven SHM [4, 10]. This is a 'hybrid' approach in that it makes use of both data-driven and model-based SHM techniques: a physics-based model is used to generate training data for a machine learning-based statistical model [4]. This statistical model can range in capability, from novelty detection in response data to full damage prognosis, depending on the system requirements and the level of training data that the physics-based model can provide [4, 10, 25, 26].

The method offers greater feasibility in an SHM context than IMD-SHM methods in that it avoids the difficulties present in developing a stable model updating problem, with predictions being used in the forward direction only. This also has the added benefit of concentrating computational resource at the development stage: once the machine learning algorithm is trained, no further computation is required beyond the running of the statistical model. Compared to purely data-driven techniques, FMD-SHM is less reliant on data availability, and the use of physics-based predictions allows for more advanced SHM tasks, such as damage prognosis, to be carried out.

The research concerning FMD-SHM is still in its infancy; however, certain investigations have achieved success in a variety of applications. An early example by Castellini and Revel aimed to prove the effectiveness of laser Doppler vibrometry to detect, locate and characterise damage in composite structures [27]. The investigators successfully used deterministic predictions generated using an FEM to train an artificial neural network (ANN) to be able to detect and to locate damage in two separate specimens. Ko et al. used an FEM to train ANNs for damage detection, location and assessment in a bridge [28]. The model used was validated in a separate paper [29], but uncertainty quantification was not taken into account. However, the method

proved effective when applied to simulated test data. In a separate investigation of the FMD-SHM method, Zapico et al. used an FEM-generated dataset to train ANNs for damage location [30], aiming to use ANNs to directly update the stiffness matrices of a separate FEM for a set of damage scenarios. Zapico and Gonzalez followed up on this investigation in 2008 with an investigation on a five-storey building structure [31]. They used modal data to predict the characteristic matrices of the damaged structure, using training data for ANNs generated by an FEM.

The first known application of the FMD-SHM technique to real-life civil structures was by Lee et al. in 2005 [32]. They applied FEM-generated training data to ANNs for damage detection in a beam and a bridge, using mode shapes as the network inputs. The method was found to be successful for damage detection in bridge structures, in both laboratory conditions and in operation (in the presence of traffic loading conditions) [32]. In an extension to this work, in which probabilistic neural networks were trained for damage localisation on bridge structures, probabilistic FMD-SHM techniques were shown to be viable [33]. The general use of ANNs trained by FEMs in a forward manner is finally classified by the same authors as 'model-based' in [34]. An additional application of the FMD-SHM method to bridges used a support vector machine (SVM) method for damage detection [35]. The results aligned, to a degree, with the investigators' previous findings using a model updating procedure [35]. An FMD-SHM method was used for damage classification on an experimental bridge structure [36]. The process was successful in classifying damage simulated in the structure by alteration of its boundary conditions. These findings are highly encouraging for the potential of FMD-SHM.

The comparative merits and drawbacks of the FMD-SHM compared to IMD-SHM and data-driven SHM were discussed in [4] based on a series of case studies. A following work focuses on developing a framework for executing FMD-SHM, and also presents investigations into a range of methods for V&V and UQ in an FMD-SHM context [10]. The discussions cover the use of emulators and methods to address model discrepancy such as Bayesian history matching and Bayesian calibration bias correction.

Pawar and Jung used the FMD-SHM method to develop an SVM for damage classification as part of a helicopter health usage and monitoring system [37]. Using load levels as inputs, the SVM was able to successfully classify damage severity according to a three-tier scale, including in the presence of random noise on the simulated values. Separately to this, FMD-SHM was also shown to be effective where an FEM was used to train SVMs for damage location in beams [38].

ANNs were trained for crack detection, localisation and assessment in aerospace fuselage using FMD-SHM methods by Sbarufatti et al. [39]. This study is notable because it focused on damage propagation in an experimental structure, and because it took account of uncertainty. This work was furthered in 2014, using Lamb waves with a forward model-driven technique [40]. In this paper, numerical simulations were used to inform ANNs, which were then applied to the assessment of damage extent in cracked plates experimentally. FMD-SHM was used in order to detect damage in an experimental 8-degree-of-freedom (DoF) system, where antiresonances were used as the input features for the ANNs [41]. This was extended with further tests on the 8-DoF structure and a damaged beam using a novel learning technique [42].

Attempts to integrate uncertainty into FMD-SHM were made by Zhou et al., who used probabilistic neural networks trained on noise-contaminated FEM predictions [43]. This showed some promise in a damage localisation context for a bridge structure. An investigation of probabilistic damage prognosis was presented in 2016 by Leser et al. [25]. This paper used a crack growth model with associated uncertainty to predict the growth of a crack in a notched specimen with associated confidence intervals via surrogate models.

Applications of FMD-SHM in practice can be found in damage detection for offshore structures [44], image recognition [45, 46] and human biology [47]. This is a limited selection, and further investigations using realistic datasets are required to lend the method further credibility within the SHM canon. Looser applications of FMD-SHM involve the use of physics-based models to *augment* data, rather than to simulate a whole training dataset. This has been applied in the case of the Z-24 bridge, where classifiers trained entirely with simulated data were compared to classifiers trained on multi-source (generated and experimental) data; classification accuracy was improved when data were fused from all possible sources [48].

It is clear that for further advancement of FMD-SHM, formal strategies to quantify uncertainty (both random and systematic) must be developed [10, 26]. Barthorpe and Gardner discussed the development of FMD-SHM in a paper presented in 2019 [26]; the authors highlighted V&V and UQ as key issues in particular. Future research required to advance the status of FMD-SHM should therefore follow two directions: the demonstration that the method can be used within a robust uncertainty framework (in order to establish quantifiable confidence in the method), and more exhaustive demonstrations using real-life datasets that the method is applicable in an SHM context.

2.5 Verification and validation

When employed for SHM purposes, physics-based models can be used for a range of tasks, from informing feature selection to generating training and testing data for statistical inference [4]. The use of the model dictates the level of trust required, with relatively coarse models considered suitable for feature selection [4]. However, if a model is to be used to inform machine learning for some level of damage identification, a higher level of precision is required, which therefore demands a more rigorous establishment of trust. This required level of trust or confidence is achieved through V&V of the model [49], and is a basic requirement for the use of models in FMD-SHM. It is widely accepted that no model is ever truly accurate [50]; therefore, the V&V process is, in truth, a model falsification exercise [51]. Models are therefore eliminated until they can be considered acceptably accurate and reliable based on some quantitative or qualitative criteria. Two comprehensive guides for V&V and UQ respectively are [7] and [52]. The former is a manual for the application of V&V in an SHM context, while the latter provides an excellent summary of the techniques developed for UQ in SHM up to 2015. A further guide for validation workflow is presented in [53]. Resources which cover V&V problems specific to SHM include use of masses as 'pseudo-faults' in data acquisition [54] and UQ of an FEM of a bridge [8].

A V&V hierarchy for FMD-SHM is proposed in [10]. The approach suggested is as follows:

1. Hypothesis testing
2. Quantification using validation metrics
3. Visual diagnostics
4. Diagnosis by deterministic validation metrics

Each layer of this hierarchy provides more information regarding the model's performance, and to its sources of uncertainty [10]. Hypothesis testing is essentially the definition of a boundary between two or more distributions, such that null and alternative hypotheses are defined. Frequentist methods for hypothesis testing include the Kolmogorov-Smirnoff Test and Maximum Mean Discrepancy Test, while Bayesian Hypothesis Testing is also possible [10]. The criteria for validation metrics have been summarised as follows [55]:

1. The metric should quantify the difference between model prediction and observational data.
2. The metric should be interpretable such that it can be used to aid model improvements.
3. The metric should be objective and consistent across all models and applications.
4. The metric should reflect the full distribution of the datasets, rather than certain statistical moments.
5. The metric should be computationally efficient, and should converge reliably for iterative procedures.

The verification and validation of predictive models is a very mature research field, with a wide range of literature covering the area; the same can be said for uncertainty quantification, a process that often goes hand in hand with V&V. However, the quantification of epistemic (as opposed to aleatoric) uncertainty, is still an outstanding issue. The observations of aleatoric and epistemic uncertainties are explained concisely in [4]. Aleatoric uncertainty refers to the uncertainty arising from *known randomness*, such as noise on input measurements. By contrast, epistemic uncertainty pertains to actual *lack of knowledge*, exemplified by errors due to discretisation and model-form, and systematic measurement errors [4].

In [9], Roache clearly articulates the differences between error and uncertainty. Error is the value found by subtracting the true value of a variable from its measured value. Uncertainty is the lack of knowledge that can be represented by the distribution of a set of values that could be 'reasonably' attributed to the measured variable. The solution error is the difference between a simulator output and the true values, while the experimental error is the difference between the measured data and the true values. Total validation error for a model is the difference between these two values and is therefore independent of the true values.

The solution error can be defined as the cumulant of the model error (δ_{model}), numerical error (δ_{num}), and input error (δ_{input}). The goal of validation is to reduce model error, the goal of verification is to reduce numerical error and the goal of calibration is to reduce input error. Each of these can be made the subject of Equation 2.1, the expanded error equation for validation error (E) [9]. The final term in this equation, δ_D, represents the measurement error.

$$E = \delta_{model} + \delta_{num} + \delta_{input} - \delta_D \tag{2.1}$$

A complete appraisal of the errors associated with modelling and simulation is given in [56]. This resource defines thirteen errors sources based on a 'generic' system of interest, which are intended to be applicable across applications [56].

Uncertainty in the modelling process causes model discrepancy for both forward and inverse methods. This is often combined with measurement uncertainty in an overall term in the model execution. [52] gives a thorough overview of methods of model updating by Bayesian inference with examples and includes references to various Markov chain Monte-Carlo sampling methods. The paper also covers the use of fuzzy set theory for modelling of epistemic uncertainty. Model uncertainty is split into three separate types in [52]:

- Parameter uncertainty
- Structure or model form uncertainty
- Code or numerical uncertainty

Approximate Bayesian computation (ABC) is a popular method for model calibration due to its ability to incorporate prior beliefs into the process without requiring the definition of a formal likelihood function (the probability distribution of the calibration data) [57]. Two methods for model calibration, accounting for model discrepancy, are introduced in [10]. Model discrepancy is defined as the difference between the predictions of a model and the observed data when the true parameters of the model are known due to estimations and discretisations of physical processes in the model [5]; it is therefore usually the result of epistemic uncertainty.

Bayesian calibration and bias correction, developed by Kennedy and O'Hagan [58], was found to be effective when applied in conjunction with an informative prior in [10]. The method was used to remove model form errors in the calibration process more effectively than Bayesian calibration without bias correction. However, when significant noise was included on the training data, the method struggled to accurately identify parameter distributions [10].

Bayesian history matching, a special case of ABC [10], was also applied successfully, allowing for effective capture of uncertainties due to model discrepancy. In history matching, parameters are rejected from the plausible input space [59]. The method

avoided issues of non-identifiability which can hinder the Bayesian calibration and bias correction method [10].

2.6 Review conclusions

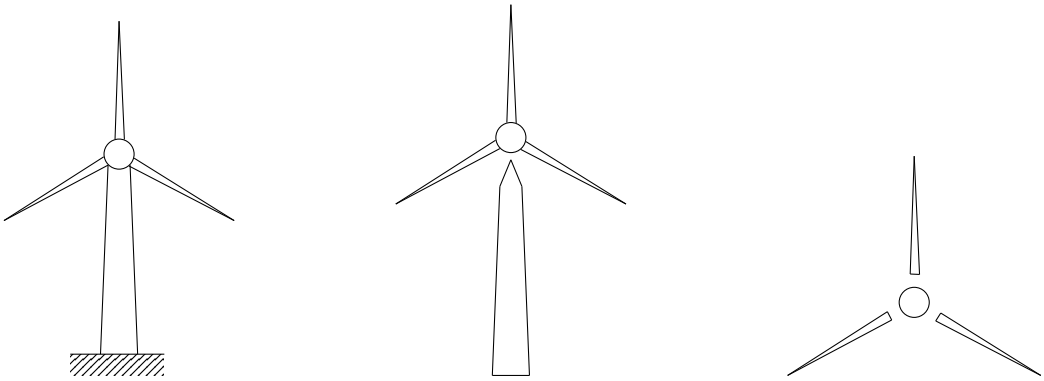
It can be concluded from this review that, while SHM techniques have become very advanced and can be widely applied, outstanding issues remain. Key among these is the lack of labelled damage data on which to develop statistical or physics-based models. This underlies key areas of interest in both data-driven and model-based SHM techniques. Further integration of risk and probability into damage identification is also required.

This thesis aims to address these issues by approaching FMD-SHM via a hierarchical modelling strategy. This is designed to enable validation to take place at the component level of structures, where labelled damage-state data may be more feasible to acquire. The framework for this and its impact within current SHM literature are described in the following chapter. The later chapters of the thesis will present tests of this framework on realistic datasets, in an attempt to underline the applicability of the techniques presented to real SHM problems in industry.

Chapter 3

Hierarchical verification and validation: Hypothesis

This chapter will introduce the main themes for the research in this thesis, focusing on motivations for the research and initial ideas for frameworks to be investigated later on; foreseen strengths and drawbacks of the proposals are also discussed. The broad aim is to develop a framework for hierarchical validation of physics-based models for SHM, which would allow for predictive damage models to be developed that do not rely on validation data from high-value structures.



(a) An assembly-level model of an onshore wind turbine

(b) A subassembly of the wind turbine containing the tower component and the turbine substructure

(c) A subassembly of the turbine substructure containing the blade and hub components

Figure 3.1: An example model hierarchy for an onshore wind turbine

3.1 Motivation for hierarchical validation

The goal of SHM is to identify structural damage in some way, ranging from detection to classification and prognosis [1]. Physics-based models can be used for SHM purposes in a variety of ways. IMD-SHM refers to model updating methods whereby a model's input parameters are updated in order to align its predictions with live structural data; this allows for damage to be inferred from the state of the model inputs [4]. Models can also be used in a forward sense, where their predictions are used to provide training data for statistical machine learning models for damage inference [4]. In order to validate a model for each of these purposes, damage-state data must be acquired for comparison with damage-state predictions. Many target structures for SHM are unique or of extremely high value because this value itself motivates the implementation of an SHM strategy. These requirements lead to a series of difficulties when conducting V&V in an SHM context, summarised below:

1. Target structures may be of prohibitively high value to carry out invasive or damaging data acquisition processes.
2. If the target structure is unique, usage requirements and other factors may restrict the data acquisition process, limiting the types of testing that can be performed.
3. The target structure may be difficult to scale or transport in a way that allows for well-designed, controllable laboratory tests.
4. The design or operating environment of the target structure may make sensor placement difficult, such as in the case of offshore wind turbines.
5. The operating condition of the structure may be difficult to replicate in order to acquire representative validation data.

Notwithstanding some successes [60, 61, 62], due to the issues outlined above, model-driven SHM has been significantly handicapped in its applicability in industry, despite a large volume of research into the effectiveness of the various techniques. This motivates research into the advancement of V&V techniques to mitigate the current difficulties.

In order to ascertain the accuracy and reliability of representative models of reality, their predictions must be ratified against actual observations from reality. This forms a significant portion of the V&V process, and is critical in establishing model-driven

systems that require any level of confidence in their predictions or outcomes. The term verification refers to the efforts to ensure that the model is accurate in its attempts to estimate a given solution, and verification therefore considers factors such as discretisation errors and errors of numerical model design [9]. Validation refers to the efforts to ensure that the model is an accurate estimation of reality and that the solutions it attempts to derive are representative of real-life observations; validation therefore considers model discrepancies, or biases, and random errors [9].

The concept of uncertainty can be separated into two contributors to overall lack of knowledge: epistemic and aleatoric uncertainty. Aleatoric uncertainty, also known as irreducible uncertainty, pertains to unavoidable uncertainties inherent to a problem that cannot be reduced with additional knowledge. Aleatoric uncertainty is generally unbiased and random, and can therefore be captured and estimated in the V&V process. Epistemic uncertainty, also known as systematic uncertainty, pertains to uncertainties due to actual lack of knowledge, for example in the case of model simplification or assumptions leading to certain physics being neglected from the problem. These issues can lead to biased uncertainties, but they cannot be accounted for solely by model discrepancy, as numerical error can also contribute significantly to overall bias [6].

Error refers to the difference between an estimation and the true solution, and is unavoidable when any level of uncertainty is present. Random errors are generally considered the more benign of the two types, and manifest in a scattering of predictions around a mean value that can be described by some statistical model. Systematic errors, which are also referred to as model discrepancy or model bias, cause a repeated offset between the prediction and the true value; these can be described by a particular function based on a set of input parameters.

Hierarchical V&V offers the potential for a complex model to be validated without the need for assembly-level data. This is achieved by using subassembly data to validate a series of submodels separately and then constructing an assembly-level model from the validated submodels. The uncertainty can be quantified at the assembly level by propagating the uncertainty from the subassembly levels upwards via dynamic substructuring, thereby establishing quantifiable confidence in the predictions of the assembly-level model. This chapter presents the ideology for hierarchical V&V in an SHM context, which aims to improve the feasibility of model-driven SHM in the long term. This is potentially achievable for the following reasons:

1. The method avoids the need to acquire damage-state data from assemblies that represent high capital investment.
2. Testing can make use of repeated subassemblies or components, particularly in the case of modularity or symmetry of components.
3. Ease of data acquisition can be improved as smaller-scale structures are required for testing.
4. The testing of simpler structures could improve ease of sensor placement.

3.2 Examples of hierarchical validation strategies

Gardner [10] directly investigates the process of validating a system model through validation of its subsystems. It is noted that it is through this capability that FMD-SHM offers an advantage over data-based methods; if damage state data of the full-scale system were required for model validation, the use of simulators would simply comprise an additional step between real-life training data and the machine learning process. The uncertainty propagation process was tested on a numerical study of a 4-DoF system, and damage location was successfully predicted at the full-system level.

Uncertainty propagation was used by Barthorpe [4] in order to pass uncertainty through increasing model rank in the V&V process. This allowed for V&V to be carried out on simpler submodels of the full system, which in turn allowed for modelling of a system in various damage states. V&V was not necessary for the full-scale model, as the confidence developed at the sub-system level was propagated upwards to provide confidence at the higher level.

Nagel and Sudret [63] presented a complete framework for the quantification of uncertainty in hierarchical models for the IMD-SHM task. Using Bayesian inference, both epistemic and aleatory uncertainty were passed through the various levels of a model using validation data from the lowest level upwards. The process was found to be highly computationally expensive and prone to issues due to high dimensionality and ill-posedness.

Li and Mahadevan investigated the use of two novel metrics for multi-level uncertainty integration in their paper [64]: model reliability as a validation metric, and data relevance. Data relevance was used to assess the relative importance of low-level component model validation when integrated into the larger assemblies, and it was

quantified using the Sobol' index [65]. The model reliability metric used was effectively a threshold on the prediction accuracy which, if exceeded, would result in the rejection of the model.

3.3 Strengths and drawbacks of hierarchical verification and validation

The key aim of hierarchical V&V for SHM is to eliminate the requirement for assembly-level damage-state data. This avoidance would significantly reduce the cost of validating models for SHM, and would in many cases make strategies feasible that would otherwise be considered prohibitively expensive.

Many common engineering structures exhibit modularity in their design, from the multiple blades of wind turbines to repeated deck sections in bridges. Validation at the assembly level is therefore highly inefficient when a single subassembly could be tested in isolation and then repeated throughout an assembly. Taking advantage of repeated components and subassemblies in this manner has the potential to enormously reduce the scale of the V&V task.

An extension of this logic could be applied to target structures that form part of a population; for example, in the case of wind farms or airline fleets. Individual structures in these populations may not be identical in specification, but they may share nominally identical components or subassemblies [17]. The ability to validate substructures of this type would offer strong value gains in the SHM of populations.

Many target structures for SHM are inherently difficult to acquire data from; this can be because of accessibility issues or simply because of the scale of the structure. Hierarchical V&V offers the potential to mitigate each of these issues. Validating below the assembly level will generally simplify the structures of interest; this would in many cases make sensor placement a less difficult task. It is also possible that the subdivision of the assembly into its submodels could be designed to maximise this benefit. In addition to this, working with smaller scale structures would allow for greater ease, or feasibility, of laboratory testing to acquire validation data. This would in turn allow for modellers to acquire better quality data, as laboratory testing (as opposed to field testing) will usually allow for greater control of experimental design.

Outside the field of V&V, hierarchical model use has its own set of potential benefits. A key advantage of submodelling is that it grants the ability to focus model resolution, and therefore computational effort, in facets of the model assembly that are integral to its behaviour [66]. This therefore allows for more parsimonious model design. Another more practical benefit is that submodelling naturally allows for a division of labour; this would allow modelling departments within companies to effectively manage large projects and workloads collaboratively.

Hierarchical V&V (and hierarchical model design in general) does offer certain difficulties. The foremost issue with using hierarchical V&V in SHM is the question of how to quantify the uncertainty in the submodels, and then propagate this uncertainty through the assembly process [10]. Handling of uncertainties in a complex model can often lead to extremely computationally intensive activities, for example in the use of random sampling methods to propagate input uncertainties through to model predictions. In addition, while aleatoric uncertainty is a relatively well-understood discipline, the quantification and handling of epistemic uncertainty is an open research question. Consideration into how this would affect the ability to assign confidence to an assembly-level model based on a set of validated submodels must be taken into account.

Another key difficulty in the use of submodelling is the question of how to model the interfaces between the substructures [66]. Joint behaviour can greatly affect the dynamic behaviour of a structure, and therefore presents a potential pitfall if not properly accounted for during the assembly process. It could also be very difficult to model joint behaviours from the assembly as boundary conditions of the submodels. On the other hand, the use of submodels does offer the potential for joints to be considered in specific detail, with bespoke interaction models able to be designed to reflect complex physical behaviours.

Finally, a potential drawback of hierarchical V&V is the effect of local versus non-local, or global, behaviours [64]. In designing and validating the submodels for SHM, modellers must make sure to focus on model behaviours that would allow for global damage detection in the assembly, and are not sensitive to local effects in the substructure only.

3.4 Proposed framework

In order to give structure to the successive discussions in this thesis, an initial proposed framework for hierarchical V&V is presented in this section. This initial framework is illustrated in Figure 3.2. This shows the steps required to fully develop a model using hierarchical V&V, with the steps indicated in separate blocks, and other inputs or endpoints to the process indicated in circles.

3.4.1 Operational evaluation

The statistical pattern recognition paradigm for SHM defined by Farrar and Worden runs as follows [2]:

1. Operational evaluation
2. Data acquisition
3. Feature selection
4. Modelling for feature discrimination

While this thesis concerns model-driven SHM strategies (as opposed to data-driven strategies), the paradigm is still useful as an overall guide. Operational evaluation refers to the process of defining the justification and motivations for the SHM strategy, the damage states of interest, the operating conditions of the target structure, and the specific limitations concerned with data acquisition from the structure [2]. It must therefore precede any further activities in the development of a model-driven SHM strategy.

3.4.2 Model conceptualisation

Based on the recommendations of the operational evaluation, an assembly-level model can be designed on a conceptual level to develop a representative version of the real-life system. This will define the scope of the model and will provide a reference for the following activities in the framework. Considerations should be made at this stage as to the required level of complexity in the model, as well as to the confidence required in its predictions and any further factors of safety that may be required. The use of the model in an SHM capacity (for damage detection, localisation, etc.) should also be decided, in order to inform its design for either updating or forward model-driven methods. Initial specifications on selection of statistical models should also be defined here in the case of FMD-SHM.

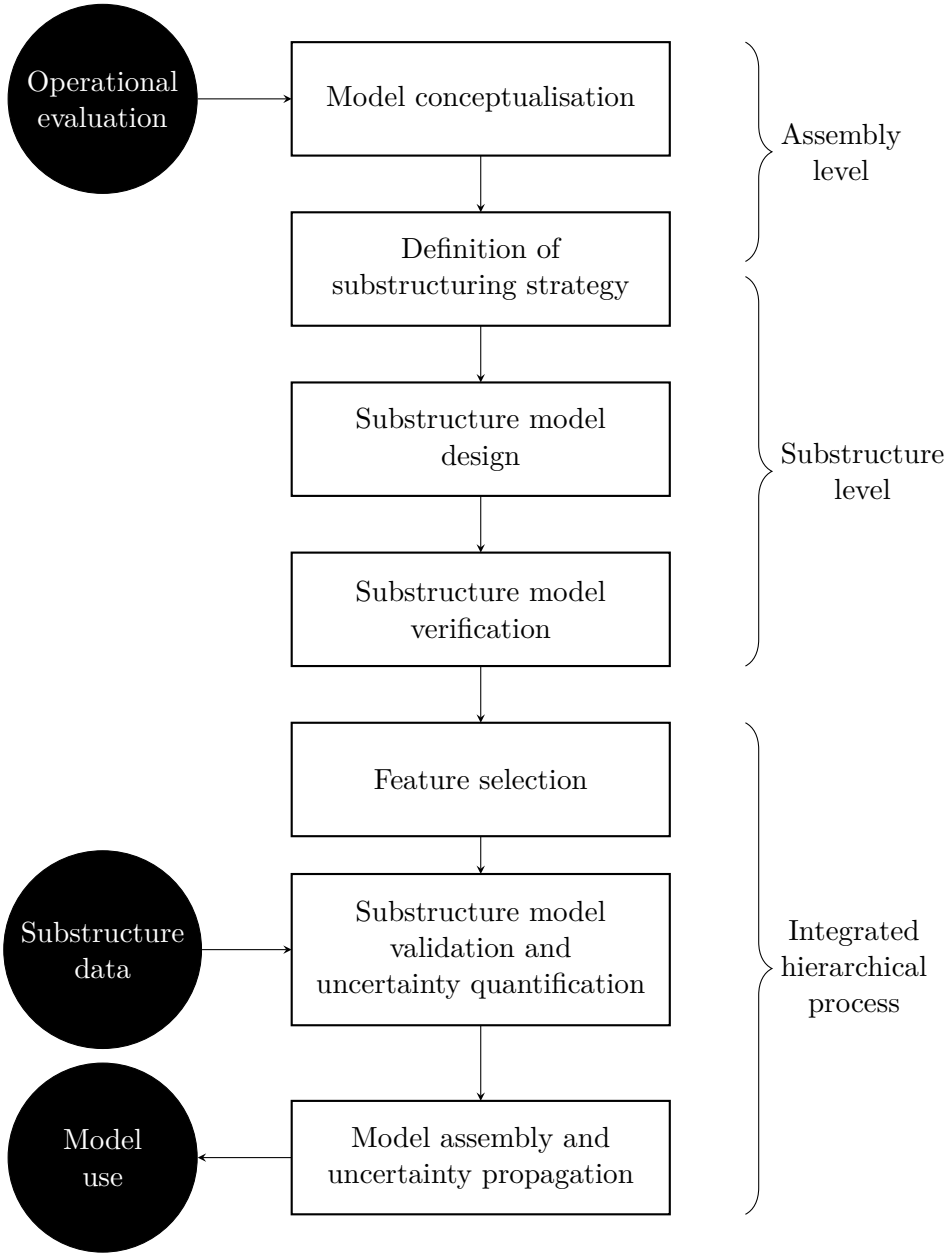


Figure 3.2: Proposed framework for hierarchical V&V

3.4.3 Definition of substructuring strategy

Once some idea of the form of the required assembly-level model is defined, the focus can move on to the substructuring strategy that would enable hierarchical V&V. The aim of hierarchical V&V is to validate an assembly-level model using substructure-level data only. The aim of the assembly substructuring strategy is therefore to define a series of submodels that will maximise the ease of data acquisition. Dynamic substructuring has been identified in this thesis as suitable for the hierarchical V&V process, and this is discussed in full in Chapter 4.

The main trade-off introduced in hierarchical V&V is the additional complexity of joint modelling; this must be considered at this stage of the process. The interfaces between substructures must be approximated closely, and any model-form errors introduced should be anticipated to inform error quantification in the model re-assembly process.

3.4.4 Substructure model design

This process entails the actual creation of a set of mathematical models to represent the assembly substructures, and implementing the submodels in computational code. This will usually involve the creation of finite element models. The models should be suitably parameterised to enable effective model calibration (and to allow for validation of the calibrated parameters), and a set of inputs to reflect the operating condition and health state of the structure should be defined. These criteria will draw on the specifications set out in the previous stages of the framework.

3.4.5 Substructure model verification

Verification refers to the efforts to ensure that the computational processes written to implement a mathematical model are accurate [9]. Checks should be carried out at this stage of the process to ensure that the submodels are free of coding errors. In addition, where FEM codes are employed, convergence studies must be carried out to ensure that the domain discretisations are suitably refined and that the code predictions are converged to an acceptable level. The uncertainty associated with computational models should also be quantified at this point as part of the verification process.

3.4.6 Feature selection

Once the submodels have been verified, feature selection can be carried out to determine the features in the model outputs that are most suitable for SHM. These features should ideally be sensitive to damage but robust to other varying factors.

Carrying out feature selection at this stage of the process allows for streamlining of the model validation process, as it will inform the data processing required for the validation data and simplify the problem domains over which the submodels should be validated. Further discussions of these ideas can be found in Chapter 5.

In the case of FMD-SHM, feature selection should reflect the requirements of the statistical model as well as the physics-based model. Further criteria for the design of the statistical model can be defined here.

3.4.7 Substructure model validation and uncertainty quantification

Validation refers to the efforts to ensure that the predictions of a representative model are accurate compared to reality [9]. This stage is therefore key to the overall hierarchical V&V process, as it will ensure the accuracy and reliability of the submodel predictions compared to structural data.

Validation should be carried out over the full operating range and at all damage states of interest in the structure. The acquisition of validation data from the structures should therefore reflect this, with a strategy being drawn up based on the operational evaluation and overall substructuring strategy. Separate parameter calibration data may be required for accurate estimation of the model parameters.

Full uncertainty quantification should also be carried out for each of the submodels at this stage. This should take into account code uncertainties that were quantified in the verification process, as well as the uncertainties associated with validation. Uncertainty should be quantified in terms of distributions on the model parameters and inputs, and in terms of any additional model discrepancy functions where required. These uncertainties would then enable probabilistic predictions using the submodels, which would quantify the confidence that can be attached to them.

3.4.8 Model assembly and uncertainty propagation

Once the submodels have been fully verified and validated, and the uncertainty in their predictions can be robustly quantified, they can be re-assembled to produce the assembly model. This process must allow for propagation of the uncertainty quantified at the substructure level to the assembly level. If this process is successful, then no further validation activities will be required. However, some awareness should be kept of the modelling of the interfaces between the substructures, as this could introduce a significant further model discrepancy if not carried out adequately.

Preliminary discussions concerning this stage of the process, focusing on dynamic substructuring and associated UP techniques, can be found in Chapter 4. A case study on a realistic dataset is presented in Chapter 6.

3.4.9 Model use

Having developed a physics-based model for SHM using hierarchical V&V, the model can be used for its given application. As has been hypothesised in this chapter, this could potentially be achieved without the requirement for assembly-level validation data, which could significantly reduce the data acquisition problem commonly associated with SHM. The application of models developed via this framework is presented in Chapter 7.

3.5 Conclusions and further research in this thesis

The framework discussed in this chapter should be considered an initial guide for hierarchical V&V in SHM only; aspects of it will be discussed and developed as part of this thesis. Significant further work is required to cement the form of the framework, but, in its current form, it should provide a jumping-off point from which its benefits can be fully realised.

Conceptually, there is a clear scope for the reduction of the lack-of-data problem currently present in model validation for SHM. These ideas are further refined in this thesis through specification of various methodologies and testing of the techniques on real-life datasets.

The hierarchical V&V framework was tested for an FMD-SHM strategy throughout this project and the results are presented in the following chapters of this thesis. The structure of interest was a laboratory-scale truss bridge, with experimental data

drawn from the bridge struts for submodel validation (Chapter 6). Assembly-level data was recorded for testing of the framework (the experiments are presented in Chapter 5, the testing of the framework is carried out in Chapter 7), with assessment of model accuracy following hierarchical validation carried out alongside typical SHM techniques.

Chapter 4

Hierarchical verification and validation: Methodology

This chapter presents the methodologies that can be employed to carry out hierarchical V&V in an SHM context, as was postulated in Chapter 3. Dynamic substructuring methods are presented and evaluated as a viable method for model assembly of predictive damage models, and uncertainty propagation techniques are implemented in tandem to investigate how uncertainty quantified at a subassembly level can be robustly translated to validated predictions at the assembly level. Through the demonstration of a numerical case study, it will be shown that dynamic substructuring is applicable to the hierarchical V&V problem. This will then provide a basis for rigorous evaluation of the methods against experimental test data in the following chapters.

A nomenclature for this chapter (and the thesis as a whole) is provided below.

Structure An engineering system with a particular operating domain that can be described by a set of physics-based laws.

Substructure A system that can be defined within a structure, often representing a particular subassembly or component, that can be described with its own set of physics-based laws.

Assembly A structure that comprises a set of subassemblies and/or components connected by joints.

Subassembly An assembly of components that are connected by joints that itself would be combined with other subassemblies to form an assembly.

Component A basic structure containing no modelled joints that can be described by its own bespoke set of physics-based laws.

Model A physics-based set of laws defined using engineering expertise to estimate the behaviour of a structure.

Submodel A discretisation of a larger model over a particular domain that can be combined using other submodels and known constraints to recover the 'parent' model.

4.1 Dynamic substructuring

Dynamic substructuring refers to a group of methods for the assembly and disassembly of physics-based models used for dynamic analyses [66]. This allows for a model of an assembly-level structure to be broken down into a set of submodels representing substructures or components of the larger assembly. The techniques are reversible, meaning that the assembly-level model is recoverable from its submodels. Dynamic substructuring includes a diverse range of techniques, including methods for reducing computational effort through model reduction and implementations of uncertainty quantification and propagation. Based on these qualifying criteria, dynamic substructuring was outlined as a means for applying hierarchical V&V to physics-based models for SHM; a demonstrative application is used to justify the choice in this chapter.

4.1.1 Motivation

Dynamic substructuring enables the solving of highly complex modelling challenges. Physics-based models of dynamical systems, usually based on FEM, are often highly onerous to execute due to the high number of degrees of freedom required to adequately represent engineering structures; this problem is compounded in SHM, where models may have to be executed many times in order to produce a probabilistic solution or solve an inverse problem. Therefore, techniques to reduce this computational load are highly desirable.

It is a generally accepted paradigm that to solve a complex, multi-level engineering problem, the problem domain can be broken down to define a set of more basic problems [66]. This process can allow the overall solution to be approached piecewise over a series of more tractable solutions, provided that a complete solution can be recovered

from these. This idea forms the basis of FEM as a technique, the widespread uptake and success of which should clearly underline the usefulness of problem discretisation.

Dynamic substructuring allows modellers to tackle systems which would be highly complex to describe adequately using a unified mathematical model. Instead, an assembly can be constructed from a series of suitable substructure models; these submodels can be developed to describe any particular set of subassemblies or components. This enables modellers to approach assemblies in a more simplified manner, allowing for more efficient allocation of computational resource, and more high-fidelity simulation of significant components or processes where required (this can be compensated for by saving computational effort on substructures that can be adequately approximated with relatively coarse solutions). Further to this, dynamic substructuring allows for the definition of key aspects of the physics of a system to be assigned in a bespoke manner throughout an assembly; for example, a certain component or joint could be modelled as a non-linear system, while the remaining components could be accurately represented using linear models. It should be kept in mind that an assembly model will always be (by definition) more complex than each of its submodels; this is because it will usually contain all of the information encoded in the submodels, as well as additional constraints that are defined in the assembly process.

4.1.2 Background

The foremost resource in the field of dynamic substructuring is the work by Allen et al. [66]. This book comprehensively covers the techniques for dynamic substructuring that are available and details the circumstances for their implementation. The Craig–Bampton method for dynamic substructuring [67] (a widely used method for the assembly of substructures and for model reduction in which the substructures are separated into their boundary and internal degrees of freedom) is outlined in [66, 68]. Dynamic substructuring was presented as a technique for aiding the analysis of offshore wind turbines in [69]; this allowed the investigators to capture local dynamic effects and their contribution to the global system with accuracy. Uncertainty quantification and propagation techniques were applied to frequency-based substructuring in [70]; this showed that the dynamic substructuring methodology was receptive to integration of uncertainty and error, underlining its credentials in an SHM context. Two theses applied the use of dynamic substructuring to problems in the automotive industry [71, 72]. The former work focuses on the use of experimental frequency-based

substructuring in the analysis of the gear train of BMW cars. The latter is more primarily concerned with applying UP to experimental frequency-based substructuring with a view to developing a strategy for system identification.

	Linear substructures	**Non-linear substructures**
Linear joints	• Well-understood in literature. • No computational difficulty due to non-linearity. • Danger of model discrepancy if non-linearity is present in system of interest.	• Reasonably well-explored in literature, but lacking benchmark methods. • Some computational difficulty added due to non-linear substructures, the extent of which depends on the type of non-linearity considered. • Dynamic substructuring allows for non-linear components to be modelled in a bespoke manner within the assembly; however, discrepancy may be present in the assembly if interaction effects are not anticipated.
Non-linear joints	• Reasonably well-explored in literature. • Significant computation effort required due to time-integrating solving methods (some novel research challenges this). • Allows for incorporation of non-linear friction interactions between substructures, which is a common interface condition in assemblies.	• Limited novel research in literature, not directly addressing assembly of non-linear substructures with non-linear joint behaviours. • Heavy computational load required for fully non-linear model. • Would allow for development of very low-discrepancy model assemblies.

Table 4.1: A summary of the state of research on non-linear dynamic substructuring, including key points of interest and considerations for each field

The majority of physics-based modelling in engineering and SHM concerns linear models, mainly because many engineering structures are often designed to operate in the linear regime, or because observable non-linearity in the real system is relatively weak and can therefore be ignored. Incorporating non-linearity into modelling processes adds significant layers of complexity to the design and implementation of the model, not to mention additional computational burden in running the model; therefore, weak non-linearity or non-linearity outside the domain of interest is often neglected. However, there are a range of engineering applications in which non-linearity must be considered, including several examples within the field of SHM. As such, the field of non-linear modelling has been explored in depth in the literature, including non-linear FEM for dynamic models. On the other hand, the incorporation of non-linearity into dynamic substructuring methods has been limited, and research into how best to incorporate non-linearity into model hierarchies and assemblies is an ongoing task.

When considering non-linearity with respect to dynamic substructuring, it is clear that two aspects of non-linearity are key to consider: non-linearity in the substructures, and non-linearity in the joints between substructures (which may themselves be linear or non-linear). A brief discussion of these two considerations is given in Table 4.1.

An introduction to non-linearity in dynamic substructuring is provided in [66], which considers methods for modelling geometric non-linearities in substructures. In addition to this, a review is provided of literature concerning the modelling of weakly non-linear joints. Craig–Bampton has been extended for application to geometrically non-linear substructures in [68], and is also included in [66].

Modelling of more complex non-linearities generally requires some sort of transient analysis, and, as such, time-step integration is often used to derive a solution. This was applied to dynamic substructures in [73] in conjunction with a Craig–Bampton substructuring process. Time-integration methods were also used in order to reduce model discrepancy in models with localised non-linearities in [74, 75, 76, 77]. Avoiding the necessity for time-integrating methods in dynamic substructuring methods for non-linear joints was investigating using a novel technique in [78].

Non-linear experimental substructuring was investigated in [74] and [75], where non-linear boundary conditions and non-linear couplings between substructures were modelled respectively. A comprehensive modelling strategy for a bridge has been carried

out where the substructured model predictions were validated against experimental data in [76]. Frictional effects were considered in [77], where a state-space formulation was also implemented in the substructuring process.

4.1.3 Methodology

Dynamic substructuring can be approached in several distinct manners, depending on the information available to the modeller and the intended application of the model. There are two mathematical processes for the assembly of substructures: primal and dual assembly. These are equivalent to each other mathematically, but each lends itself to different techniques and situations [66].

In assembly of a set of substructures, conditions need to be defined which describe the interface behaviour at the joints. The two key defining conditions are DoF compatibility, and interface force equilibrium. The simplest compatibility constraint is to set the responses to be equal at the interface; however, this can be loosened to improve accuracy in the modelling of joints.

The compatibility condition is enforced by defining a matrix, B (the compatibility matrix). B is a signed Boolean matrix; in the case of rigid connections, it is defined such that its product with x (the response vector of the substructures) is the zero vector (Equation 4.1). The dimensions of B are the number of interface connections in assembly by the number of unassembled DoFs.

In addition to DoF compatibility, the interface forces must satisfy the constraint of equilibrium in assembly. This constraint is enforced by the matrix L (the localisation matrix), which is defined such that the product of its transpose with g, the vector of interface forces, is equal to zero (Equation 4.1) in the case of rigid connections at the interfaces. It should be noted that internal forces are present in all masses but do not generally contribute to global dynamic behaviour (or are neglected from most models). However, the internal forces at the interface of an assembly must be combined to describe a new set of internal forces within the new lumped mass.

The localisation matrix is an unsigned Boolean matrix whose dimensions are the number of unassembled DoFs by the number of DoFs in the assembly. L will also map the global vector of assembled degrees of freedom, x_{global}, to x (Equation 4.2). It can therefore be used to remove redundant information from the assembled equation of motion, and it can also be shown that the product of B and L is the null space (for non-zero solutions).

$$\begin{cases} M\ddot{x} + C\dot{x} + Kx = f + g \\ Bx = 0 \\ L^T g = 0 \end{cases} \tag{4.1}$$

For a substructuring problem under the assumption of rigid joint connections, the assembly can be represented using the three-field formulation (see Equation 4.1). The mass M, damping C and stiffness K matrices for the assembly are assembled from the substructure matrices in a block-diagonal form, while the force f and response x vectors are concatenated vertically. Other quantities shown in these equations are the interface force vector g, the compatibility matrix B and the localisation matrix L.

$$x = Lx_{global} \tag{4.2}$$

Starting with the three-field formulation, to apply primal assembly, Equation 4.2 is substituted into Equation 4.1 to eliminate redundant response entries from the equation of motion. Following this, the equation is premultiplied by L^T to eliminate the vector g. This yields Equation 4.3.

$$\tilde{M}\ddot{x}_{global} + \tilde{C}\dot{x}_{global} + \tilde{K}x_{global} = L^T f \tag{4.3}$$

Direct physical assembly of substructures can be carried out simply using the primal assembly method. To derive the assembled and updated parameter matrices ($\tilde{M}$, $\tilde{C}$ and $\tilde{K}$) the parameter matrices from the full equation of motion (Equation 4.1) are premultiplied by L^T and post-multiplied by L. This process is similar to the assembly of submodels in FEM. Following this transformation, solutions for the assembled equation of motion can be derived as normal. This method has the benefit of being very physically informative – the compatibility and constraint matrices can be derived analytically, given knowledge of the joint coordinates – but can become computationally difficult when combining substructures with many DoFs.

In dual assembly (also known as the Lagrange multiplier method [66, 71, 72]), a new vector, λ, is defined. This is a vector of Lagrange multipliers containing the magnitudes of g, according to Equation 4.4.

$$g = -B^T \lambda \tag{4.4}$$

When Equation 4.4 is applied to the three-field formulation, it can be reduced to Equation 4.5. In this case, the interface force equilibrium condition is satisfied by definition.

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \ddot{x} \\ \lambda \end{pmatrix} + \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \dot{x} \\ \lambda \end{pmatrix} + \begin{bmatrix} K & B^T \\ B & 0 \end{bmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{4.5}$$

Frequency-based substructuring refers to the use of dynamic substructuring methods in the frequency domain, and typically uses the dual assembly process. Frequency-based substructuring is often used in experimental substructuring processes [70, 71, 72]. For conciseness, Equation 4.5 is simplified by substituting in Z, the structural impedance of the assembly.

$$\begin{bmatrix} Z & B^T \\ B & 0 \end{bmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{4.6}$$

Using the admittance, Y (the inverse of impedance, Z), this system of equations can be rearranged to derive Equation 4.7, which gives the response of the assembly. The admittance of the subsystems, Y, is a block diagonal construction of the substructure admittances. The assembled admittance, x, is therefore of the same number of dimensions as Y, with duplicate entries at either side of the interface; this means that the matrix holds redundant information. The redundant entries can be removed to derive the global admittance for the assembly.

$$x = \left[Y - YB^T \left(BYB^T \right) BY \right] f \tag{4.7}$$

The modal domain can also be used for dynamic substructuring methods. This allows modellers to reduce the number of modes used in assembly based on the idea that the higher order modes contribute less energy to the full solution. Therefore, an accurate solution can be estimated with significant potential reductions in computational load, at the expense of an error term due to the approximation [66].

$$x = \Phi\eta \tag{4.8}$$

To apply dynamic substructuring in the modal domain, the response vector, x, is replaced with η, the response in modal coordinates using Equation 4.8. The discarding of higher-order modal contributions introduces a residual error term, r, to the

equation of motion; this can be eliminated by premultiplying the equation by Φ^T, the matrix of mode shapes that are retained in the solution.

$$\begin{cases} M_m\ddot{\eta} + C_m\dot{\eta} + K_m\eta = f_m + g_m \\ B_m\eta = 0 \\ L_m^T g_m = 0 \end{cases} \tag{4.9}$$

The whole process yields an updated three-field formulation (Equation 4.9), from which primal or dual assembly can be carried out. The reduced constraint matrices, B_m and L_m, are defined according to Equation 4.10 and Equation 4.11.

$$B_m = B\Phi \tag{4.10}$$

$$\Phi L_m = L \tag{4.11}$$

In any dynamic substructuring procedure, the assembly is carried out by applying physical constraints to the substructures that define the joints in some way. Therefore, it is critical that the modeller fully understands the specification and degrees of freedom concerned with this definition (the Craig–Bampton method targets this by identifying the interface DoFs prior to any substructuring, in order to maintain relative ease of interface compatibility). In addition, assumptions that are made to simplify the dynamic substructuring assembly process, such as the assumption of rigid connections (which has been used in the above derivations), must be stated, and attempts to quantify the discrepancy they introduce are highly recommended.

4.2 Uncertainty propagation

Three methods for UP through a dynamic substructuring process have been investigated in this thesis and are summarised in this section: Monte-Carlo sampling, Latin hypercube sampling [79] and linear perturbation. These methods were selected as initial candidates for UP as they are relatively straightforward to understand and implement; more advanced techniques may be required where necessary (various direct implementations of UP within dynamic substructuring are available), for example in the case where computational limitations may make sampling-based methods unfeasible.

Monte-Carlo sampling is based on efforts to sample from a distribution on a purely random basis, and is therefore the preferred method when sufficient computing resource is available to run enough samples to achieve a converged distribution. Latin hypercube sampling is a stratified sampling method that aims to reach convergence faster than Monte-Carlo; this is achieved by dividing the prior distribution into subregions or strata of equal probability within each marginal distribution and randomly sampling once from each those subregions; this space-filling method is analogous to placing rooks on a chessboard such that no piece threatens another piece (in this analogy the chessboard represents a Latin square – a two-dimensional distribution). An illustrative example of Latin hypercube sampling from two uniform distributions in the region $[0, 1]$, x_1 and x_2, is given in Figure 4.1.

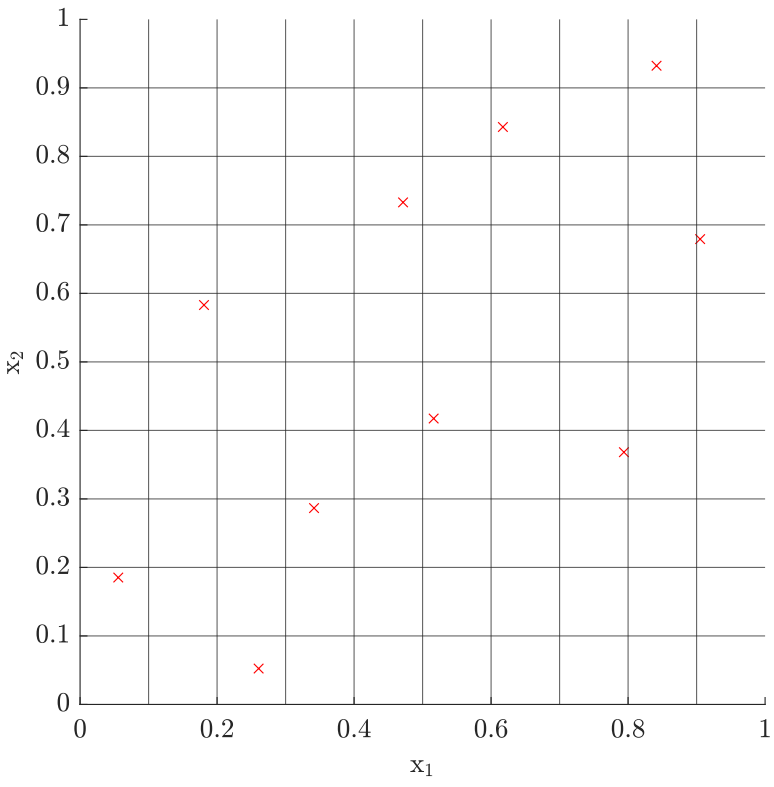


Figure 4.1: Example of Latin hypercube sampling from a two-dimensional joint distribution, also known as a Latin square, with 10 samples

Linear perturbation uses the mean and the perturbed mean only to evaluate the distribution of the predictions, and is therefore the least computationally intensive

of the three methods investigated in this thesis. Generally, the mean value for each parameter value is perturbed by one standard deviation from its distribution [6]. The values of the mean and perturbed parameters and predictions can be used to construct the sensitivity matrix, S, where $S_{i,j}$ contains the entry $\frac{\partial f_i}{\partial Q_j}$ (with f denoting the vector of model predictions and Q the vector of parameters). The covariance matrix of the model outputs can then be estimated according to Equation 4.12, in which V denotes the covariance matrix of the model parameters.

$$var[f(Q)] = S^T V S \tag{4.12}$$

This method of linear perturbation provides an *estimate* of the distributions of model predictions based on parametric uncertainties. This estimate is based on a truncation of the Taylor series that fits the model; therefore its accuracy is limited based on the linearity of the model over the problem domain.

Direct evaluation of uncertainty was also considered for application to dynamic substructuring. This has the virtue of giving the exact uncertainty on model outputs based on the input uncertainty. However, this method quickly becomes highly complex when a variety of inputs and outputs of interest are considered. It was therefore not thought to be readily applicable to problems in SHM, as in this case, and was not implemented in this investigation.

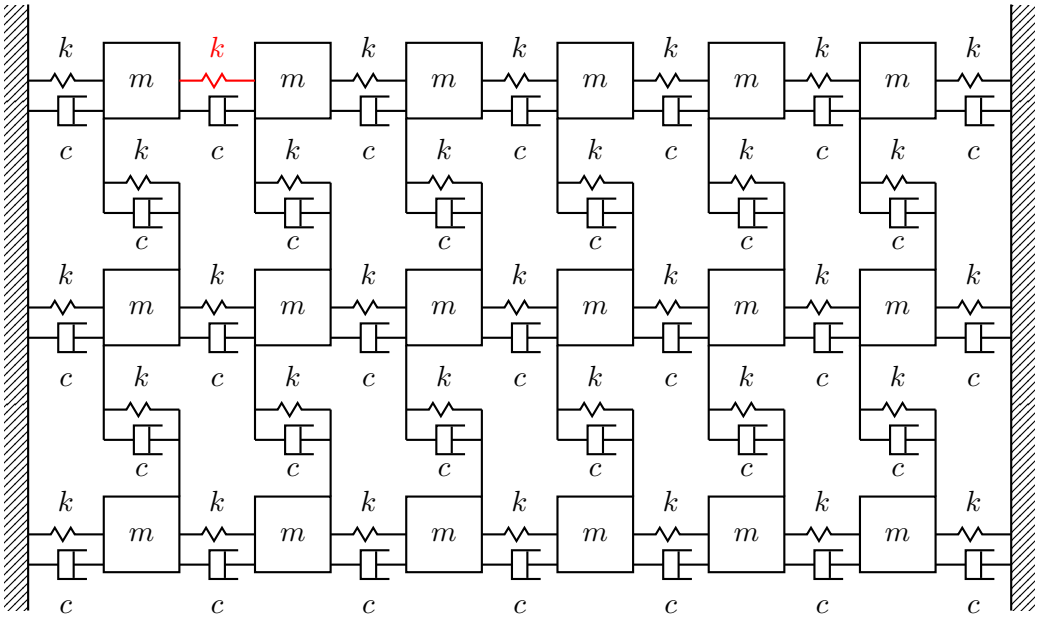
Various methods for UP bespoke to dynamic substructuring are also available. An analytical approach to UP for frequency-based substructuring methods can be found in [80]; this uses a truncated Taylor series approximation to propagate a normal distribution through the assembly of frequency response functions and compares the results to an equivalent Monte-Carlo simulation. A similar linear perturbation approach has also been applied to the propagation of uncertainty through FRF assembly in the aerospace sector [81]. Another, earlier, perturbational method for UP to estimate uncertain FRFs, where Monte-Carlo sampling used to draw from uncertain FRF predictions can be found in [82]. The results in this paper compared favourably to a full Monte-Carlo sampling scheme. UP for Craig–Bampton in has also been investigated in an aerospace context [83]. A perturbation-based method was investigated, and the results were compared against Monte-Carlo sampling. This paper used an extended version of the Craig–Bampton Stochastic Method proposed in [84]. The impact of perturbation magnitude was also investigated [83]. Sampling-based methods were applied to a Craig–Bampton substructuring scheme in [85]; this highlighted issues with

sampling methods when large numbers of model parameters are used. An example of UP in model reduction methods for aerospace applications can be found in [86]. Here, meta-modelling was used to reduce burden of sampling from reduced-order models for UP, in order to predict modal analysis results from the model through eigenvalue computation. A review of uncertainty handling in dynamic substructuring can be found in [87].

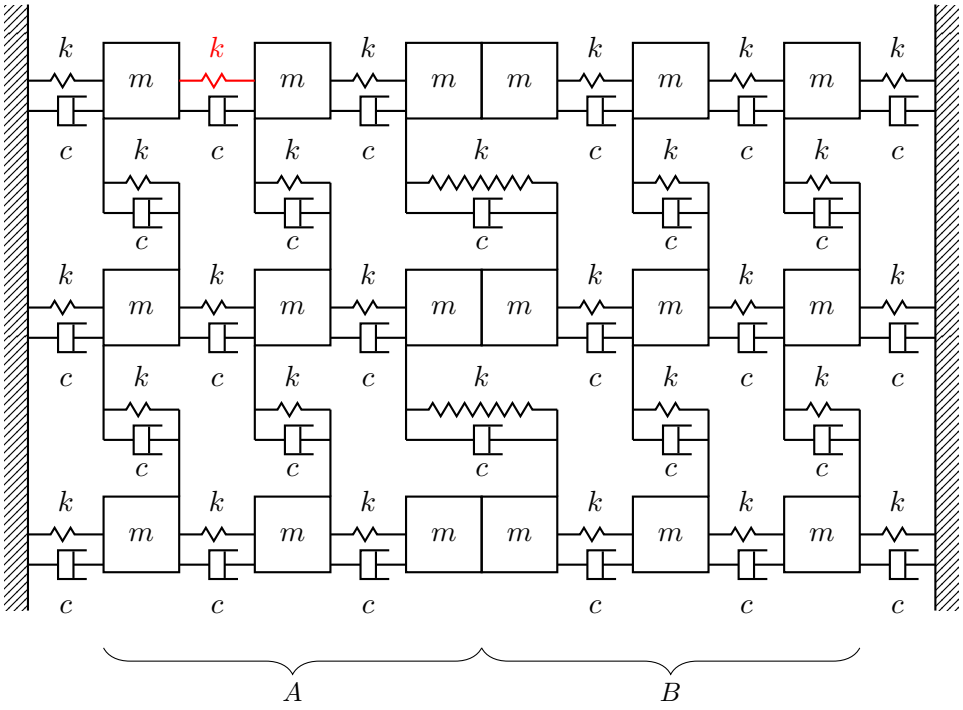
4.2.1 Case study: Uncertainty propagation in a plate assembly

The methods of Monte-Carlo sampling, Latin hypercube sampling and linear perturbation are evaluated for UP in this case study. These methods were selected as they were relatively straightforward to implement, which would allow for timely progression of the following investigations. Given the relatively small number of DoFs in the models involved in this thesis, sampling methods were not expected to be prohibitively expensive to execute.

The full assembly model represents a plate that is constrained at opposite ends, from two substructures representing each half of the plate (themselves fixed at one end and free at the other). The assembly is illustrated in Figure 4.2. The propagation of parametric uncertainty of the substructures, as well as model discrepancy at the assembly level are taken into account. Damage conditions could be simulated in the models by reducing the stiffness of any of the springs. This reflects that damage in structures can often cause a local reduction in stiffness; natural frequencies are often useful as features to monitor for these effects. The case study compares and contrasts the model predictions when using the three methods of UP, and also investigates the propagation of uncertainty in the plate in the presence of increasing damage.



(a) The true plate assembly



(b) The assembled substructures A and B

Figure 4.2: Plate assembly estimated using dynamic substructuring and compared to true model with damage location marked in red

The plate substructures were modelled as mass-spring-damper structures with nine DoFs; the equation of motion is given in Equation 4.1. The model parameters were sampled from normal distributions as summarised in Table 4.2; this was intended

to reflect a scenario in which V&V had been carried out to construct appropriate distributions for the key parameters of the submodels. The choice of the normal distribution was made for convenience and illustrative purposes; the fitting of probability distributions to describe uncertainty is an important element of UQ, where assumptions of a particular distribution can prove problematic [6] (the choice of a flexible method for UQ – such as ABC – can mitigate this issue in practice [88]). The normal distribution can cause problems in some cases because it is possible to sample parameter values that are not physically possible (for example, stiffness or mass cannot be less than zero); other distributions such as a uniform distribution can be more effective in situations like these. Proportional damping was modelled for the dampers. The 'true' assembly was modelled using the mean values for each parameter; this was used to evaluate the performance of the assembled substructure model. The models were used to predict the natural frequencies of the plate assembly using the eigenvalue solution of the characteristic equations of the models; the uncertainty associated with these predictions was then compared for the three UP methods.

	Mean	**Standard deviation**
m	1kg	0.1%
c	10Ns/m	10%
k	10000N/m	1%

Table 4.2: The parameters used for the plate assembly

The dynamic substructuring method of primal assembly in the physical domain was used, where the parameter matrices of the substructures are assembled and updated directly using the L matrix. The assembly process was carried out by using the assumption of rigid joints, which made definition of the B matrix trivial – the L matrix could then be derived numerically. The rigid-joint assumption introduced a discrepancy to the benchmark model, which contained an additional set of springs. This was intended to reflect that in a more realistic scenario some model bias introduced in the assembly process would be unavoidable.

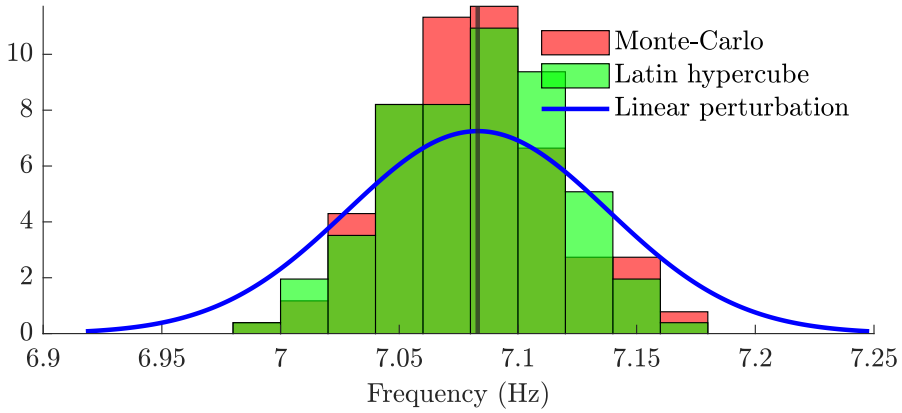


Figure 4.3: The histograms/estimated distribution for the first mode of the assembly (with true value marked as a vertical black line) in its undamaged state

Each of the UP methods were successfully applied to the dynamic substructuring problem. The sampling methods were run based on 128 samples. This number of samples was selected for numerical efficiency – verification activities would be required to set this value in the implementation of these methods. Typically, samples are taken in the order of at least 10^3 in order to assume convergence. The results are compared in Figure 4.3, which gives the predicted distributions for the first natural frequency of the substructured plate in its healthy state for the three UP methods, with the true solution marked. The sampling-based results are presented as histograms, while the propagated mean and variance from the linear perturbation were used to directly compute the normal distribution for this method. The sampling-based methods produce output distributions that appear close to normal, but with skewness values of 0.1217 and -0.0209 for the Monte-Carlo and Latin hypercube methods respectively; these results indicate some deviation from the normal distributions applied to the model parameters. The kurtosis values of the results presented in Figure 4.3 were 2.9992 and 2.7484 respectively.

The perturbation method is comparable in accuracy to the sampling-based methods but has a higher degree of uncertainty, with greater standard deviation observable in the predictions. This overestimate of uncertainty indicates that the linear perturbation approach is not suitable for this application, most likely because of invalid assumptions of linearity through the model assembly and eigenvalue analysis; this could potentially be improved by using more bespoke perturbation-based methods

for dynamic substructuring, such as in [80, 81, 82]. Of the two sampling methods, Latin hypercube has a lower offset in mean prediction from the true solution due to its expected outperformance of Monte-Carlo sampling in solution convergence (the errors for these were 7.7e-5% and 4.1e-2% respectively). It is clear, however, from Figure 4.3 (and from the error values) that both of the sampling methods were fairly accurate – using greater numbers of samples should further increase this accuracy.

The two sampling methods can also be compared on their performance in terms of estimating a converged output distribution for the uncertainty on the model predictions. This is illustrated in Figure 4.4 for the first mode of the assembly in its undamaged condition. This figure, plotting the Kullback–Leibler divergence metric for the two sampling methods, shows that the Latin hypercube method converges to a stable distribution much more quickly than the Monte-Carlo method. The sampling processes were repeated five times to produce this figure. The Kullback–Leibler divergence metric is a measure of the similarity between distributions, which will tend to zero as the distributions converge [89]. The Kullback–Liebler divergence for two distributions, f_1 and f_2, $(D_{\mathrm{KL}}(f_1 \| f_2))$ can be evaluated for x samples according to Equation 4.13 [89].

$$D_{\mathrm{KL}}(f_1 \| f_2) = \sum_x f_1(x) \log \left(\frac{f_1(x)}{f_2(x)} \right) \tag{4.13}$$

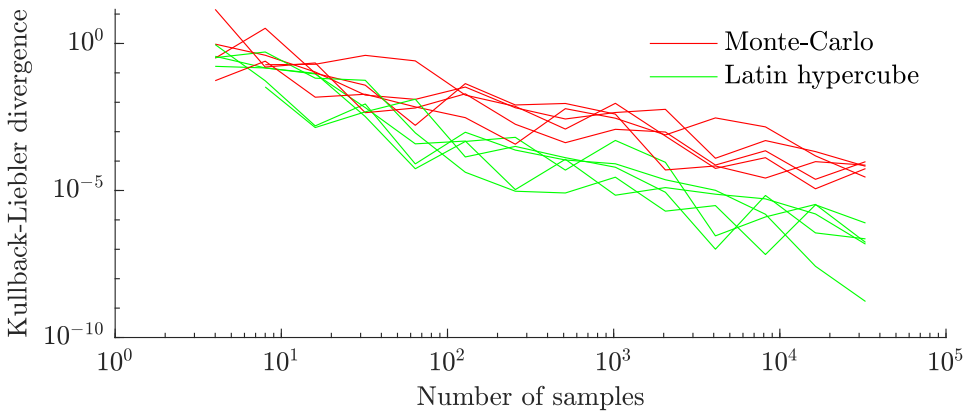


Figure 4.4: The Kullback–Leibler divergence for the first mode of the plate assembly in its undamaged condition based on Monte-Carlo and Latin hypercube sampling methods

This case study succinctly demonstrates the relative merits and drawbacks of the three UP methods investigated: Monte-Carlo sampling, Latin hypercube sampling and linear perturbation methods. Of the two sampling methods, Monte-Carlo is preferable because it is based on true random sampling from the parameter distributions. However, where computational resource is limited, the Latin hypercube method could outperform Monte-Carlo, as the stratified nature of its sampling process allows it to achieve convergence with far fewer samples. If computational resource is extremely critical, the linear perturbation method can be used. This method requires two model evaluations only, but provides a much more rudimentary estimate of the output distribution – the assumption of linearity in the assembly process and modal analysis was a likely cause of inaccuracies here (the measurable skewness and kurtosis in the sampling results underline this issue). For this case study, it was shown that the uncertainty ascribed to the model predictions was much greater for the linear perturbation method compared to a converged sampling scheme. Based on these results, Latin hypercube sampling was used in the following research work, as it was judged to represent the best compromise between computational load and the ability to estimate a set of uncertain model predictions. Implementing a sampling scheme to propagate uncertainty was straightforward to implement, and this made it an attractive option going forward.

The plate was then investigated in the presence of damage. Damage was simulated in the model by reducing the stiffness value of the spring connecting masses one and two in Substructure A; the damping at this position was also reduced by the same proportion (the damage location is highlighted in Figure 4.2. The stiffness was reduced proportionally in increments of 20% from the undamaged condition to full damage. This would be comparable to an edge crack in the plate of increasing severity; a damage extent of 100% represents a total loss of stiffness in the spring, which would be equivalent to a full break in a structural member.

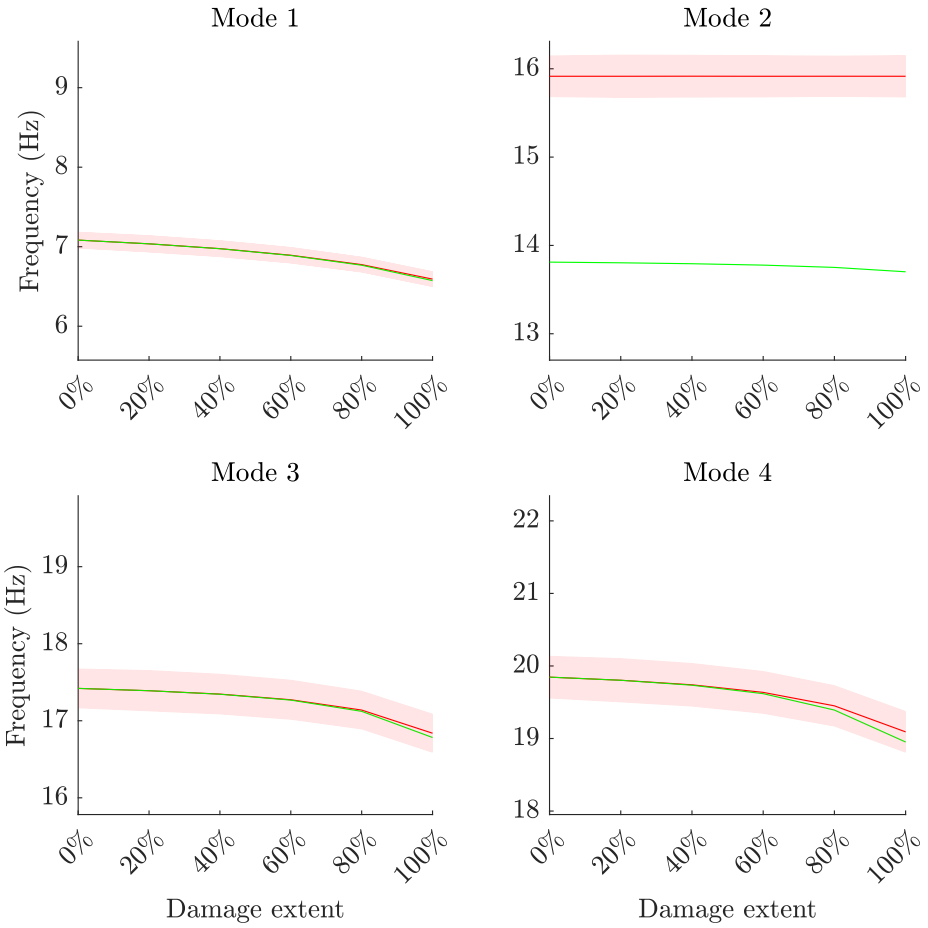


Figure 4.5: The effect of damage extent on the first six modal frequencies; true solution in green, dynamic substructuring solution in red (with $\pm 3\sigma$ indicated for the dynamic substructuring solution)

The results of the dynamic substructuring approach are compared with the true solution (obtained from the full assembly model) across the damage range in Figure 4.5, which shows how the first four modal frequencies develop with increasing damage in the structure. The results given in Figure 4.5 show good agreement between the substructured solution and the true assembly, with the exception of mode 2. This was caused by the discrepancy between the two models – the two substructures were rigidly fixed together while the true assembly contained a set of springs and dampers at this location (see Figure 4.2).

The successful propagation of uncertainty across damage states and the observable sensitivity to damage suggest that the dynamic substructuring method will be highly

applicable to the FMD-SHM problem, as statistical damage recognition models could clearly be trained using these uncertain predictions. The problem of model discrepancy in dynamic substucturing, as can be seen for mode 2 in Figure 4.5, is a concern. It could be argued that, given the lack of damage sensitivity in this feature, this mode would not be utilised in a model-driven SHM strategy; however, in the case where significant discrepancy existed between damage sensitivity exists some mitigation of the discrepancy is necessary. This could be addressed by carrying out a basic assembly-level validation exercise to identify significant discrepancies using the undamaged model; alternatively, selecting features such as the proportional change in natural frequencies to indicate damage (rather than their absolute values) could reduce the impact of these discrepancies.

4.3 Conclusions

This chapter presented the basic theory required to carry out hierarchical V&V for SHM. This can be divided into two sections: the first being the assembly methods required to couple validated submodels and the second being UP methods that can be used in conjunction with the coupling process.

Dynamic substructuring was identified as providing the means for the assembly of validated submodels; a series of dynamic substructuring techniques were presented and primal assembly was successfully demonstrated in a case study. Linear perturbation and sampling-based methods were investigated for UP techniques to be used alongside dynamic substructuring. Latin hypercube sampling was identified as the most appropriate for this work. It should be noted that while sampling methods were not expected to be computationally prohibitive for the models used in this thesis, this will not always be the case. Where sampling is not a feasible approach, more advanced techniques based on perturbations or model reduction should be investigated.

The case study presented showed that these methods should be adaptable to a realistic SHM problem by investigating the propagation of uncertainty through a plate assembly in the presence of damage in the plate. Probabilistic substructure models such as these could be applied as part of an FMD-SHM framework by using their predictions to generate a training dataset for machine-learning based damage identification algorithms; this would be dependent on the selection of appropriate damage-sensitive features to predict using the models. The following chapters will investigate key aspects of the hierarchical V&V framework and demonstrate the methods outlined here

on a more realistic test case based on the initial success of the case study in this chapter.

The integration of bespoke joint mechanics using dynamic substructuring was not investigated in this chapter. The potential inaccuracies caused by this are illustrated in the predictions for mode 2 of the plate assembly in Figure 4.5. Various methods of modelling joints between substructures are available in dynamic substructuring, which would be expected to reduce model discrepancy in creating assembly-level models from a set of submodels. The impact of this would be a key area of future research following this thesis.

Chapter 5

Feature selection and sensitivity analysis

This chapter investigates the feature selection stage of the hierarchical V&V framework (see Figure 3.2). The hypotheses are tested on an experimental dataset drawn from a laboratory-scale truss bridge and provide the basis for the following research work carried out for this thesis. Feature selection is critical to the successful application of a hierarchical V&V strategy for SHM since the features are used not only for comparisons between model predictions and experimental data for validation purposes, but also for the application of models to SHM tasks. Analyses of the implications of these requirements are presented in this chapter as they have not been directly addressed in the literature.

This chapter aims to present a conceptual methodology for the application of feature selection in a hierarchical modelling strategy for FMD-SHM. This will be informed by sensitivity analysis literature in order to determine which features show sensitivity to damage, as is required for SHM strategies, but will be bespoke to the hierarchical model form.

The feature sets identified in these tests will then be applied to model validation and SHM tasks in Chapter 6 and Chapter 7 respectively; the suitability of the selected features will be evaluated in these chapters when they are applied in practice. This will demonstrate the potential of feature selection at this stage of the framework to increase the efficiency of the validation process. It may also be shown that models can be used for feature selection prior to validation as part of the FMD-SHM process.

5.1 Background and definitions

Damage-sensitive features underpin the application of any online SHM strategy. They can be defined as a given subset of a recorded dataset that shows non-negligible variation when damage is present in a structure. Useful features generally have to be acquired from raw recorded data through some extraction process, which will usually reduce the dimensionality of the dataset in order to focus on aspects of it that are of interest.

The discussion of features constitutes a large area of interest in the literature surrounding SHM [2]. Feature *extraction*, as a discipline, refers to the efforts to determine a series of subsets (features) from a given set of data. In effect, this motivates a series of strategies for data or signal processing based on the requirements from the data. Dynamic data, such as those collected for SHM purposes, are usually recorded by accelerometers measuring the response of a target structure in the time domain. The time domain response is highly complex and difficult to analyse, so generally a discrete set of lower-dimension features are extracted. The overall aim of feature extraction in SHM is to acquire the features of the response data that most clearly, robustly and parsimoniously indicate the presence of damage in a target structure.

Feature *selection* refers to the efforts to discriminate between the features that have been extracted from a set of data, in order to identify those that are of most use to the modeller. This is effectively carried out through some form of sensitivity analysis; note that SA represents a broad set of methods and ideas – these will be discussed in more detail later in this chapter. Modellers are usually required to make predictions in a parsimonious manner, and, as such, they are incentivised to make the most efficient use of the predictive features available. It should also be considered that for a full set of features, more complex features may make certain other features obsolete, as they contain all the information that is present in less complex features.

A large range of features have been previously utilised for SHM, and the various options are of differing usefulness depending on the situation. Features that are relevant to SHM include (but are not limited to) response in the time domain; responses in the frequency domain such as the FRF (complete spectra or sections of particular interest); modal properties such as natural frequencies, mode shapes and antiresonances; and features derived from other modal properties, such as mode shape curvatures and modal strain energies. Global features, such as natural frequencies, are often suited to tasks such as damage detection, while features with more local dependencies, such

as mode shapes, are generally more applicable to tasks like damage localisation and classification. Different features have different demands in terms of data acquisition; natural frequencies and FRFs can be recorded with very few sensors, while detailed mode shapes could potentially require a lot of sensors. Feature dimensionality is also a consideration; features of high dimensionality can be more informative, but this can also make them more difficult to utilise in computational models.

An example of a feature identification strategy in the field of wind energy can be found in [90]. An attempt to benchmark a wide range of features for bridge damage classification has been made in [91]. A review of feature selection metrics is given in [92], which also introduces a new relevance metric. A further feature selection study, in this case applied to acoustic emission features for condition monitoring, can be found in [93].

5.1.1 Sensitivity analysis for feature selection

SA refers to a set of methods which aim to determine the links between variance in a set of outputs to the variance in the inputs of a system [94]. It is therefore critical in determining a set of features to extract and select from experimental data, as well as analysing the predictive performance of physics-based models. In both of these cases, the aim of SA in an SHM context is to determine, with confidence, a set of features which have good sensitivity to damage but minimal sensitivity to other confounding variables, such as environmental or operational variables (EOVs). SA can also be used to determine suitable features on which to validate models, and to identify significant inputs to carry out any UQ processes.

At the highest level, SA methods can be divided into qualitative and quantitative methods [95]. Qualitative SA underpins the logic of many decision-making processes, but, formally speaking, it refers to the application of engineering judgement when presented with non-empirical evidence of output sensitivity. A simple example of this would be to acquire scatter plots of model outputs over a range of input values and to observe which plots show the greatest variance based on a particular input of interest.

Quantitative SA can be further divided into local and global methods [95]. Local methods tend to rely on a perturbation analysis, in which a model is evaluated at a nominal value and a perturbation of that value. A linear extrapolation is then fitted to the results of these evaluations to estimate the output sensitivity in terms of partial derivatives [94]. The method is useful in that it can be used to empirically

estimate parameter sensitivity using a small number of datapoints only. However, it is highly limited when model responses are not expected to be linearly correlated with their inputs. Perturbation analyses can still be applied over a local range of the problem domain, provided that modellers are aware of the potential inaccuracies of the method and select an appropriate perturbation magnitude.

Global quantitative SA represents the most rigorous method for SA, with the aim of giving empirical estimates for the sensitivity of model outputs over the full problem domain [94]. These can be estimated by using sampling methods on the parameter distributions and evaluating the model over the sample space [94]. The drawback of these methods is that their execution can become very computationally expensive, as multiple model evaluations are required to achieve a converged estimation of the output sensitivity. Methods to reduce the computational effort involved in global SA include the use of emulators [96, 97, 98], which reduce a complex model to a simple function over a given domain, and Bayesian methods, which allow for faster convergence by making use of prior knowledge and previous evaluations as the analyses progress. A full mathematical outline of the Bayesian approach to global SA is given in [96].

Probabilistic SA refers to attempts to determine the cause of variance in the case where uncertainty is present as to the true set of model inputs [96]. In this case, the input values are drawn from a probability distribution that describes the uncertainty associated with the true values. In probabilistic SA, the model output variances can be broken down into a set of main effects (due to single input) and interactions (due to the joint effects of multiple inputs) [96]. Plots of the main effects and joint effects can then be used to inform a qualitative SA process in a much reduced domain space.

$$S_i = \frac{\mathrm{var}(E(Y|X_i))}{\mathrm{var}(Y)} \tag{5.1}$$

$$S_{Ti} = \frac{\mathrm{var}(Y) - \mathrm{var}(E(Y|X_{-i)})}{\mathrm{var}(Y)} = 1 - S_{-i}. \tag{5.2}$$

The variance-based approach to global SA can be applied using two metrics, known as the *main effect index* (S_i) – also known as the Sobol' index [65] – and the *total effect index* (S_{Ti}, which refer to input i of d) [96]. The main effect index is calculated by dividing the posterior variance (given a particular input distribution) in the model output by the overall variance in the model output (Equation 5.1 [96]). The total

effect index applies the same operation to the total variance in the model output given a particular input distribution (Equation 5.2 [96]). The total variance can be calculated by subtracting the variance in model output given all input distributions apart from the input of interest from the overall output variance. Reasoning on the variance-based metrics can also make use of the joint effects of two or more inputs on the output [96]. Through variance decomposition, it can be shown that the following conditions will hold, in the case of independence between the input variables (for d number of inputs) [96].

$$S_i \leqslant S_{Ti} \tag{5.3}$$

$$\Sigma_{i=1}^{d} S_i \leqslant 1 \leqslant \Sigma_{i=1}^{d} S_{Ti} \tag{5.4}$$

An exemplar SA based on the Sobol' indices applied the method to the transport of pesticides in soil [99]. Examples of Bayesian SA include applications to a model of the aortic valve in biomechanics [100] and a non-linear airship model in FEM [101]. A good example of global SA applied in an SHM context can be found in [98], which applied the process to tunnel models.

5.2 Handling of features in forward model-driven structural health monitoring

FMD-SHM represents a hybrid of two separate techniques for carrying out SHM [4]. The technique combines elements of law-based SHM with those of data-driven techniques, and in doing so aims to circumvent some of the difficulties associated with each.

In FMD-SHM, a physics-based model is created whose predictions are used in a *forward* manner to generate training data for a statistical machine learning model [4]. Based on the requirements of the SHM strategy, the capability of the statistical model could feasibly range from damage detection as a novelty detection exercise, through damage location, to estimation of damage extent [1]. These capabilities could then be extended to estimation of remaining life prognosis with the incorporation of an additional predictive damage progression model [10].

In the past, SHM techniques have focused on using either a data-driven or a physics-based approach in relative isolation (notwithstanding research into the use of physical laws to constrain the training processes of data-driven methods [16, 102, 103]) [4]. Each of these methods has been affected by issues relating to the lack of data; data-driven methods require large training sets, while physics-based methods are dependent on data for model validation purposes. In addition, while the requirement for model validation data can represent a one-off investment, traditional physics-based strategies have used inverse techniques, in which damage is determined through a parameter updating process that matches the model predictions with live data. This process presents a serious computational challenge as well as raising issues of solution uniqueness.

FMD-SHM obviously maintains a reliance on data to validate the physics-based model. However, given that the model updating problem is avoided in the implementation of the strategy, many of the computational issues associated with purely physics-based SHM are removed. The method also allows for careful management of the lack-of-data problem through hierarchical model development. Validating the model at a subassembly or component level allows for significant simplification of the data acquisition process, and potential reductions in the quantity of data required where components of subassemblies are modularised within the full assembly.

Due to the two-stage nature of FMD-SHM (first developing and running the physics-based model, second training and applying the statistical model), the method demands a certain duality when considering the use of features in the process. The first consideration concerns the selection of features on which to validate the predictions of the physics-based model. The second consideration – which should be viewed with equal, if not greater, importance than the first – concerns the selection of features based on which the statistical model can make suitable damage inferences. Given that the statistical, or data-driven, model would utilise a set of measurable inputs from the real-life structure (such as mode shapes or natural frequencies) to infer the presence of damage, the physics-based model must be used to generate these inputs based on a pre-determined domain of damage scenarios and EOVs. Therefore, the input and output feature spaces of the data-driven and law-driven models will be opposites of each other.

The validation of the physics-based model used in an FMD-SHM strategy must ensure and assess its ability to accurately predict a range of damage-sensitive features based on a set of input variables and model parameters. The input variables would cover

the full range of health states of the structure to be detected or classified by the statistical model, as well as other EOVs related to the structure where required. Clearly, as the complexity of the considered damage scenarios increases and as more EOVs are included, the dimensionality of the validation space of the model increases. This quickly increases the difficulties associated with the validation task; therefore, utilising features that cover the full variance caused by damage and EOVs with minimal dimensions are ideal.

The selection of features on which to validate the physics-based model must be carried out based on a trade-off between available computational resource and the difficulties associated with the data acquisition process, and the desire for maximum data fidelity. For example, a particular mode shape may be strongly indicative of a particular health state of the structure, but may require a large number of sensors in unfeasible locations on the structure. Data acquisition difficulties are well recognised and do not require an exhaustive discussion here; the same can be said for computational limitations (see introductory chapters for discussion of the challenges facing SHM and the motivations for this research). On the other hand, increased data fidelity at this stage would increase the ability of the model to train a sensitive damage detection algorithm or develop an advanced classification model that can perform many levels of Rytter's Hierarchy [1].

To maximise the effectiveness of features to be used to train the statistical damage identifier, a set of three criteria can be applied:

1. Feature selection should maximise the sensitivity to damage across the damage states of interest. Ideally this sensitivity should highlight the presence of damage as early as possible.

2. Feature selection should minimise the influence of EOVs.

3. Feature selection should minimise the dimensionality of the features or feature space.

This stage of the selection process is much more closely aligned with feature selection processes commonly found in literature [104, 105]. A review of feature selection methods for machine learning processes can be found in [106]. Since the outputs of the statistical model are damage labels of some kind, the aim of feature selection is to identify the input features which maximise variance in the outputs. This is the inverse of feature selection for model validation, in which case the inputs vector is set and the aim is to identify the most sensitive output features.

5.3 Relevance to hierarchical models and validation

A key method discussed in this thesis which could improve the feasibility of model-driven SHM, in either a forward or inverse sense, is hierarchical V&V. This strategy involves the division of an assembly-level model into a series of submodels representing subassembly or component structures which are substructures of the full assembly. The main benefit of this sub-division is a potential alleviation of the data acquisition problems commonly associated with the model validation process (see Chapter 3 for a full discussion of this).

While there are considerable advantages to constructing models for SHM in a hierarchical manner, the process introduces a set of further considerations in terms of feature selection. The most pertinent of these is the issue of feature locality. Many features that are global to a submodel would behave in a local manner in the assembly, which would raise issues in the ability of the SHM system to detect damage in the global structure. Therefore, in the design of a hierarchical modelling strategy, features should be selected that have sufficient local sensitivity to be used to validate the submodels, but would also have sensitivity to global damage when the submodels are executed at the assembly level.

A strategy for the handling of features in a hierarchically validated FMD-SHM context would be to carry out a feature selection process using the full model hierarchy prior to the substructuring and validation processes. The benefits of this are twofold: firstly, the prior physical knowledge written into the nominal un-validated models can be used to elicit sets of features which are appropriate for use at the assembly level and the substructure levels (i.e. features which are sensitive to damage and fulfil any other particular criteria depending on context); secondly, efforts can be made to ensure that the feature sets employed at the assembly level for training the statistical model and at the substructure level for validating the physics-based submodels are relevant to each other (this would lend credibility to the substructure-level validation process). This would allow for a set of features that are damage-sensitive in the global assembly-level model to be determined, which would inform and streamline the validation process at the substructure level. If the model is to be incorporated into an FMD-SHM strategy, a further SA could be carried out following the hierarchical validation process to finalise the features that would be used to train the statistical

model. This feature selection strategy is investigated in this chapter through a case study summarised in the following sections.

5.4 Case study: Feature selection for hierarchical validation of a truss bridge

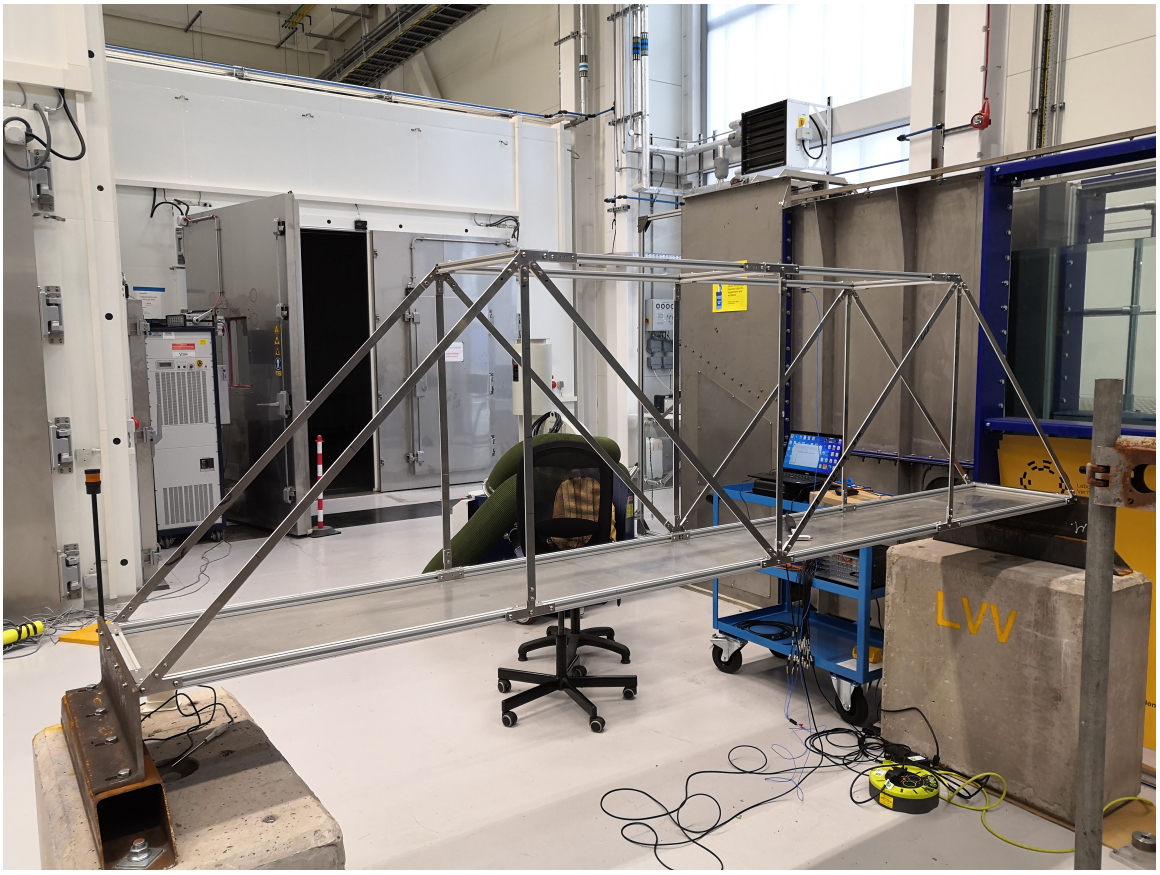


Figure 5.1: The laboratory-scale truss bridge used for this case study

This case study aims to demonstrate how feature selection would be carried out prior to validation for a model assembly in an FMD-SHM strategy. The structure of interest is a laboratory-scale truss bridge, shown in Figure 5.1, which was constructed at the Laboratory for Verification and Validation in Sheffield. The aims of the study are to indicate how an informed feature selection process can make use of a hierarchical model in its nominal form to optimise the validation process by utilising the engineering knowledge that was used to build the model, leading to features sets that can be used for submodel validation and for assembly-level SHM tasks.

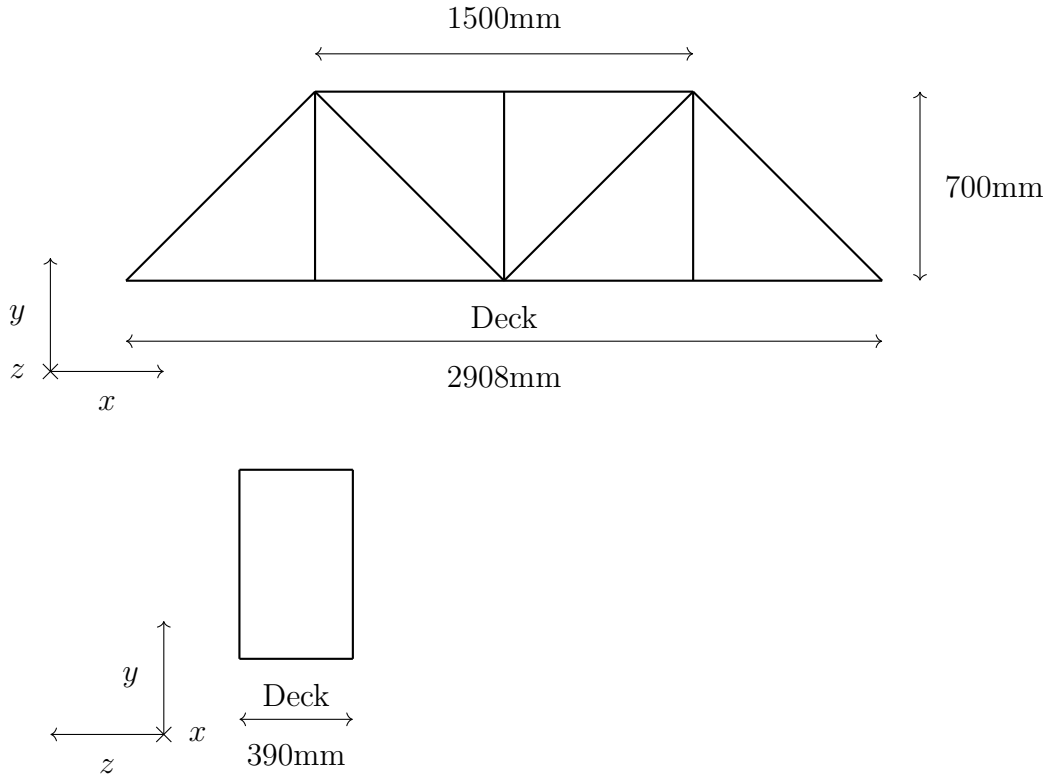


Figure 5.2: A schematic of the bridge structure with key dimensions marked

The bridge is a truss bridge of standard design, and it can be compared to many real-life structures around the world. SHM studies have been carried out on similar structures in the past, most notably on the KW51 railway bridge in Leuven [13, 107, 108]. The main load-bearing components of bridges of this type are the trusses, or struts, which are therefore of key interest in SHM strategies. The advantage of this laboratory structure is that the individual components are of low value; therefore, a range of tests can be carried out on the bridge in its damaged states, where required for this research.

The bridge is 700mm in height, 390mm in width and 2908mm in length; see Figure 5.2. The struts and deck are cut from aluminium plate of 4mm and 3mm depth respectively. The longer, diagonal, struts are 1006.52mm in length and the shorter, vertical, struts are 700mm in length; both are 20mm in height. The deck border and upper frame are constructed from alumimium Rexroth, which is of 20mm in depth and height; the upper frame is of 1500mm in length. The deck is of 390mm width and 2908mm length, which are the same as the bridge as a whole.

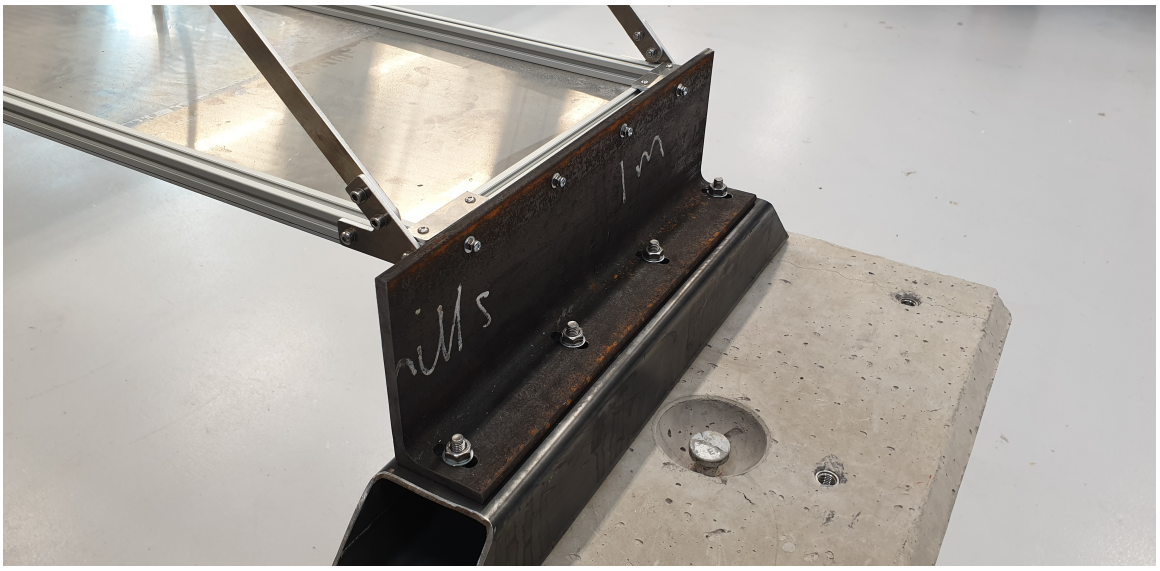


Figure 5.3: The boundary conditions of the bridge installed for testing

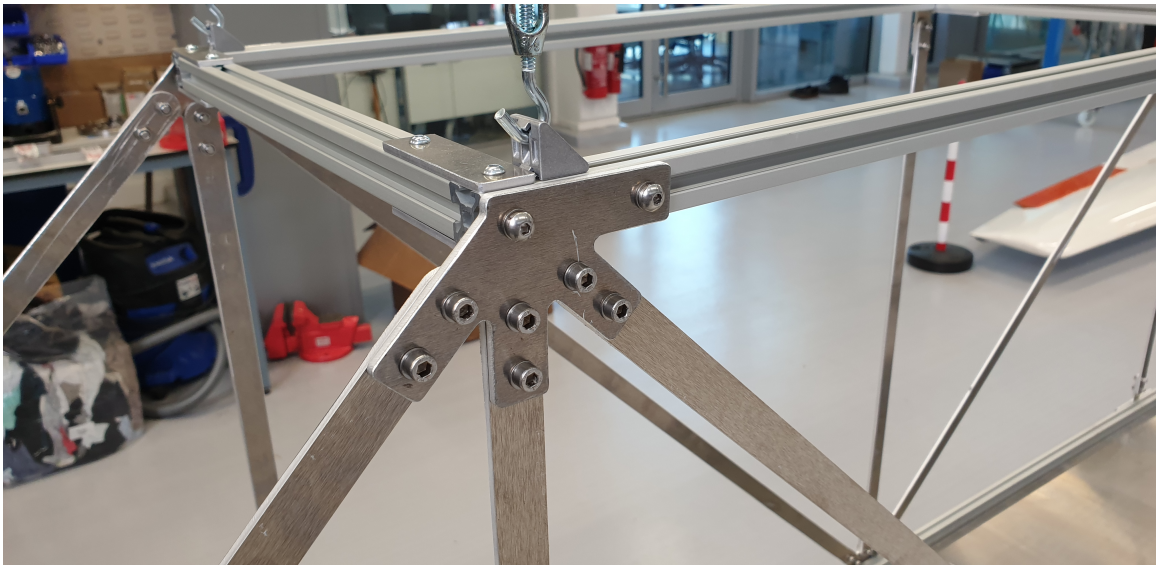


Figure 5.4: Example of the joint components of the bridge connecting the struts to the deck and upper frame

5.4.1 Experimental data

The experimental datasets used in this case study are drawn from a suite of tests that followed up a dataset presented in 2021 [109], following similar methodology to acquire healthy- and damaged-state data from the bridge structure shown in Figure 5.1. The bridge was fixed at each end, with all degrees of freedom constrained, to cast-iron mounts, which were in turn attached to heavy concrete blocks (Figure 5.3). For the purposes of these tests and the following investigations, these will be considered rigid boundary conditions.

The previous tests were carried out in two phases. The first phase entailed a roving hammer tap test in order to identify the mode shapes of the structure under excitation in its nominal, undamaged state. This dataset was then used for matching the experimental modes to those predicted by the numerical model. The second phase entailed shaker-excited damage-state testing to identify the effect that increasing damage had on the natural frequencies of the structure. Damage was introduced by saw cut to the mid-point of each of the three vertical struts on one side of the bridge at 2.5mm intervals up to maximum depth of 17.5mm, at which point the strut was replaced before commencing the test on the next strut.

Similar tests were carried out for this case study, with the addition of a test to identify the impact of boundary condition uncertainty on the results. The first main objective of the new test set was to aid the accuracy of mode-matching by the modal assurance criterion (MAC) to the model predictions; this was done by carrying out a more high-fidelity roving hammer test with the bridge in its nominal condition. The previous set of tests used 54 tap locations: 27 on the deck, one at the midpoint of each strut and 13 on the upper frame [109]. This was insufficient for providing mode shape data on which to discriminate between modes with similar mode shapes, because the midpoint of the struts was a node for many mode shapes. In addition, the struts showed the most deviation in mode shape between modes and were the key components of interest, so additional tap locations along the struts were desirable.

The present tests used 76 tap locations, with three locations excited along each strut at each quarter-length; the tap locations at the extreme ends of the deck were removed for these tests as the bridge was fixed at these locations. The formula for the MAC, which gives an indication of the similarity between mode shapes φ_1 and φ_2, ranging from 0 to 1 (where 0 indicates no similarity and 1 indicates two identical mode shapes), is given in Equation 5.5. This definition of the MAC is used throughout this thesis.

$$\mathrm{MAC} = \frac{\left|\{\varphi_1\}^T \{\varphi_2\}\right|^2}{\left(\{\varphi_1\}^T \{\varphi_1\}\right)\left(\{\varphi_2\}^T \{\varphi_2\}\right)} \tag{5.5}$$

The secondary objective of the new tests was to acquire a more globally descriptive set of damage-state data for the bridge; the damage tests were carried out on the struts that were symmetrically 'unique' in the structure, i.e. the two diagonal and vertical struts located in one 'corner' of the bridge – these were the four struts in the far corner in Figure 5.1, closest to the laptop. These struts were labelled 1, 5, 2, 6 (see Figure 5.5). The previous tests used in [109] investigated the progression of damage in the three vertical struts on one side of the bridge using shaker excitation. The present dataset should be more globally informative as, due to the lines of symmetry of the structure in the xy- and yz-planes (see Figure 5.5), damage in any individual strut would result in a similar effect on the response as one of the four conditions considered here.

The final aim was to investigate the uncertainty implications of the joints on the structure response. This was carried out by applying a known nominal torque to the bolts fastening each strut to the assembly using a torque wrench, and carrying out repeat tests in between removing and reattaching a particular strut.

5.4.1.1 Methodology: Healthy-state testing

Tap testing was used to identify the mode shapes for the lower modes of the bridge structure using the roving hammer method (a PCB Piezotronics 086C03 impact hammer was used in these tests). This allowed for a large number of test points to be used without adding many accelerometers, which could impact the dynamic response of the structure. Prior to testing, the ambient temperature was measured at 17°C using a local digital thermometer (visible in Figure 5.1 underneath the deck on the near side of the bridge) in the vicinity of the bridge – this remained constant throughout the test period. The bolts that fixed the struts to the bridge deck and Rexroth were tightened using a torque wrench to 8Nm.

The dynamic response of the bridge was recorded in the range of 0–128Hz. Reducing the scope of the tests to this frequency range had two benefits. Firstly, it allowed the use of an impact hammer with a very soft tip. This meant that double-impacts could be avoided, which were found to be an issue when exciting the struts due to their extreme flexibility. In addition to this, a frequency resolution of 0.0625Hz could be

achieved for the FRFs by limiting the recorded frequency spectrum to 128Hz and the sampling frequency to 256Hz. This in turn allowed for greater precision in identifying the natural frequencies within the spectra. This was desirable, as many of the bridge's natural frequencies were very close to each other in the frequency domain.

A triaxial accelerometer was used to record the response of the bridge to the impacts. This was attached to the Rexroth on the upper frame of the bridge, as the mobility of the structure was significant in that location, and many of the lower natural frequencies had mode shapes which involved displacement of this part of the structure. The accelerometer was fixed to the upper Rexroth on the near side of the bridge, between the first two joints in the x-direction; it is visible in Figure 5.1. The accelerometer used was a PCB Piezotronics 356B21.

Five repeat impacts were carried out at each damage location in order to reduce the noise in the results through averaging. Additional pre-processing measures to increase the cleanliness of the data were carried out by windowing (an exponential window was used) the recorded excitation and response data.

Data acquisition was performed using the Siemens LMS system, with modal analysis carried out using the PolyMAX algorithm which uses a curve-fitting method to isolate modal characteristics from the recorded FRFs. The final chosen modes were extracted concurrently with the modes used in the damage-state testing to ensure compatibility between the two datasets.

5.4.1.2 Methodology: Damage-state testing

Tap tests were carried out on the bridge across a range of damage conditions using the same impact hammer as was used in the roving hammer testing. A single tap location was used with multiple accelerometers attached in the y- and z-directions at each joint (see Figure 5.5). The accelerometers used were PCB Piezotronics 353B18s. As with the roving hammer tests, repeats and windowing were used to reduce the noise level in the recorded data.

Prior to testing, the ambient temperature was measured at 17.3°C – this had increased to 17.4°C by the end of the damage tests and remained constant at this temperature during the final tests. The bolts that fixed the struts to the bridge deck and Rexroth were tightened using a torque wrench to 8Nm. The same measured spectrum was used as for the roving hammer tests and the same feature extraction method was used to acquire the modal data from the recorded FRFs.

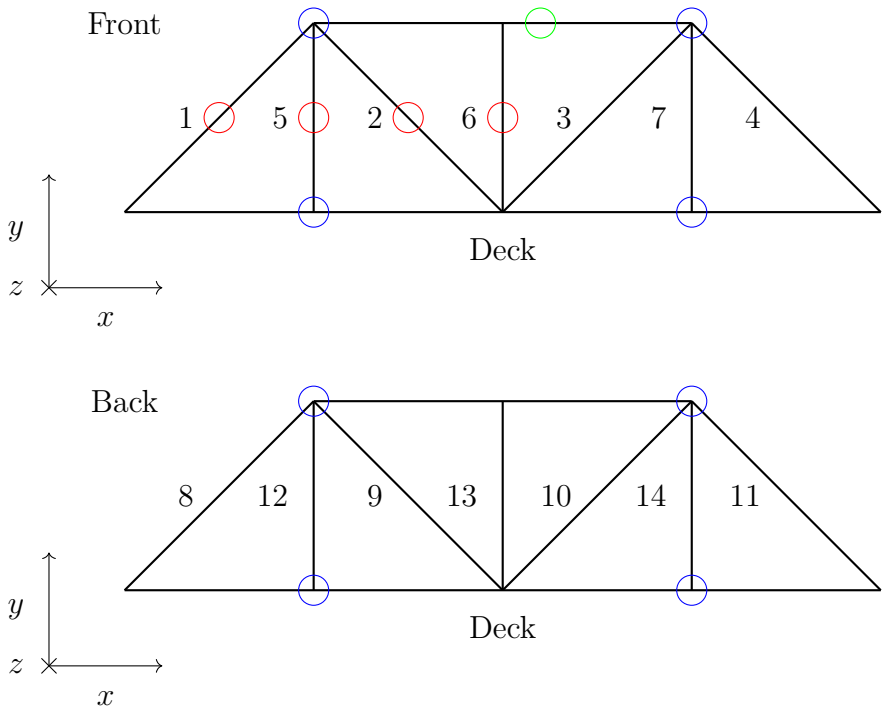


Figure 5.5: A schematic of the bridge structure, with damage locations marked in red, tap location in green and accelerometer locations marked in blue; strut numbers also marked

Damage was introduced to Struts 1, 2, 5 and 6 by saw cut at the midpoint at 2.5mm intervals, up to a maximum 'crack' depth of 17.5mm. When each damage run was completed the strut of interest was replaced with a new strut. The damage locations and tap location are illustrated in Figure 5.5.

In order to assess the uncertainty caused by the boundary conditions and provide a number of test points describing the bridge in its normal condition, Strut 1 was removed and reattached three times, with tap tests carried out following each reattachment. These final tests were carried out immediately after the damage-state testing. The methodology was otherwise the same as for the damage state testing, described above.

5.4.1.3 Results

The FRFs were stored for each of the above-described test sets, from which modal data were extracted using the PolyMAX curvefitting algorithm. This method utilises a numerical approach to isolate the natural frequencies, damping ratios and mode shapes from manually selected resonance peaks on the FRFs. The key parameters

guiding this process are the tolerances set for the natural frequencies, damping ratios and mode shape vectors for each peak, the maximum number of DoFs allocated to the curve fit and engineering judgement. These tolerances describe the stability of a modal fit for a given resonance peak, where the solution would be considered stable if it was within each tolerance criterion when compared to the solution with one fewer DoF. The tolerances set for this analysis were 0.1% for the frequency, 5% for the damping ratio and 0.5% for the mode shape, meaning that the exported solutions are accurate to these boundaries. These values were set to tightly control the extracted natural frequencies, since these were of interest as features in the following analyses.

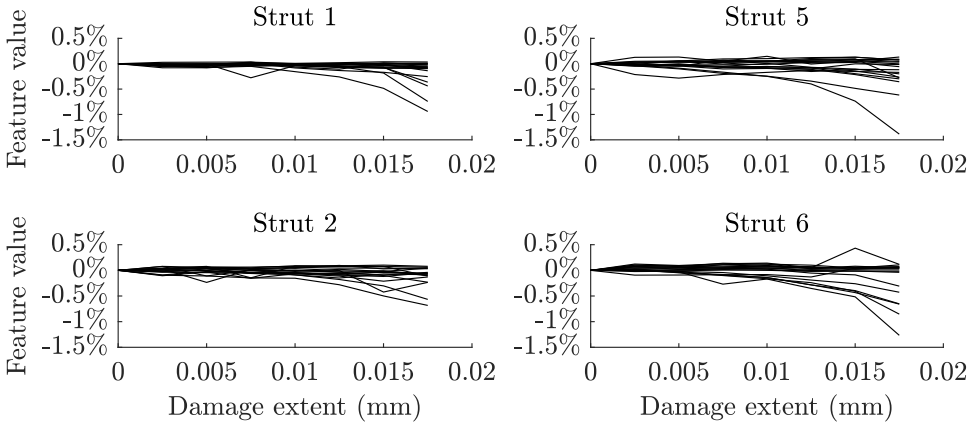


Figure 5.6: The experimentally extracted modal features for damage in each strut of the bridge across the full range of damage in each strut

Following feature extraction, the datasets were matched to each other using the natural frequency values, which resulted in an experimental dataset of 18 natural frequencies for each test. These are plotted as features across the full range of damage for each strut in Figure 5.6; this figure shows that, of the extracted features, some are sensitive to damage but others are relatively insensitive – this indicates the requirement for further feature selection activities. The plotted feature is proportional deviation from the undamaged natural frequency for each mode to allow comparison between different natural frequencies across the full range of damage.

5.4.2 Model specifications and verification activities

The bridge used for this case study is representative of common designs implemented in the real world [108] and is composed of three main sets of substructures: the upper

frame, the deck and the struts (where each strut is a component). The struts of a truss bridge are the main load-bearing components of the structure, and were therefore the key components of interest for damage detection. The struts were cut from 5083-'O'/H111 grade aluminium plate. The deck was composed of two components: a plate section (also 5083-'O'/H111 grade aluminium) and an aluminium Rexroth border. The upper frame was also constructed of aluminium Rexroth sections. Joint components (a series of brackets and connecting bolts – see Figure 5.4) were also part of the assembly, however they were neglected from the model in this research. Submodels were constructed using ANSYS to represent each of these substructures resulting in four submodels: the upper frame, the deck, the diagonal strut and the vertical strut. The assembly could be constructed from these submodels using primal physical dynamic substructuring.

The strut submodels were constructed using ANSYS BEAM188 elements. BEAM188 is a two-node element in the ANSYS library designed for analysis of beam structures. Each of the nodes has six degrees of freedom as standard (plus an optional warping degree of freedom) and the element can be evaluated by linear, quadratic or cubic laws based on Timoshenko beam theory [110]. BEAM188 is a 1D line element with cross-section data specified separately to make it 3D.

The upper frame was also constructed using BEAM188 elements. The geometrical cross-section parameters of these elements are very flexible, which made modelling of complex cross-section geometries such as the Rexroth (which the upper frame was constructed out of) straightforward.

The deck comprised two components: a flat plate and a Rexroth boundary. The Rexroth was modelled as above and conjoined in ANSYS to the plate, which was modelled using SHELL181 elements. SHELL181 is designed for analysis of thin shell structures. Each element has four nodes, each of which has six degrees of freedom as standard. SHELL181 is a 2D area element, where the thickness is defined separately (it is suitable for laminate and homogeneous shells). The number of integration points within each element is optional, the default being three.

In order to evaluate the natural frequencies and modes shapes of these submodels, an eigenvalue solution was obtained using the mass and stiffness matrices of the structures. Due to the nature of the deck as a thin plate, the aspect ratio of the elements used in this substructure were very high, which led to poor conditioning of the mass matrix of the deck and assembly. This then meant that the inversion of

the mass matrix could cause potential inaccuracies, a process which was required in computing the eigenvalue solution. The Cholesky decomposition was used in order to carry out the Krylov–Schur algorithm for computing a few eigenvalues as part of the MATLAB 'eigs' function. In this method, a matrix is decomposed into the product of a lower triangular matrix with its own transpose, as shown in Equation 5.6. This decomposition was assessed by comparing the product of the right-hand side to the original matrix, which was found to have very low error.

$$A = LL^T \tag{5.6}$$

To verify the accuracy of the beam elements used for the strut and Rexroth sections of the model, a nominal case was set up to compare the solutions of a modal analysis to a set of equivalent analytical solutions. For simple beam structures, such as the bridge struts, the mode shapes and associated natural frequencies can be derived analytically from first principles. The natural frequency (in Hertz) for a beam is defined as follows [111]:

$$f_i = \frac{\lambda_i^2}{2\pi L^2}\sqrt{\frac{EI}{m}}; i = 1, 2, 3, \ldots \tag{5.7}$$

where L is the beam length, E the Young's modulus, I the second moment of area and m the mass per unit length. The parameter λ_i is a dimensionless factor that is a function of the mode number (i) and boundary conditions of the beam. The assumptions made for this solution are as follows:

- The beam is of uniform cross-section with dimensions much less than the length of beam.
- The material is linear, homogeneous and isotropic.
- The beams can only deflect normal to the undeformed axes.
- No axial loads are applied.
- The rotation and translational motions of the beam are not coupled.

The natural frequencies for the rotational modes were calculated separately, under the same assumptions as the translational modes. The formula for these is given below [111]. For this formula, C represents the torsional constant of the cross-section, G the shear modulus, μ the density and I_P the polar area moment of inertia.

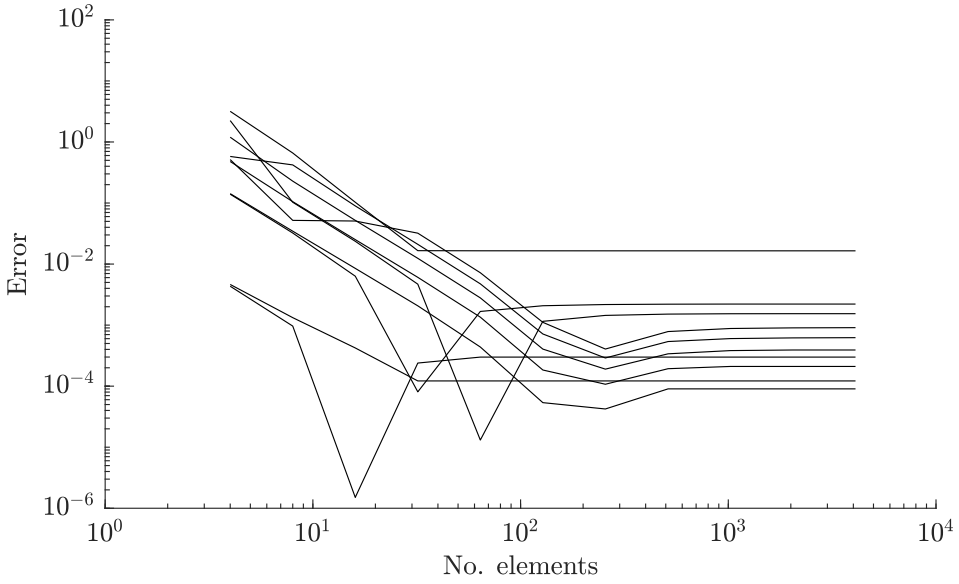


Figure 5.7: The error between the analytical solution and the model predictions of the first ten modes for the BEAM188 elements

$$f_i = \frac{\lambda_i}{2\pi L}\sqrt{\frac{CG}{\mu I_P}}; i = 1, 2, 3, \ldots \tag{5.8}$$

The case modelled the longer diagonal strut as a cantilever beam. The Young's modulus was set to 72GPa, the Poisson's ratio to 0.33 and the density to 2650kgm^{-3}. Figure 5.7 shows the error between the numerical solution and the analytical solution for the first ten modes of the cantilever beam. It can be seen that the error remains low overall and converges with element size to the order of 10^{-3}, with the exception of mode 10. The discrepancy in mode 10 is likely due to a mismatching between the modes due to the difficulty in separating the torsional modes from the modes that cause displacement in the y-dimension (this is due to the moment of inertia being very similar in these two directions of motion).

As with the beam elements, a nominal case was set up to compare the shell element formulations to a set of analytical solutions. The shell elements were used to construct the plate in the deck substructure, which was bordered in the model by a Rexroth beam structure. As for beams, the rectangular plate is a simple member and its natural frequencies can be derived analytically. The formula for these is as follows [111]:

$$f_i = \frac{\lambda_i^2}{2\pi a^2}\sqrt{\frac{Eh^3}{12\gamma\left(1-\nu^2\right)}}; i = 1, 2, 3, \ldots \tag{5.9}$$

where a is the plate length, E the Young's modulus, h the plate thickness, γ the mass per unit area and ν the Poisson's ratio. The parameter λ_i is a dimensionless factor that is a function of the mode number, the boundary conditions and the plate geometry (the relative width compared to length). The assumptions made for this solution are as follows:

- The plate is flat and of constant thickness (which is much less than the length or width of the plate).
- The material is linear, homogeneous and isotropic.
- The deflections are small and flexural with no rotary or shear contributions.
- There are no in-plane loads on the plate.

The case modelled a plate constrained at each end of length 2500mm, with width 100mm and thickness 3mm. The Young's modulus was set to 71GPa, the Poisson's ratio to 0.33 and the density to 2700kgm^{-3}. Figure 5.8 shows the error between the numerical solution and the analytical solution for the first six modes of the plate. It can be seen that the error remains low overall and converges with element size to the order of 10^{-2}.

A grid convergence analysis was carried out on the submodels using the grid convergence index (GCI) [9]. The GCI uses the Richardson Extrapolation to provide an indication of the level of numerical convergence of an FEM compared to the estimated value of the exact solution.

The Richardson Extrapolation is a popular method for determining a theoretical 'exact' value for a numerical problem based on the observed convergence of a set of results. It was developed in 1910 under the name 'h^2 extrapolation' and is commonly used in assessment of the level of convergence of numerical solutions. To derive the Richardson Extrapolation, the exact solution to a numerical problem is modelled using the following expansion [112]:

$$f = f_{\text{exact}} + g_1 h + g_2 h^2 + \cdots \tag{5.10}$$

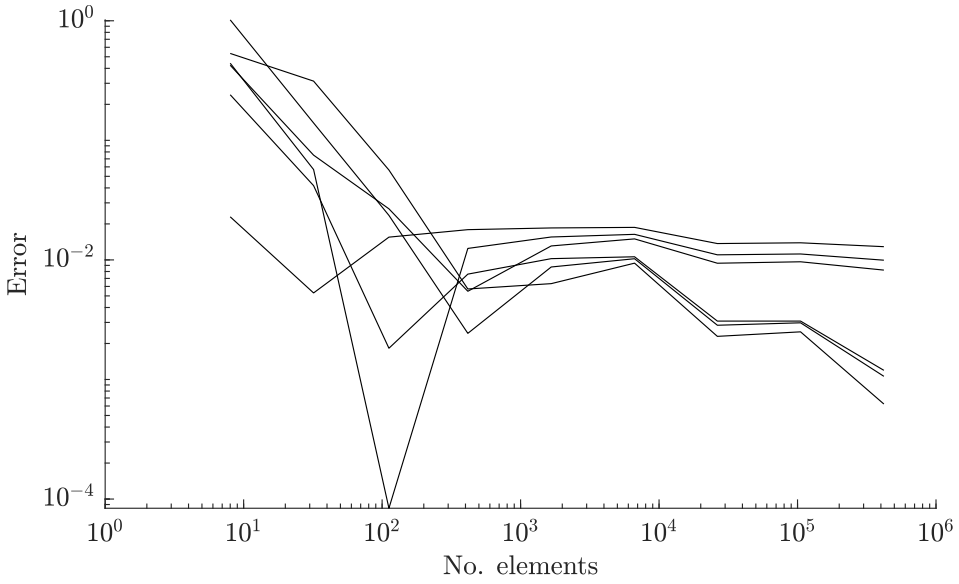


Figure 5.8: The error between the analytical solution and the model predictions of the first six modes for the SHELL181 elements

This representation will give the numerical solution based on a particular element size or grid spacing, h, where the coefficients g_n are independent functions. For second order methods, g_1 is equal to zero. For small grid sizes, higher order terms can also be neglected, meaning that for two solutions (f_1 and f_2) with grid sizes h_1 and h_2 (representing relatively fine and coarse grids respectively) the exact solution can be estimated as follows [112]:

$$f_{\text{exact}} \approx f_1 + \frac{h_2^2 f_1 - h_1^2 f_2}{h_2^2 - h_1^2} \tag{5.11}$$

This equation can be further simplified by substituting in r, the ratio between h_1 and h_2. A further extension can also be made to extend the extrapolation to p^{th} order methods [112].

$$f_{\text{exact}} \approx f_1 + \frac{f_1 - f_2}{r^p - 1} \tag{5.12}$$

A mathematically gratifying example that demonstrates the effectiveness of the Richardson Extrapolation for estimating a true, or converged, solution is the numerical derivation of π, the ratio of the perimeter of a circle to its diameter. A circle of nominal

Substructure	Parameter	Nominal value
Struts	Crack width	5mm
	E	72GPa
	η	0.33
	ρ	2650kgm^{-3}
Deck	E	71GPa
	η	0.33
	ρ	2700kgm^{-3}
Rexroth	E	70GPa
	η	0.34
	ρ	2500kgm^{-3}

Table 5.1: The nominal parameters of the bridge substructures

radius can be approximated numerically by defining a regular internal polygon, whose vertices intersect with the perimeter of the circle. Certain iterations of this approximation can be evaluated with extremely basic geometric knowledge: the triangle, square and hexagon (where the grid density is 3, 4 and 6 respectively). Using the square and hexagon iterations, the estimated perimeters are 5.657 and 6 respectively, for a circle of unit radius; the estimated value of π is 2.828 and 3 respectively in these cases. This is a second order method ($p = 2$), so the Richardson Extrapolation can be evaluated from these two estimations (where $r = 1.5$), leading to an estimated value of 3.137 for π. This is clearly very close to the commonly-accepted two-decimal place approximation of π – 3.14. The equivalent number of iterations required to achieve a solution of this accuracy is 35, representing a significant saving in computational effort when employing the Richardson Extrapolation of 94%.

The GCI is effectively an evaluation of the error between the f_i and the Richardson Extrapolation, and can be evaluated as follows:

$$\text{GCI} \approx F_S \frac{|\epsilon|}{r^p - 1} \tag{5.13}$$

where ϵ represents the difference between f_i and f_{exact}. F_S represents an optional factor of safety on the estimate of the GCI, which can be used to account for the fact that the Richardson Extrapolation only provides an estimate of the exact solution; this was evaluated using the first two grids for each submodel. Because this study was not linked to any particular risk, the factor of safety was ignored ($F_S = 1$). The order of convergence, p, was estimated to have a value of 2.

For the grid convergence analysis of the strut submodels, the diagonal (longer) strut was used in a cantilever setup. The nominal material parameter values were used (see Table 5.1), and no load or damage conditions were applied to the model. The results of this analysis are shown in Figure 5.9, which plots the convergence of the first ten natural frequencies of the strut with solution times recorded at each grid point. This indicates good convergence of the modal analysis with increasing grid refinement for the first ten modes; the optimum refinement being around 10^2 elements. The mode shapes for the submodel in this configuration are plotted in Figure 5.10. The tenth mode had a rotational mode shape, and hence shows no translational deflection in Figure 5.10.

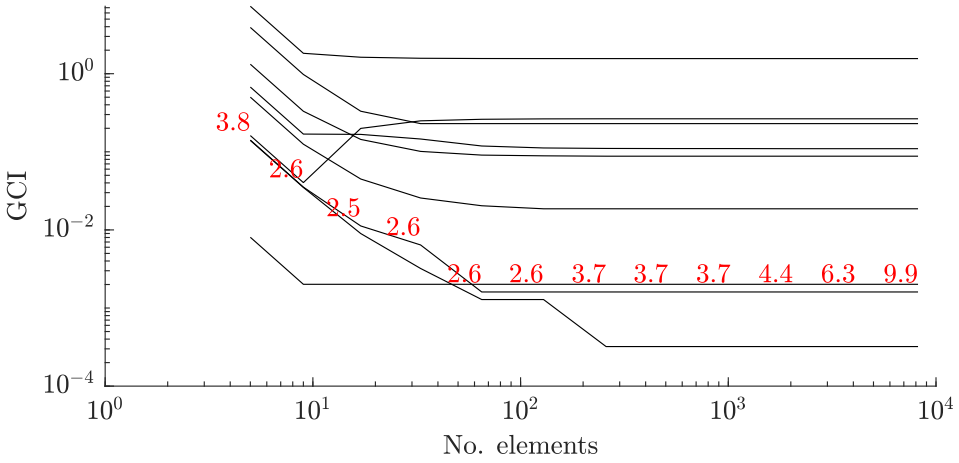


Figure 5.9: The GCI for the diagonal strut submodel, with solution times (in seconds) marked in red

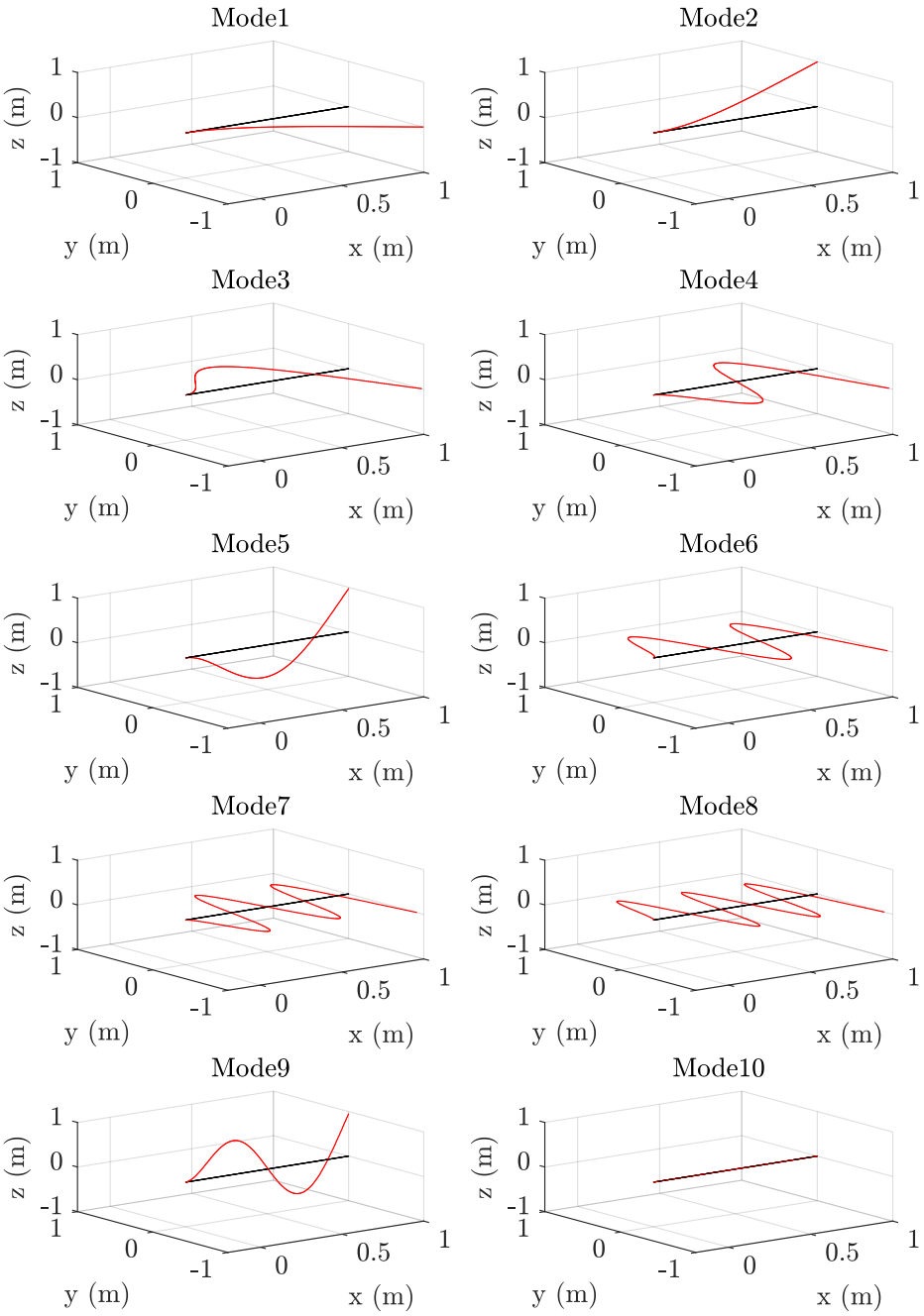


Figure 5.10: The normalised mode shapes for the strut constrained at its origin in red, with undeformed geometry in black

The upper frame submodel consisted of five lengths of Rexroth beam joined to form a figure-of-eight, with free-free boundary conditions. This meant that the modal analysis produced six rigid-body modes, which were discounted from the analysis. The grid convergence results are shown in Figure 5.11, with model evaluation times in seconds added for each grid. The figure indicates good convergence of the modal analysis with increasing grid refinement for the flexural modes of the substructure. Based on the time taken to evaluate the model compared to its convergence, it is clear that the optimum number of elements for this substructure is around 10^3 – above this, the accuracy of the solutions does not increase due to round-off error. The mode shapes for the submodel in this configuration are plotted in Figure 5.12).

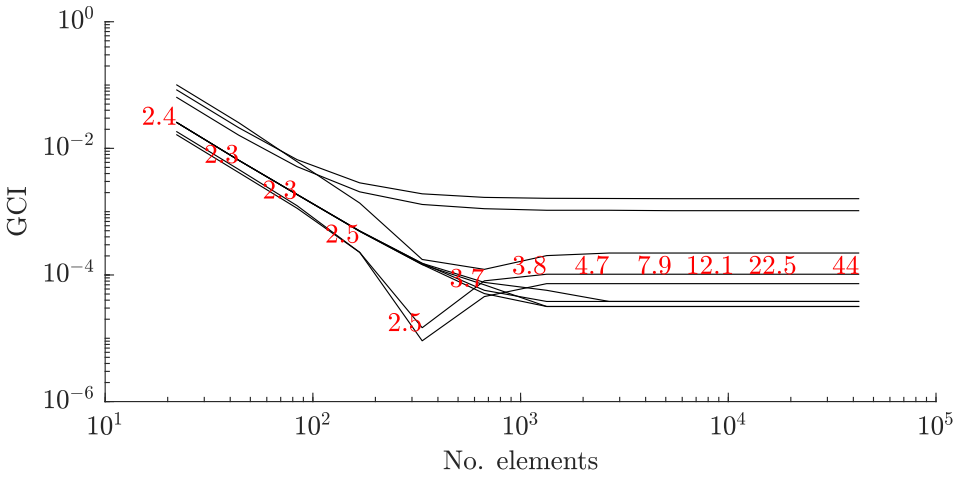


Figure 5.11: The GCI for the bridge's upper frame submodel, with solution times (in seconds) marked in red

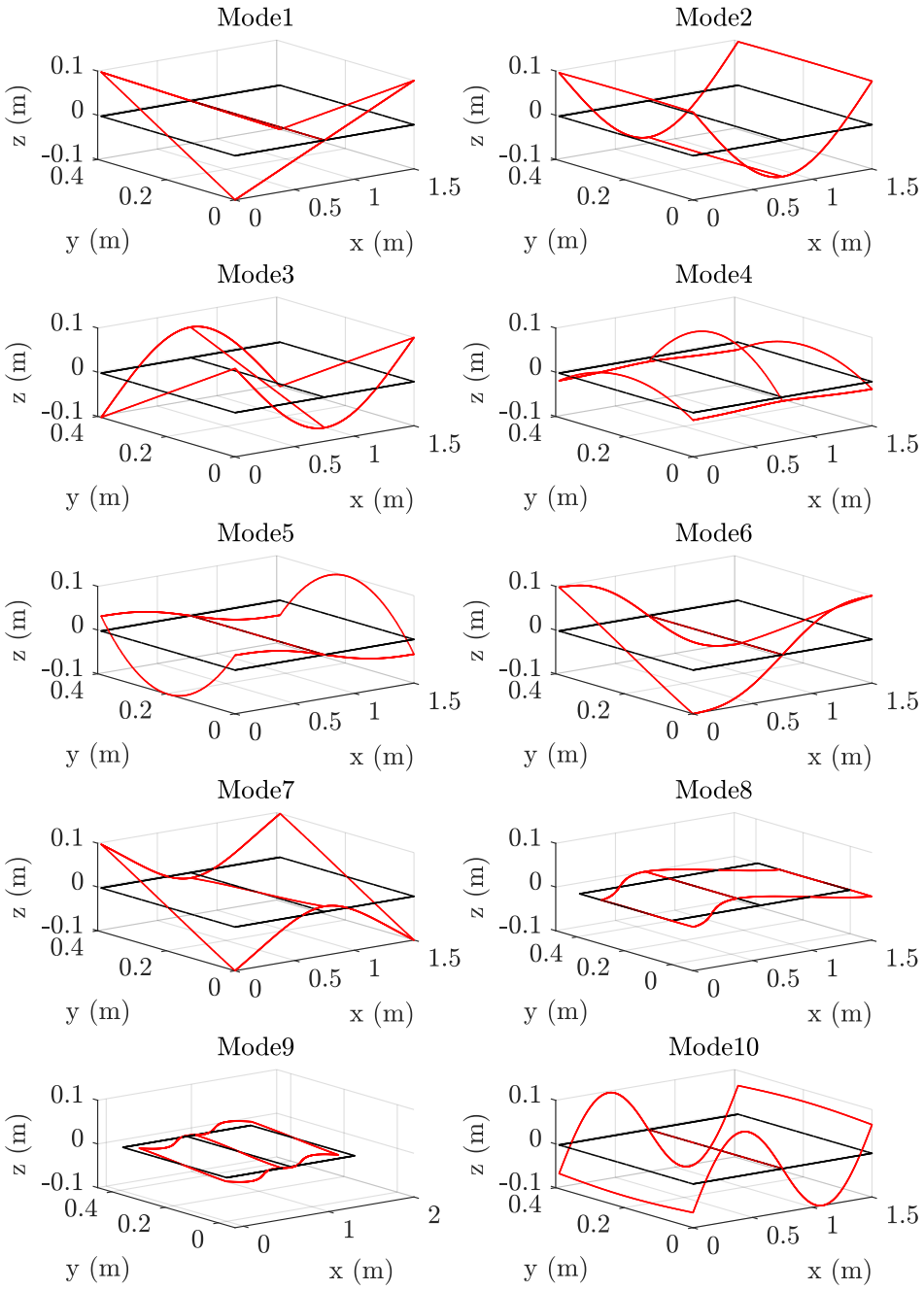


Figure 5.12: The normalised, scaled mode shapes for the upper frame in red, with undeformed geometry in black

The deck submodel consisted of four lengths of Rexroth beam joined at the ends to form a rectangle, which bounded a thin rectangular plate. The submodel was fixed at each end, which replicated its boundary conditions as part of the bridge assembly. The grid convergence results are shown in Figure 5.13. This indicates good convergence of the modal analysis with increasing grid refinement for the first ten modes of the model. As previously, the optimum grid refinement for the deck substructure seems to be around 10^3 elements, above which most of the solutions do not converge any further. The mode shapes for the submodel in this configuration are plotted in Figure 5.14).

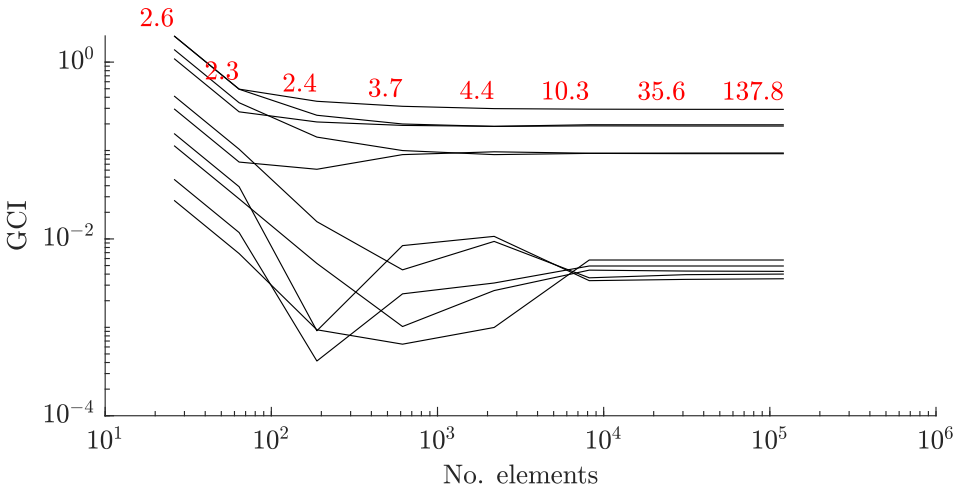


Figure 5.13: The GCI for the bridge deck submodel, with solution times (in seconds) marked in red

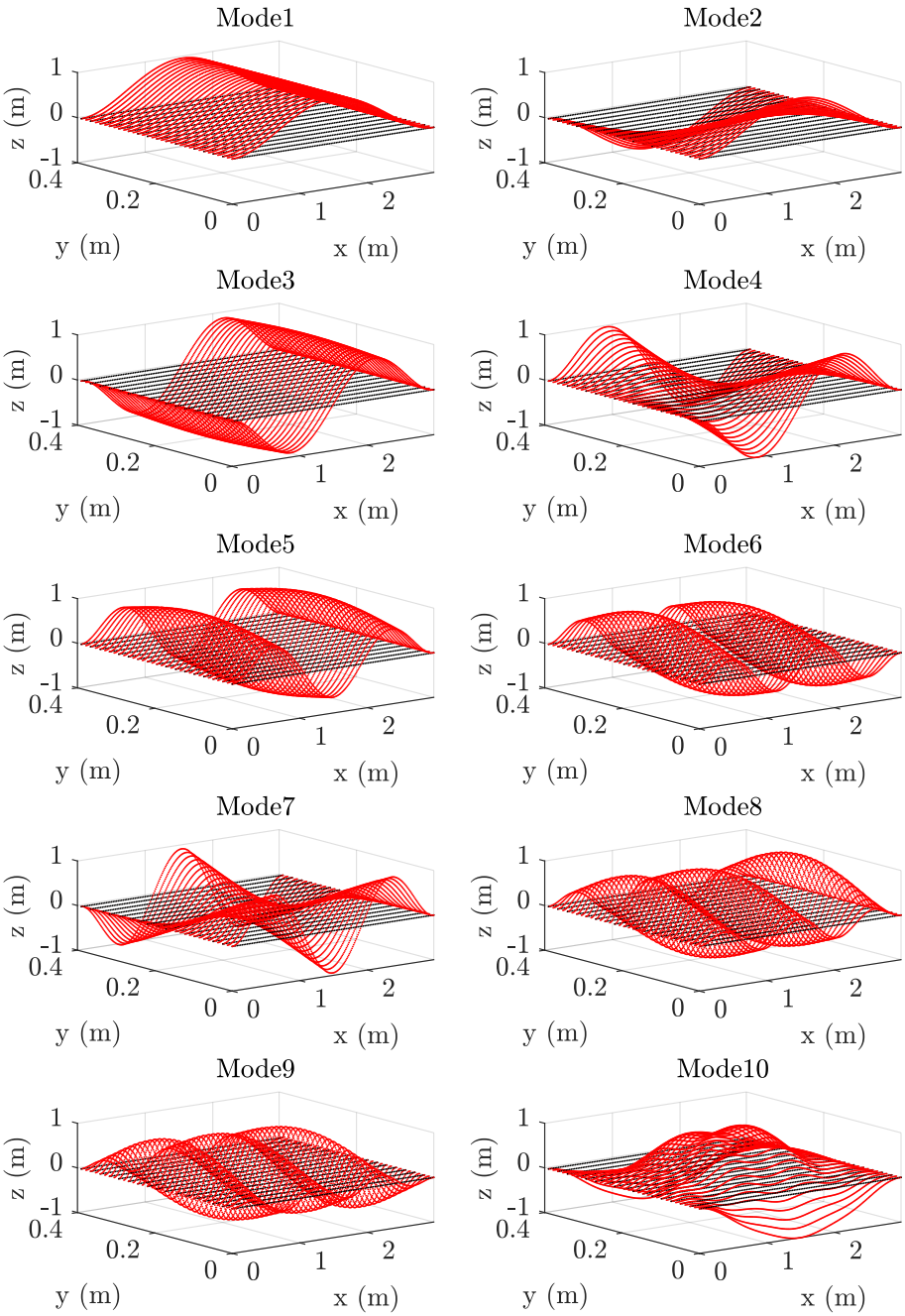


Figure 5.14: The normalised mode shapes for the deck constrained at each end, with undeformed geometry in black

Following this grid convergence analysis, the element sizes were set to 11mm, 25mm and 50mm for the struts, upper frame and deck respectively. These were chosen by taking the element size at which the convergence for each substructure reached the asymptotic range and then doubling it to save on computational effort. This step back from satisfactory convergence would add some additional numerical uncertainty to the solution, but was required for this study to aid the timely development of the methods. Based on the verification activities summarised here in conjunction with many checks and code iterations, the numerical model implementations were considered accurate for following studies.

5.4.3 Feature selection

The aim of feature selection at this stage of an FMD-SHM strategy using hierarchical validation is twofold: firstly, a set of damage-sensitive features that can be used to train a statistical model for damage detection in real-life structural data is required; secondly, a set of features on which it is appropriate to validate the submodels which make up the assembly should be identified.

The features generated by the assembly-level model to train the statistical model for damage detection should be selected based on their variance when the input damage to the model is varied (where large observable variance is preferable).

The features selected on which to validate the individual strut models must exhibit sensitivity to damage in order to allow for the damage model to be validated at this level. Engineering judgement must be exercised to ensure that the validation features at the substructure level are relevant to the selected features at the assembly level.

Natural frequencies were identified as suitable features for use in this study, due to their low dimensionality. In addition to this, the natural frequencies can be shown to be sensitive to damage in structures and substructures, and are therefore appropriate to hierarchical model designs, as they can be expected to indicate damage at both the assembly and substructure level. Finally, a key advantage of the use of natural frequencies as damage-indicating features is that they require few sensors in order to measure, provided that the sensors are not placed on any significant nodes of the structure. A key drawback of using natural frequencies as features in vibration-based SHM is that they do not give good information on damage location compared with other features such as mode shapes. Any structure has an infinite range of natural frequencies, a discrete number of which will exist in a particular part of the frequency

spectrum which is measurable in the response. A further subset of these will be sensitive to damage states of interest, as the natural frequency is directly related to the stiffness of the structure.

The feature extraction process in this case study was carried out by finding the eigenvalue solution to the model's equation of motion and was verified as discussed in the previous section. Extracting the natural frequencies from the experimental data was carried out using the PolyMAX curvefitting algorithm to the FRFs of the structure, and this is also summarised in this chapter.

Feature selection was carried out in this study by using the physics-based model in its nominal, un-validated condition to identify which of the natural frequencies at the lower end of the frequency spectrum were sensitive to the damage states of interest. The damage was simulated by using an element stiffness reduction technique implemented by reducing the cross-section of the damaged element in the model generation – the input was the crack depth, which was the amount by which the cross-section of the damaged element was reduced. The parameters of the model were its material parameters and the crack width, which are summarised in Table 5.1.

The first stage of the feature selection process was to use the model to generate a set of natural frequency predictions across the full range of damage. The damage states were midpoint cracks in each of the struts, ranging from the healthy condition to severe damage at a crack depth of 17.5mm at intervals of 2.5mm. The first 50 natural frequencies were predicted using the model, which ranged up around 100Hz. Of each of these natural frequencies, the MAC was used to assess which modes remained 'stable' across the full range of damage – i.e. which natural frequencies kept a consistent mode shape throughout the described range of inputs and did not switch with other modes as damage progressed. The threshold for the MAC, below which the modes were considered to have changed significantly, was set to 0.9. The majority of the predicted modes satisfied this criterion and were therefore retained for further analysis. Assuring that the modes remained comparable to each other across the full range of damage meant that they could be expected to retain a fit to the experimental data after being matched to data from the structure in its undamaged state.

Following the selection of the above subset of 'stable modes' from the model predictions, mode-matching between the model and the experimental data was required to identify which of the predicted modes would be useable for inference of health state

from the experimental data. This was carried out by using the MAC (Equation 5.5) to assess the similarity between the experimental and predicted mode shapes and by comparing the predicted natural frequencies with the experimental data. Given that there were many more predicted modes than were extracted from the experimental data, the matches were then finalised by selecting matches that maximised the MAC and minimised the error between the predicted and the experimental natural frequencies. The MAC was set to a minimum of 0.5 and the error between natural frequencies was set to a minimum of 10%. This yielded a further reduced subset of matched modes. The tolerance on the MAC was set relatively low since the mode shapes of the experimental data were only evaluated at a small number of locations.

Based on the criteria above, differing numbers of features were matched to the experimental data based on the number of available stable modes from the first stage. The matched modes were then fed back into the model in order to determine which showed the most sensitivity to damage in each strut. The sensitivity was assessed by finding the percentage difference in natural frequency as damage progressed compared to the undamaged natural frequency for a given mode. Two modes were then selected for each strut. This selection was carried out by ordering the modes by their sensitivity at the highest level of damage and discarding the half of the feature set that was the least sensitive. Following this, the two modes with the highest MAC were selected from each remaining subset; this process constitutes a basic quantitative sensitivity analysis. Further selection criteria could be applied at this point, such as tests for feature robustness to EOVs, as was investigated in [113] – this paper also applied Bayesian SA to investigate the impact of uncertainty in the feature predictions. This feature selection process was effectively carried out for four damage classes: the growth of a midpoint crack in each of Struts 1, 5, 2 and 6. This should be possible to generalise to any given structure, provided that the damage progression can be modelled for each class of damage to be detected in the structure.

Strut no.	Extracted mode	Predicted mode	MAC	Sensitivity
1	2	2	0.88	0.25%
	8	12	0.79	0.14%
5	10	14	0.80	0.06%
	11	15	0.73	0.38%
2	2	2	0.88	0.18%
	5	7	0.53	0.10%
6	10	15	0.80	0.36%
	12	17	0.67	0.17%

Table 5.2: The selected modes from the assembly based on damage sensitivity and mode-matching

The resulting subset of modes for this analysis are shown in Table 5.2. These are plotted across the full range of damage in Figure 5.15, which shows their damage sensitivity in both the model predictions and the experimental data. Given that the predictions shown are plotted before any validation has taken place, their accuracy should not be expected; however, comparable levels of sensitivity are observable between the model predictions and experimental data.

The selected feature set has the clear benefit of being damage sensitive, low dimensional and matched with confidence; this makes computational tasks associated with model validation and damage inference potentially very efficient. However, there are some foreseeable drawbacks: the mode-matching is rudimentary due to the limited number of sensors leading to low resolution experimental mode shapes; the natural frequencies are not well-separated, apart from the very first few, and planes of symmetry in the structure mean that many modes are very similar to each other; and damage localisation can be very difficult using natural frequencies because they are global to the structure.

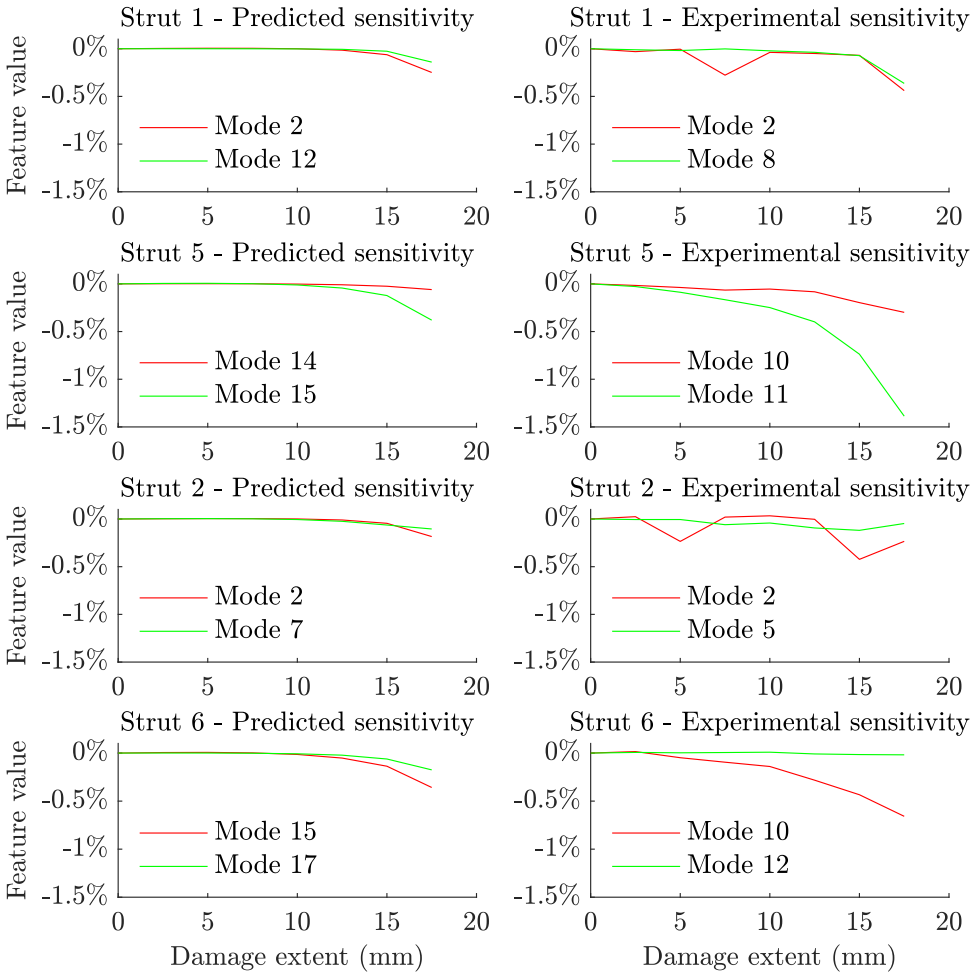


Figure 5.15: The selected modes in the model predictions and experimental data over the full range of damage prior to validation

Having determined a set of candidate features at the assembly level, a similar feature set was required on which to validate the strut models. Given that the first few natural frequencies of the struts in isolation covered the full range of frequencies extracted from the bridge, these were considered a reasonable feature set for submodel validation. To ensure that the modal behaviour was comparable between the struts in isolation and the struts when built into the assembly, a mode matching was carried out between the two cases using the MAC. A key difficulty in this was that the main local mode shape for the struts in the assembly (at the lower modes) was an s-bend as

the upper Rexroth was displaced. This was not captured in the isolated strut model, as the strut was constrained at both ends – difficulties in replicating the boundary conditions when a substructure is built into a larger assembly would be expected to persist when testing the substructure in isolation. The struts' mode shapes within the assembly are highlighted in Figure 5.16, which illustrates the displacement of the upper frame that dominates the low modes of the bridge.

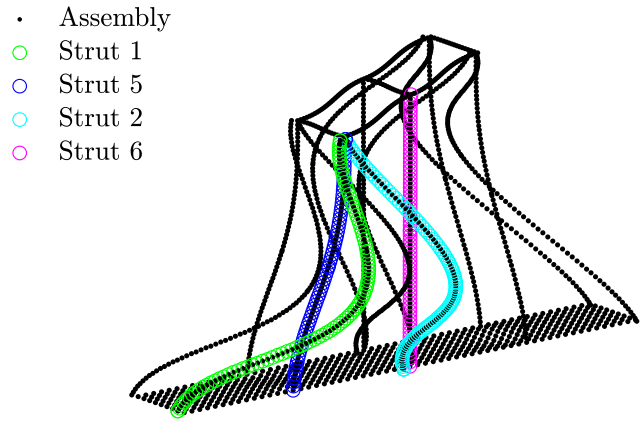


Figure 5.16: The second mode shape of the assembled bridge model, with individual struts highlighted

Figure 5.17 shows the results of the matching process. It can be seen that the first few modes of each strut in isolation can be matched to many modes of the struts as part of the assembly. Modes 1, 2 and 4 seem to cover the majority of modal behaviour of the diagonal struts in the assembly (Struts 1 and 2), while for the vertical struts (5 and 6) the first two modes seem to most closely match the modal behaviour in the assembly. It can also be seen that there are very low levels of agreement between the two sets of mode shapes for certain modes of the assembly – this indicates that validating the struts in isolation may not translate well to those modes in the assembly.

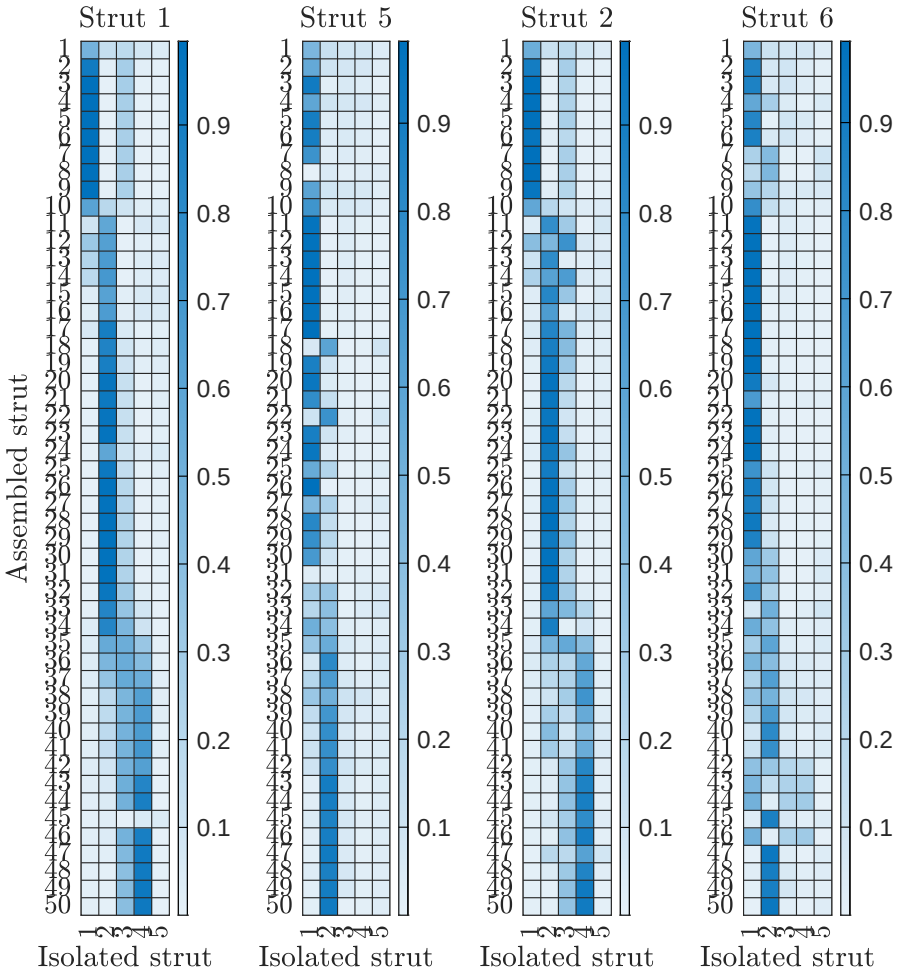


Figure 5.17: Heatmaps of the MAC between the predicted mode shapes of the assembled and isolated struts of the bridge, with colour keys marked

On the whole, Figure 5.17 shows that the first few modes of the struts in isolation fully cover the displacements that would be expected when the struts are built into the full bridge. This gives some credibility to the suitability of these features for validation of the submodels.

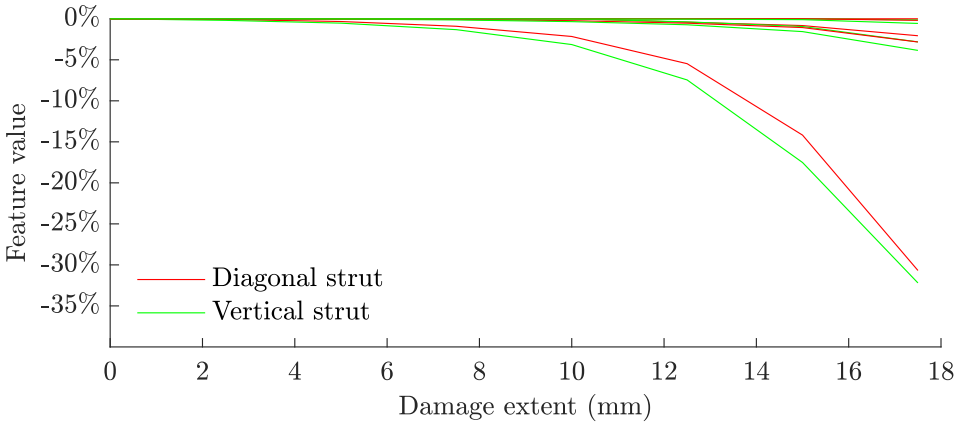


Figure 5.18: The predicted features selected to validate the diagonal and vertical strut submodels in isolation for damage prediction as part of the assembly model

The candidate validation feature sets (first five natural frequencies) for the two strut submodels are plotted in Figure 5.18. The features plotted are the proportional changes in the natural frequency as the damage progresses. Sensitivity to damage is clear for certain modes, although some modes are clearly less sensitive. These features should be suitable targets on which to validate the predictive damage models in the strut submodels, where part of the validation process will involve ensuring that the modes which are insensitive to damage match their equivalents in the experimental data.

Based on the efforts to ensure relevance between the features in the assembly that could be used in an SHM strategy, and the features selected for the submodels for validation, it is hoped that an accurate set of assembly-level models can be developed through hierarchical V&V. This is tested in the following chapters, where the features selected for submodel validation are applied in Chapter 6 and the features selected for the assembly are tested in Chapter 7. If the feature selection process has been successful, each of these chapters will be able to demonstrate that the features selected here are appropriate for use on the tasks carried out.

5.5 Conclusions

Feature extraction and selection are integral to the implementation of SHM, where raw recorded data are usually of high complexity and dimensionality, and are therefore difficult to interpret in unprocessed forms. Feature extraction refers to the signal processing tasks carried out to acquire subsets of the data that can be interrogated for SHM purposes. Damage sensitivity is usually the main requirement of a feature for SHM; secondary requirements include robustness to other variables and low dimensionality. Feature selection refers to the discriminatory processes applied to the feature set to determine the optimal features, and is carried out by SA.

This project concerns the use of models in a hierarchical, forward model-driven capacity. This introduces certain additional considerations for the handling of features. In traditional data-driven SHM, the feature selection process seeks to determine the input variables that cause the greatest variance in the output damage labels. In IMD-SHM, feature selection seeks to determine the output features which show the greatest dependency on the damage-indicating inputs.

Forward model-driven SHM combines these two interests and seeks to use the feature set that is most sensitive to damage inputs to the physics-based model, and that causes the greatest predictive variance in the statistical model. When integrated into a hierarchical modelling strategy, further considerations must be made to ensure that the features selected for the submodels will still be useful at the assembly level.

SA provides the means to carry out informed feature selection. Qualitative SA refers to analyses based on engineering knowledge or intuition, or visual methods of inspecting the relationships between input/output relationships. Quantitative SA attempts to attach a numerical estimate to analyses. SA can be carried out locally or globally over the problem domain; local SA is more straightforward to carry out, but is only valid in certain circumstances, while global SA represents the more rigorous approach.

A demonstrative global SA was carried out on the model of a bridge and on the submodels of its struts with the aim of determining a small feature set that would be suitable for SHM purposes in a hierarchical, FMD-SHM framework – this could be expanded in future work to take into account uncertain model predictions and using more rigorous quantitative metrics. A candidate feature set of natural frequencies was evaluated on its sensitivity to damage and matched to experimental data, and from this it was possible to determine that suitable modes for implementation of the model

in an SHM context could be identified – this feature set will be evaluated for an FMD-SHM problem in Chapter 7. Further to this, use of the models in their assembled and isolated forms allowed for a set of features to be identified that would be useful for validation of the submodels (the first five natural frequencies) and could be shown to be relevant to the global assembly – this feature set will be used in Chapter 6 for validation of the strut submodels. Some issues were identified in the mismatching of certain features between the assembly and the submodels due to difficulties in replicating the boundary conditions of the submodels when built into the assembly.

The process of feature selection ahead of validation represents one of the key advantages of this method, as it allows for an optimisation of the validation task. Model updating and selection can be very onerous computational tasks, so the ability to select meaningful features on which to carry out these tasks is invaluable. If the tasks carried out using the features selected here in the following chapters of this thesis are successful, this will underline the potential for carrying out feature selection for hierarchical models in this manner. Further work on this process would entail the selection of features that are insensitive to EOVs while retaining sensitivity to damage – this was demonstrated for a model of the bridge in [113] and could be applied to substructured models in a similar manner.

Chapter 6

Quantification of uncertainty in a hierarchical model structure

This chapter presents a case study in the validation and uncertainty quantification of a hierarchical model of the truss bridge. In order to attain an assembly-level predictive damage model for the bridge, the individual strut submodels were validated against experimental data. This entailed calibration of the material parameters of the submodels, and calibration, selection and validation of the crack models within the strut submodels.

Following the substructure-level validation described above, the uncertainty quantified in these processes was propagated to the assembly level using dynamic substructuring to enable probabilistic prediction of the response of the full assembly to damage, as was proposed and investigated in Chapters 3 and 4. The aim of this case study was to demonstrate that the assembly-level model of the bridge could be validated, across a pre-determined range of damage states, without using assembly-level validation data from the structure in its damage states. This is achieved in this case study by carrying out the damage-state validation tasks at the substructure level (using the strut submodels) and is tested against assembly level experimental data following the validation and UQ activities.

6.1 Submodel validation

The second phase of the case study in this thesis is validation. Validation follows feature selection, the process by which certain features are identified which are most suitable for further use. A full discussion of feature extraction and selection within a

hierarchical validation, SHM context can be found in Chapter 5; the feature selection process for this case study is written up in full in that chapter. Feature selection was carried out using the un-validated assembly-level model of the bridge and assembly-level data from the bridge in its healthy condition – these data would be relatively low cost to acquire for similar SHM strategies in real-life applications.

Validation is generally carried out in conjunction with, or immediately following, verification. Verification aims to ensure the suitability of the numerical implementation of a model, with common activities involving bug-fixing and grid convergence analysis. Verification of the submodels that make up the bridge assembly has been carried out and is also summarised in Chapter 5.

The aim of validation at this stage is to ensure sufficient accuracy of the predictions of a set of submodels which will be used to construct an assembly model. These predictions will be of the first few natural frequencies of the submodels, and the proportional changes in the natural frequencies as damage progresses. They will be local to the struts of the bridge and should exhibit variance as the damage states are varied in order to allow for the damage model to be validated. In addition, the features should be relevant to the predictive assembly model. Activities to ensure that the described feature set fulfils these criteria were carried out in Chapter 5.

The validation process will allow quantifiable confidence to be attached to further predictions in the implementation of the model, which can then be used to make predictions at the assembly level with associated confidence via dynamic substructuring. This should be demonstrable by comparing the assembly-level predictions for the model against experimental data before and after the validation process, where an increase in accuracy after validation would be expected.

6.1.1 Experimental data

A set of dynamic data was acquired from a single strut of each type (diagonal and vertical) for use in validating the strut submodels. The dynamic response of the struts to shaker excitation was recorded across a range of damage conditions and static loads. Previous static analyses of the bridge showed that the load applied to the struts when built into the full assembly was dependent on the location of the strut on the bridge as well as any additional loads on the bridge itself (see Appendix A); therefore it was important to ensure that the strut damage models would be accurate across a large range of static load conditions.

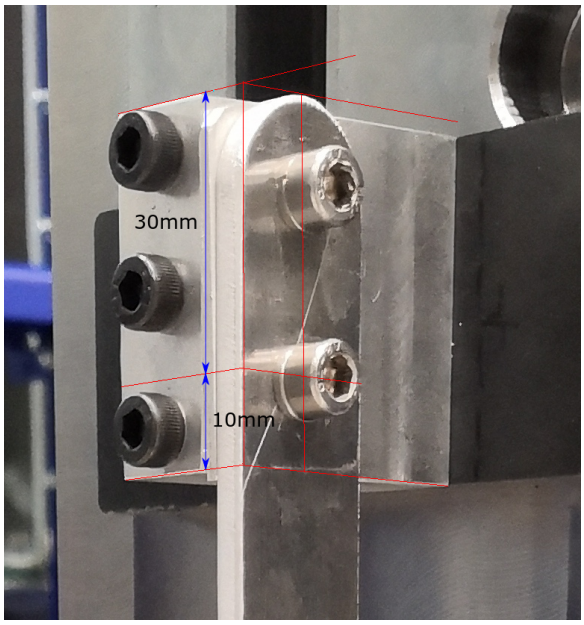


Figure 6.1: A close-up of the boundary condition on the struts for testing in isolation

The struts were clamped to a rigid base (which was itself set on damped steel feet) in such a way that their top end was constrained in all degrees of freedom and the bottom end was constrained in all degrees of freedom except for movement along the length of the strut; this was to allow for the application of static load. The bolts fixing the strut to its mountings were tightened to a torque of 10Nm. The boundary conditions for the struts are pictured in detail in Figure 6.1. This shows that the strut was clamped at each end up to 30mm from each extremity; this was used as the boundary condition for the following calibration and validation work. The contact that extends an additional 10mm along one side of the clamped boundary condition was not modelled; any non-linearity ensuing from this additional detail was assumed to be negligible but will cause some level of model discrepancy in the results.

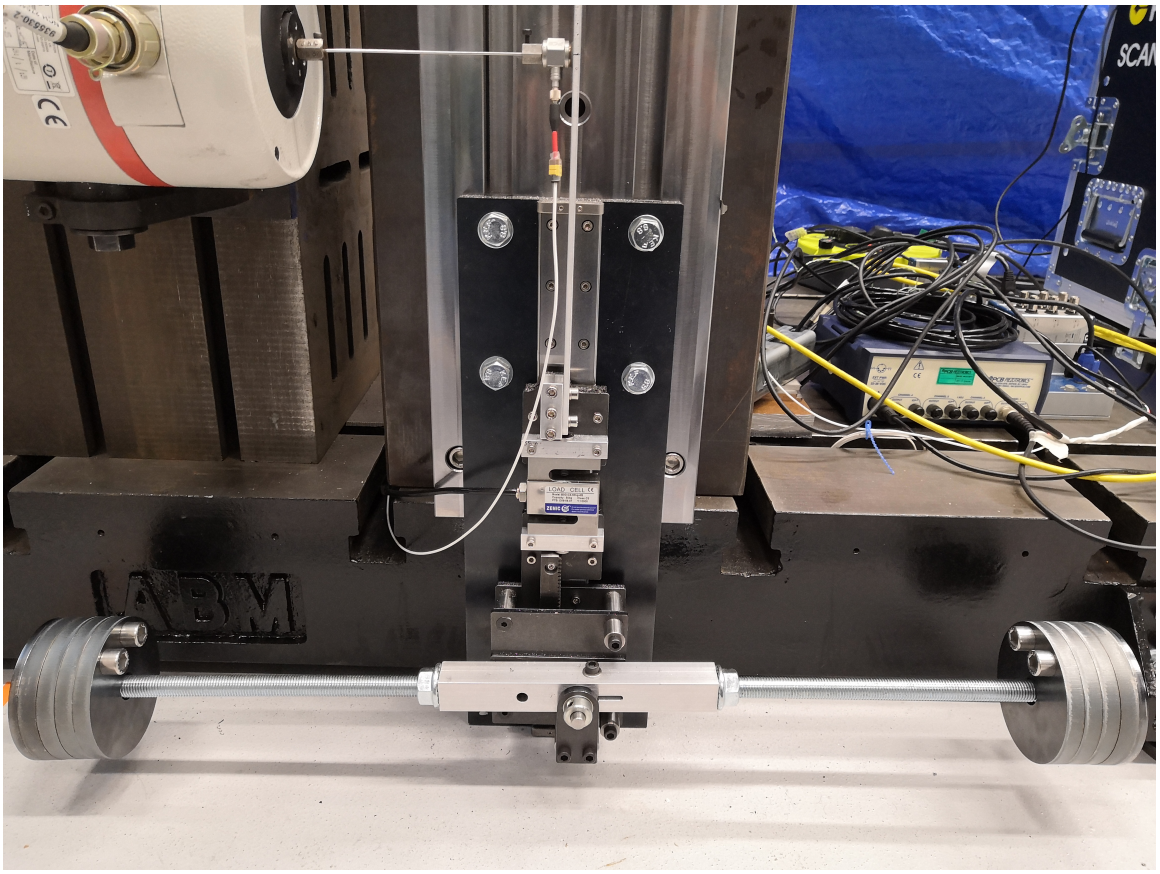


Figure 6.2: A close-up of the load cell and mechanism for applying static loads to the strut

Loads were applied to the struts by means of a wheel and pinion, as shown in Figure 6.2. The loads were controlled using weights that applied a moment to the wheel, which in turn applied vertical load to the pinion. Anticlockwise moment on the wheel produced a tensile (positive) load, and clockwise moment produced a compressive (negative) load. The load applied by this system was measured using the load cell shown in Figure 6.2. The true load applied to the strut included the weight of this load cell itself and the surrounding componentry; calibration to find the true load for a given output on the load cell was carried out in the static tests, which are summarised in Appendix A.

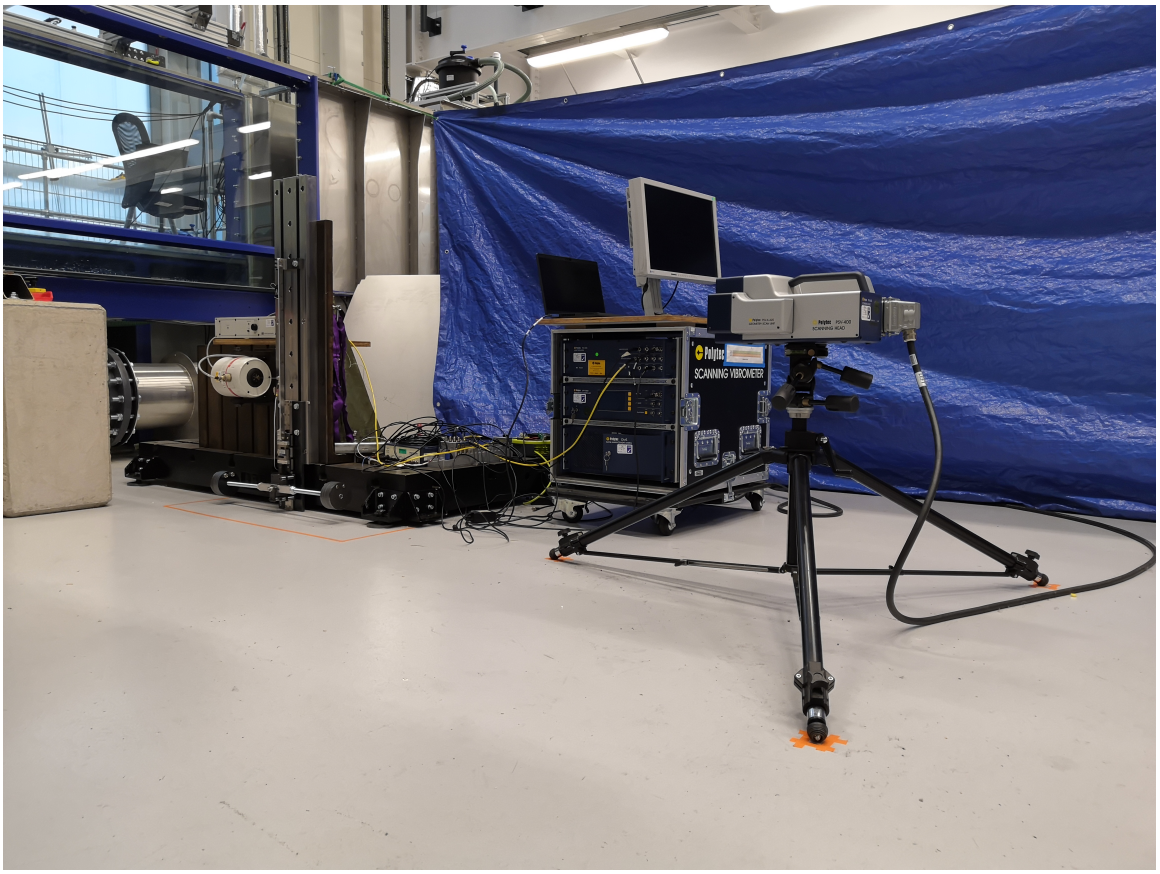


Figure 6.3: The strut set up, with scanning laser vibrometer in the foreground

An electrodynamic shaker was used to excite the struts with a white noise signal via a stinger. The stinger was attached to the strut away from the measurement locations and away from any significant integer divisions along the strut length in order to avoid any significant nodes. The strut response was captured using a scanning laser vibrometer. The full experimental setup is shown in Figure 6.3. The vibrometer measured the strut accelerations at a sample rate of 2560Hz in order to capture data in the range 0–1000Hz. A total of 32768 samples were taken for each test point. On the diagonal strut, 18 measurement points were recorded at 100mm intervals (measured from the top end of the strut). At each location two points were recorded, one on each side of the central axis of the strut. For the vertical strut, 22 points were recorded in the same manner, but at 50mm intervals. The arrangement is illustrated in Figure 6.4. The points were labelled in a clockwise manner, starting with point 1 at the top left. Damage was introduced as a saw cut at 4mm intervals at the midpoint of the diagonal strut and 150mm from the top end of the vertical strut; these locations are indicated in Figure 6.4. Each test was repeated three times.

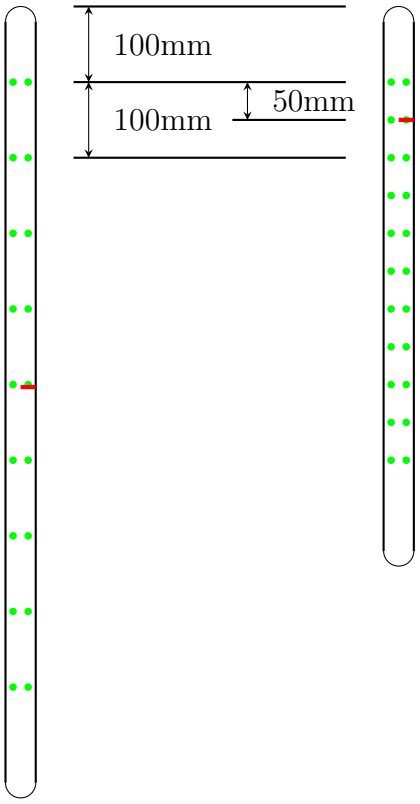


Figure 6.4: The damage (red) and measurement locations (green) for each strut

The strut response was assumed to be linear since the displacement of the strut by the shaker was small. This assumption could have been tested by investigating the reciprocity condition in the FRFs or by varying the amplitude of the excitation; however, these tests were not performed as they would have required an additional set of experiments to be carried out. Similarly, the stiffening effect of the stinger attachment could have been investigated by comparing the results to an equivalent impact hammer test; this was not carried out for the same reasons as above, and because difficulty had been encountered in preliminary tests in exciting the strut with an impact hammer.

Modal testing and analysis is a mature field within structural dynamics; as such, a range of resources outlining best practice are available which cover the methods and techniques used in this thesis [114]. H_1 and H_2 FRF estimations were extracted from the time domain data using the fast Fourier transform; Hann windows were applied to the time-series data and 16 averages were taken to reduce noise. This allowed for FRFs to be plotted with a frequency step of 0.4167Hz. Examples of these FRFs are plotted in Figures 6.5 and 6.6, which show sensitivity in the data to both

static load and damage. Despite the signal processing activities carried out, there is a significant noise noticeable in these FRFs; it also seems that the noise level increases with damage extent and with compressive load.

The noise could be caused by reflectance of the strut causing difficulties for the laser vibrometer. Spray coating to reduce this effect is available but was not used as the additional mass could affect the dynamic response of the strut. It is also possible that the distance between the vibrometer and the measurement points was not set optimally; the acquisition signal quality is dependent on this distance, which was difficult to measure, and the additional parallax effect of measuring locations along the length of the strut may have exacerbated this issue. The use of more averages could have reduced the noise levels, but this would come at the cost of reduced FRF resolution. FRF smoothing or filtering techniques could also have been employed and may have aided the following modal analysis [115].

Figure 6.5 shows the summed FRFs of all load conditions at each damage location. The H_1 estimator contains less noise than the H_2 estimator (this is indicative of lower noise levels on the input measurement compared to the response measurement, which is understandable given that the output was measured using the laser vibrometer while the input was measured at the shaker); however, the H_2 FRFs contain much more clearly defined resonance peaks. This was also the case in Figure 6.6, which shows the summed FRFs of all damage conditions at each load. The H_2 estimations were therefore selected for use in the following feature extraction activities to acquire the natural frequencies and mode shapes from the experimental data. A full set of FRFs for these tests can be found in Appendix B.

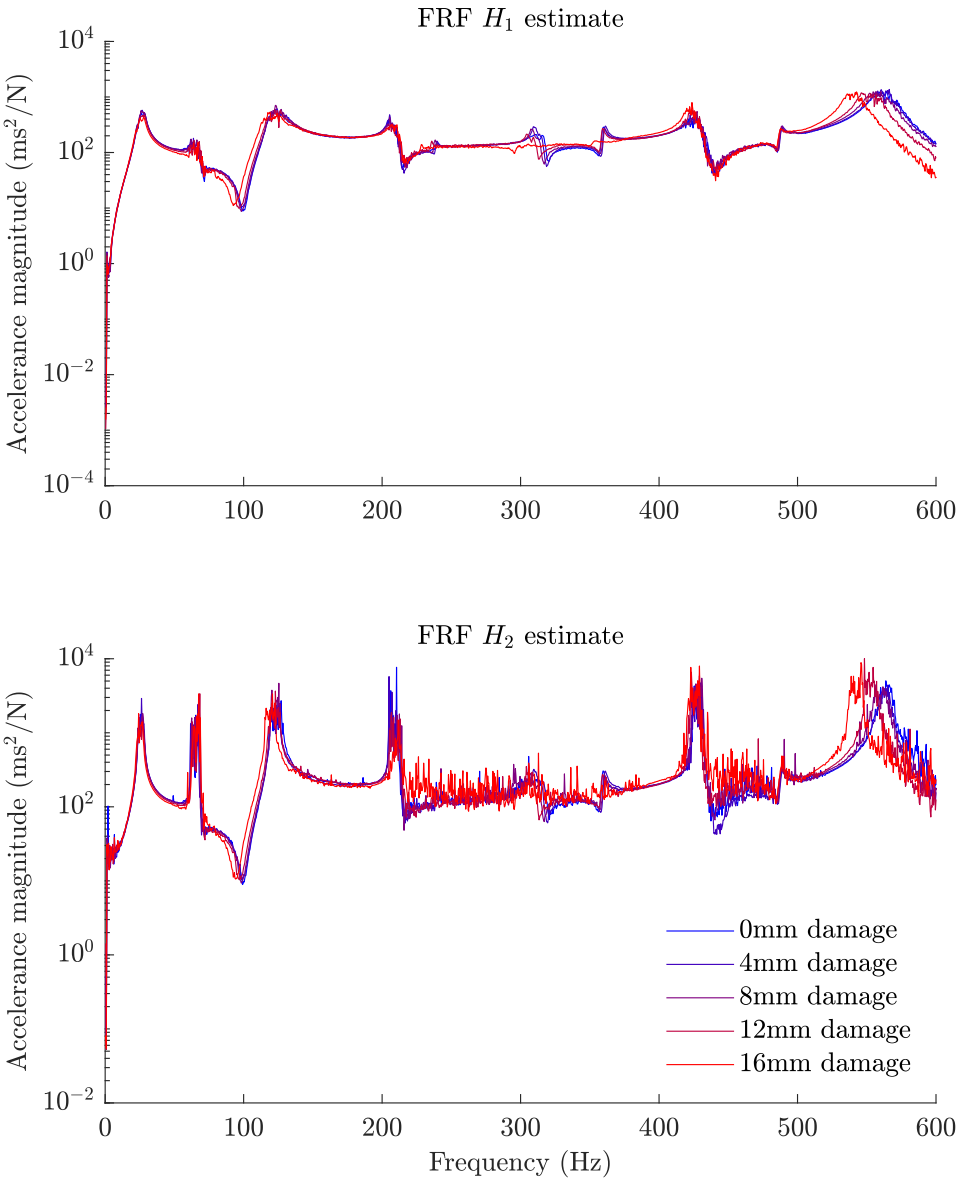


Figure 6.5: The FRFs extracted from the diagonal strut data plotted across damage conditions and summed across load conditions

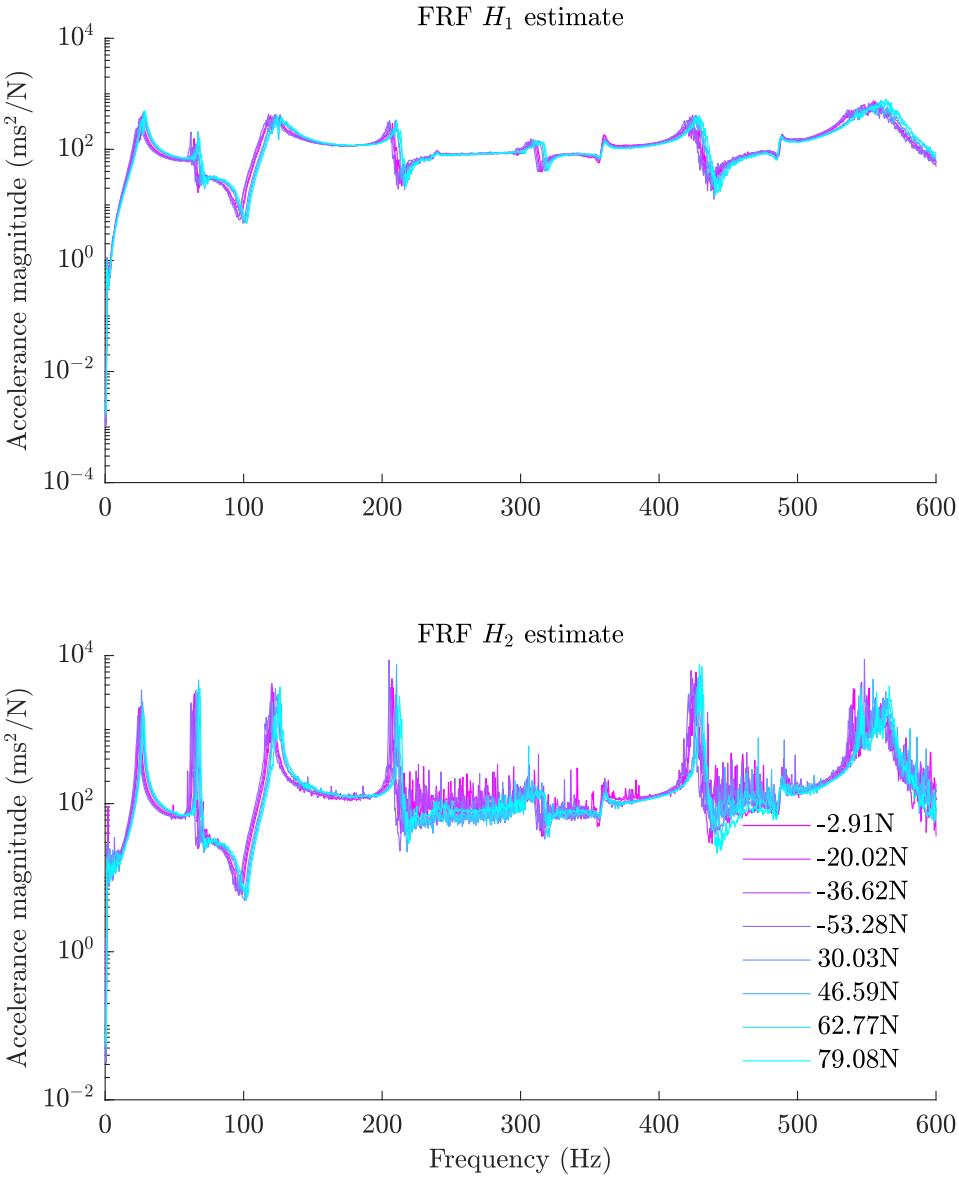


Figure 6.6: The FRFs extracted from the diagonal strut data plotted across load conditions and summed across damage conditions

6.1.2 Mode-matching

Mode-matching was required in this case study when carrying out validation tasks in order to ensure that the features being validated from the numerical predictions were being compared fairly to their equivalents found in the experimental data. The aim of mode-matching at this stage was to determine which natural frequencies predicted by the model could be confidently matched with natural frequencies found in the experimental data for each strut in isolation. This would then allow for a feature set to be determined on which the strut models could be calibrated and validated.

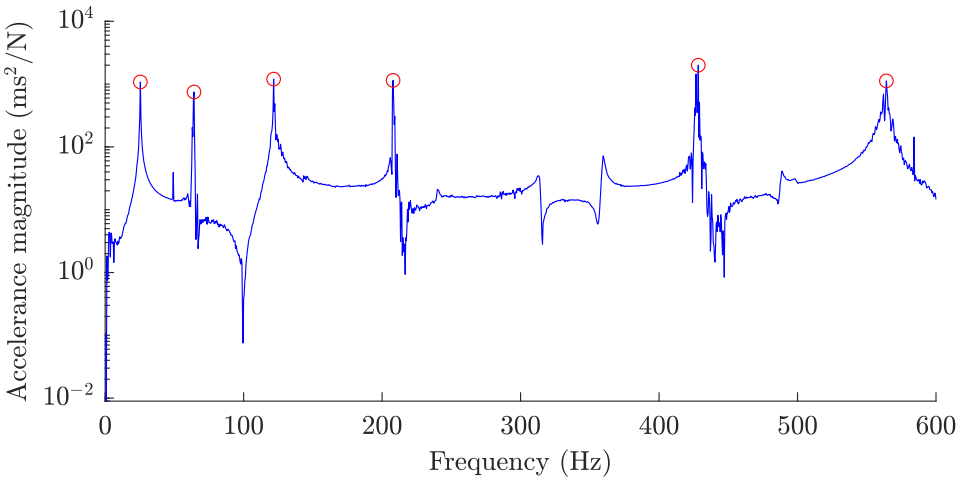


Figure 6.7: An example FRF from the diagonal strut with resonance peaks marked

Extracted natural frequencies (Hz)		
Mode no.	**Diagonal strut**	**Vertical strut**
1	25.85	49.09
2	63.86	131.11
3	122.29	273.80
4	207.82	436.04
5	427.27	640.16
6	563.47	

Table 6.1: The natural frequencies extracted from the experimental data for the two struts in their undamaged state at compressive loads of 3.54N and 19.07N respectively

The natural frequencies extracted from the experimental data were identified in the FRFs summed across all measurement points, as shown in Figure 6.4. A single-degree-of-freedom curvefitting method was then used to extract accurate natural frequencies

for each load case and damage scenario. The resonance peaks (as marked in Figure 6.7) were used to inform the curvefitting algorithm. Six natural frequencies were extracted for the diagonal strut and five natural frequencies were extracted for the vertical strut at each load and damage scenario. An example set of these at compressive loads of 3.54N and 19.07N, for the diagonal and vertical struts respectively, with no damage are summarised in Table 6.1. The full modal analysis results for these tests are presented in Appendix B.

Predicted natural frequencies (Hz)		
Mode no.	**Diagonal strut**	**Vertical strut**
1	22.99	49.08
2	63.59	136.36
3	115.22	247.03
4	124.88	268.47
5	206.73	445.26
6	309.26	667.22
7	316.60	676.32
8	432.63	909.14
9	576.96	
10	617.97	

Table 6.2: The natural frequencies predicted by the model for the two struts in their undamaged state at compressive loads of 3.54N and 19.07N respectively

The natural frequencies in the range 0–1000Hz were generated by the model for each strut; this resulted in 10 natural frequencies being estimated for the diagonal strut and 8 for the vertical strut. These were generated under each load case for each strut taken from the experimental conditions; equivalent examples are shown in Table 6.1 to those in Table 6.2.

Given that the vibrometer recorded the response in the z-direction only (Figure 6.3), it was to be expected that fewer modes would be identifiable in the experimental data, in the same frequency range, than would be predicted by the model. Due to the well-separated nature of the natural frequencies in this range of the frequency spectrum, it was possible to make the following matches using the predicted natural frequencies. Analysis of the predicted mode shapes at these frequencies showed that these modes did have mode shape displacements in the z-direction, and the MAC indicated that the matches were legitimate based on these mode shapes (although certain estimates of the MAC were very low due to measurement noise on the experimentally-derived mode shapes). This led to the set of matched modes summarised in Table 6.3.

Strut type	Match number	Extracted mode	Predicted mode	MAC	Natural frequency error
Diagonal	1	1	1	0.22	10%
	2	2	2	0.36	1%
	3	3	4	0.49	2%
	4	4	5	0.44	0%
	5	5	8	0.31	1%
	6	6	9	0.45	3%
Vertical	1	1	1	0.20	0%
	2	2	2	0.75	4%
	3	3	4	0.93	−2%
	4	4	5	0.01	2%
	5	5	6	0.45	4%

Table 6.3: Summary of the matched modes between the experimental data and model predictions

6.1.3 Calibration and validation of material parameters

The calibration of the strut models in their undamaged condition was carried out in order to determine the uncertainty in the material parameters of the models and to ensure the accuracy of their predictions. The calibrated models could then be used in the validation of the damage model applied to both struts, which would allow for confidence to be established in the predictions of the models under damage conditions. Parameter calibration is a common exercise in model-based engineering methods; for discussions on the subject related to structural dynamics, the reader is referred to [18].

The calibration of the material parameters was carried out by minimising the error between the predicted and experimental values for the first four matched modes of the vertical and diagonal struts. The material parameters of the strut models were Young's modulus, density and Poisson's ratio. The nominal values of these for the struts were 72GPa, 2650kg/m^3 and 0.33 respectively. The material parameters of the two struts were assumed to be the same, and were therefore drawn from the same underlying distributions, given that the components were cut from larger sheets of the same material. Previous analyses of the bridge under static loading (see Appendix A) showed that the struts would experience a range of static loads in operation of the bridge, both compressive and tensile (along the length of the struts). Therefore, the calibration was carried out at a range of static loads. The distributions of the

parameters were estimated by fitting a normal distribution to the values derived by using the MATLAB 'fminsearch' function at each load point and each strut. The optimisation function carries out an unconstrained multivariate optimisation process using the Nelder–Mead Simplex algorithm [116].

Initial attempts were made to calibrate the Young's modulus and density of the struts concurrently, with the Poisson's ratio assumed to be relatively accurate (and also to have negligible impact on the results). However, this caused issues because the relationships between the target feature – natural frequencies – and the two calibration parameters were inverses of each other; this is a fairly well-known issue in calibration of dynamic models [18]. Therefore, the final calibration was carried out on the Young's modulus of the struts only, and the Poisson's ratio and density were kept at their nominal values.

It should be noted here that for an ideal calibration process, a bespoke design of experiments would be carried out prior to experimental data acquisition. This would use multiple struts in tests designed purposely to measure the quantities of interest in order to form a relatively large distribution. For example, a set of static load tests on a large set of struts could have been used to measure the Young's modulus, and another set of struts could have been weighed in order to measure the material density. In this set of tests, only two struts were used: one vertical and one diagonal. The vibration data recorded were intended for validation of the dynamic models, and was not particularly suitable for material parameter calibration. This limited the capability of the calibration process; however, some improvements were made on the estimation of the model parameters, and some uncertainty quantification could be carried out on the parameters.

$$E_{calibration} = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| f_i^{predicted} - f_i^{experimental} \right|}{f_i^{experimental}} \tag{6.1}$$

The error function for the optimisation algorithm was the mean of the error magnitude between the first four predicted and experimentally-derived natural frequencies (Equation 6.1, where $N = 4$). The results of this optimisation are shown in Table 6.4, which shows some divergence between the nominal and optimised parameter values at each load point. The magnitude of this divergence is low, which would be expected given the expectedly reliable nature of manufacturer-quoted material values, and – as

such – the error for the optimised values is only slightly lower than for the nominal values.

Strut type	**Tensile load (N)**	**$E_{calibration}$ – Nominal**	**Optimised Young's modulus (GPa)**	**$E_{calibration}$ – Optimised**
Diagonal	-3.54	3.53%	72.6	3.33%
	-21.03	3.29%	72.5	3.11%
	-37.04	3.58%	72.5	3.40%
	-53.87	3.81%	73.1	3.50%
	29.20	3.08%	72.6	2.88%
	45.87	2.79%	72.8	2.56%
	62.63	2.98%	72.5	2.82%
	80.35	2.77%	72.7	2.57%
Vertical	-19.07	2.02%	69.1	1.99%
	-51.91	2.30%	69.0	2.16%
	44.48	2.12%	69.2	1.81%
	78.01	2.06%	68.9	1.55%

Table 6.4: The results of the optimisation process for calibrating the Young's modulus of the struts

The underlying distribution for the Young's modulus was estimated by fitting a normal distribution across all the test points. The mean and standard deviation fitted to these results were 71.5GPa and 2.52% respectively. Given that the manufacturer-quoted values for the material parameters would be expected to be relatively accurate, this small difference from the nominal value of 72GPa is unsurprising. The uncertainty on the parameters was quantified with the aim of taking into account the overall uncertainty in the model predictions, such as the uncertainty in modelling the boundary conditions in these tests as mentioned previously. When these values were tested against the fifth natural frequency for each strut, this yielded the results given in Table 6.5. These results show that the error for the fifth mode, which was left out of the calibration process, is comparable in magnitude to the error for the previous modes, which gives validity to the calibration process. Therefore, these results were used as the basis for further work.

Strut type	Tensile load (N)	Error – Mode 1	Error – Mode 2	Error – Mode 3	Error – Mode 4	Error – Mode 5
Diagonal	-3.54	11.40%	0.80%	1.74%	0.89%	0.88%
	-21.03	9.77%	1.83%	1.53%	0.74%	0.99%
	-37.04	10.12%	1.90%	2.31%	0.74%	0.75%
	-53.87	11.56%	1.25%	2.21%	0.97%	0.90%
	29.20	9.50%	1.68%	0.98%	0.80%	0.75%
	45.87	8.87%	1.28%	0.76%	0.87%	0.97%
	62.63	9.72%	0.68%	1.24%	0.90%	0.97%
	80.35	8.95%	1.01%	0.89%	0.82%	0.95%
Vertical	-19.07	0.39%	3.61%	2.31%	1.73%	3.84%
	-51.91	0.16%	4.32%	2.41%	1.81%	4.36%
	44.48	0.17%	3.69%	2.28%	1.60%	3.39%
	78.01	0.54%	2.89%	2.27%	1.80%	3.86%

Table 6.5: The predictive error for the mean calibrated Young's modulus tested against the first five natural frequencies for each test point

6.1.4 Damage model calibration and validation

The overall aims of the damage model validation tasks were to acquire a calibrated damage model with demonstrable accuracy in predicting the change in the natural frequencies of the struts over the full range of damage states.

The primary source for crack modelling in beams is [117]; this paper reviews four open crack modelling methods and compares them against experimental data – the vast majority of damage models for beams can be categorised as one of these overall techniques. The methods can be categorised as follows:

- Element stiffness reduction methods
  - This method reduces the stiffness of the element where the crack is located on the beam.
  - The two key inputs are the location of the element and the magnitude of the stiffness reduction.
  - The stiffness reduction can be related to the crack depth by reducing the depth of the element at the crack location by the crack depth – this has the benefit of adding physical interpretability to the model inputs.

  - An additional parameter is crack width (the length of the element at the crack location). This gives greater control of the model but means that the beam requires remeshing for different parameters.

- Discrete spring methods
  - This method replaces the continuum beam model with a spring of variable stiffness at the crack location.
  - The inputs to the model are the spring stiffness, which is based on the depth of the true crack, and the crack location.
  - The physical interpretability of this model is low, and the damage magnitude is only controllable through the spring stiffness, which does not allow for independent control of crack depth and width.
- Element removal methods
  - This method is based on beam models with 3D meshes.
  - Elements can be removed entirely from these meshes to best match the removal of material observed in the real crack.
  - This method is the most physically representative and can be used to very precisely model the geometry of a crack, but it has relative high computational cost due to its dependence on fine meshes and remeshing.
- Stiffness distribution methods
  - This method uses a law to describe the distribution of the stiffness across the whole length of a cracked beam.
  - Many distributions can be used; the Christides and Barr [118] and Sinha [119] distributions were tested in [117]. A Gaussian distribution was used in [120].
  - The accuracy of the method is entirely dependent on the distribution law selected; significant disparities were observed between the accuracy of the Sinha distribution and the Christides and Barr distribution in [117].

Three candidate models were identified for this case study, described in the following sections. Given that the strut submodel was constructed as a solid beam of 2D line elements, element removal methods were not applicable in this case. Discrete spring methods were also not investigated for this case study. Two element stiffness

reduction models were tested, and one stiffness distribution model. These models were developed incrementally until a satisfactory level of accuracy was achieved in preliminary testing; this allowed for models of different types and differing levels of effectiveness to be tested against each other throughout the overall framework.

6.1.4.1 Model 1

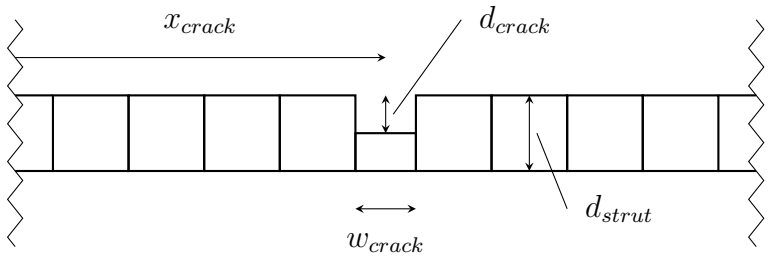


Figure 6.8: Graphical representation of Model 1, with inputs and parameters marked (x, d and w represent distance, depth and width respectively)

Model 1 was essentially an element stiffness reduction damage model. The inputs to the model were crack depth and location on the strut. The element at the crack location was then reduced in stiffness by reducing its cross-sectional depth by the depth of the crack, with the centroid offset from the undamaged elements, as shown in Figure 6.8. The key tuneable parameter of the model was crack width (w_{crack}), which determined the size of the element for stiffness reduction. The input parameters were crack location (x_{crack}) and crack depth (d_{crack}).

6.1.4.2 Model 2

Model 2 was developed based on an observed lack of sensitivity to damage at greater crack lengths in Model 1 during initial testing. It was therefore designed as an extension of Model 1 with an additional tuneable parameter that controlled the Young's modulus of the element at the damage location, α (the model inputs were also crack depth and location). This parameter was intended to increase the flexibility of the model in calibration, and to allow it to take into account potential work hardening or softening effects as the crack developed. The contribution of this parameter is described in Equation 6.2.

$$E_{cross-section} = E_{strut}\left(\frac{d_{strut} - d_{crack}}{d_{strut}}\right)^{\alpha} \tag{6.2}$$

6.1.4.3 Model 3

Model 3 was a stiffness distribution damage model which used a Gaussian distribution to describe the Young's modulus at each node of the model, as was demonstrated in [120]. The inputs were the crack location (the mean of the Gaussian distribution) and depth (which determined the peak value of the probability density function) and the tuneable parameter was crack width (which was set as the standard deviation of the Gaussian distribution). The form of the function is given in Equation 6.3.

$$E_{node} = E_{strut}\left(1 - \frac{d_{crack}}{d_{strut}} e^{-\frac{1}{2}\left(\frac{x_{node}-x_{crack}}{w_{crack}}\right)^2}\right) \tag{6.3}$$

6.1.4.4 Damage model parameter calibration

The parameters of the above damage models were calibrated using ABC [57]. This algorithm draws samples of the model parameters from a set of prior distributions, and then tests the models based on the samples against experimental data. An error metric is defined such that any sample points that produce an error greater than the threshold on this metric are discarded, while those that produce a lower error are retained and used to form an estimate of the posterior distribution. This makes ABC a likelihood-free method for estimating the posterior distributions of model parameters; this was the reason ABC was selected for this research, as it removed the task of fitting a formal likelihood distribution to the experimental data while enabling the incorporation of prior knowledge on the model parameters. A further advantage is that the posterior estimate can be added to over time by continuing to draw samples from the prior. ABC also gives an exact limit for model discrepancy based on the threshold set on the error metric [121]. Initial indications of model performance can be drawn from ABC, as 'accurate' models with well-chosen prior parameters will pass the error threshold at a greater rate than 'inaccurate' models or models with poorly chosen priors; however, these inferences must then be backed up by validating the posterior models against independent test data.

ABC is a popular method for model selection and parameter estimation in structural dynamics, with examples of its application being found in [88, 122]. It can also be applied to non-linear system identification [122]. An early example of the use of ABC applied the method to parameter estimation for population genetics [57].

Model	Parameter	Distribution	Mean	Standard deviation
1	cr_w	Normal	5mm	1mm
2	cr_w	Normal	5mm	1mm
	α	Normal	0	0.1
3	cr_w	Normal	5mm	1mm

Table 6.6: The prior parameter distributions for each model

The prior parameter distributions for each model were set according to Table 6.6. The values for crack width were based on observations of the saw cuts during experimental work, with a standard deviation of 20% to allow for some adjustment in the calibration process (an additional criterion that the crack widths could not be less than 0 truncated the prior; however this was very unlikely and did not occur in any of the samples). The prior value for α was set to zero since it was not known whether the saw cuts would introduce a softening or stiffening effect to the remaining cross-section.

The models were tested at a range of load points and damage conditions, with certain damage cases left out in order to validate the results of the ABC process for each model. The test points used for each set are summarised in Table 6.7.

Strut type	**Damage extent (mm)**	**Tensile load (N)**							
Diagonal	0	-3.54	-21.03	-37.04	-53.87	29.20	45.87	62.63	80.35
	4	-2.46	-19.11	-35.95	-52.66	29.44	46.22	62.63	78.51
	8	-2.68	-20.12	-37.10	-53.90	30.55	46.98	63.26	79.43
	12	-2.86	-19.86	-36.63	-53.01	30.28	45.30	63.14	78.45
	16	-3.04	-20.02	-36.40	-52.98	30.68	48.58	62.20	78.63
Vertical	0	-19.07		-51.91		44.48		78.01	
	4	-18.18		-51.00		44.85		77.54	
	8	-19.34		-52.07		44.94		77.41	
	12	-19.52		-52.18		45.48		75.70	
	16	-18.52		-51.06		44.92		77.11	

Table 6.7: The test points at which the experimental data were recorded for each strut (green shading indicates training datapoints; red shading indicates testing datapoints)

The training set utilised the proportional change in natural frequency from the undamaged condition at damage extents of 8mm and 16mm, while the validation set used the same feature at 4mm and 12mm. This makes the validation test points

interpolative rather than extrapolative from the test points, which is less challenging for model validation. However, this choice was made because having greater damage extents in the training data would allow for more effective calibration of the models to the impacts of greater damage extents. Further work based on this research could potentially investigate the impact of separating the training and validation data in this way.

$$E_{ABC} = \frac{1}{N_{damage} + N_{loads}} \sum_{j=1}^{N_{damage}+N_{loads}} \left[\frac{1}{N_{modes}} \sum_{i=1}^{N_{modes}} \mid \Delta f_{i,j}^{predicted} - \Delta f_{i,j}^{experimental} \mid \right] \tag{6.4}$$

The error function upon which the threshold for the ABC posterior estimation was set was the average of the difference between the predictions and experimental data at the damage extents described above (Equation 6.4), averaged across all damage and load cases, and across the first five matched modes. Δf represented the proportional change in natural frequency from its undamaged value.

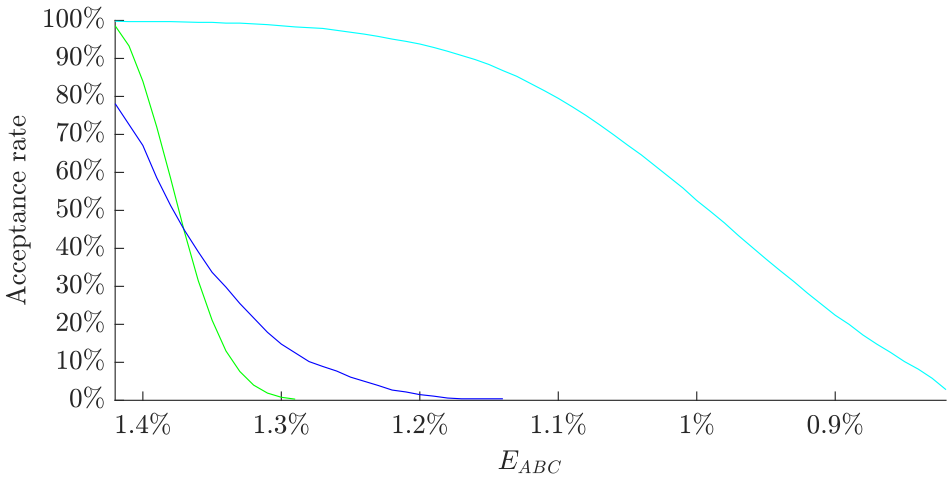


Figure 6.9: ABC acceptance rates for Model 1 (green), Model 2 (blue) and Model 3 (cyan)

The acceptance rates for the three models based on this error metric for prior distributions of 1000 samples are given in Figure 6.9. This indicates that Model 3 performs the best of the three considered models, with Model 2 slightly out-performing Model

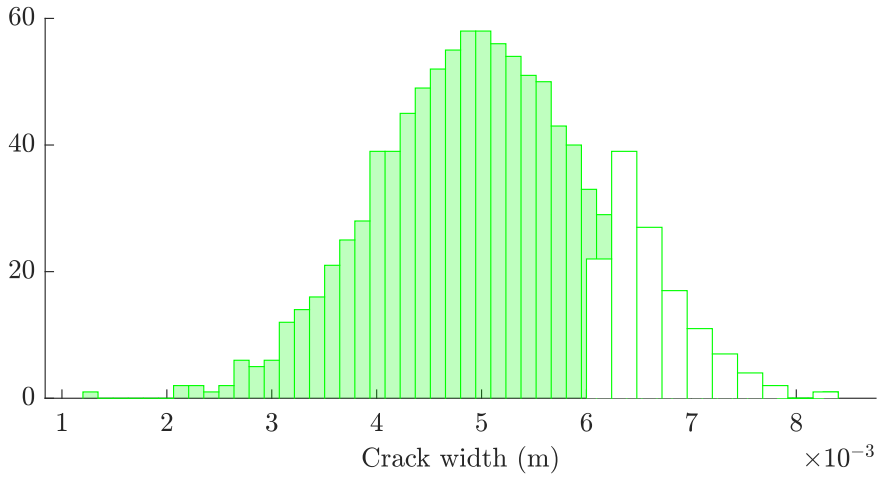


Figure 6.10: Histograms of the prior samples (filled) and posterior samples (clear) of crack width for Model 1

1. This was then to be confirmed using the validation set based on posterior estimates for each model.

In order to allow for fair comparison between the posterior distributions estimated for each model, a minimum acceptance ratio was set at 10%. This yielded the posterior parameter distributions shown in Figures 6.10, 6.11 and 6.12. Models 1 and 3 yielded highly skewed posterior distributions for the crack width parameter compared to the prior distributions; as this was the only tuneable parameter for these models, the nature of ABC means that one tail of the prior will be selected, which causes skewness in the posterior. Model 2 has two parameters, which allows for a less skewed estimation of each, as can be seen in Figure 6.11. This means that the skewness in the model parameters is unlikely to be representative of physical laws, but further analysis would be required to confirm this. In each case, the posterior distributions contained crack widths from the higher end of the prior distribution, indicating that the prior model parameters resulted in underpredictions in the sensitivity to damage in the struts (the α posterior for Model 2 also indicates this trend as positive values were selected for the posterior).

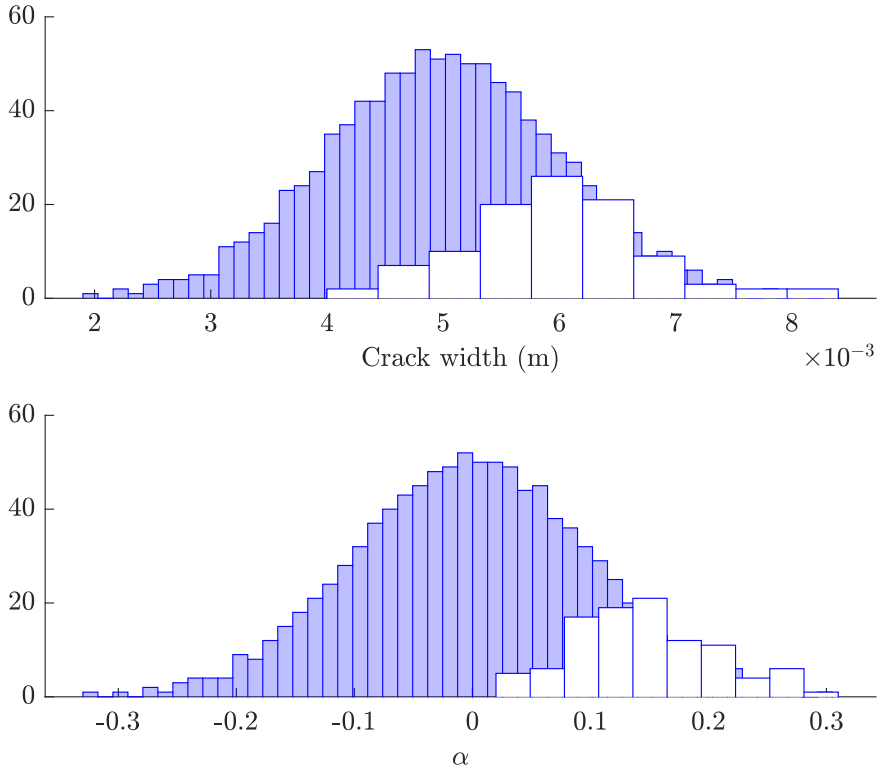


Figure 6.11: Histograms of the prior samples (filled) and posterior samples (clear) of crack width and α for Model 2

6.1.4.5 Posterior validation

As discussed in the methodology, the damage extents of 4mm and 12mm were left out of the error function for estimating the parameter posteriors in order to provide validation data for the ABC results. The prior and posterior parameters were evaluated at these test points in order to validate the ABC process carried out at the other test points. Ten samples were taken from the prior and posterior datasets for this process. These samples are plotted across the damage range in Figures 6.13, 6.14 and 6.15 for the vertical strut for each of the models.

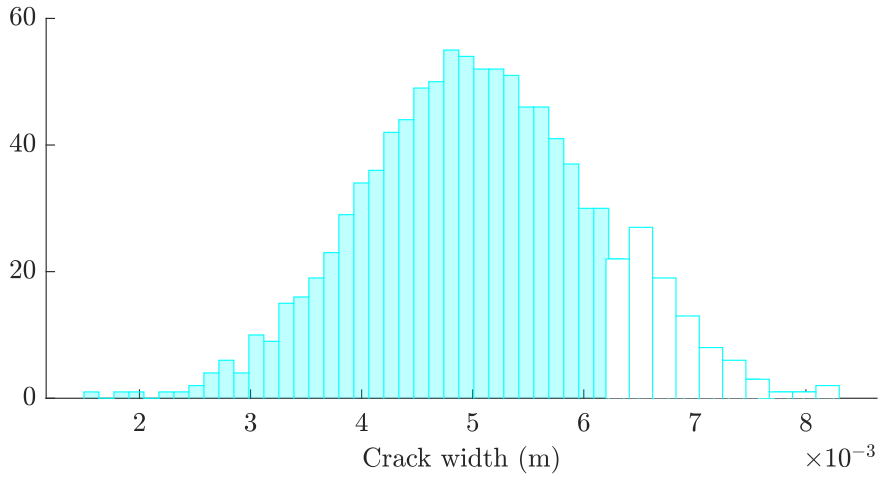


Figure 6.12: Histograms of the prior samples (filled) and posterior samples (clear) of crack width for Model 3

It is clear from these that some improvement in accuracy has been made, as the posterior samples sit closer to the data than the prior samples; however some inaccuracies remain, which are particularly noticeable at higher damage extents. This indicates that, despite the calibration processes carried out, the models do not fully capture the true sensitivity to damage in the struts. Notably, the predictions are clearly highly robust to the static load conditions that the struts were subjected to for all modes and damage models.

Model 1 clearly does not capture the full sensitivity to damage for modes 3 and 4; however, the posterior does offer a more accurate set of predictions than the prior model in this case. The same can be said for Model 2, which seems to provide a greater estimate of the uncertainty on its predictions. Both models predict the fifth mode well, and capture the lack of damage sensitivity on modes 1 and 2 (although the level of uncertainty on these modes does not seem to have been fully quantified given the variation seen in the experimental data for these modes). Given the similarity in form between models 1 and 2, they would be expected to perform relatively similarly; this is clear from the results in Figures 6.13 and 6.14.

Model 3 performs better than Models 1 and 2 when predicting modes 3 and 4, but is comparatively less accurate when predicting mode 5. Model 3 captures sensitivity to minor damage much better than Models 1 and 2, which tend to underpredict damage sensitivity at crack depths of 4mm. Like Models 1 and 2, Model 3 captures the lack

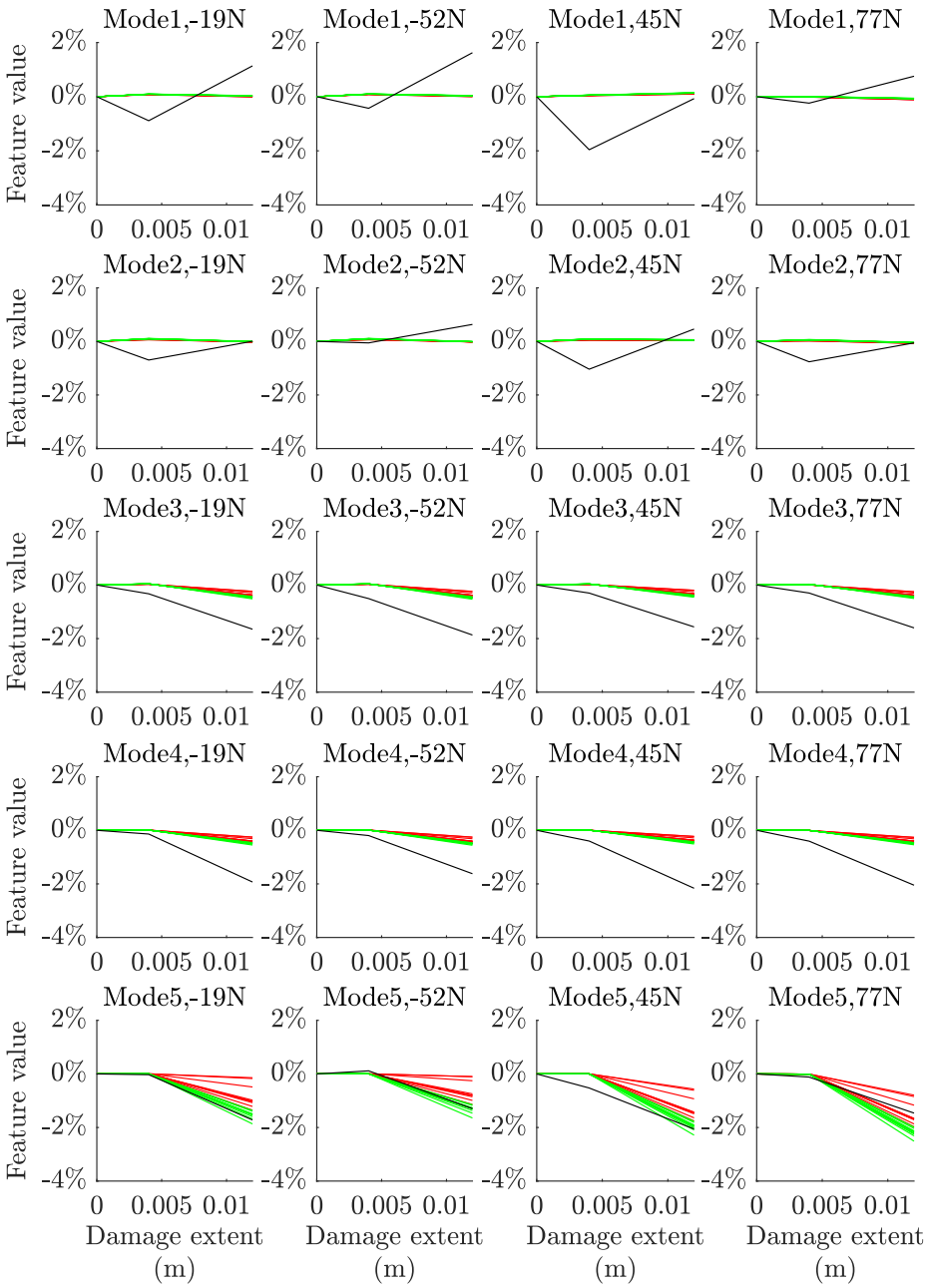


Figure 6.13: The change in natural frequency at the validation test points for the vertical strut, with experimental data in black, prior samples in red and posterior samples in green for Model 1

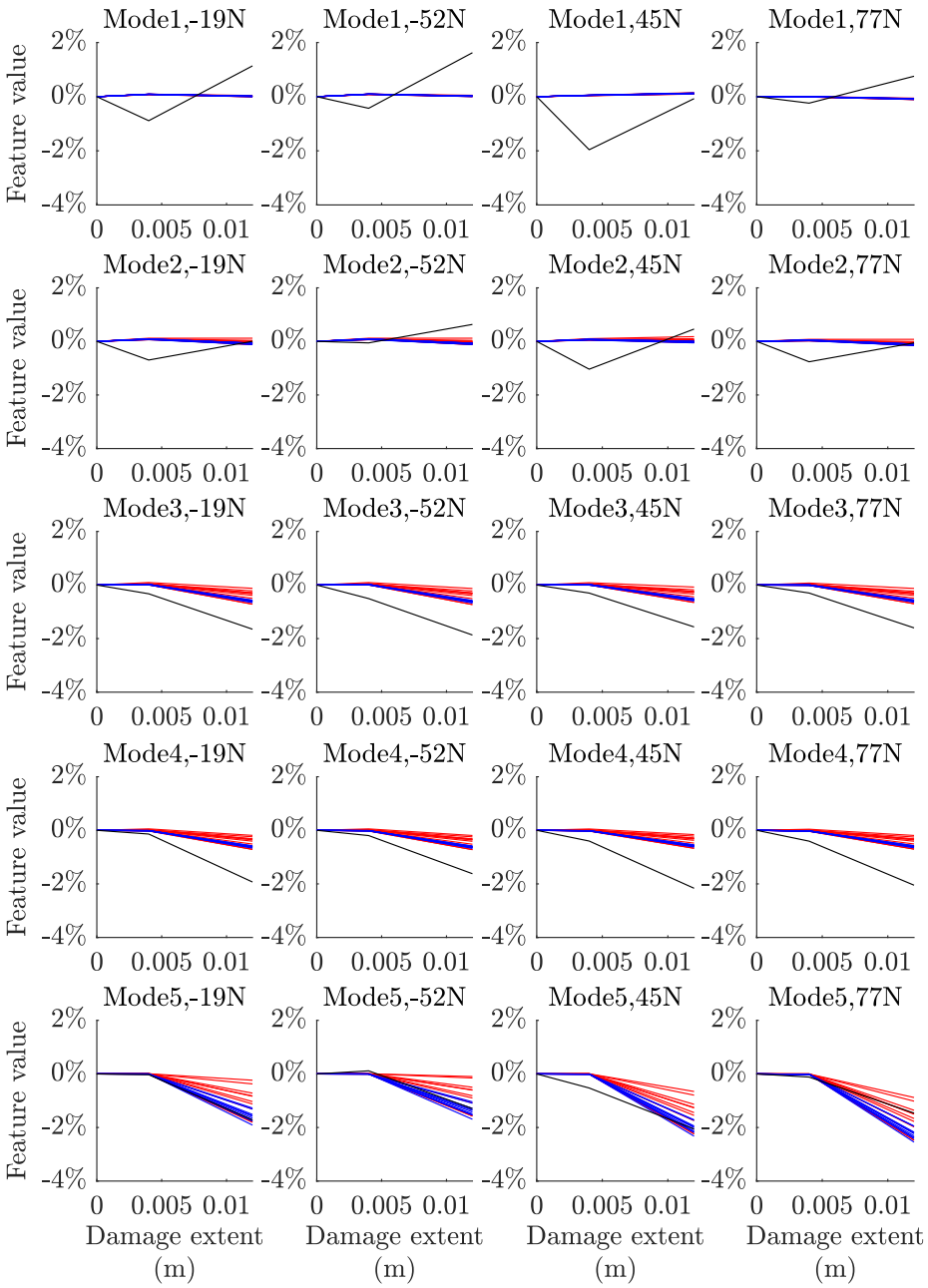


Figure 6.14: The change in natural frequency at the validation test points for the vertical strut, with experimental data in black, prior samples in red and posterior samples in blue for Model 2

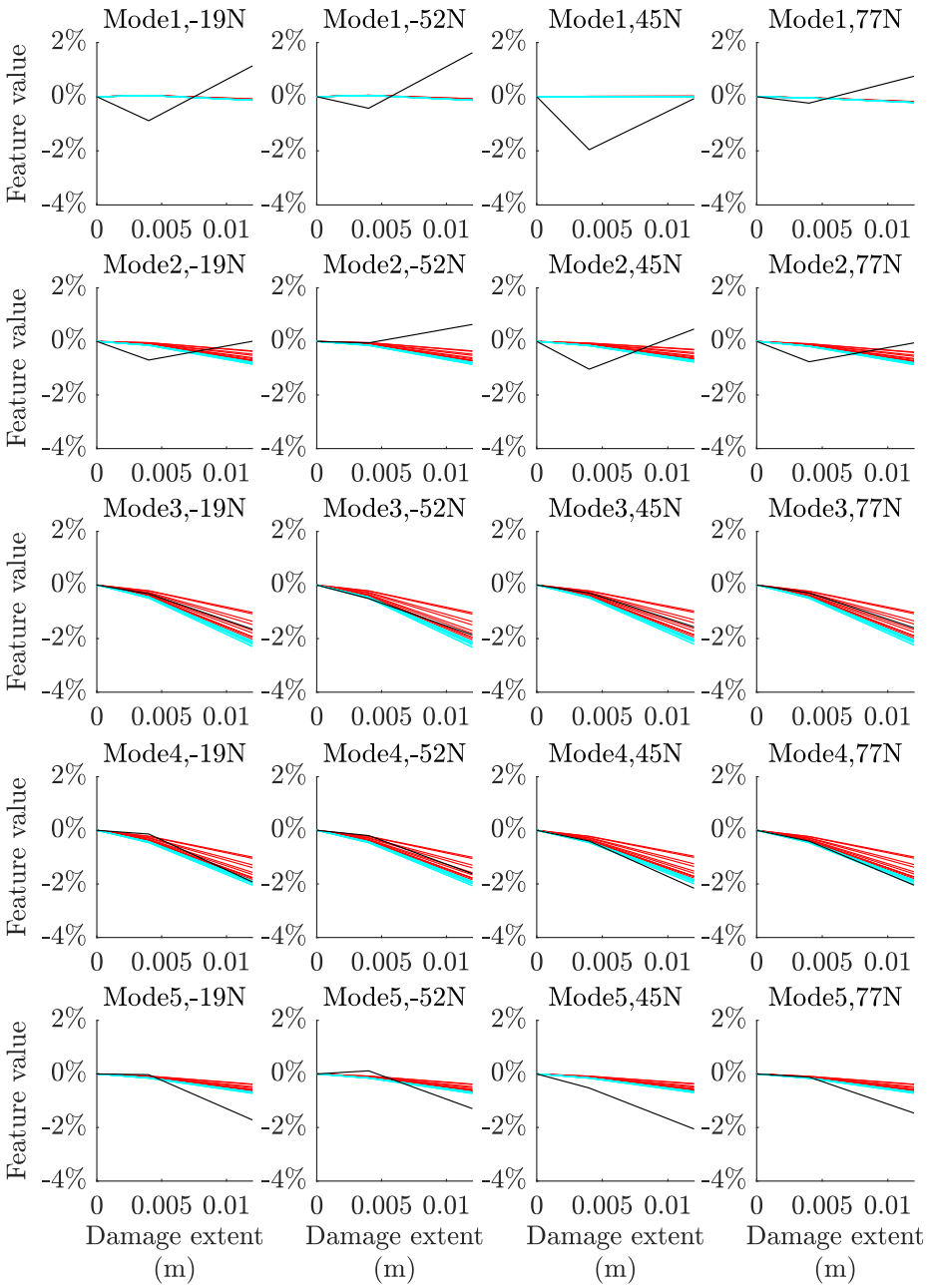


Figure 6.15: The change in natural frequency at the validation test points for the vertical strut, with experimental data in black, prior samples in red and posterior samples in cyan for Model 3

Strut / Model	Diagonal		Vertical	
	Prior	**Posterior**	**Prior**	**Posterior**
1	0.4682%	0.4614%	0.4857%	0.4475%
2	0.4664%	0.4520%	0.4798%	0.4318%
3	0.4219%	0.4096%	0.3710%	0.3722%

Table 6.8: The error between the predicted features and experimental features for the three strut models before and after validation

of damage sensitivity in modes 1 and 2, but seems to do a better job of robustly quantifying the level of uncertainty on mode 2.

Due to the skewed nature of the posterior parameter distributions for the models (Figures 6.10, 6.11 and 6.12) it is difficult to meaningfully quantify the accuracy in their predictions compared with the uncalibrated prior predictions. Nevertheless, a crude estimate of the mean error between the predictions and the experimental data can be made by taking the mean across each of the modes, load cases, damage conditions and sample points, as was carried out for the ABC process. This yielded error estimates (for the sample cases plotted in Figures 6.10, 6.11 and 6.12) which are given in Table 6.8. These results indicate that Model 3 outperformed the other two models, with Model 1 performing the worst, as was indicated in Figure 6.9. The error was reduced in the ABC process at all test points, except for the vertical strut with Model 3, where a slight increase was observed in the error; this increase was low and was likely due to Model 3 being initially accurate under its prior parameters – in addition to this, the error function did not reward the models for capturing uncertainty and would return a high error in the case of highly uncertain predictions (see Equation 6.4).

This seems to indicate validity in the ABC process, since the improved prediction accuracy through calibration was replicable in the test data, and the models are clearly accurate at the substructure level. As such, the calibrated models will be referred to following these comparisons as having been 'validated'; this does not reflect realistic scenarios where quantitative validation criteria would be defined prior to testing, but instead reflects that the models have been shown to have a reasonable level of accuracy and that the calibrated predictions were effective when compared against independent validation data.

Based on the calibration and validation tasks carried out at the substructure level in this section, Model 3 would be selected for use at the assembly level. However, all

three validated models were tested at the assembly level in order to investigate if the conclusions drawn from testing at the substructure level would prove accurate at the assembly level.

This validation can be considered an initial success, as it avoided the use of damage-state data at the assembly level, drawing these data from the struts in isolation instead (the benefits of this are discussed in Chapter 3). If this can be shown to be impactful on the model predictions at the assembly level, this will demonstrate the potential for hierarchical validation within an SHM context.

6.2 Uncertainty propagation and results

UP was carried out in conjunction with the model assembly process following the sub-model validation activities in Section 6.1 in order to allow assembly-level predictions to be made with the associated uncertainty that was quantified through validation.

The outcomes of the validation process were a set of posterior distributions for the parameters of the three strut-level damage models. These were incorporated into the assembly-level model of the bridge via dynamic substructuring directly. In addition to this, the uncertainty quantified on the material parameters of the struts in calibration was applied to the undamaged struts in the assembly by drawing samples from the material parameter distributions.

The validated model predictions had been shown to be accurate at the substructure level; testing of the model assembly using these validated submodels against assembly-level experimental data would give a strong indication of the legitimacy of the hierarchical V&V framework. If a demonstrable improvement in performance can be achieved at the assembly level, this will indicate that substructure-level validation is an effective way to validate assembly-level models in an SHM context.

6.2.1 Assembly methodology

The full assembly model was generated from a set of submodels via primal dynamic substructuring in the physical domain. This was carried out under the assumption of rigid joints between the substructures at the nearest node location in the submodels. The rigid-joint assumption is not accurate for this structure: the joints actually contained a set of non-rigid bracket components. These components were neglected from the model; their mass and stiffness contributions would lead to some model

discrepancy in the assembly model. However, this was assumed to be relatively insignificant to the model predictions. Other inaccuracies in the assembly include the geometric discrepancies caused by assuming that the struts were fixed directly at their extremities to the deck and upper frame.

These simplifying assumptions were made on the basis that a relatively simple substructuring regime could be implemented quickly and effectively for this thesis, allowing for discussions to be made on the overall framework. Many other methods for dynamic substructuring are available [66], meaning that assembly inaccuracies could potentially be reduced in future work by using more advanced techniques. What is key in this thesis is to demonstrate that the hierarchical V&V framework can be effective for SHM.

The inputs to the model were the damage extents, the strut at which the damage was located, and the location of the damage along the strut. For each set of inputs, a full set of struts were generated of the same length as the posterior distribution of the damage model parameters. The strut containing the damage was assigned the crack model parameters and material parameters from the posterior distribution, while the remaining struts were assigned material parameters sampled randomly from the distributions estimated during calibration using Latin hypercube sampling. These submodels were then coupled with the nominal submodels of the deck and upper Rexroth (which had not been calibrated or validated – another unquantified source of uncertainty) to create a set of assembly models of the same length as the damage model posterior. This allowed for a stochastic set of model predictions to be made for a given set of inputs,.

Two sets of assembly-level experimental data were used at this stage. The first was used for mode-matching between the model and the data; this process is summarised in Chapter 5. The second set of assembly data was taken across a range of damage states of the bridge. This allowed for the validated models to be tested against real-life data. The damage extents were introduced at intervals of 2.5mm from the undamaged state to a maximum of 17.5mm at the midpoints of the struts. The struts tested for damage were Struts 1, 5, 2 and 6, as shown in Figure 5.5. A full summary of the experimental data can also be found in Chapter 5.

The stochastic and nominal model predictions were generated at each of these test points for comparison against the experimental data. This would allow for any improvements in prediction accuracy through calibration to be observed, and would

allow for the level of uncertainty assigned to the model predictions to be assessed.

6.2.2 Results

The predictions across the damage extents are plotted against experimental data for the three validated models and their nominal equivalents in Figures 6.16, 6.17 and 6.18. The features of interest were identified in Chapter 5, with two features identified as being sensitive to damage in each strut. There are two immediate findings that can be taken from these plots. The first is that a significant leap in accuracy was *not* achieved in calibration, although some improvement is clear for certain cases (for example, the predictions for damage in Strut 5 using Model 2, or Strut 1 using Model 3). The second is that ABC allowed for good quantification of uncertainty in the assembly-level model predictions – this is informative when employing the model without assembly-level test data.

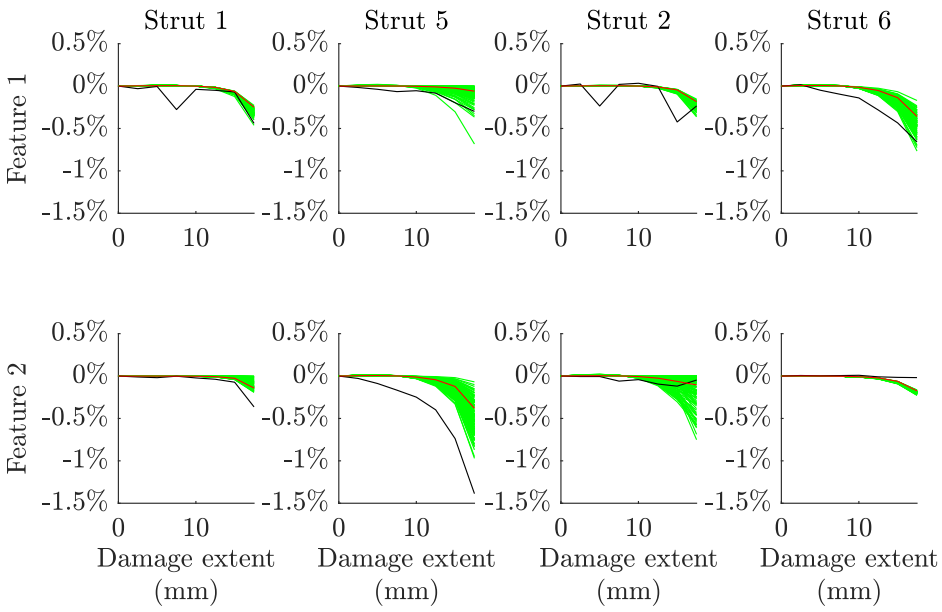


Figure 6.16: Predicted features at the assembly level compared with experimental data (black) for the nominal (red) and validated (green – full posterior) parameter sets for Model 1

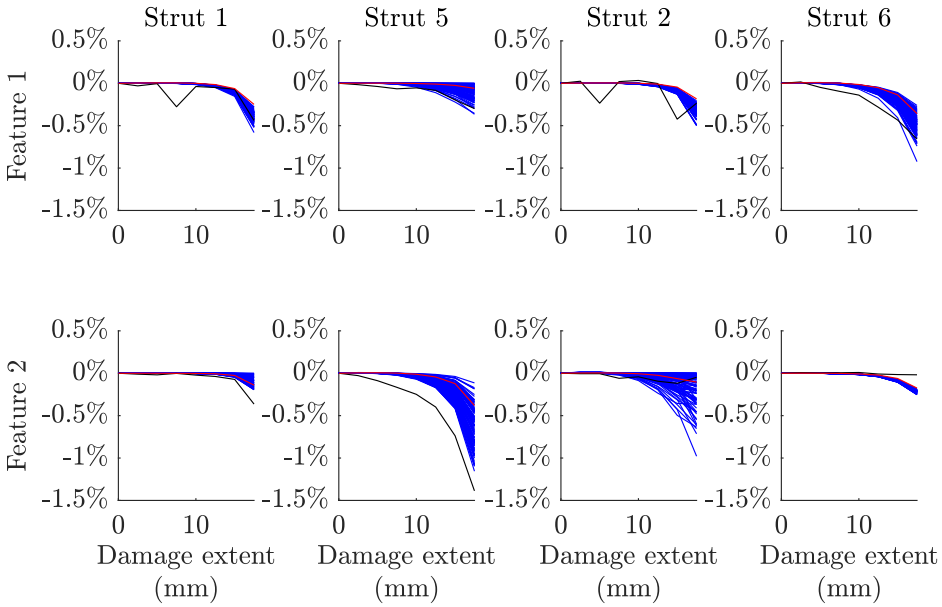


Figure 6.17: Predicted features at the assembly level compared with experimental data (black) for the nominal (red) and validated (blue – full posterior) parameter sets for Model 2

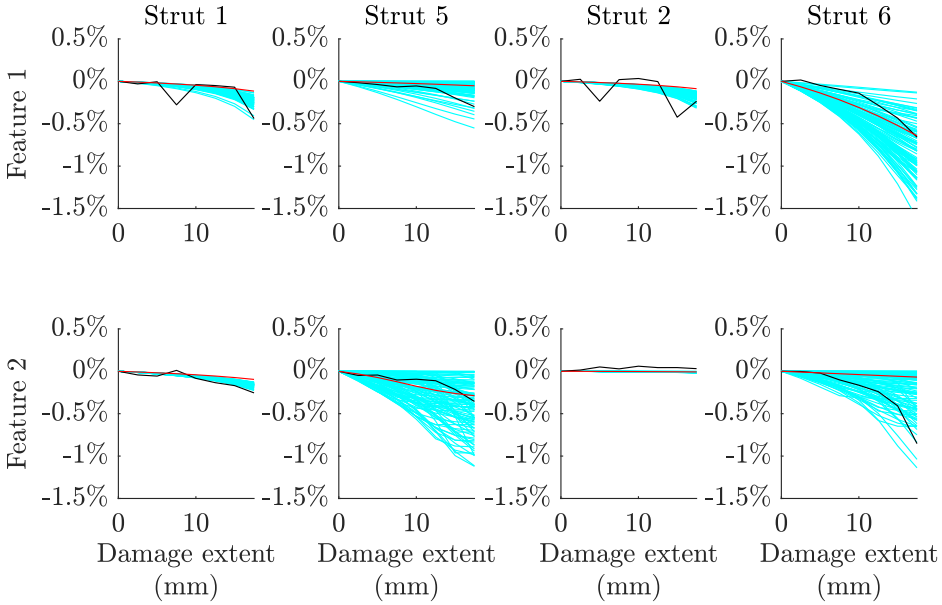


Figure 6.18: Predicted features at the assembly level compared with experimental data (black) for the nominal (red) and validated (cyan – full posterior) parameter sets for Model 3

As with the evaluation of the error in validation, the skewed nature of the posterior parameter distributions made effective error estimation for the uncertainty propagation results difficult. However, as at the validation stage, averages were taken of the errors between the model predictions and experimental data to attempt to discern the accuracy levels of the three models at the assembly level; these error results are summarised in Table 6.9. The first and last damage extents were excluded from this error estimate; the first (no damage) because the error was zero by definition at this point (the feature of interest being the change in natural frequency from its healthy value) and the last because this damage extent (17.5mm) was outside the range of the validation data.

Model / **Strut**	**Nominal**		**Validated**		
	1/2	**3**	**1**	**2**	**3**
1	0.0355%	0.0347%	0.0372%	0.0379%	0.0393%
5	0.0949%	0.0684%	0.0918%	0.0845%	0.0690%
2	0.0518%	0.0575%	0.0525%	0.0521%	0.0593%
6	0.1000%	0.0769%	0.0987%	0.0956%	0.0835%
Average	*0.0705%*	*0.0594%*	*0.0701%*	*0.0675%*	*0.0628%*

Table 6.9: The error between the predicted features and experimental features for the three assembly models before and after validation

Table 6.9 confirms, as was suggested at the strut validation stage, that Model 3 is the most accurate of the crack models investigated, followed by Model 2. As with the substructure-level validation process, the observed increases in prediction accuracy here compared to the prior models are low. This confirms the predictions made at the substructure level, and lends credibility to the hypothesis that validation and model selection activities carried out on assembly submodels can be used to confer validity of the assembly-level model.

However, it should be noted that while Model 3 does have the lowest mean error, it is not significantly lower than Models 1 and 2, as it was at the submodel level. This indicates that the model may not generalise as well as Models 1 and 2, which makes sense, given that there is less physical basis for the design of the model; Models 1 and 2 use a reduction in element cross-section that is physically not dissimilar to a real crack, while Model 3 used a distribution of the stiffness of the beam to describe the damage scenario.

The average prediction error actually increased after validation for Model 3 at the assembly level. Given that the increase in error was small, this does not necessarily

invalidate the model (Figure 6.18 implies relatively good predictive performance), as the scalar estimate of the error does not reward the model for capturing high levels of uncertainty in its predictions. Given that the model performed well with its nominal settings, it is not surprising that the validated stochastic predictions actually yield an increased mean error value. The slight loss of prediction accuracy was also observed in validation of the ABC calibration for the models at the substructure level – it is to be expected that this effect would also be carried through to the assembly level.

6.3 Conclusions

This chapter presented an investigation into how an assembly-level model can be validated at the substructure level and then used to make predictions with confidence at the assembly level. The structure in question was a laboratory-scale bridge structure, where predictive damage submodels of the bridge struts were calibrated and validated against experimental data in isolation, before being reassembled into the structure using dynamic substructuring. This allowed for accurate predictions to be made across a range of damage conditions for the assembly-level model, with the level of uncertainty in the predictions also clear. Model selection carried out on the basis of predictive accuracy at the substructure level also proved successful, as the selected model (a stiffness distribution damage model) performed with the greatest accuracy at both the substructure and the assembly level.

Given the discrepancies that were noted in both the submodels and the assembly, the performance of the validated assembly level models is pleasing. A more detailed model design (for example, a model that includes the joint components) would be expected to improve further on these results. While any model discrepancy added by neglecting the joint components (and other simplifying assumptions) from the assembly did not appear to add significant inaccuracies to the assembly-level predictions, a full evaluation of this in future work would be very informative. In addition, a more rigorous calibration and validation process at the substructure level – with a bespoke dataset for material parameter calibration and larger dataset for damage model validation – would allow for more accurate model development.

It should be noted that quantitative evaluation of the error in the predicted posterior distributions was difficult, as the posterior distributions were highly skewed. The estimates of error in this chapter require further testing in order to truly validate the conclusions drawn. Therefore, further examination of the three damage models will

be carried out in Chapter 7, where the models will be compared on their accuracy when applied to SHM tasks, such as damage detection. Since model 3 was determined to be the most accurate of the three damage models in this chapter, this would be expected to be the most successful when applied in an SHM context. Evaluation of the model performance on SHM testing criteria will give further information on the success of the hierarchical V&V process carried out in this chapter.

Chapter 7

Demonstration of a hierarchical model on a structural health monitoring problem

This chapter presents an investigation into the application of FMD-SHM techniques based on the bridge structure and models developed in the previous chapters. The aims of the investigation are to demonstrate the viability of model-driven SHM techniques based on hierarchical validation, and to identify key areas for improvement and further investigation in future research.

Two SHM tasks are approached in the following research: damage detection and damage assessment. Damage localisation is not covered in this chapter due to the experimental data and the feature sets identified not being suitable to this task. However, based on the validation of the models carried out in Chapter 6, the models could feasibly be adapted to produce a new feature set (such as mode shapes) which would be more suitable to damage localisation.

Machine learning techniques for classification and regression will be introduced in order to be applied to the FMD-SHM tasks. SVMs are applied to a damage detection problem and Gaussian processes are applied to damage assessment. These algorithms are trained using FMD-SHM techniques, and also using more traditional data-driven techniques for comparison. The specific choices of machine learning algorithms in this chapter are less significant than the relative performance of the models trained by each method; the key test is to determine whether SHM systems trained through FMD-SHM and hierarchical V&V methods are competitive with more traditional data-driven methods.

7.1 Damage detection

Damage detection refers to the task of flagging data from a structure or system when they meet certain criteria as being indicative of the presence of damage in the structure. In SHM this is generally carried out through continuous monitoring of a pre-selected set of damage-sensitive features, with the onset of damage being indicated by the passing of a given threshold on those features. This is effectively a binary classification problem and is well suited to treatment using statistical classification models.

In FMD-SHM, a physics-based model is used to generate a set of training data for a statistical classifier; this classifier would then be able to label incoming data from the real structure as being from the structure in its healthy condition or in any of its damaged states. A simple strategy for FMD-SHM would proceed as follows:

1. Use the model to make system-level predictions across health states of interest with any associated uncertainty.
2. Define an acceptable global damage threshold.
3. Label the predicted data as belonging to either the undamaged class or damaged class.
4. Train a classifier using predicted training data and apply it to predict the global health state of the structure when presented with new test data. Options include:
    - Building a non-probabilistic classifier: train a classifier using labelled training data; when test data are presented, report the predicted class label.
    - Using a probabilistic approach: learn the distribution of the training data; when test data are presented, report the probability of class membership.

While the above is conceptually simple, challenges nonetheless arise and a number of decisions must be taken by the system designer. The setting of a damage threshold (which operates as the decision boundary for the classifier) on acceptable damage is key among these. The compromise here is between maximising the *sensitivity* (true positive rate) to damage without compromising the *specificity* (true negative rate). This must be based on rigorous risk analysis and safety evaluation in industrial applications.

In principle, the FMD-SHM damage detection paradigm described above could be extended to both damage localisation and identification of damage extent by expanding the set of health state labels that are considered. These tasks could be carried out using discrete health state classification, or regression methods could be used to arrive at a continuous valued prediction for both tasks – the choice between these methods would depend on the application.

Classification refers to the task of assigning a label to a particular datapoint, which signifies membership of a class (such as the undamaged or damaged classes in damage detection); in the field of SHM this usually means to take a data feature, such as a natural frequency, and to label it as being recorded from a structure in a particular health state. The algorithms that carry out this task are called classifiers, and can be trained using machine learning techniques.

Machine learning techniques can be separated into supervised and unsupervised learning processes. Classifiers that are trained by an unsupervised process learn the class of a particular set of data; when exposed to fresh data, they should then be able to distinguish whether or not it came from the original class of data. The training data do not have to be labelled, or they can be labelled as being from one class only – in SHM applications, this is usually the undamaged health state. Unsupervised classification models are also known as novelty detectors.

Classifiers trained by a supervised process are exposed to labelled data from at least one class; they should then be able to assign fresh data to each of the classes they have learned. For damage detection, two classes are defined on either side of the specified damage threshold; for damage localisation or assessment, the label space is generally more complex and will usually contain more than two classes.

In SHM, classifiers can be used for damage detection [123, 124], damage localisation [123, 124] and damage assessment [124]. Damage detectors can be trained using either supervised or unsupervised learning processes, while classifiers for localisation or assessment require a supervised learning process (where the labels would be attached to potential damage locations or extents respectively). The output of labels based on features within recorded data, in an SHM context, allows engineers to make informed decisions on the management of a structure – for example, to continue operations, to schedule an inspection or to take the structure out of service. These decisions are often enhanced by developing processes to analyse the risk attached to various

classifier outputs; this can be done by using various cost functions bespoke to the application and often relies on a probabilistic classification scheme.

Two key types of classifier models are discriminative and generative classifiers. Discriminative models aim to use the training data to learn the boundaries between sets of labelled data, and are therefore often suited to supervised learning methods [125]. Generative classifiers, on the other hand, attempt to infer the probability distributions that the datasets are drawn from; they are therefore often applied to unsupervised learning problems [125].

A further important pair of classifier types are probabilistic and deterministic classifiers [126]. Deterministic classifiers refer to methods that will always label a particular datapoint with a particular class, while probabilistic classifiers will assign a particular datapoint a probability of membership to each available class. The deterministic classifier is therefore often quicker and easier to use, but probabilistic classifiers can be more informative, particularly in a risk-based context.

Several machine learning-based algorithms exist which are appropriate for classification. These include the nearest neighbour algorithm, decision trees, support vector machines and Bayesian classification (among many others). SVMs were selected for use in this research due to their efficiency in learning quickly from large datasets. An additional benefit of SVMs was that through their kernelised design they can be very flexible to the application and are appropriate for both novelty detection and binary classification tasks. SVMs were demonstrated in an FMD-SHM study in [109], and have been applied to various SHM tasks throughout the literature.

SVMs are powerful tools for binary classification. They are based on a discriminative method and are non-probabilistic by nature. They can be used to produce a probabilistic output using the Platt method [127]; this proceeds by fitting a sigmoidal probability function to the scores returned by the classifier rather than simply classifying them by some particular threshold. They can also be extended to multi-class classification tasks by using multiple binary classifiers to separate different regions of the feature space.

SVMs function by attempting to maximise the margin between two classes of data, where the margin is the perpendicular distance from the decision boundary to the nearest datapoints (support vectors) on each side – see Figure 7.1 for a visual representation of this. The form that the decision boundary takes is governed by the kernel; widely used options include linear, polynomial and radial basis function (RBF)

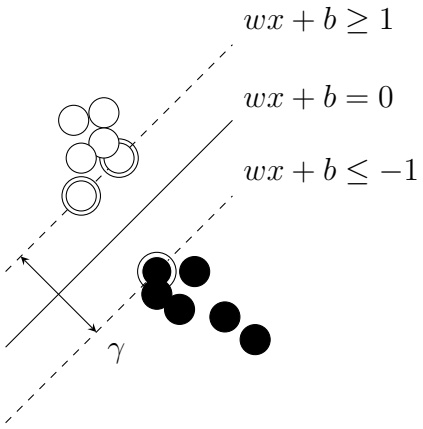


Figure 7.1: Graphical representation of a generic linear support vector machine with labels for the margin γ, data x, and kernel parameters w and b

kernels. The choice of kernel dictates the parameters that are adjusted to maximise the margin; for example for a linear kernel, the decision function would take the following form:

$$t_{new} = \text{sign}\left(w^T x_{new} + b\right) \tag{7.1}$$

In this example, the (binary) classes are denoted by -1 and 1, and the parameters to be learned from the data are w and b. The decision boundary for SVMs is set orthogonal to the margin, γ, which is designed to maximise the perpendicular distance between data from each class. The learning process for an SVM is therefore based on learning the parameters that maximise the margin. The support vectors are a subset of datapoints nearest to the decision boundary that are used to compute and maximise the margin. This makes training SVMs very effective for large datasets, as only a relatively small amount of training data from the full set is required. The margin between two support vectors from different classes, x_1 and x_2, can be derived according to Equation 7.2.

$$2\gamma = \frac{1}{\|w\|} w^T \left(x_1 - x_2\right) \tag{7.2}$$

Given that Equation 7.1 yields a binary output, the argument is invariant to linear scaling and can be fixed such that the output is ± 1 for the two support vectors x_1 and x_2. This allows Equation 7.2 to be simplified to the following:

$$\gamma = \frac{1}{\|w\|} \tag{7.3}$$

Equation 7.3 gives us a simple set of parameters to learn in order to maximise the margin for a linear SVM. Given that any new training datapoints must be classified as greater than 1 or less than -1, the following constrained optimisation problem can be found:

$$\begin{gathered} \underset{w}{\operatorname{argmin}} \frac{1}{2}\|w\|^2 \\ \text{subject to } t_n\left(w^T x_n + b\right) \geq 1 \end{gathered} \tag{7.4}$$

The clear issue with basic SVMs is the identification of a suitable set of support vectors when the training data overlap, which can lead to violation of the constraint condition in Equation 7.4. In addition, where a 'hard' margin is enforced (which means no data are allowed to overlap the decision boundary), the SVM can be vulnerable to overfitting to a few support vectors. This is avoided by applying a 'soft' margin, which allows for an error term for each support vector. This term quantifies the distance between the margin and the support vector for n support vectors, allowing them to sit within the margin or on the wrong side of the decision boundary. This error term is included in the constraint equation for maximising the margin, as shown in Equation 7.5.

$$t_n\left(w^T x_n + b\right) \geq 1 - \xi_n \tag{7.5}$$

The contribution of the error term for each support vector to the optimisation function is controlled by the multiplier C, which forms a key hyperparameter when training SVMs. The new constraint equation for linear SVMs with soft margins is given as follows:

$$\begin{gathered} \underset{w}{\operatorname{argmin}} \quad \frac{1}{2} w^T w + C \sum_{n=1}^{N} \xi_n \\ \text{subject to } \xi_n \geq 0 \text{ and } t_n\left(w^T x_n + b\right) \geq 1 - \xi_n \end{gathered} \tag{7.6}$$

SVMs can be trained using model predictions for FMD-SHM. The ability to generate labelled predictions of the features from anywhere in the damage space is a major advantage of FMD-SHM and can be used to tailor the learning process for the classifier

(subject to the model predictions offering an adequate fit to real experimental data). As an example of how classifiers can be tailored to data, for the damage detection case it may be reasonable to assume that the normal condition data can be approximated by a Gaussian distribution. This assumption can be tested using stochastic model predictions of the undamaged case, with fit to a Gaussian distribution assessed using an appropriate metric (such as Kullback–Liebler divergence). If the Gaussian assumption on normal state data can thus be satisfied, it would be reasonable to adopt an RBF (alternatively referred to as a 'Gaussian') kernel to distinguish damaged- and undamaged-state data at the detection stage. In all cases cross-validation is required to set the hyperparameters of the classifier. This allows for an optimised classifier to be trained, and also gives an indication of the expected performance of the classifier when exposed to new data.

A number of applications of SVMs to SHM problems can be found in the literature [4, 109]. A real-world example concerns the health monitoring of harbour infrastructure using SVMs [128].

7.1.1 Damage detection methodology

The training data generated using the physics-based models in this study matched the test points from the assembly-level damage testing described in Chapter 5. This entailed predictions of the bridge in its healthy condition and with damage introduced to Struts 1, 5, 2 and 6 (Figure 5.5), where damage was introduced to the midpoint of each strut in increments of 2.5mm up to a maximum extent of 17.5mm. For each damage case, the posterior predictions were evaluated for each damage model, as was demonstrated in Chapter 6. As discussed previously in Chapter 5, the struts used in this dataset – 1, 5, 2 and 6 – represent the four structurally 'unique' struts in the assembly, as the bridge can be divided into four symmetrically identical quadrants along the xy- and yz-planes (see Figure 5.5). This means that a damage detector trained on this dataset should generalise to detecting damage in any of the struts of the bridge.

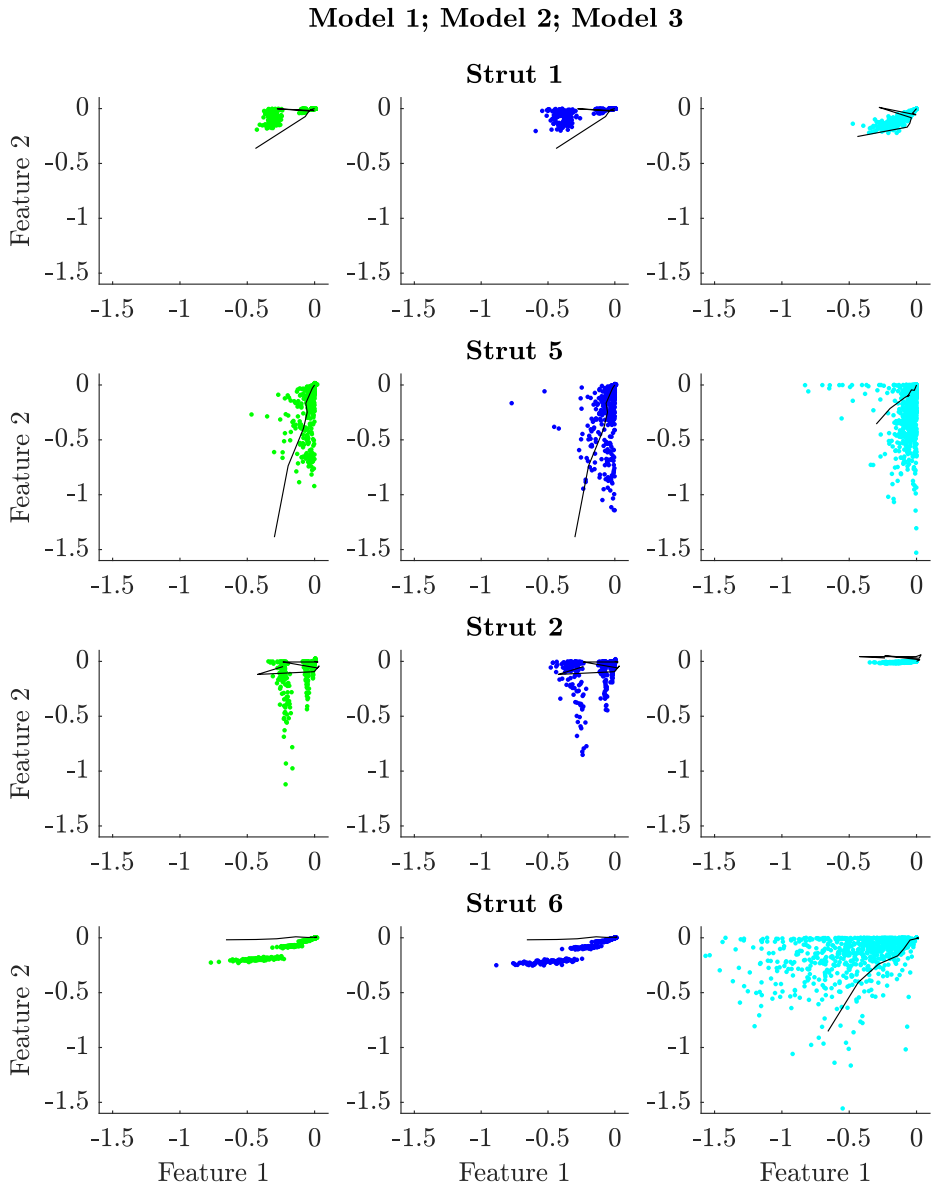


Figure 7.2: The stochastic model-generated training data for progressive damage in each strut using pre-validation selected features for damage detection with experimental data marked in black

The initial feature set was selected in Chapter 5 (the process followed in this chapter was repeated for the nominal version of Model 3); these features are plotted in Figure 7.2, which shows the domains on which the SVM damage detectors could be trained.

As damage progresses, the input data would be expected to move across the feature space from [0, 0] towards the bottom left of the plots; this can be observed for both the experimental data and model predictions. Using the training data shown in Figure 7.2, a classification boundary would be learned for each strut based on an acceptable damage threshold. As the data were shown to be relatively well described by the model predictions in Chapter 6, these classifiers would be expected to perform with reasonable accuracy when tested on the experimental data. However, based on the observable accuracy of the model predictions in Figure 7.2, some classification problems are to be expected: the fit to experimental data is relatively poor for Models 1 and 2 when predicting Struts 5 and 6. In addition to this, the trends of the experimental data are not always smooth increases in the magnitude of the feature values as damage progresses – this can be seen in the experimental data for Strut 1 and Strut 2.

The global damage detector for the bridge was constructed by training local SVM damage detectors for each strut based on each set of generated training data. The four SVMs were then individually exposed to the test data and the four sets of classification scores were summed to derive a global set of classification scores. This global set of scores was then compared to a decision boundary in order to label to the test data as either healthy or damaged. The global damage detector is visualised in Figure 7.3. The strategy was based on the technique employed in [109], which was successfully applied to a similar dataset.

The SVMs were trained and optimised using the MATLAB function 'fitcsvm'. The hyperparameter optimisation process used cross-validation minimisation to select the optimal kernel, kernel scale and C value for each SVM.

In order to evaluate the success of the FMD-SHM method on this dataset, two more traditional data-driven methods were developed for comparison. The first of these was a novelty detector, which used a one-class SVM for each strut. These SVMs used an RBF kernel and C value of 1 by default. The training data for the novelty detector were drawn from the assembly-level dataset, utilising the undamaged-state datapoints drawn from the damage-state testing and boundary condition testing (see Chapter 5). The novelty detector was therefore not dependent on damage-state data from the assembly, but would be limited to the damage detection task only, while the model-based strategy could feasibly be extended to other SHM tasks. The detector relied on the physics-based model for feature selection. Setting of the decision boundary

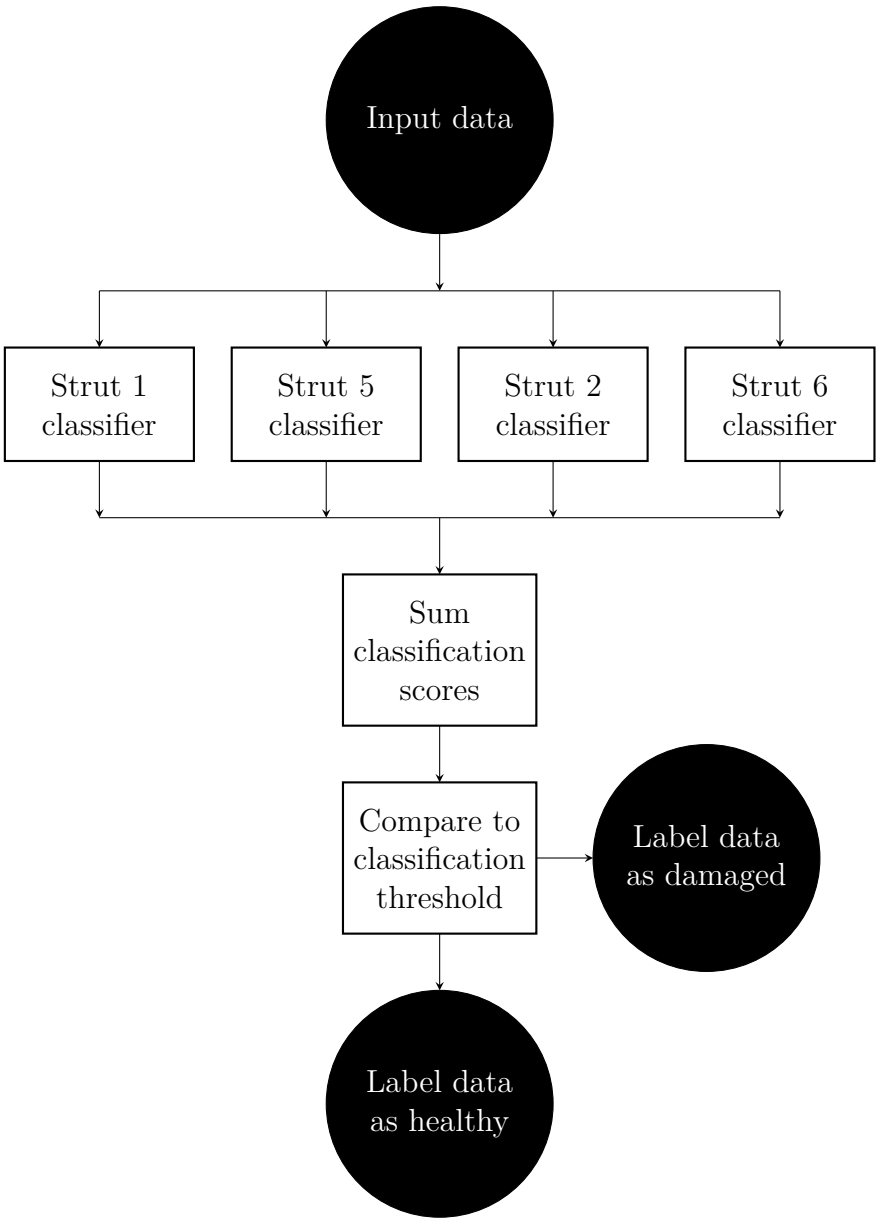


Figure 7.3: Illustrative structure of the global classifier designed for this case study constructed from four local strut classifiers

would also be very difficult in industrial applications if models or assembly-level test data were not available.

A supervised data-driven damage detector was also trained using the damage-state test data. The SVMs were trained on the same dataset as the novelty detector, with the addition of the damage-state data at the extents of 5mm, 10mm and 15mm. These SVMs were optimised similarly to those used for the FMD-SHM method. This detector was clearly dependent on assembly-level data across damage states – a major drawback for many SHM strategies. For both the novelty detector and the supervised damage detector, the individual strut SVMs were combined to create a global damage detector in the same way as the FMD-SHM damage detector (Figure 7.3).

7.1.2 Results

The test dataset used in this chapter was the damage-state data recorded from the full bridge assembly. This dataset covered the dynamic response of the bridge across a range of damage scenarios, where single-site damage was introduced to Struts 1, 5, 2 and 6 as a saw cut to the midpoint of the strut. Damage was introduced at increments of 2.5mm up to a maximum extent of 17.5mm. This dataset is described in full in Chapter 5.

Classifiers are generally assessed on their performance in terms of accuracy metrics, such as sensitivity – also known as the true positive rate (TPR) – and specificity. Metrics such as these aim to capture the ability of the classifier to correctly label datapoints as being from the class that they are drawn from in reality. The TPR (Equation 7.7) and false positive rate (FPR) (Equation 7.8) for classification are dependent on the classification threshold, where increasing the threshold will increase the rate of true positive labels, but simultaneously increase the FPR. A 'good' classifier will maximise the ratio between the TPR and the FPR, and the setting of the classification threshold is clearly key to this.

$$\text{TPR} = \frac{\text{True positive labels}}{\text{True positive labels} + \text{False negative labels}} = 1 - \text{FNR} \tag{7.7}$$

$$\text{FPR} = \frac{\text{False positive labels}}{\text{True negative labels} + \text{False positive labels}} = 1 - \text{TNR} \tag{7.8}$$

One way to rigorously evaluate the performance of a classifier, given a particular set of test data, is through the receiver operating charactistic (ROC); this covers the TPR

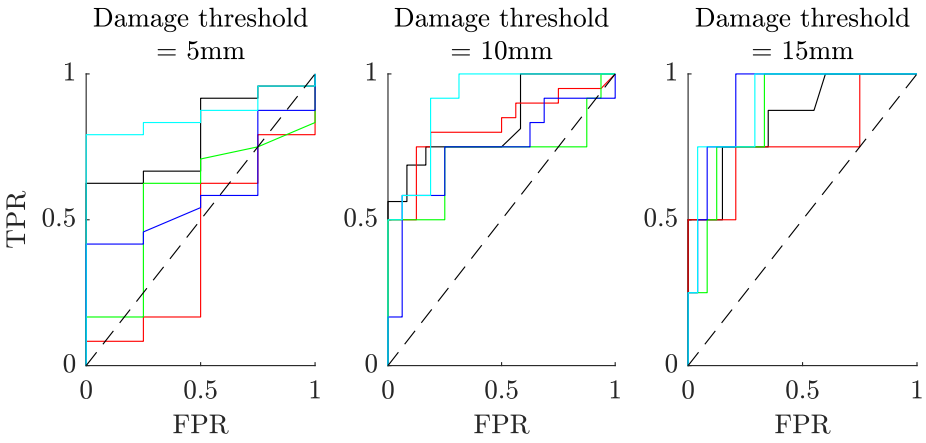


Figure 7.4: ROC curves for the novelty detector (black), supervised detector (red), Model 1 detector (green), Model 2 detector (blue) and Model 3 detector (cyan), with the boundary for informative classification marked with a dashed line

and the FPR over a range of classification thresholds. A 'good' classifier would have a ROC curve as far above the unit gradient as possible (TPR = FPR), meaning that a high TPR can be achieved over the full range of classification thresholds without compromising the FPR. The ROC curves are plotted for the classifiers generated in this paper in Figure 7.4. The damage thresholds indicated for each subfigure define any crack of depth equal to or greater than the threshold as damaged. This figure shows strong performance for the model-based methods compared to the data-based methods, with performance for all methods increasing as the damage threshold is increased; this would be expected because, as can be seen in Figure 7.2, the data become increasingly separable as the severity of damage increases.

The area under the ROC curve gives a useful scalar metric for a classifier, where the minimum value for an informative classifier would be 0.5 (this is the value for a classifier where the TPR is equal to the FPR at all classification thresholds); the maximum value for area under the ROC curve is 1, which would indicate perfect classification performance. These values are plotted against damage threshold for the classifiers in Figure 7.5. The figure gives a clearer indication of classification performance overall than Figure 7.4; it can be observed that the model-based methods perform strongly compared to the data-based methods, with all models outperforming the data at the highest damage threshold. Model 3 is the best-performing of the damage models, although Models 1 and 2 are just as effective at higher damage thresholds.

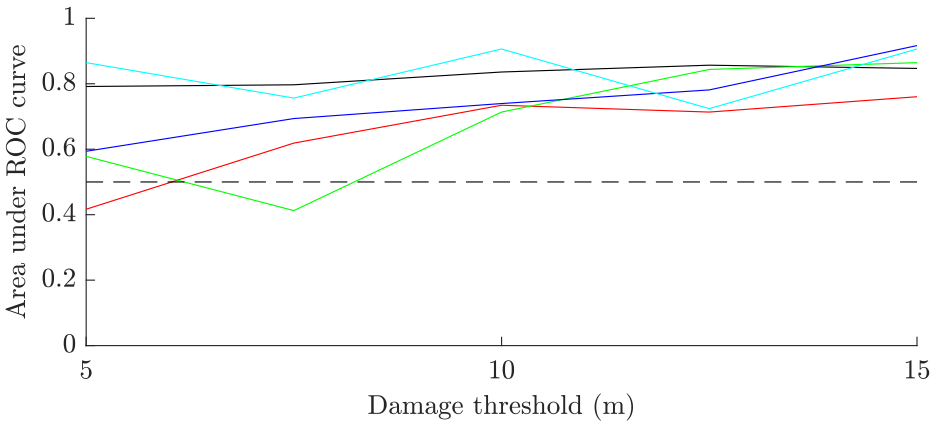


Figure 7.5: The area under the ROC curves at a range of damage thresholds for the novelty detector (black), supervised detector (red), Model 1 detector (green), Model 2 detector (blue) and Model 3 detector (cyan), with the boundary for informative classification marked with a dashed line

It is interesting to note that, based on the results presented, the novelty detector consistently outperformed the binary classifier trained using the damage-state data. This is an unexpected result, given that the supervised learning method was allowed more training data than the novelty detector. However, it should be remembered that the setting of an optimised damage threshold on the novelty detector would be very difficult in the absence of damage-state data; this task would be more feasible for the supervised learning method. In addition to this, only a small amount of experimental training data were available, so these results would be expected to be sensitive to erroneous datapoints that are affected by experimental uncertainty.

The initial conclusions that can be drawn from Figures 7.4 and 7.5 are promising. The FMD-SHM detectors perform reasonably well compared with the data-driven methods. However, it is possible that the FMD-SHM detectors could be further improved by selecting an updated feature set using the validated models. The initial feature set was selected using the un-validated, nominal condition models; given the increased accuracy and knowledge of the uncertainty in the model predictions, a more informative feature set may be found that would improve these results further.

Two new feature sets were selected using the validated models to investigate this hypothesis. Feature set 1 was selected prior to validation, and is described in Chapter 5. The matched features were selected based on damage sensitivity and the MAC.

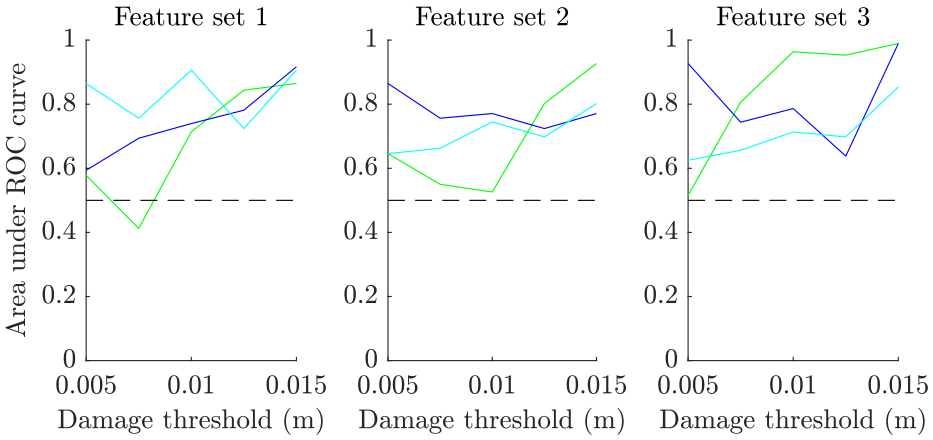


Figure 7.6: The area under the ROC curves at a range of damage thresholds for the Model 1 detector (green), Model 2 detector (blue) and Model 3 detector (cyan) using three features sets

The features were ordered in terms of damage sensitivity (the proportional change in natural frequency at the greatest damage extent) and the less sensitive half of the set was discarded. The two features were selected that maximised the MAC between the model predictions and the experimental data within the remaining features. This process was repeated using the validated models for feature set 2. 'Minimum sensitivity' was used as a measure of damage sensitivity; this was the proportional change in natural frequency at the greatest damage extent for the posterior sample that showed the least sensitivity to damage. Using the minimum sensitivity was intended to value robustness in the damage sensitivity of the features. For feature set 3, damage sensitivity was prioritised over the MAC, where the two features selected from the matched mode set were simply the ones that showed the greatest minimum sensitivity.

The classification results for the three feature sets are plotted across damage thresholds in Figure 7.6, where set 1 represents the original features selected before validation, set 2 was selected using the validated models prioritising prediction robustness and set 3 was selected using the validated models prioritising damage sensitivity. It can be seen that the reselected feature sets 2 and 3 offer a slight improvement in classification performance for Models 1 and 2. This was not the case for Model 3, for which the reselected feature sets proved less useful than the original features selected prior to validation. This is not necessarily very surprising given the strong performance of the original feature set. It is noticeable that the classification results are relatively

robust to feature selection, which offers some credibility to the overall model performance, as good accuracy can clearly be obtained for a large range of predictions. It is possible that in future research, model predictions could be generated to test various candidate feature sets, similarly to using model predictions to set an optimal decision boundary. This could potentially enable users to maximise the effectiveness of a set of validated models, again without the requirement for experimental damage state data.

7.1.3 Concluding remarks on damage detection task

The results of the damage detection task for this case study are very encouraging. The FMD-SHM method proved successful using the training data from the hierarchically-validated models, and performed very favourably when compared to traditional data-driven methods.

The data-driven methods were reliant on the model themselves, which was utilised to select the features on which they were trained. Furthermore, while the novelty detector performed particularly well, it should be noted that it would be very difficult to set an optimal classification boundary without further utilising the models.

The good performance of the FMD-SHM method was particularly pleasing given that it was achieved – through hierarchical validation – without the use of any damage-state assembly data (these data were required for the supervised learning data-driven model). Further credibility for hierarchical validation can be drawn from the fact that the best-performing damage model for the classification task was Model 3; this was identified at the strut level – see Chapter 6 – as the most accurate of the three models (followed by Model 2 and finally Model 1 – this prediction also proved to correspond to assembly-level classification performance).

The features selected prior to validation in Chapter 5 proved to be useful for damage detection after validation, which backs up the hypothesis that models can be used in a relatively coarse form for the feature selection task in SHM. As was suggested in the hierarchical validation framework, fresh features could be selected using the models after validation - these new feature sets improved the performance of classifiers trained using Model 1 and Model 2.

A further application of the validated models in classifier design would be to set an optimal classification threshold. Given the demonstrable accuracy of the models based on the hierarchical validation process, this could again be carried out without

the requirement for any assembly-level damage-state data. Instead, cross-validation of the model predictions could be used to determine an ideal classification boundary with confidence. This could be carried out based on a risk-based framework to design a suitable cost function to classification metrics such as the true positive rate and the false positive rate.

Additional extensions of the work presented here would be to use the stochastic predictions to train a probabilistic damage detector; this would be more informative in a risk-based context, as mentioned above.

7.2 Damage assessment

Damage assessment refers to the task of determining the severity of a type of damage. Assessment can be carried out as a classification task, where damage is assigned some discrete severity label, or by regression, where the severity is assigned a value from a continuous distribution. SHM techniques generally proceed by monitoring a data feature and assigning that feature to a class of damage extent (in the case of classification), or to a particular value describing damage extent such as crack depth (in the case of regression).

Rytter's Hierarchy progresses from damage detection to localisation prior to assessment [1]. Localisation and categorisation of damage is generally required before the severity of damage can be determined for a particular type of damage. The work presented in this chapter relied on the use of natural frequencies as features; due to their global nature, this presents significant difficulties when carrying out damage localisation. Localisation is particularly difficult for the truss bridge due to the lines of symmetry in the structure and the number of natural frequencies very close to each other in the frequency spectrum. Therefore, the study of damage assessment in this chapter is carried out based on the assumption that the location and type of damage is known.

A range of statistical techniques are available for regression tasks; these include linear and polynomial regression, decision trees and SVM techniques [125]. The regression model selected for this task was the Gaussian process. GPs are highly flexible, and could therefore be readily fitted to the form of the training data. Other methods were initially tested, such as linear regression and SVM regression; however, GP regression appeared to perform the most strongly at this stage. A similar use of GPs on bridge model predictions can be found in [113].

GPs function on the assumption that the underlying form of the predictions of a system based on its inputs can be represented by a collection of normal (Gaussian) distributions. They are flexible in form and, given sufficient training data, can be fitted to any given underlying function; this is dependent on the mean and covariance functions (m and k; Equation 7.9) of the GP, which operate similarly to the kernel for other machine learning algorithms. A key advantage of GPs is that they contain an inherent uncertainty function alongside their mean prediction for a given input; this allows for confidence to be attached to any models used for damage assessment. GPs are Bayesian functions, where the prior form (usually with a zero mean function) is updated according to training data in order to arrive at a posterior [129] – this reduces the potential for overfitting [130]. The key assumption in the definition of the covariance function is that it is smooth over the domain [129]; it should also be noted that the length scale of the covariance function is a major hyperparameter [129].

$$\begin{aligned} m(\mathrm{x}) &= \mathbb{E}[f(\mathrm{x})] \\ k(\mathrm{x}, \mathrm{x}') &= \mathbb{E}\left[(f(\mathrm{x}) - m(\mathrm{x}))\left(f(\mathrm{x}') - m(\mathrm{x}')\right)\right] \end{aligned} \tag{7.9}$$

The prior form of the GP is governed by the initial mean and covariance, with any associated hyperparameters, and is given in Equation 7.10 [10]. Optimisation and selection processes are then required to fit the prior model hyperparameters to given training data, and to choose appropriate mean and covariance functions. For a detailed explanation of the mathematical basis and learning processes for GPs, the reader is referred to [129].

$$f(\mathrm{x}) \sim \mathcal{GP}\left(m(\mathrm{x}), k(\mathrm{x}, \mathrm{x}')\right) = \mathcal{N}\left(m(\mathrm{x}), k(\mathrm{x}, \mathrm{x}')\right) \tag{7.10}$$

An example of a GP regression learning process is shown in Figure 7.7. This figure shows that, as more training data are added to the learning process, the regression model is able to more accurately estimate the form of the data, and uncertainty is reduced. It can also be seen that outside the range of training data ($x > 10$ in Figure 7.7), the model is unable to make accurate predictions; however, the increased uncertainty in this region gives a useful indication of this lack of accuracy.

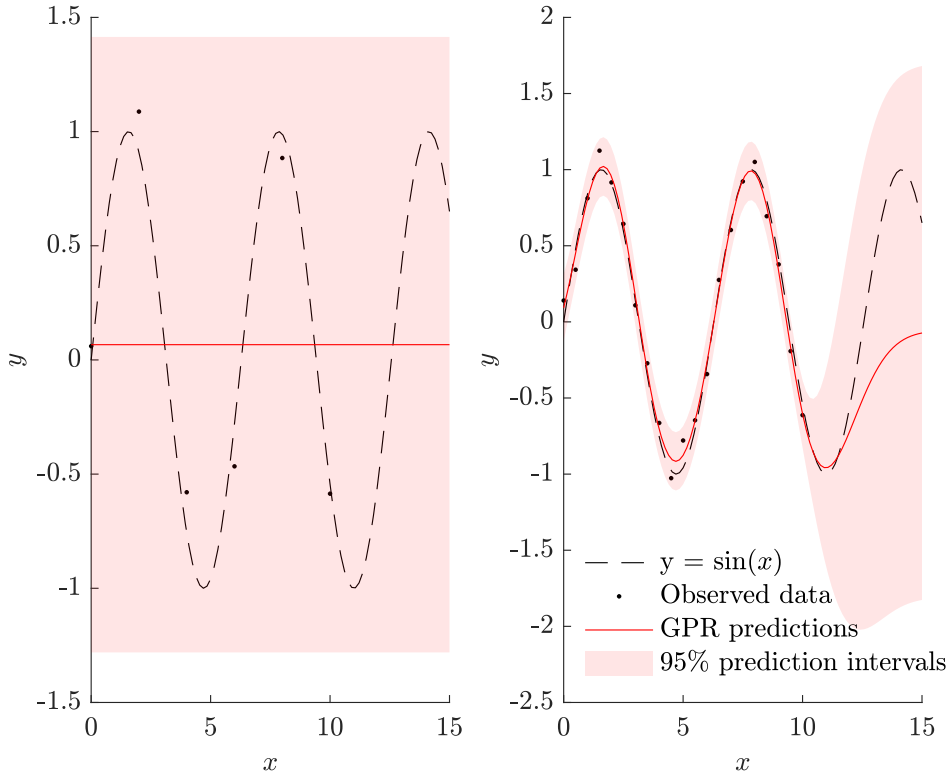


Figure 7.7: Examples of a Gaussian process regression model for $\sin(x)$ with more (right) and less (left) training data

GPs are a popular choice for modelling in SHM and structural dynamics contexts. They have been applied for use as emulators to save on computational effort in running complex models [10], and for model sensitivity analysis [96, 100, 101]. An application of GP models to damage location and assessment can be found in [131].

7.2.1 Damage assessment methodology

The validated models were tested on their ability to train a regression model to assign a value to the depth of a crack at the midpoint of each strut of the bridge. This was based on the assumption that the strut containing the crack and the location of the crack was known, and that the damage – more generally – was known to be an open edge crack, as explained in Chapter 5.

The GP regression models trained for this chapter were fitted using the MATLAB function 'fitrgp'. The optimisation algorithm for this function was quasi-Newton

optimisation. Bayesian optimisation was used to select the covariance function and the hyperparameters; this was carried out using the MATLAB function 'bayesopt' on its default settings.

The training data used for the damage assessment task were the same as those used for the detection task; feature set 1 is shown in Figure 7.2. The label set was expanded in the training data for this task; instead of the labels being damaged or healthy, as was the case for training damage detectors, the labels were the damage extents at which the training data were generated. Each of the three feature sets used for damage detection were applied to this task. This constitutes an expansion in the dimensionality of the learning space, but with no additional training data provided compared to the damage detection task. As such, a lower level of of accuracy should be expected compared to the damage detection results (this is in addition to the fact that damage assessment is inherently a more challenging task for statistical models).

7.2.2 Results

The results of the GP regression models are plotted in Figure 7.8. These show the predicted crack depths based on the feature inputs derived from the experimental data (the same test dataset that was used for the damage detection task). All three damage models were tested on each strut, and the three feature sets described earlier in this chapter were each tested.

It can be seen from Figure 7.8 that the success of the FMD-SHM method for damage assessment is mixed. The rate of damage progression is well captured by the statistical models in some cases, such as for Strut 1, but for others the accuracy of the regression models is poor. In a number of cases, the predicted damage extents are not physically possible (the predicted crack depth being either less than zero or greater than the depth of the strut). Model 2 appears to be the most robustly accurate across the test data, and it is also apparent that the models are able to more accurately predict crack depth at greater damage extents.

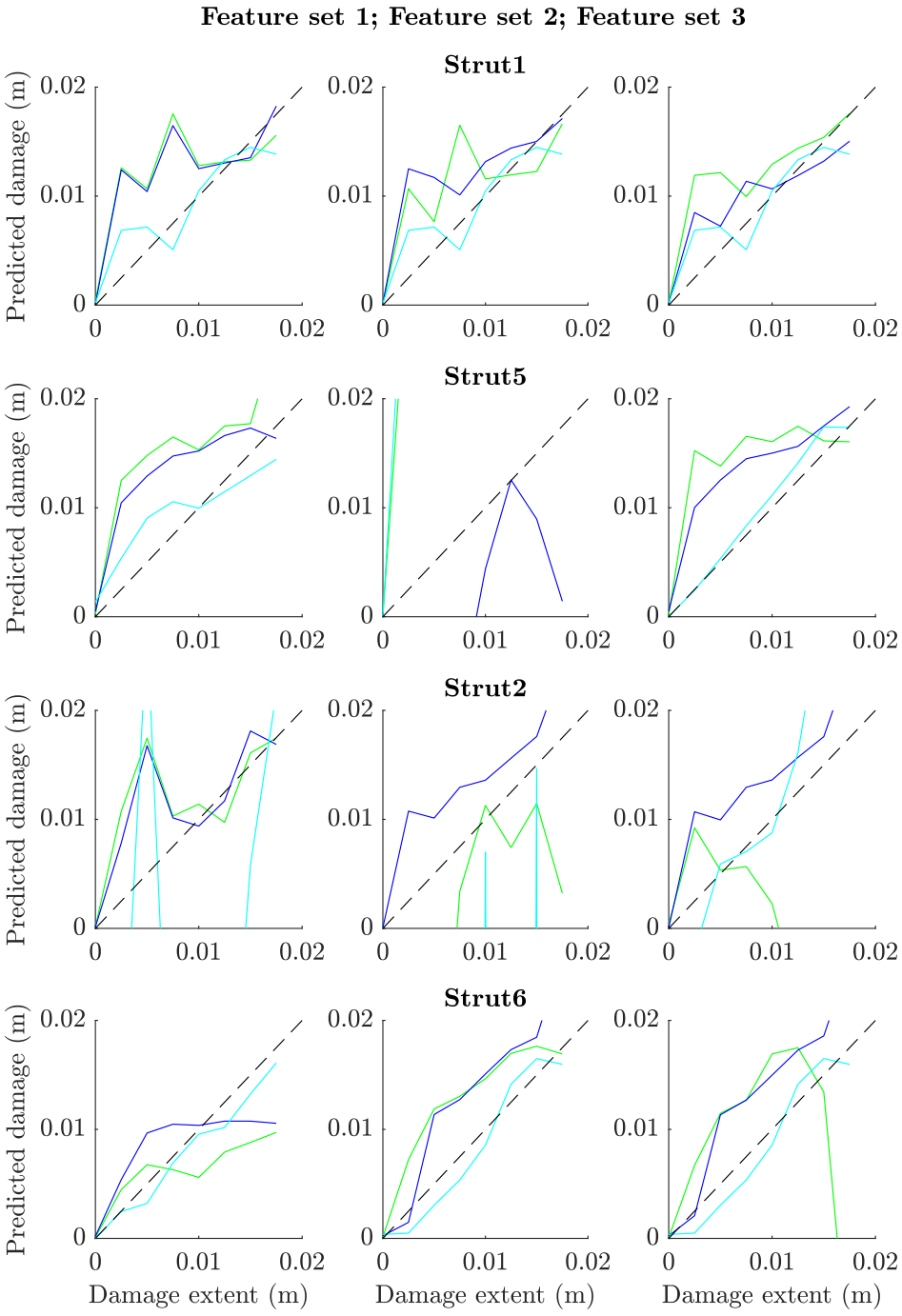


Figure 7.8: The predicted damage extent compared to the true value for Model 1 (green), Model 2 (blue) and Model 3 (cyan), with the ideal regression model marked as a dashed line

7.2.3 Concluding remarks on damage assessment task

The damage regression task was attempted using FMD-SHM techniques based on the three validated damage models developed in Chapter 6. GP regression models were trained using the model-generated datasets, and were successful in predicting the depth of cracks in the struts in most cases – given knowledge of the damage location. Poor predictive performance, such as the Model 3 regression model for damage in Strut 2 based on feature set 1, can be linked to poor feature matching and selection; Figure 7.2 shows that the damage sensitivity in this feature space was low and that the model predictions do not correlate well with the experimental data for feature 2. This indicates that the poor performance of certain sets of predictions is due to inadequate feature matching and selection, although model discrepancy could also cause inaccuracies in these cases.

Models 1 and 2 generally perform better than Model 3 for this task; this is likely due to the greater separability of their predictions making the task of learning a regression model much easier. As can be seen in Figure 7.2, there is significant overlap in the training data generated using Model 3.

Most of the results show that accurate prediction of damage extent is difficult when the crack length is small, but accuracy is much better at greater damage extents. This is to be expected, given that examination of the selected features has shown that lesser damage extents have a relatively small impact on the natural frequencies, while much greater sensitivity was observed at greater damage extents (see Figures 6.16, 6.17 and 6.18). This makes the data at low damage extents very sensitive to noise, and the stochastic model predictions have significant overlap, making separability of the data difficult, which in turn makes it difficult to confidently assign a particular estimate of crack depth.

The GP models could potentially be significantly improved by developing a more advanced probabilistic set of beliefs. It is known that the predicted outputs of the models cannot physically be less than 0mm or greater than 20mm; this could therefore be incorporated into the model prior to improve the hyperparameter learning process. This would, however, present its own challenges. The use of probabilistic predictions, rather than the simply using the mean function, would be informative in this case, as the uncertainty on GP predictions will naturally increase at regions of the problem domain away from any previously-seen training data (see Figure 7.7); this would give a useful indication of potential inaccuracies in these regions.

Further improvements to this investigation could also likely be made with more training data; if a test was focused on a single strut then the test data could be recorded with more sensor locations to allow for more accurate feature matching. This would allow for more confidence to be attributed to the model predictions, and it would be easier to determine the true sources of error in the model predictions (these could be due to model discrepancies or to mismatching of features). Similarly, generating more training data at a larger range of damage inputs would be expected to improve the performance of the trained statistical regression models.

7.3 Conclusions

This chapter was aimed to demonstrate the potential of FMD-SHM compared to traditional SHM techniques, and to use this demonstration to evaluate the success of the hierarchical validation process carried out in Chapter 6. This was carried out by setting up two SHM tasks: detection of damage in a truss bridge, and assessment of damage in the struts of the bridge given prior knowledge of the damage location.

The damage detectors were based on SVMs trained using the validated model predictions; these were compared to an SVM novelty detector and an SVM trained using labelled damage-state data drawn directly from the real structure. It should be noted here that, while SVMs were selected for this demonstration as a suitable classification algorithm, the key investigation was to compare the performance of the FMD-SHM method to data-driven SHM methods, so the actual choice of classification algorithm was not critical. The FMD-SHM method performed well compared to these benchmark methods, but did not rely on any experimental assembly-level damage-state data. It was also found that the model selection carried out in Chapter 6 was legitimate for the task, as the models performed as predicted – compared with each other – at the substructure level. This indicates that hierarchical validation is very appropriate for physics-based models in SHM, given the potential benefits outlined in Chapter 3. Furthermore, the feature set selected in Chapter 5 proved to work well for the assembly-level damage detection task. This indicates that feature selection at the pre-validation stage of the framework is an appropriate strategy, although further feature selection following validation can also be carried out where desired.

The task was then extended to damage assessment, where Gaussian process regression models were trained to predict the depth of a midpoint crack in Struts 1, 5, 2 and 6. This was a higher-level task than damage-detection, and the accuracy of the models

was lower as a result. Despite this, where the model predictions were demonstrably accurate, the regression models trained using these predictions were largely successful. This indicates that the FMD-SHM framework can be legitimately applied to more difficult SHM problems such as this, but that the model validation activities required for successful prediction must be more rigorous than for lesser tasks like damage detection.

Future work would include the application of the FMD-SHM framework to damage localisation. This was not carried out in this research due to the difficulty in identifying a suitable feature set. Mode shapes are generally recognised as a useful feature for damage localisation, and the model can be used to predict the mode shapes of the structure. If a set of experimental data containing detailed mode shapes of the bridge was acquired across a range of damage states, this could be used to test the FMD-SHM process for damage localisation.

A further extension to this study would be to carry out the classification and assessment tasks from a more probabilistic perspective; the validated models would be highly tractable to this challenge, as they provide a set of stochastic predictions for a given input. A probabilistic SHM study could then be incorporated into a risk-based framework, and further investigations into the setting of key modelling parameters – such as the acceptable damage threshold for damage detection – could be carried out.

Chapter 8

Discussions and conclusions

8.1 Discussions

This thesis presented investigations into a novel framework for hierarchical validation of physics-based models for SHM, with a focus on FMD-SHM. This was motivated by the lack of data problem common to many existing SHM methods. Data-driven SHM methods are reliant on training data to develop statistical damage models via machine learning. These training sets are often very difficult to acquire, as they usually require the collection of data from the structure of interest in its damaged states. FMD-SHM provides the means for these data to be provided through model simulations, but predictive damage models are themselves dependent on labelled damage-state data from the structure in order to carry out model validation.

Hierarchical V&V aims to reduce the burden of the data acquisition process for model validation by drawing the validation data from components or subassemblies of the larger structure, rather than from the full assembly. This would be expected to be a more feasible process, as it would not require the introduction of damage to high-value assembly-level structures. Instead, the submodels representing these subassemblies and components of the assembly could be validated separately over an appropriate range of damage and operational conditions. These validated models could then be reassembled to recover the assembly-level model, which would now be useable with quantifiable confidence in its predictions.

Two aims were set out for this thesis: firstly to present an initial framework for the implementation of hierarchical V&V in an SHM context, and secondly for this framework to be tested through realistic case studies – this was also intended to provide a further demonstration of the FMD-SHM method, which is still a relatively

novel process within SHM literature. The initial framework was presented in Chapter 3 (Figure 3.2).

The experimental dataset used for the case studies in this thesis was drawn from a laboratory-scale truss bridge (Figure 5.1). This was considered a realistic target structure given its similarity to larger bridges found in worldwide civil infrastructure. Validation data were acquired from the struts of the bridge (the load-bearing components) across a range of damage and static load conditions – these experiments were described in Chapter 6. Testing data for the SHM tasks (which were also used for matching the model predictions to features in the experimental data) were gathered from the full assembly in a range of damage states – this was described in Chapter 5.

One noteworthy comment to be made here is that one of the postulated benefits of hierarchical V&V was that data acquisition from component structures would be easier than from the full assembly. This is due to the increased simplicity of components compared with the larger assembly, their smaller size and the increased ease of sensor placement. While these proposals would still be expected to stand in many cases, difficulties were encountered in gathering vibration data from the struts in isolation which were not a problem when working with the larger assembly. This was because of the lightweight and highly flexible nature of the struts. Their flexibility made the use of an impact hammer to excite the struts very difficult, as avoiding double-impacts was virtually impossible. This meant a shaker was required to excite the struts via a stinger – again, since the struts were so flexible, this had a non-negligible stiffening effect on the struts, the extent of which was difficult to estimate (comparison to equivalent tests with impact hammer excitation may have helped with this). Because the struts were so light, the use of accelerometers to measure the vibration had an observable effect on the results due to the increased mass of the system. This problem was avoided by using a scanning laser vibrometer to capture the response. These issues were specific to the bridge system investigated, and therefore they would not be expected to be encountered on other structures or substructures. However, each structure would be expected to have its own idiosyncrasies, and many of these may present data acquisition problems. Careful design of experiments when capturing validation data from the substructures must take these potential issues into account.

The theory and techniques for hierarchical V&V were presented in Chapter 4. Dynamic substructuring was identified as a suitable methodological basis for the assembly and disassembly of models in an SHM context, and sampling methods were

identified as suitable for the propagation of uncertainty through these assembly processes where uncertainty can be quantified in the parameters of the submodels. Latin hypercube sampling was found to outperform Monte-Carlo sampling in terms of distribution convergence. These ideas were tested on a numerical study of a plate assembly.

Feature selection in a model hierarchy for FMD-SHM was carried out on the bridge dataset in Chapter 5. Here, the assembly model was used to select a set of damage-sensitive modes prior to validation having been matched to assembly-level data from the bridge in its healthy condition. These were then used to select a set of appropriate modes on which to carry out the strut-level validation task. The features selected in this chapter proved to be effective for the SHM tasks investigated in Chapter 7, despite being chosen using the un-validated model. Carrying out feature selection at this stage allowed for significant streamlining of the validation process, as a low-dimensional feature space appropriate for validation was able to be identified for the strut submodels.

Chapter 6 presented a case study on the UQ and UP processes for hierarchical V&V of the bridge. The material parameters were calibrated for the struts in their undamaged condition, and the damage model parameters were calibrated using ABC. Three predictive damage models were investigated, with Model 3 proving to be the most accurate of the three when validated against independent test data. In a real-life application, the other two models would be discarded in the model selection process; however, they were kept in this study in order to investigate the legitimacy of the model validation selection activities when applied to the assembly-level model. The validated strut models were then built back into the larger assembly models, which allowed for probabilistic predictions to be made for the bridge in a range of damage conditions. These were shown to be accurate when compared to assembly-level test data for each model and compared to the nominal, un-validated models. Model 3 also performed the best at the assembly level, followed by Models 2 and 1; given that this was predicted at the substructure level, this lends credibility to the hierarchical V&V process.

The accuracy of the models varied based on which feature was being predicted. This could be down to erroneous mode-matching, or the influence of model discrepancy. A philosophical question arises here as to the impact of model inaccuracies on certain predicted features, when other predicted features are accurate and could be selected for use in an SHM strategy. In light of the paradigm that all models are wrong, while some models are useful [50], it can be shown that even models that *can be*

useful may not be useful for *any* given prediction. The presence of certain erroneous predictions should highlight the risk inherent in using predictive models and highlights the requirement for rigorous validation.

Applications of the FMD-SHM method to damage detection and damage assessment were presented in Chapter 7. The validated models were used to train binary SVMs for global damage detection in the bridge; the classification models were then compared to data-driven methods in terms of performance on detecting damage in the assembly-level test data. The model-trained detectors performed well in comparison to the traditional data-driven methods, despite not making use of any assembly-level damage-state data in the development of the classifiers. Additionally, the models identified as performing best at the strut level (Model 3) proved to be the most accurate in the damage detection task, which underlines the credibility of the hierarchical V&V technique.

The damage assessment task proved more difficult when using the FMD-SHM method; GP regression models were trained using the validated models for this task. However, some success was observable in these results, and it is clear that a more refined study of this task could be successful in an FMD-SHM context. Investigations were carried out into the potential of selecting an updated feature set based on the calibrated models (the original feature set being selected using the models in their nominal settings). This had a mixed impact on the results: where some models were improved by using the new features (Models 1 and 2), this was not observed for Model 3, which performed similarly across feature sets for each SHM task.

Based on the investigations summarised above, it is clear that the FMD-SHM approach was broadly successful. This project was the first to present a large-scale study of the method on a realistic dataset, and the results of this showed good performance compared to a data-driven novelty detector and a supervised learning data-driven classifier for damage detection. The hierarchical V&V method was also successful when applied in an SHM context; improved performance was observable following the substructure-level validation tasks, and model selection at this level was shown to be a legitimate activity based on assembly-level testing. Feature selection was shown to be a reasonable task to be carried out prior to rigorous validation; however, it is clearly possible to ratify the outputs of this process once validation has been carried out.

Some limitations of the methods explored were also identifiable. The quantification of uncertainty added throughout the assembly process is an outstanding issue, and the level of accuracy required to achieve sound assessment of damage extent was not achieved to a level that would ensure trust in the method at this stage. In addition to this, the matching of model outputs to experimental data is still clearly limited by issues associated with data acquisition, particularly for structures such as the truss bridge where many modes exist with similar mode shapes and natural frequencies.

In light of these discussions, an updated version of the initial framework proposed in Chapter 3 has been developed. This is illustrated in Figure 8.1. The framework is still intended as a high-level guide to hierarchical V&V; however, it summarises the overall activities required in order to achieve the outcomes that were presented in this thesis.

The updated framework includes an additional input of assembly-level data to the feature selection task. This reflects the fact that selection of features must include some matching of predicted features to those that can be extracted from experimental data. These data do not need to be from the assembly in its damaged conditions, so they would be expected to be relatively straightforward to acquire in most contexts. The acquisition of substructure data for submodel calibration, validation and UQ requires a bespoke design of experiments to maximise the effectiveness of these processes (different datasets may be required for each stage of the overall process). This is not considered to be a part of the central framework, as it feeds into the substructure data acquisition bubble; however, any future research focusing on this element of the process may lead to arguments for its inclusion into the framework. High-quality experiments for the acquisition of model validation data are certainly key to successful validation activities. A more traditional process of parameter calibration through direct testing of material parameters would highlight the remaining inaccuracies as model discrepancy; this would then motivate further research into the propagation of this discrepancy and how it impacts on the application of the assembly-level model.

A choice has been added to the framework following the submodel validation stage, representing model selection based on some pre-defined accuracy or validation criteria. If the candidate submodel fails these criteria, then new submodels will be developed until the criteria are met and model assembly can be carried out. This process was carried out informally in Chapter 6, as Models 2 and 3 were developed following observed model discrepancy in the model predictions from Model 1. The setting of

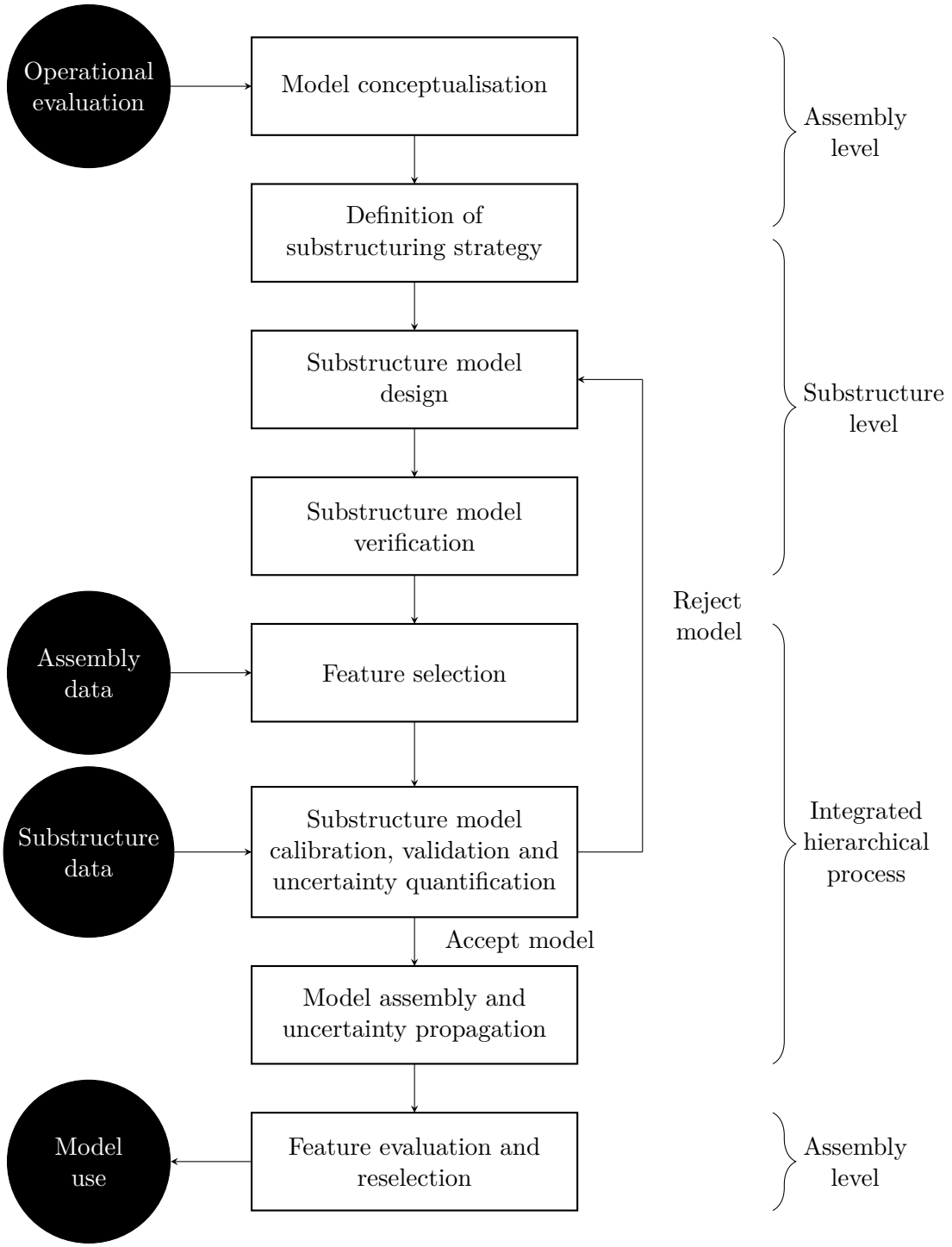


Figure 8.1: Updated framework for hierarchical verification and validation based on investigation in this thesis

validation criteria for this selection process presents an area for future work following this thesis.

An additional block has been added following model assembly and UP for evaluation of the features selected after the submodels have been selected and validated. This is to ensure that, having been calibrated, the features selected on which to operate the model are still the most effective for the given SHM application.

8.2 Future work

Certain areas for future research are identifiable based on the findings and limitations of the work presented in this thesis. The first among these concerns a potential additional block in Figure 8.1; this would constitute the use of the physics-based model to set decision criteria for the statistical models used in FMD-SHM. Based on the confidence attached to the model predictions, it would be possible to use these to evaluate the parameters for damage classifiers in a risk-based context. For example, given a maximum threshold for acceptable crack length in a damage detection problem (as was presented in Chapter 7), it could potentially be possible to identify the optimal position on the ROC curve from the model predictions. Testing of this idea on experimental data would be a natural continuation from the case study in this thesis. An additional set of future research into the required activities for 'operational evaluation' in the context of this framework would be useful – this would include significant decisions on the requirements for the SHM strategy, such as the setting of a maximum permissible crack length at which to define the structure as being damaged, which would have major ramifications into further activities such as feature selection in this framework.

Another clear continuation from the thesis would be to apply the FMD-SHM method to the problem of damage localisation. This was not carried out in this research because the models had been set up to generate natural frequency data, which has limited effectiveness when try to make inferences on local damage in the structure, as is required in damage localisation. However, it would require minimal adjustment to generate a set of mode shape predictions (mode shapes being generally accepted as useful features for damage localisation). Given an appropriately-designed set of test data, the study of damage localisation based on the FMD-SHM method using hierarchical V&V would be very informative. Similarly to the problem of localisation, the identification of multi-site damage was not addressed in this thesis. However, once

a validated assembly-level model for damage prediction has been developed through the hierarchical V&V framework, this could potentially be addressed by generating the required training data for a more complex FMD-SHM method. As with the localisation problem, the key foreseeable challenge here is in the ability to generate and detect features that robustly discriminate between damage in different locations, including combinations of those locations. A case study investigating these issues with a new experimental dataset would highlight any key issues with the process, but would be expected to be reasonably successful based on the findings of this thesis.

The application of hierarchical V&V to population-based SHM is an intriguing prospect. The basis for population-based SHM is that knowledge can be transferred across similar structures in a population in order to develop damage inference machines for each population member [17]. Given that structures in a population would be expected to share many components, the individual validation of submodels of these components across the range of damage and operational conditions could be very effective, as the submodels could then be integrated into a large range of nominally similarly assembly-level models representing structures across the population. Demonstrations of this could be carried out using the bridge structure used in this study by applying the validated strut models to different configurations of the assembly (i.e. an assembly without vertical struts, or with different boundary conditions).

The quantification of uncertainty in the assembly process is an outstanding issue in the hierarchical V&V process. However, it should be possible to address the question in future research without compromising the aims of the framework – to develop predictive damage models without the use of assembly-level damage data. Given that the damage models themselves (i.e. the stiffness distribution model or the element reduction models outlined in this thesis) are validated at the substructure level, these would be expected to be valid at the assembly level also. The uncertainty added in the assembly process would be expected to be independent of the health state of the assembly. This hypothesis could be tested by carrying out a validation of the model assembly in its healthy condition only in order to quantify the model discrepancy added in the assembly process – this could then be incorporated into the assembly model as a discrepancy function. The model would then be tested across its damage states to ensure that the discrepancy did not change as damage increased. A study such as this would add further credibility to the hierarchical V&V method, and would motivate a further update to the framework, which would then incorporate the assessment of model discrepancy after model assembly.

The modelling of joints is closely related to the ideas presented above. Joints were neglected from this study, but are known to have contributed to dynamic effects in assemblies. An investigation into the modelling of joints in assemblies, as well as tests of more advanced dynamic substructuring techniques (such as the use of non-rigid connections between substructures), in a hierarchical V&V context would yield further information as to the level of accuracy and confidence that can be achieved through hierarchical V&V.

Further to the points above, the issue of non-linearity in dynamic systems was not addressed in this thesis. This merits further investigation, since many common joint types exhibit non-negligible non-linearities, and incipient damage can also be expected to introduce non-linearity to a system. Investigations into the effects of this on execution of the hierarchical V&V framework may be required in future.

Finally, an additional problem in the implementation of SHM strategies in industry is the influence of confounding variables which are difficult to separate in the data from the effects of damage. A common variable here is temperature, which would have a significant stiffening or softening effect on many structures. In theory, FMD-SHM is a good method to use in tackling this issue, as physics-based models can be used to generate a wide range of training data across EOVs relatively cheaply, through the use of surrogate models if necessary. Alternatively, validated models would be very useful in identifying features that are robust to the effects of EOVs [113]. However, in either of these applications, it is clear that this would require model validation across these EOVs. A study to investigate the application of a hierarchical V&V strategy in the presence of additional varying factors such as temperature would be a significant contribution to the field.

8.3 Conclusions

Hierarchical V&V offers the means for assembly-level models of structures to be developed with the ability to make probabilistic predictions across damage-states for SHM, without the requirement for assembly-level damage data. This is a potentially very significant step forward in the application of physics-based models to SHM problems, where the acquisition of damage-state data for model validation has often proved to be a significant barrier to their realistic implementation.

It has been shown in this thesis that it is possible to make confident inferences on the health states of structures using models validated at subassembly levels, with

a case study being carried out that compared this technique to more established data-driven techniques. The FMD-SHM method was shown to be applicable to both damage detection and damage assessment tasks in a truss bridge, and significant extensions of these case studies have been highlighted for future research.

A framework for the application of hierarchical V&V was presented and investigated in an FMD-SHM context, where it was shown that use of the models involved could maximise the efficiency of processes such as feature selection and model validation. As confidence is established in the models moving through the hierarchy, the greater knowledge that is acquired concerning their accuracy and associated uncertainty can then be leveraged to optimise the performance of the models when applied to SHM tasks.

References

[1] A. Rytter. *Vibration Based Inspection of Civil Engineering Structures*. PhD thesis, Department of Building Technology and Structural Engineering, Aalborg University, Aalborg, Denmark, April 1993.

[2] C. R. Farrar and K. Worden. *Structural Health Monitoring: A Machine Learning Perspective*. Wiley, Chichester, UK, 1st edition, 2013.

[3] P. Cawley. Structural health monitoring: Closing the gap between research and industrial deployment. *Structural Health Monitoring*, 17(5):1225–1244, January 2018.

[4] R. J. Barthorpe. *On Model- and Data-Based Approaches to Structural Health Monitoring*. PhD thesis, University of Sheffield, Sheffield, UK, November 2010.

[5] P. Gardner, T. J. Rogers, C. Lord, and R. J. Barthorpe. Learning of model discrepancy for structural dynamics applications using Bayesian history matching. In *Journal of Physics: Conference Series. XIII International Conference on Recent Advances in Structural Dynamics (RASD 2019)*, Valpe, France, 2019.

[6] R. C. Smith. *Uncertainty Quantification: Theory, Implementation and Applications*. SIAM, Philadelphia, USA, 1st edition, 2014.

[7] American Society of Mechanical Engineers. Guide for verification and validation in computational solid mechanics, March 2006.

[8] M. Nishio, J. Marin, and Y. Fujino. Uncertainty quantification of the finite element model of existing bridges for dynamic analysis. *Journal of Civil Structural Health Monitoring*, 2(3–4):163–173, November 2012.

[9] P. J. Roache. *Fundamentals of Verification and Validation*. Hermosa Publishers, Socorro, USA, 2009.

[10] P. Gardner. *On Novel Approaches to Structural Health Monitoring*. PhD thesis, University of Sheffield, Sheffield, September 2018.

[11] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review. Technical Report LA-13070-MS, May 1996.

[12] H. Sohn, C. R. Farrar, F. M. Hemez, and J. J. Czarnecki. A review of structural health monitoring literature 1996–2001. In *Structural Health Monitoring*, 2002.

[13] A. C. Neves, K. Maes, I. Gonzalez Silva, and R. Karoumi. Validation of an artificial neural network data-based novelty detection method. *Structural Engineering International: Journal of the International Association for Bridge and Structural Engineering*, 2020.

[14] D. Sen, A. Aghazadeh, A. Mousavi, S. Nagarajaiah, R. Baraniuk, and A. Dabak. Data-driven semi-supervised and supervised learning algorithms for health monitoring of pipes. *Mechanical Systems and Signal Processing*, 131:524–537, 2019.

[15] M. Azimi, A. D. Eslamlou, and G. Pekcan. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 20(10), 2020.

[16] D. J. Pitchforth, T. J. Rogers, U. T. Tygesen, and E. J. Cross. Grey-box models for wave loading prediction. *Mechanical Systems and Signal Processing*, 159, 2021.

[17] K. Worden, L. A. Bull, P. Gardner, J. Gosliga, T. J. Rogers, E. J. Cross, E. Papatheou, W. Lin, and N. Dervilis. A brief introduction to recent developments in population-based structural health monitoring. *Frontiers in Built Environment*, 6, 2020.

[18] M. I. Friswell and J. E. Mottershead. *Finite Element Model Updating in Structural Dynamics*. Springer Science and Business Media, Dordrecht, Netherlands, 1st edition, 1995.

[19] M. I. Friswell. Damage identification using inverse methods. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 365(1851):393–410, December 2007.

[20] D. T. Bartilson, J. Jang, and A. W. Smyth. Sensitivity-based singular value decomposition parametrization and optimal regularization in finite element model updating. *Structural Control and Health Monitoring*, 27(6), 2020.

[21] B. Weber, P. Paultre, and J. Proulx. Consistent regularization of nonlinear model updating for damage identification. *Mechanical Systems and Signal Processing*, 23(6):1965–1985, 2009.

[22] B. Moaveni, J. P. Conté, and F. M. Hemez. Uncertainty and sensitivity analysis of damage identification results obtained using finite element model updating. *Computer-Aided Civil and Infrastructure Engineering*, 24(5):320–334, July 2009.

[23] A. Teughels and G. De Roeck. Structural damage identification of the highway bridge Z24 by FE model updating. *Journal of Sound and Vibration*, 278(3):589–610, 2004.

[24] F. Ponsi, E. Bassoli, and L. Vincenzi. Bayesian and deterministic surrogate-assisted approaches for model updating of historical masonry towers. *Journal of Civil Structural Health Monitoring*, 2022.

[25] P. E. Leser, J. D. Hochhalter, J. E. Warner, J. A. Newman, W. P. Leser, P. A. Wawrzynek, and F. Yuan. Probabilistic fatigue damage prognosis using surrogate models trained via three-dimensional finite element analysis. *Structural Health Monitoring*, 16(3):291–308, April 2017.

[26] P. Gardner and R. J. Barthorpe. On current trends in forward model-driven SHM. In *Structural Health Monitoring*, Stanford, USA, 2019.

[27] P. Casellini and G. M. Revel. An experimental technique for structural diagnostic based on laser vibrometry and neural networks. *Shock and Vibration*, 7(6):381–397, 2000.

[28] J. M. Ko, Z. G. Sun, and Y. Q. Ni. Multi-stage identification scheme for detecting damage in cable-stayed Kap Shui Mun Bridge. *Engineering Structures*, 24:857–868, 2002.

[29] J. M. Ko, Z. G. Sun, and Y. Q. Ni. Modal analysis of cable-stayed Kap Shui Mun Bridge taking cable local vibration into consideration. *Advances in structural dynamics*, 1:529–536, 2000.

[30] J. L. Zapico, M. P. Gonzalez, and K. Worden. Damage assessment using neural networks. *Mechanical Systems and Signal Processing*, 17(1):119–125, October 2003.

[31] M. P. Gonzalez and J. L. Zapico. Seismic damage identification in buildings using neural networks and modal data. *Computers and Structures*, 86(3–5):416–426, February 2008.

[32] J. J. Lee, J. W. Lee, J. H. Yi, C. B. Yun, and H. Y. Jung. Neural networks-based damage detection for bridges considering errors in baseline finite element models. *Journal of Sound and Vibration*, 280(3–5):555–578, February 2005.

[33] J. J. Lee and C. B. Yun. Damage localisation for bridges using probabilistic neural networks. *Journal of Civil Engineering*, 11(2):111–120, 2007.

[34] J. J. Lee and C. B. Yun. Damage diagnosis of steel girder bridges using ambient vibration data. *Engineering Structures*, 28(6):912–925, 2006.

[35] D. Vines-Cavanaugh, Y. Cao, and M. Wang. Support vector machine for abnormality detection on a cable-stayed bridge. In *SPIE*, 2010.

[36] M. Malekzadeh, G. Atia, and F. Catbas. Performance-based structural health monitoring through an innovative hybrid data interpretation framework. *Journal of Civil Structural Health Monitoring*, 5(3):287–305, 2015.

[37] P. M. Pawar and S. N. Jung. Support vector machine based online composite helicopter rotor blade damage detection system. *Journal of Intelligent Material Systems and Structures*, 19:1217–1228, October 2008.

[38] S. B. Satpal, A. Guha, and S. Banerjee. Damage identification in aluminum beams using support vector machine: Numerical and experimental studies. *Structural Control and Health Monitoring*, 23:446–457, 2015.

[39] C. Sbarufatti, A. Manes, and M. Giglio. Performance optimization of a diagnostic system based upon a simulated strain field for fatigue damage characterization. *Mechanical Systems and Signal Processing*, 40(2):667–690, November 2013.

[40] C. Sbarufatti, G. Manson, and K. Worden. A numerically-enhanced machine learning approach to damage diagnosis using a lamb wave sensing network. *Journal of Sound and Vibration*, 333:4499–4525, 2014.

[41] V. Meruane and J. Mahu. Real-time structural damage assessment using artificial neural networks and antiresonant frequencies. *Shock and Vibration*, 2014:1–14, February 2014.

[42] V. Meruane. Online sequential extreme learning machine for vibration-based damage assessment using transmissibility data. *Journal of Computing in Civil Engineering*, 30(3), 2015.

[43] X. Zhou, Y. Ni, and F. Zhang. Damage localisation of cable-supported bridges using modal frequency data and probabilistic neural network. *Mathematical Problems in Engineering*, 2014:1–10, June 2014.

[44] J. Guo, J. Wu, J. Guo, and Z. Jiang. A damage identification approach for offshore jacket platforms using partial modal results and artificial neural networks. *Applied Sciences*, 8(11):2173–2188, June 2018.

[45] A. Dimou, G. Jager, and P. Fragos. Adaptive edge enhancement in SAR images training on the data vs. training on simulated data. In *Proceedings of 2001 International Conference on Image Processing*, pages 493–496, Thessaloniki, Greece, 2001.

[46] H. Hammer, K. Hoffman, and K. Schulz. On the classification of passenger cars in airborne SAR images using simulated training data and a convolutional neural network. In *Proceedings of Image and Signal Processing for Remote Sensing XXIV*, Berlin, Germany, 2018.

[47] J. M. J. Huttunen, L. Karkainen, and H. Lindholm. Pulse transit time estimation of aortic pulse wave velocity and blood pressure using machine learning and simulated training data. *Public Library of Science Computational Biology*, 15(8), August 2019.

[48] E. Figueiredo, I. Moldovan, A. Santos, P. Campos, and J. C. W. A. Costa. Finite element-based machine-learning approach to detect damage in bridges under operational and environmental variations. *Journal of Bridge Engineering*, 24(7), 2019.

[49] P. Arendt, D. Apley, and W. Chen. Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. *Journal of Mechanical Design*, 134(10):1–12, 2012.

[50] G. E. P. Box and N. R. Draper. *Empirical Model-Building and Response Surfaces*. John Wiley and Sons Ltd., Chichester, UK, 1987.

[51] G. T. Webb, P. K. Vardanega, and C. R. Middleton. Categories of SHM deployments: Technologies and capabilities. *Journal of Bridge Engineering*, 20(11), 2015.

[52] E. Simoen, G. De Roeck, and G. Lombaert. Dealing with uncertainty in model updating for damage assessment: A review. *Mechanical Systems and Signal Processing*, 56–57:123–149, May 2015.

[53] B. M. Wilson and A. Koskelo. A practical validation assessment workflow. *Journal of Verification, Validation, and Uncertainty Quantification*, 5(1), March 2020.

[54] E. Papatheou, G. Manson, R. J. Barthorpe, and K. Worden. The use of pseudo-faults for novelty detection in SHM. *Journal of Sound and Vibration*, 329(12):2349–2366, June 2010.

[55] P. Gardner, C. Lord, and R. J. Barthorpe. An evaluation of validation metrics for probabilistic model outputs. In *ASME 2018 Verification and Validation Symposium*, Minneapolis, USA, 2018.

[56] J. Kaizer. Progress toward a complete set of errors for modeling and simulation. *Journal of Verification, Validation and Uncertainty Quantification*, 5(3), September 2020.

[57] M. A. Beaumont, W. Zhang, and D. J. Balding. Approximate Bayesian computation in population genetics. *Genetics Society of America*, 162(4):2025–2035, December 2002.

[58] M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 63:425–450, 2001.

[59] P. B. Holden, N. B. Edwards, J. Hensman, and R. D. Wilkinson. ABC for climate: Dealing with expensive simulators. 2015.

[60] G. H. James, R. L. Mayes, T. Carne, T. Simmermacher, and J. Goodding. Health monitoring of operational structures – initial results. In *36^{th} Structures, Structural Dynamics and Materials Conference*, New Orleans, USA, 1995.

[61] T. Simmermacher, R. L. Mayes, G. M. Reese, and G. H. James. The effects of finite element grid density on model correlation and damage detection of a bridge. In *36th Structures, Structural Dynamics and Materials Conference*, New Orleans, USA, 1995.

[62] N. Stubbs, J. T. Kim, and C. R. Farrar. Field verification of a nondestructive damage localization and severity estimation algorithm. In *SPIE*, 1995.

[63] J. Nagel and B. Sudret. A unified framework for multilevel uncertainty quantification in Bayesian inverse problems. *Probabilistic Engineering Mechanics*, 43:68–84, January 2016.

[64] C. Li and S. Mahadevan. Role of calibration, validation, and relevance in multi-level uncertainty integration. *Reliability Engineering and System Safety*, 148:32–43, 2016.

[65] I. M. Sobol'. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1):271–280, 2001.

[66] M. S. Allen, D. Rixen, M. van der Seijs, P. Tiso, T. Abrahamsson, and R. L. Mayes. *Substructuring in Engineering Dynamics: Emerging Numerical and Experimental Techniques*. Springer, Cham, Switzerland, 1st edition, 2020.

[67] R. R. Craig and M. C. C. Bampton. Coupling of substructures for dynamic analyses. *American Institute of Aeronautics and Astronautics*, 6(7):1313–1319, July 1968.

[68] R. J. Kuether and M. S. Allen. Craig–Bampton substructuring for geometrically nonlinear subcomponents. In *Proceedings of the Society for Experimental Mechanics Series*, pages 167–178, 2014.

[69] S. N. Voormeeren, P. L. C. van der Valk, and D. J. Rixen. Practical aspects of dynamic substructuring in wind turbine engineering. In *Proceedings of the Society for Experimental Mechanics Series*, pages 163–185, 2011.

[70] F. Trainottii, M. Haeussler, and D. J. Rixen. A practical handling of measurement uncertainties in frequency based substructuring. *Mechanical Systems and Signal Processing*, 144(9), October 2020.

[71] S. Voormeeren. Coupling procedure improvement and uncertainty quantification in experimental dynamic substructuring. Master's thesis, Mechanical Engineering and Marine Technology, Technische Universiteit Delft, Delft, Netherlands, October 2007.

[72] D. de Klerk. *Dynamic response characterization of complex systems through operational identification and dynamic substructuring*. PhD thesis, Technische Universiteit Delft, Delft, Netherlands, March 2009.

[73] T. Simpson, D. Giagopoulos, V. Dertimanis, and E. Chatzi. On dynamic substructuring of systems with localised nonlinearities. 2020.

[74] J. Ferreira and D. Ewins. Nonlinear receptance coupling approach based on describing functions. In *Proceedings of SPIE – The International Society for Optical Engineering*, 1996.

[75] J. Ferreira and D. Ewins. Multi-harmonic nonlinear receptance coupling approach. In *Proceedings of SPIE – The International Society for Optical Engineering*, 1997.

[76] O. S. Bursi, G. Abbiati, E. Cazzador, P. Pegon, and F. J. Molina. Nonlinear heterogeneous dynamic substructuring and partitioned FETI time integration for the development of low-discrepancy simulation models. *International Journal for Numerical Methods in Engineering*, 112(9):1253–1291, 2017.

[77] J. Brunetti, W. D'Ambrogio, and A. Fregolent. Friction-induced vibrations in the framework of dynamic substructuring. *Nonlinear dynamics*, 103(4):3301–3314, 2020.

[78] S. I. Z. Estakhraji and M. S. Allen. Extension of the harmonic balance method for dynamic systems with Iwan joints. *Mechanical Systems and Signal Processing*, 166:108434, 2022.

[79] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 1979.

[80] S. N. Voormeeren, D. de Klerk, and D. J. Rixen. Uncertainty quantification in experimental frequency based substructuring. *Mechanical Systems and Signal Processing*, 24(1):106–118, 2010.

[81] D. Kammer, D. C. amd Krattiger. Propagation of uncertainty in substructured spacecraft using frequency response. *American Institute of Aeronautics and Astronautics*, 51(2):353–361, 2013.

[82] B. R. Mace and P. J. Shorter. A local modal/perturbational method for estimating frequency response statistics of built-up structures with uncertain properties. *Journal of Sound and Vibration*, 242(5):793–811, 2001.

[83] V. V. Yotov, M. Remedia, G. S. Aglietti, and G. Richardson. Improved Craig–Bampton stochastic method for spacecraft vibroacoustic analysis. *Acta Astronautica*, 178:556–570, 2021.

[84] M. Remedia, G. S. Aglietti, and G. Richardson. A stochastic methodology for predictions of the environment created by multiple microvibration sources. *Journal of Sound and Vibration*, 344:138–157, 2015.

[85] H. A. Jensen, V. A. Araya, A. D. Muñoz, and M. A. Valdebenito. A physical domain-based substructuring as a framework for dynamic modeling and reanalysis of systems. *Computer methods in applied mechanics and engineering*, 326:656–678, 2017.

[86] T. Chatterjee, S. Adhikari, and M. I. Friswell. Uncertainty propagation in dynamic sub-structuring by model reduction integrated domain decomposition. *Computer methods in applied mechanics and engineering*, 366:113060, 2020.

[87] M. S. Allen, D. C. Kammer, and R. L. Mayes. Uncertainty in experimental/analytical substructuring predictions: A review with illustrative examples. In *International Conference on Noise and Vibration Engineering (ISMA)*, pages 2–13, Leuven, Belgium, 2010.

[88] A. Ben Abdessalem, N. Dervilis, D. Wagg, and K. Worden. Model selection and parameter estimation in structural dynamics using approximate Bayesian computation. *Mechanical Systems and Signal Processing*, 99:306–325, 2018.

[89] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, March 1951.

[90] L. R. Salazar, J. A. Cobano, and A. Ollero. Small UAS-based wind feature identification system part 1: Integration and validation. *Sensors*, 17(1):8, January 2017.

[91] T. Buckley, B. Ghosh, and V. Pakrashi. A feature extraction & selection benchmark for structural health monitoring. *Structural Health Monitoring*, 2022.

[92] W. Gao, L. Hu, P. Zhang, and H. Jialong. Feature selection considering the composition of feature relevancy. *Pattern Recognition Letters*, 112:70–74, September 2018.

[93] D. Song, F. Xu, and T. Ma. Crack damage monitoring for compressor blades based on acoustic emission with novel feature and hybridized feature selection. *Structural Health Monitoring*, 21(6):2641–2656, 2022.

[94] A. Saltelli, K. Chan, and E. Scott. *Sensitivity Analysis*. Wiley, Chichester, UK, 2nd edition, 2008.

[95] P. Zhang. A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model. *Applied Soft Computing*, 85:105859, 2019.

[96] J. E. Oakley and A. O'Hagan. Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of the Royal Statistical Society*, 66(3):751–769, 2004.

[97] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93(7):964–979, 2008.

[98] W. Liu, X. Wu, L. Zhang, Y. Wang, and J. Teng. Sensitivity analysis of structural health risk in operational tunnels. *Automation in Construction*, 94:135–153, October 2018.

[99] T. Hong and S. T. Purucker. Spatiotemporal sensitivity analysis of vertical transport of pesticides in soil. *Environmental Modelling & Software*, 105:24–38, July 2018.

[100] W. Becker, J. Rowson, J. E. Oakley, A. Yoxall, G. Manson, and K. Worden. Bayesian sensitivity analysis of a model of the aortic valve. *Journal of Biomechanics*, 44:1499–1506, 2011.

[101] W. Becker, J. E. Oakley, C. Surace, P. Gili, J. Rowson, and K. Worden. Bayesian sensitivity analysis of a nonlinear finite element model. *Mechanical Systems and Signal Processing*, 32:18–31, 2012.

[102] B. Sohlberg and M. Sernfalt. Grey box modelling for river control. *Journal of Hydroinformatics*, 4(4):265–280, October 2002.

[103] P. Lindskog and L. Ljung. Tools for semi-physical modeling. *IFAC Proceedings Volumes*, 26(8):1199–1204, 1994.

[104] P. J. Jones, M. Catt, M. J. Davies, C. L. Edwardson, E. M. Mirkes, K. Khunti, T. Yates, and A. V. Rowlands. Feature selection for unsupervised machine learning of accelerometer data physical activity clusters – A systematic review. *Gait & Posture*, 90:120–128, 2021.

[105] O. W. Ahmed, R. Qahwaji, T. Colak, P. A. Higgins, P. T. Gallagher, and D. S. Bloomfield. Solar flare prediction using advanced feature extraction, machine learning, and feature selection. *Solar physics*, 283(1):157–175, 2011.

[106] J. Cai, J. Luo, S. Wang, and S. Yang. Feature selection in machine learning: A new perspective. *Neurocomputing*, 300:70–79, July 2018.

[107] K. Maes and G. Lombaert. Monitoring railway bridge KW51 before, during, and after retrofitting. *Journal of Bridge Engineering*, 26(3), March 2021.

[108] K. Maes, L. Van Meerbeck, E. P. B. Reyners, and G. Lombaert. Validation of vibration-based structural health monitoring on retrofitted railway bridge KW51. *Mechanical Systems and Signal Processing*, 165, 2022.

[109] R. J. Barthorpe, A. J. Hughes, and P. Gardner. A forward model driven structural health monitoring paradigm: Damage detection. In *Proceedings of the Society for Experimental Mechanics Series*, pages 119–126, Cham, Switzerland, 2022.

[110] S. P. Timoshenko. On the transverse vibrations of bars of uniform cross-section. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 43(253):125–131, 1922.

[111] R. D. Blevins. *Formulas for Natural Frequency and Mode Shape*. Krieger, Malabar, Florida, 1st edition, 2001.

[112] L. F. Richardson. The approximate arithmetical solution by finite differences of physical problems including differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society A*, 210(467):307–357, 1911.

[113] J. Wilson, P. Gardner, G. Manson, and R. J. Barthorpe. The role of features in a hierarchical modelling strategy for forward model-driven structural health monitoring. In *ISMA 2022–International Conference on Noise and Vibration*

Engineering and USD 2022–International Conference on Uncertainty in Structural Dynamics (ISMA-USD 2022), Leuven, Belgium, 2022.

[114] D. J. Ewins. *Modal Testing: Theory, practice and application*. Research Studies Press Ltd., Baldock, UK, 2 edition, 2000.

[115] S. R. Krishnan and C. S. Seelamantula. On the selection of optimum Savitzky–Golay filters. *IEEE transactions on signal processing*, 61(2):380–391, 2013.

[116] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, January 1965.

[117] M. I. Friswell and J. E. T. Penny. Crack modelling for structural health monitoring. *Structural Health Monitoring*, 1(2):139–148, 2002.

[118] S. Christides and A. D. S. Barr. One dimensional theory of cracked Bernoulli–Euler beams. *International Journal of Mechanical Science*, 26(11/12):639–648, 1984.

[119] J. K. Sinha, M. I. 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.

[120] M. Bruns, B. Hofmeister, T. Grießmann, and R. Rolfes. Comparative study of parameterizations for damage localization with finite element model updating. In *29th European Safety and Reliability Conference*, pages 1125–1132, Hannover, Germany, 2019.

[121] R. D. Wilkinson. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Statistical applications in genetics and molecular biology*, 12(2):129–141, 2013.

[122] A. Ben Abdessalem, N. Dervilis, D. Wagg, and K. Worden. Model selection and parameter estimation of dynamical systems using a novel variant of approximate Bayesian computation. *Mechanical Systems and Signal Processing*, 122:364–386, 2019.

[123] T. Bergmayr, S. Höll, C. Kralovec, and M. Schagerl. Local residual random forest classifier for strain-based damage detection and localization in aerospace sandwich structures. *Composite Structures*, 304(1):116331, 2023.

[124] N. Mechbal, J. S. Uribe, and M. Rébillat. A probabilistic multi-class classifier for structural health monitoring. *Mechanical Systems and Signal Processing*, 60-61:106–123, 2015.

[125] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, USA, 1st edition, 2006.

[126] S. Rogers and M. Girolami. *A First Course in Machine Learning*. CRC Press Inc, Boca Raton, USA, 2nd edition, 2017.

[127] J. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers*, 10, June 2000.

[128] A. Bolourani, M. Bitaraf, and A. Nekouvaght Tak. Structural health monitoring of harbor caissons using support vector machine and principal component analysis. *Structures*, 33:4501–4513, 2021.

[129] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, USA, 1st edition, 2005.

[130] K. Worden, W. J. Staszewski, and J. J. Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111, 2011.

[131] H. Teimouri, A. S. Milani, J. Loeppky, and R. Seethaler. A Gaussian process–based approach to cope with uncertainty in structural health monitoring. *Structural Health Monitoring*, 16(2):174–184, 2017.

Appendix A

Further experimental work

A.1 Static dataset: Experimental procedure

Experiments were carried out in order to acquire a complete dataset representing the response of the struts of a bridge under a range of static load conditions. The bridge in question was a laboratory-scale truss bridge and was housed at the Laboratory for Verification and Validation, a specialist laboratory owned and operated by the University of Sheffield for investigations of dynamic structures. The bridge, shown in Figure A.1, was a truss bridge configured with diagonal and vertical struts. The bridge material was aluminium and comprised three main types of component: the deck, the struts, and a number of Rexroth sections (these spanned the outer perimeter of the deck and connected the struts at the top).

The struts of the bridge were easily identifiable as critical components to its performance; this is well-known in bridge design theory. In addition, the struts were easily removable, and were simple components for modelling and test design. This made them ideal candidates for demonstration of the hierarchical V&V process postulated in this project.

A.1.1 Experimental taxonomy

In order to describe the experiments and results with clarity, the following taxonomy is defined regarding the bridge in terms of its components, loading, and damage condition.

As illustrated in Figure A.2, the struts of the bridge were labelled numerically. To maintain clarity, the load points on the bridge deck were labelled alphabetically, as

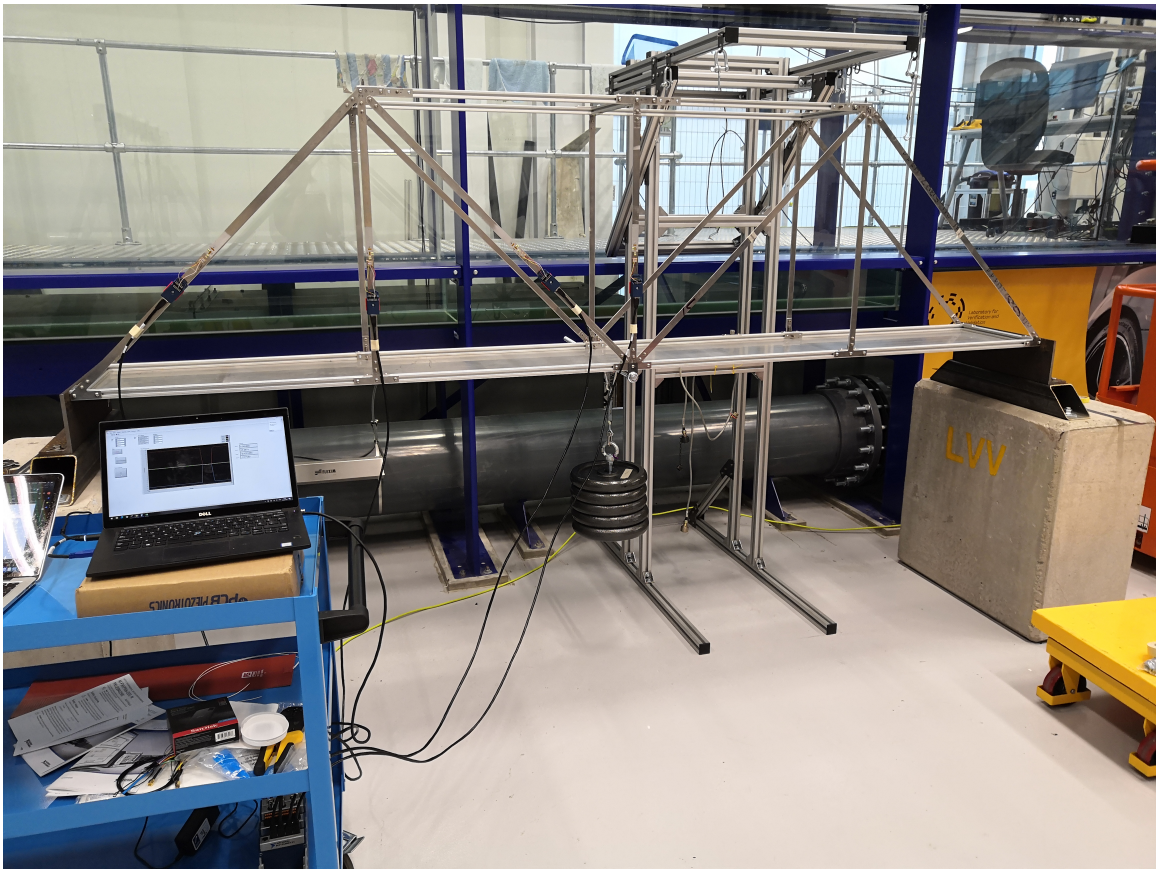


Figure A.1: The truss bridge at the LVV, loaded for static testing

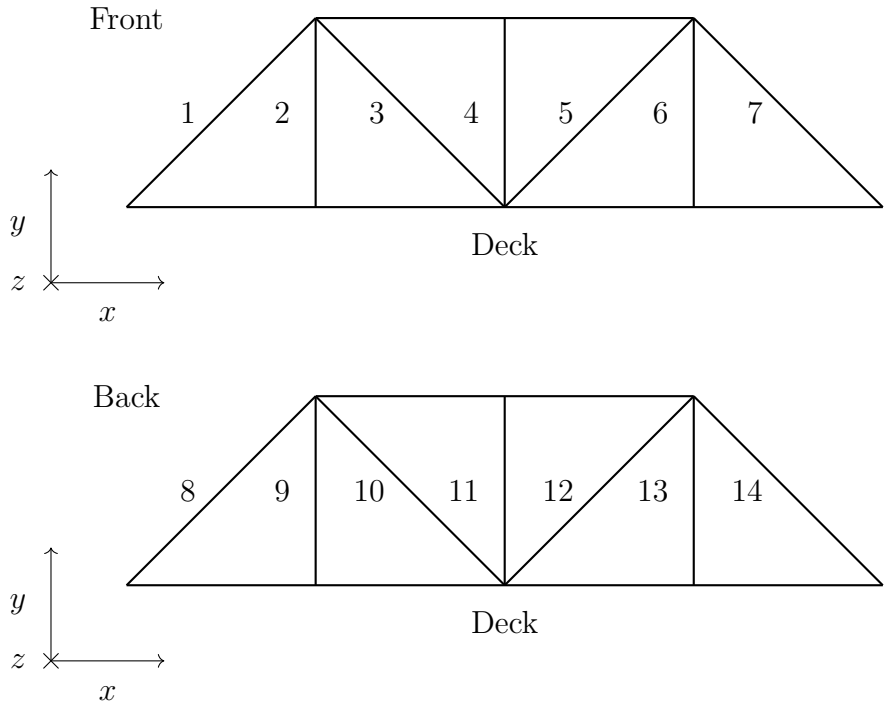


Figure A.2: The taxonomy of strut labels for the LVV bridge in this dataset

shown in Figure A.3. This was intended to avoid confusion between the two sets of labels.

The damage level present in the bridge or strut is indicated by $D^{extent}_{location}$. This indicator will vary slightly depending on whether the data are relevant to the whole bridge or to a single strut. An example of a damage state in the full bridge would be $D^{100\%}_{3}$, which would indicate complete damage (failure) in Strut 3 on the bridge. An example of a damage state for a single strut would be $D^{50\%}_{200\text{mm}}$, which would indicate a state between complete health and failure at a distance of 200mm from a reference point. Static loads are defined similarly to damage, by $L^{magnitude}_{location}$.

A.1.2 Distribution of load: Estimate

An estimate was made of the distribution of tensile loads in the struts of the bridge. The load distribution was calculated under the rigid-body assumption, for the L^{F}_{K} condition and is shown in Figure A.4. This simple calculation showed the struts that would be expected to experience compressive or tensile loads. It was also used to inform the selection of strain gauges for the experimental tests carried out afterwards.

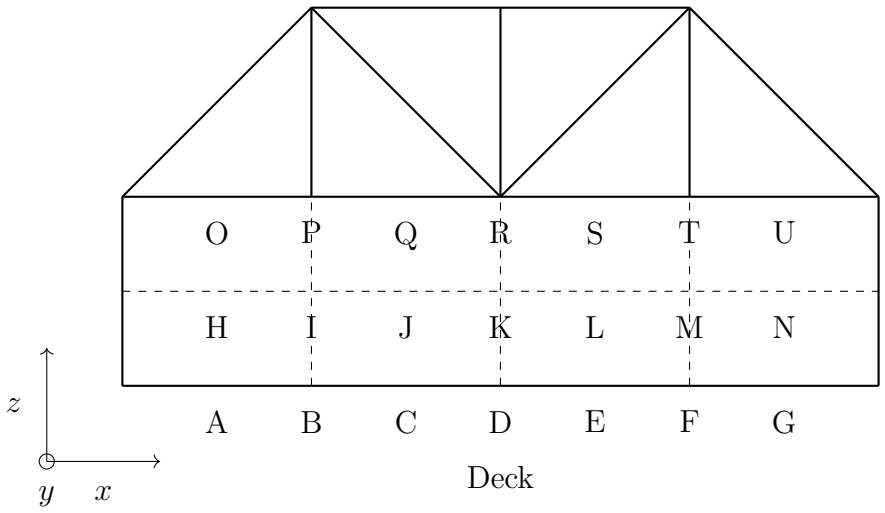


Figure A.3: The taxonomy of load labels for the LVV bridge

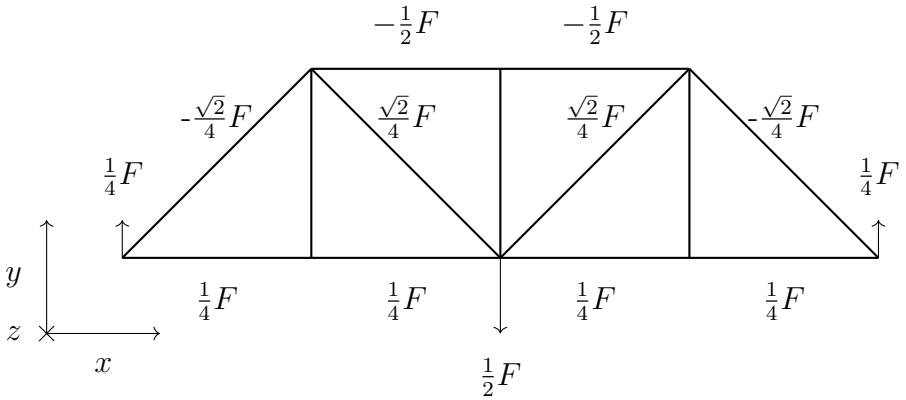


Figure A.4: The tensile load distribution estimates for a central static load under the rigid body assumption

Figure A.5: The rig used for strain gauge calibration

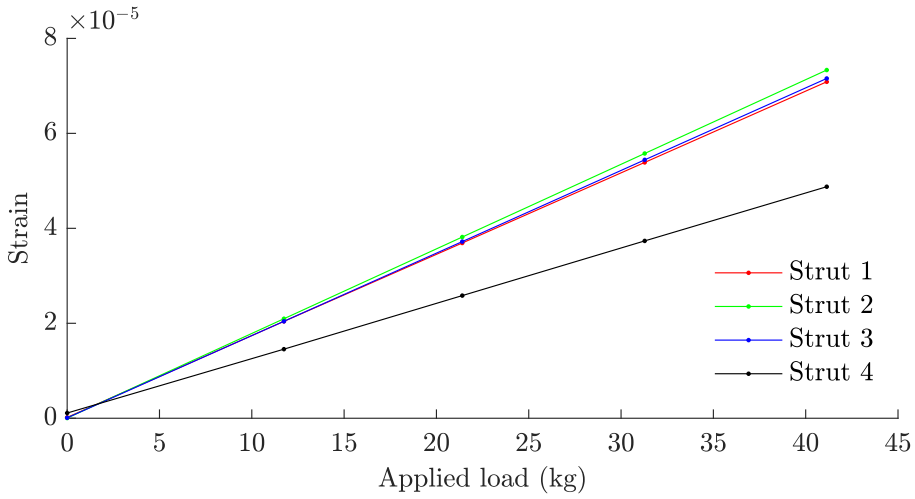


Figure A.6: The results of the strain gauge calibration tests

A.1.3 Strain gauge calibration

For each of the four struts investigated in these tests, the strain gauges used were calibrated to ensure accurate measurement. This process was carried out by applying a series of known loads to the struts and recording the gauge response. The struts were pinned at one end, while dead weights were attached at the other end. 10kg weights were used, up to a maximum load of 40kg. The setup used is shown in Figure A.5.

As can be seen in Figure A.6, the strain response to load was linear over the range tested, making the following numerical calibration tasks trivial. It should be noted that while the gradient of Strut 4 under loading was different to the others, this was not an issue given that it was still constant. The least squares fit was then calculated for each strut on the assumption of linear correlation. The error caused by using this approximation is shown in Figure A.7. It should be noted that the error remained very low (less than 1%) at all values, particularly when higher loads were used. These errors were due to measurement noise on the strain gauges and random errors associated with mass of the weights used in loading.

A.1.4 Single-strut tests

Tests were carried out to record the strain transmitted by the struts when loaded in isolation. The strain response was recorded for a range of static loads, both tensile

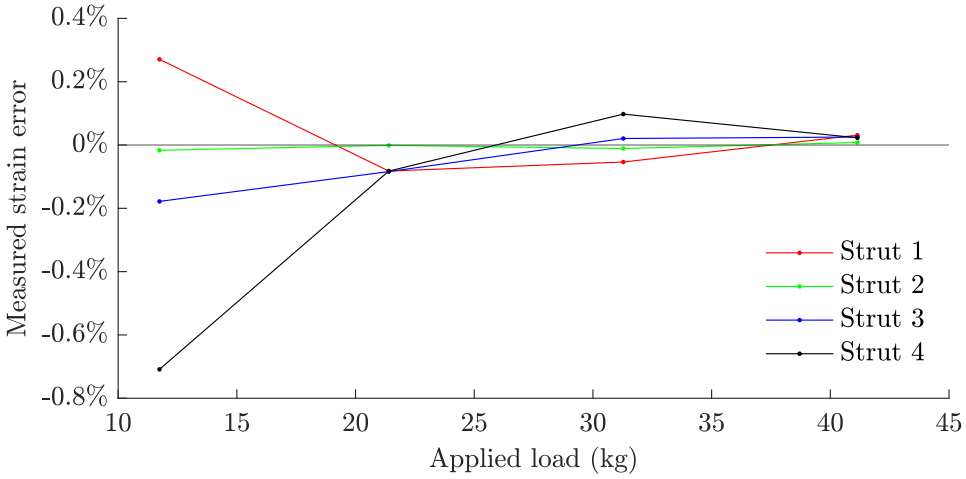


Figure A.7: The strain gauge calibration error associated with the linear fit

and compressive, with the struts in their healthy states. Both the diagonal (longer) and vertical (shorter) strut types were tested.

A.1.4.1 Load cell calibration

A load cell was used to measure the forces applied to the strut in these tests. This was calibrated beforehand, with four measurement points taken to ensure a linear response. The results of the calibration tests are shown in Figure A.8. The figure shows strong linearity, providing a reliable gradient for load cell calibration. The intercept was ignored, as the load cell was zeroed before each test after calibration.

Figure A.9 shows the error due to the gradient fitting. It can be observed that the uncertainty was very low, remaining within 0.1% for each measurement point. There is no discernible discrepancy and the variation seems to be caused by random noise in the measurement process.

A.1.4.2 Methodology and results

Figure A.10 shows the rig used for these tests. The strut was fixed to a (relatively) rigid cast iron base structure, which was itself set on damped steel feet. The strut was clamped at the top end with all DoFs fixed. At the bottom end, the strut was clamped in such a way that all except the vertical DoFs were fixed. Static loads could be applied to the strut by means of a pinion and wheel at the bottom of the strut. The loads were controlled using weights that applied a moment to the wheel, which

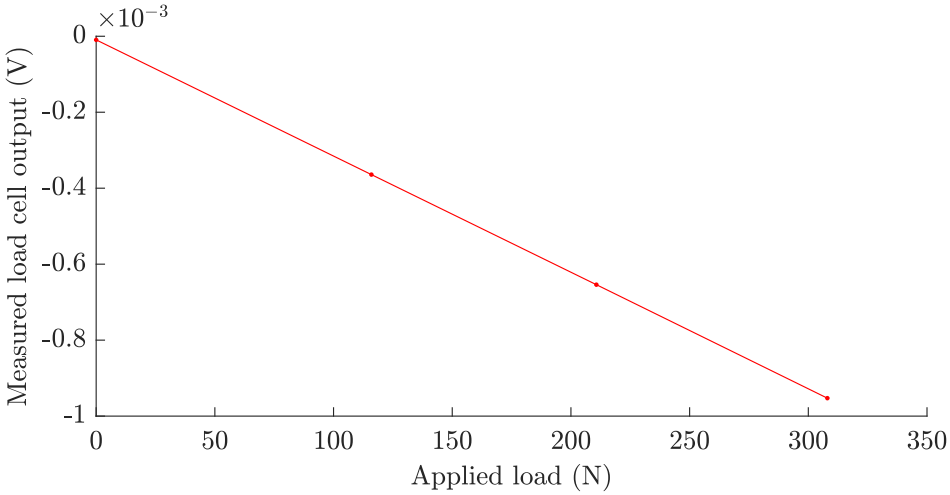


Figure A.8: The results of the load cell calibration tests

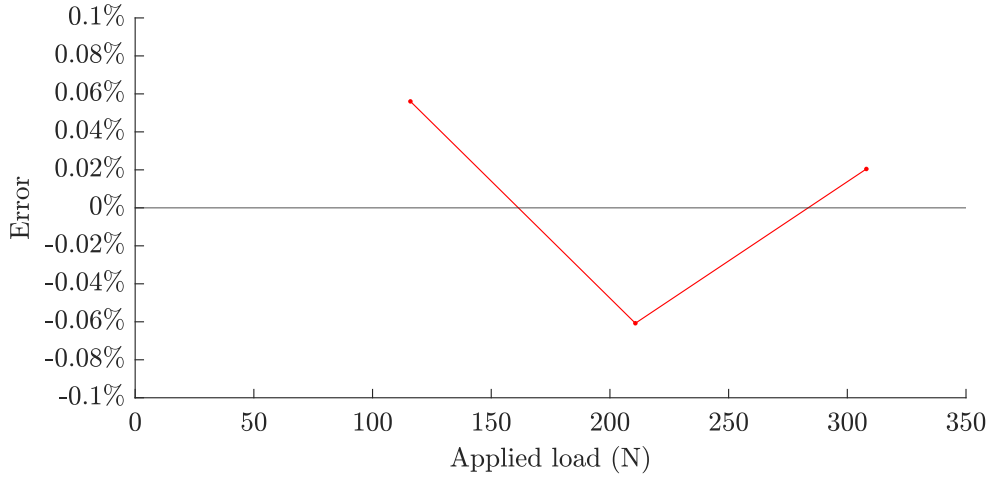


Figure A.9: The load cell calibration error associated with the linear fit

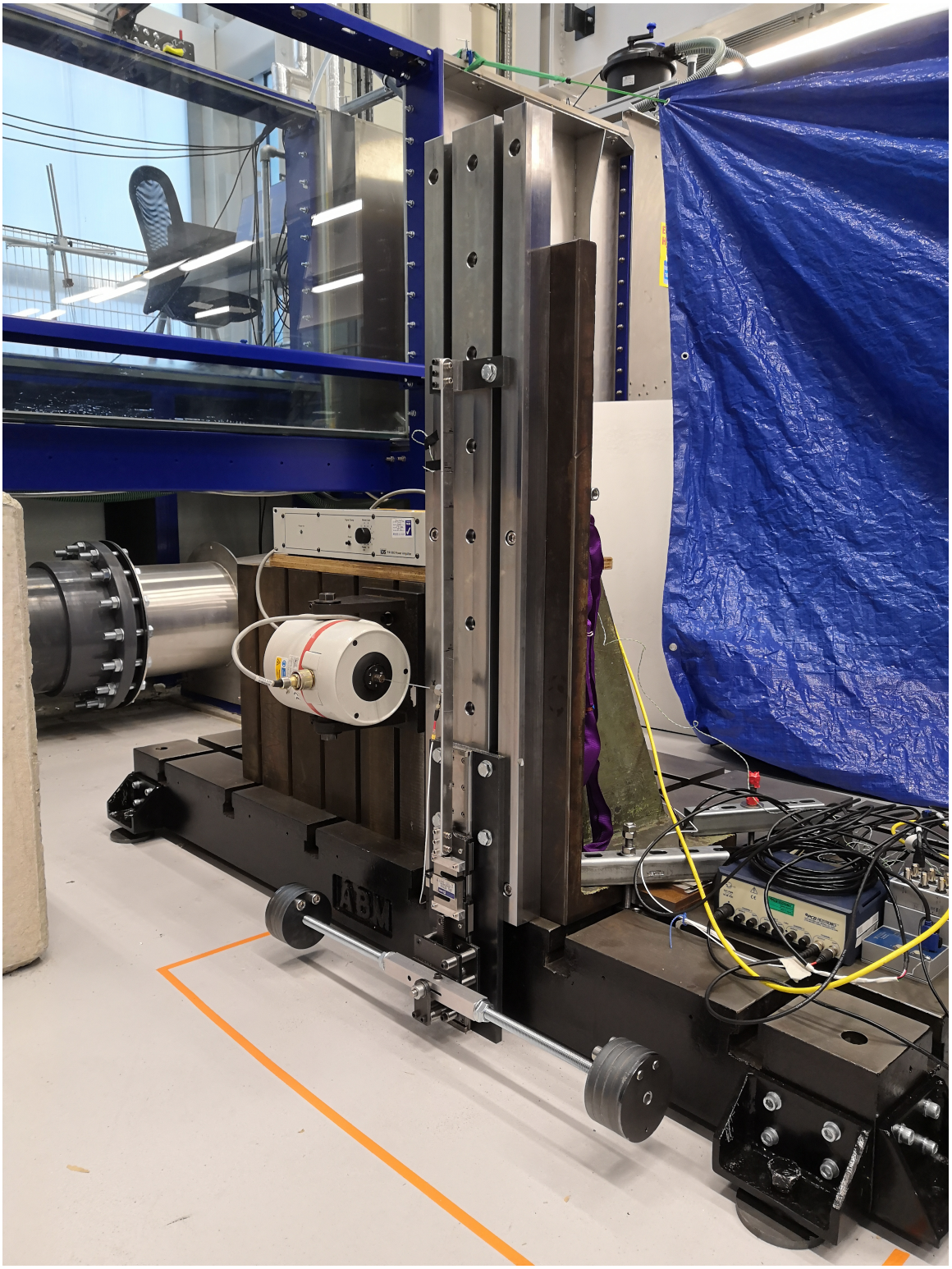


Figure A.10: The rig used for testing the struts in isolation (the shaker was detached for the static test phase)

in turn applied vertical load to the pinion. This vertical load was tracked by the load cell shown. An anticlockwise moment on the wheel produced a tensile (positive) load, and a clockwise moment produced a compressive (negative) load.

A small dataset was recorded which showed a linear response to load, across the range tested. A zero offset was recorded due to the weight of the load cell and runner below the strain gauges on the strut. This was not considered problematic, as the response gradient is all that is required for any model validation processes (for this simple case). A linear fit to the results showed that the gradient of the relationship was 0.9995 and the intercept was at 11.8973N; these could be used to find the actual load applied to the strut from any recorded load in following tests using this rig.

A.1.5 Full bridge tests

For these tests, strain gauges were placed on Struts 1 to 4, which represent each structurally unique section in the bridge. A variable deadweight mass of up to 50kg was placed at each load point (for example see Figure A.11). The attachment shown was designed to be as rigid as possible; this was in order to establish a purely vertical load condition and to reduce any pinching effect applied to the deck.

The bridge was fixed at each end, with all degrees of freedom constrained, to cast-iron mounts. These were in turn attached to heavy concrete blocks. For the purposes of these tests and the following investigations, these were considered rigid boundary conditions.

A.1.5.1 Results: Undamaged condition

The first run of tests designed were carried out with the bridge in its undamaged, healthy condition. The loads at Struts 1 to 4 were recorded for a range of loads at each load point. The results of this test were satisfactory, showing linear responses to load at each load point (although at some high loads non-linear effects were observed). Additionally, the results for load point K closely matched the predicted results from the rigid-body analysis. An example response is shown in Figure A.12. This exhibits the linear response to load, and in addition shows the development of some non-linearity that was observed for some load points at high load (see Strut 2 in particular). The cause of this non-linearity was buckling effects in the compressively-loaded struts.

When the bridge was loaded along its sides, it was found that the struts on the loaded side carried the entirety of the load (with linear response similar to the centrally-

Figure A.11: The attachment used for loading the bridge using deadweight masses

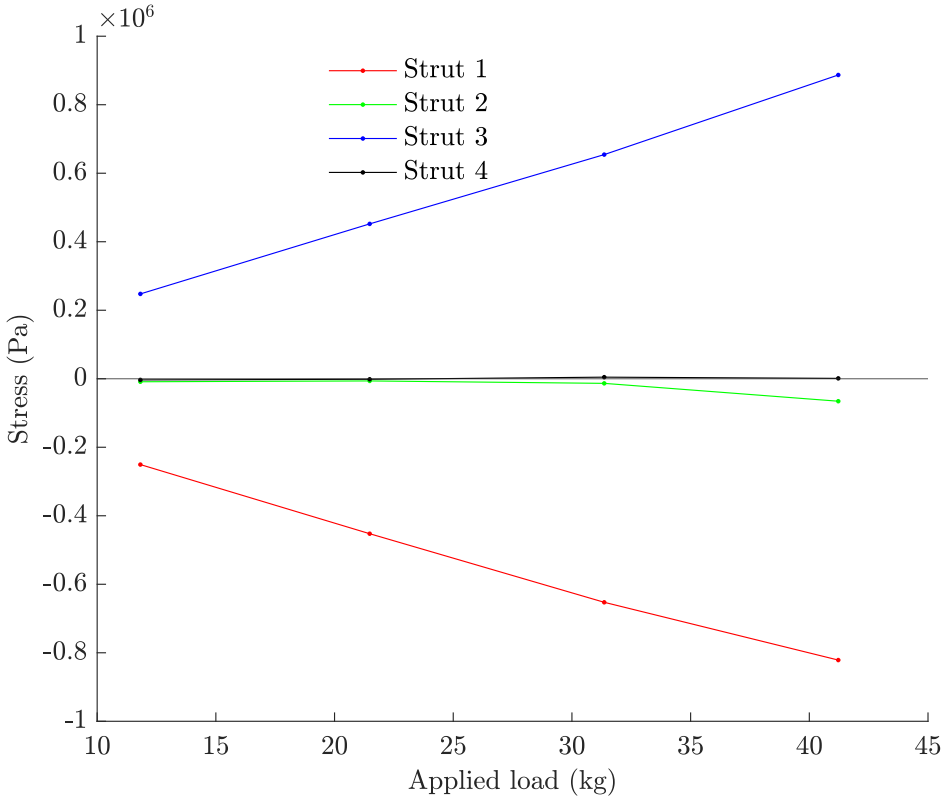


Figure A.12: The bridge response to static load for condition L_M in its healthy state

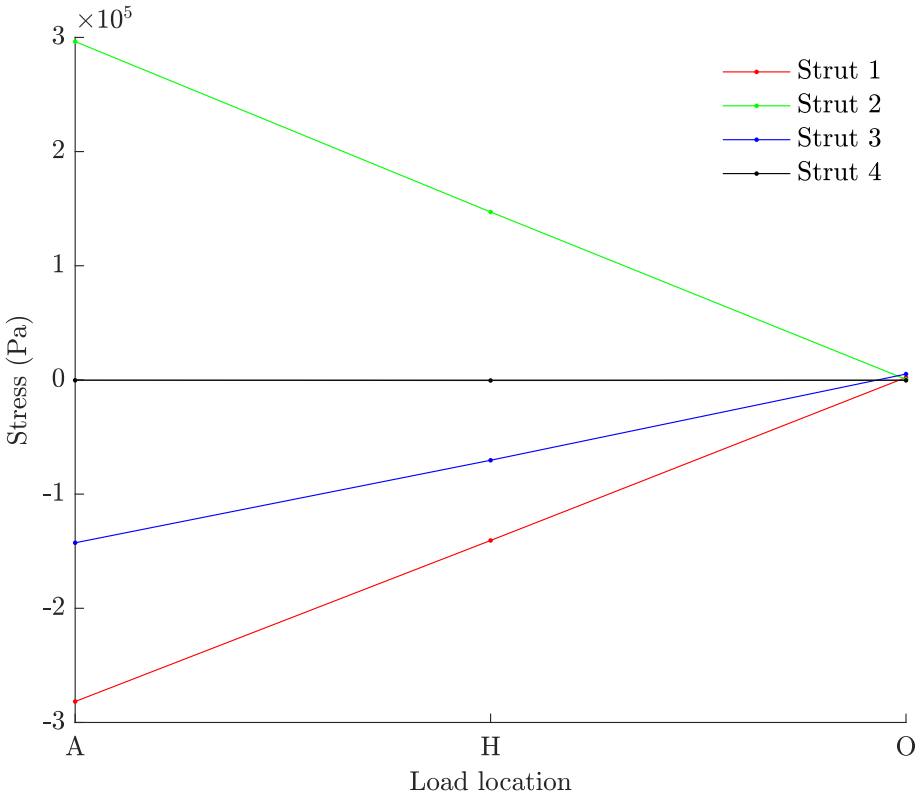


Figure A.13: The bridge response to static load for condition $L^{10kg}_{A,H,O}$ in its healthy state

loaded cases), while those on the unloaded side carried negligible strain. There was a linear relationship between the position of the load along the width of the bridge, and the load transmitted by the struts (see Figure A.13). The data in this figure were adjusted to normalise the load values to 10kg as results were drawn from different tests.

Figure A.14 shows the loads transmitted by the struts for the given (adjusted) load at each point along the length of the bridge. It can be seen that Strut 4 transmits negligible load at all locations. The load transmitted tends to decrease as the load point moves away from the sensors.

A.1.5.2 Results: Damaged conditions

In order to investigate a worst-case damage scenario, tests were carried out as before with single struts removed entirely. This situation effectively replicates the fully

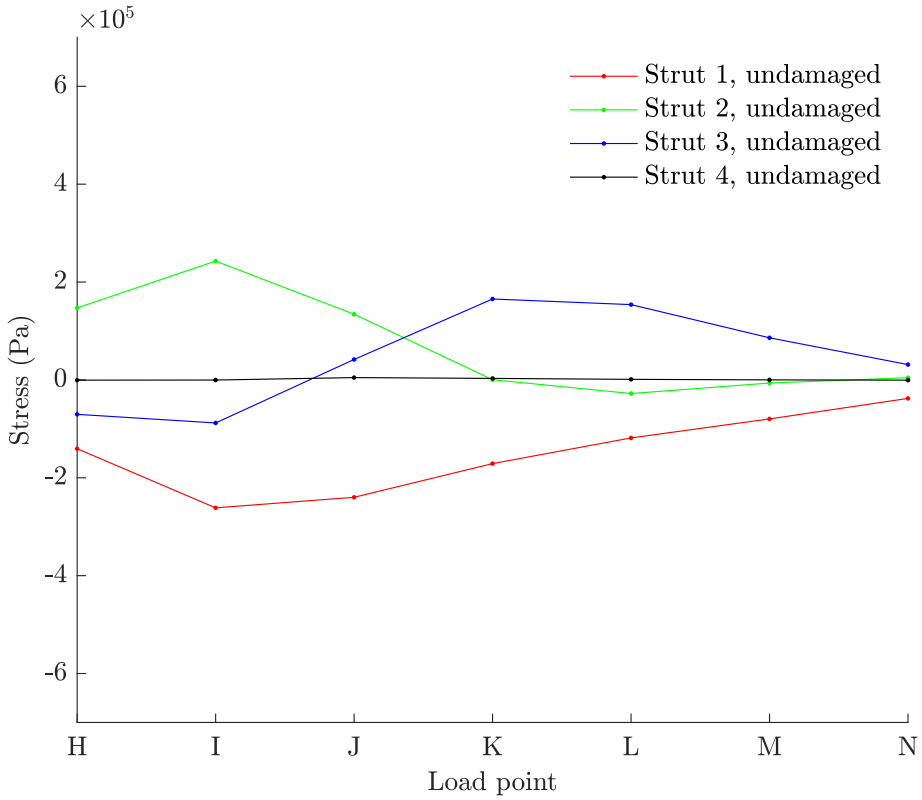


Figure A.14: The bridge response to static load for condition $L_{H-N}^{10\text{kg}}$ in its healthy state

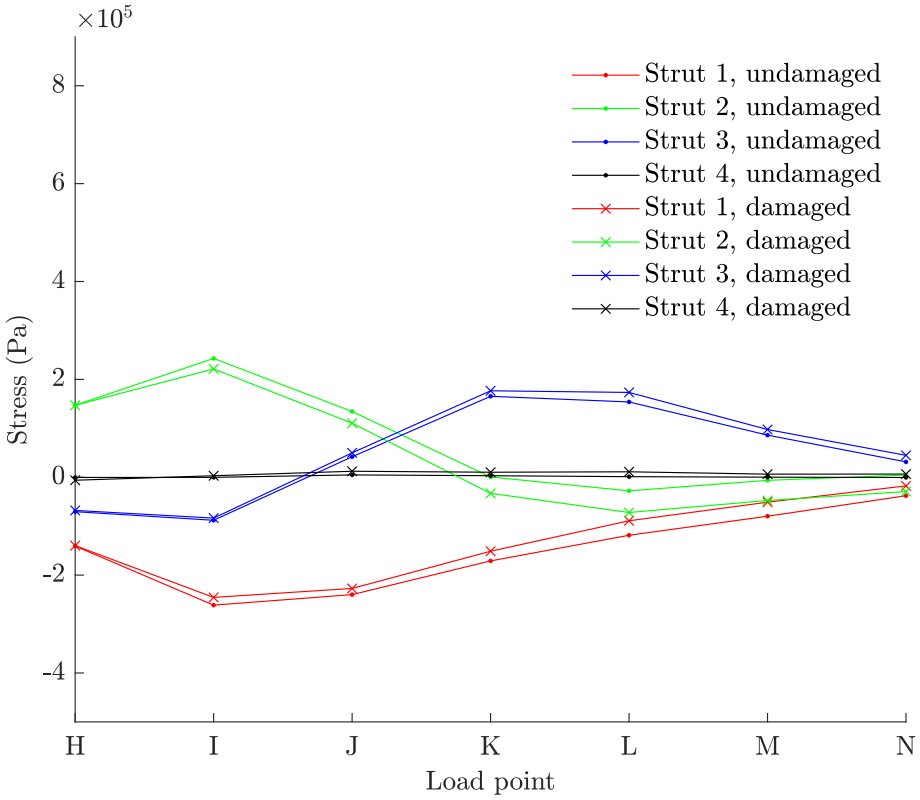


Figure A.15: The bridge response to static load for condition $L_{H-N}^{10\text{kg}}$, $D_7^{0\%}$ and $D_7^{100\%}$

damaged scenario. Attempts to investigate gradated damage states were carried out, however the response showed no sensitivity to these until the struts were removed entirely. Tests were carried out in which the diagonal struts were removed, as these contributed far more to load-bearing than the vertical struts (see Figure A.14). For the damage-state tests, the non-linearity became far more significant compared to the healthy condition tests, therefore the loads were limited to 10kg for damage-state testing.

Figure A.15 shows the difference in transmitted loads for the healthy and damaged bridge conditions, in this case for damage induced by removing Strut 7. It can be seen that there is a discernible difference in the measured response for all load points (excluding point H). It is also noteworthy that in the damaged scenario the vertical Struts 2 and 4 transmit more significant loads than previously.

The results of these tests confirm the predictions from the rigid body analysis of the

bridge: the diagonal struts carry the bulk of applied bridge loads, while the vertical struts contribute little or negligible structural support. The load distribution is altered when the location of the weight changes; this has also been recorded. Another important finding is that the struts behave linearly under load (excluding very high loads at specific locations). This means that buckling analysis can be discounted in the models for the operating range specified; this included loads of up to 50kg for the bridge in its healthy condition and 10kg in its damaged conditions.

Appendix B

Additional results from experimental work

Further results from the strut-level experimental tests presented in Chapter 6 are given in this appendix. These include a full set of FRFs and modal data for all tests.

B.1 Frequency response functions

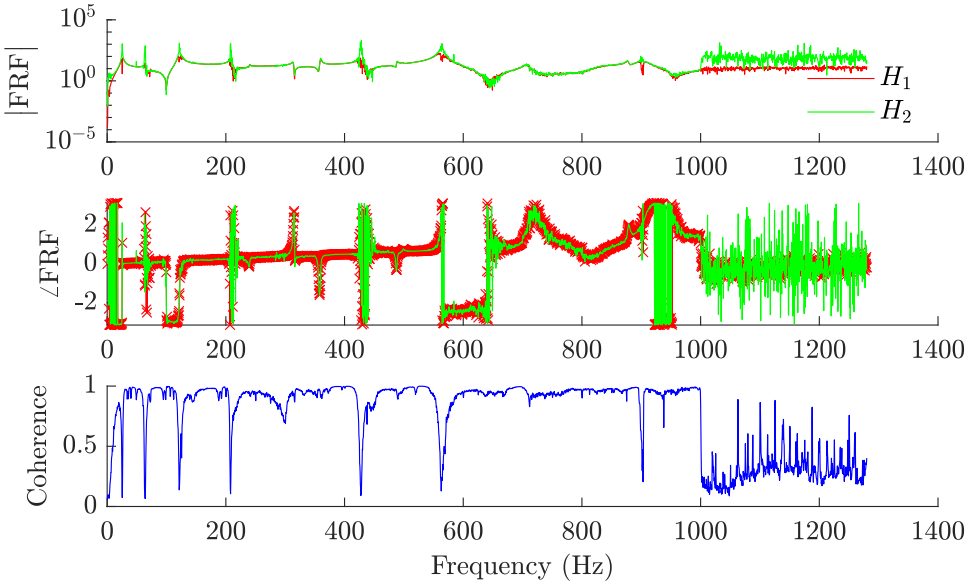


Figure B.1: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of -15.45N

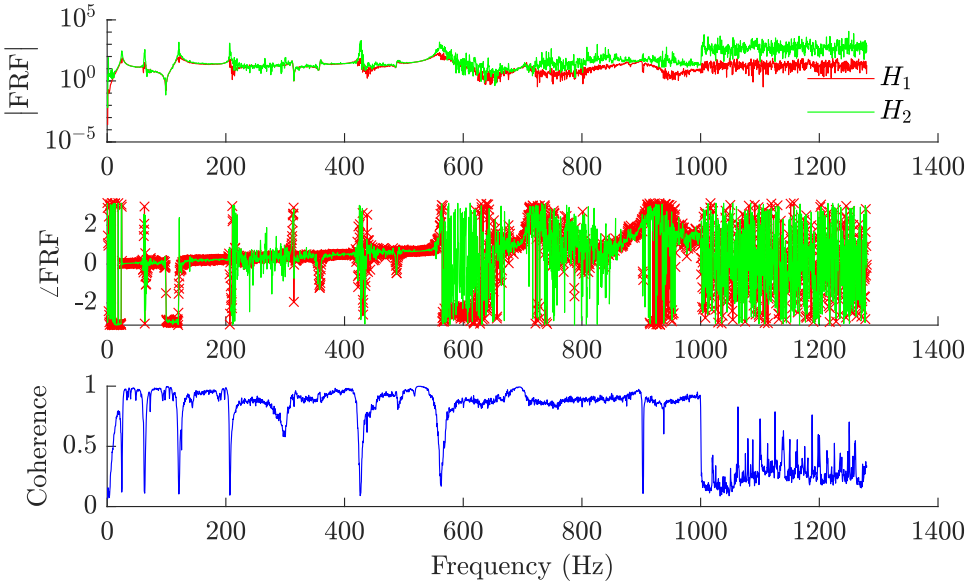


Figure B.2: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of -32.94N

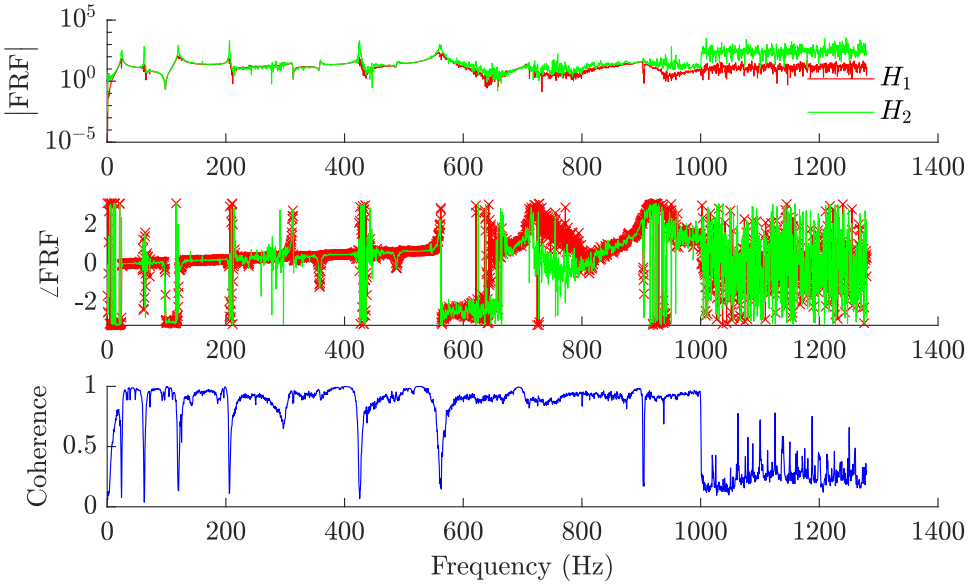


Figure B.3: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of -48.96N

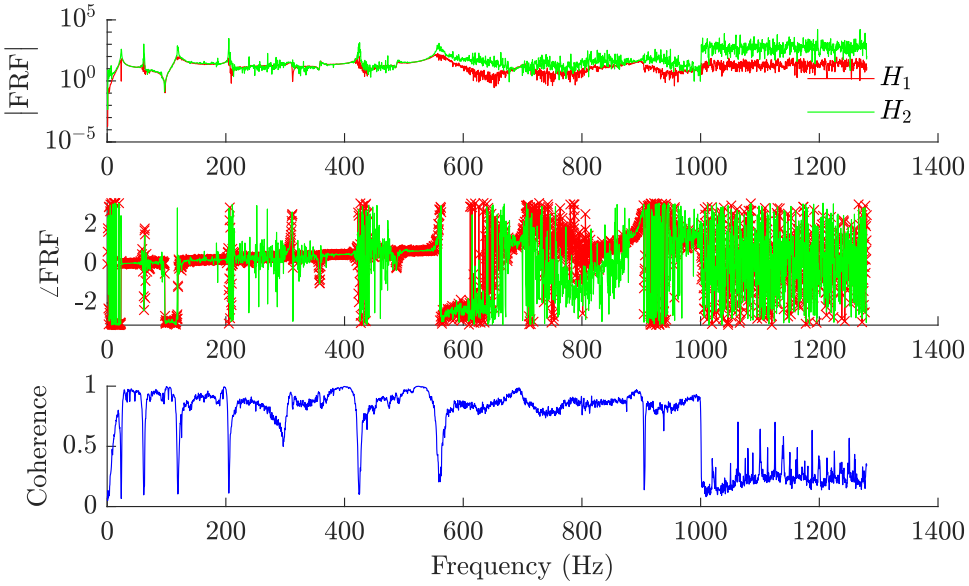


Figure B.4: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of -65.80N

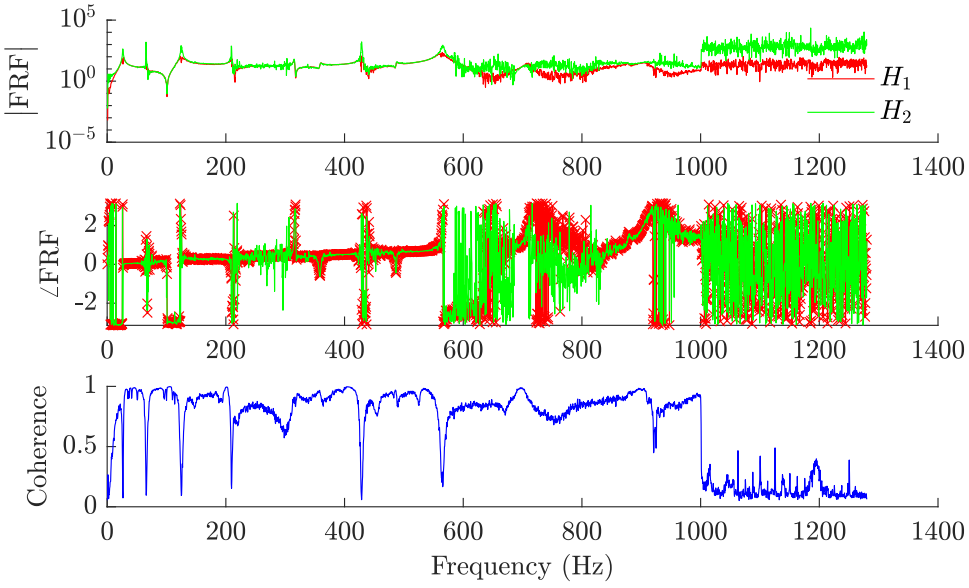


Figure B.5: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of 17.31N

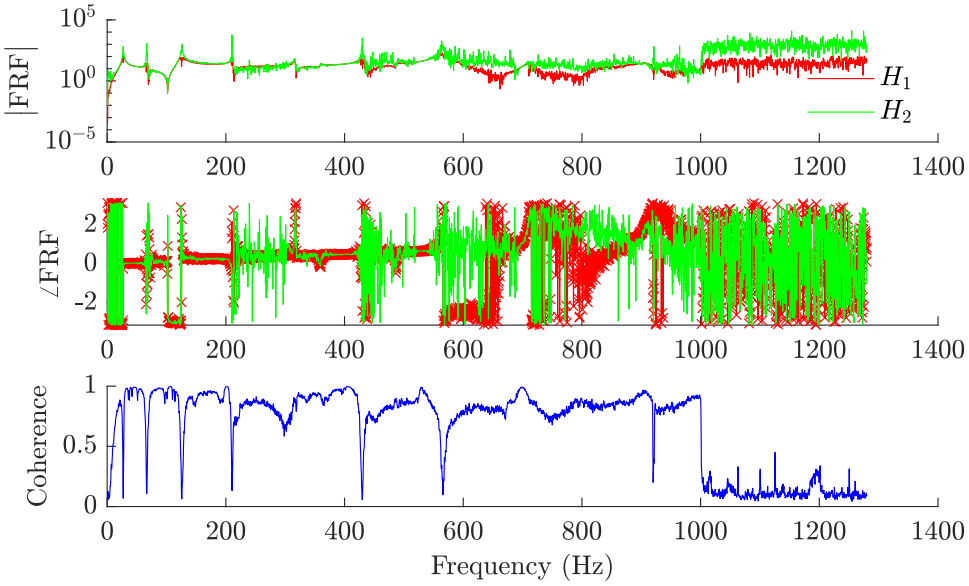


Figure B.6: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of 33.99N

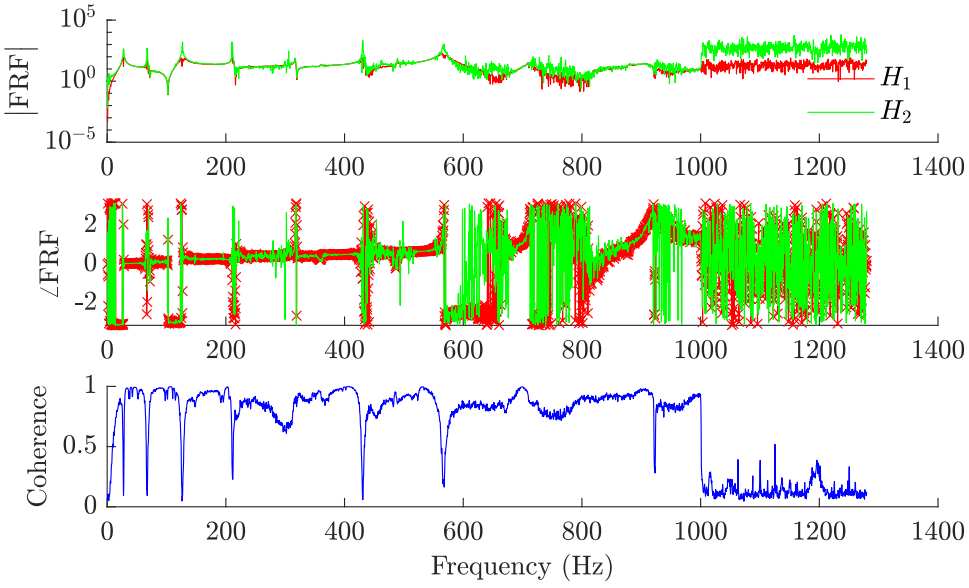


Figure B.7: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of 50.76N

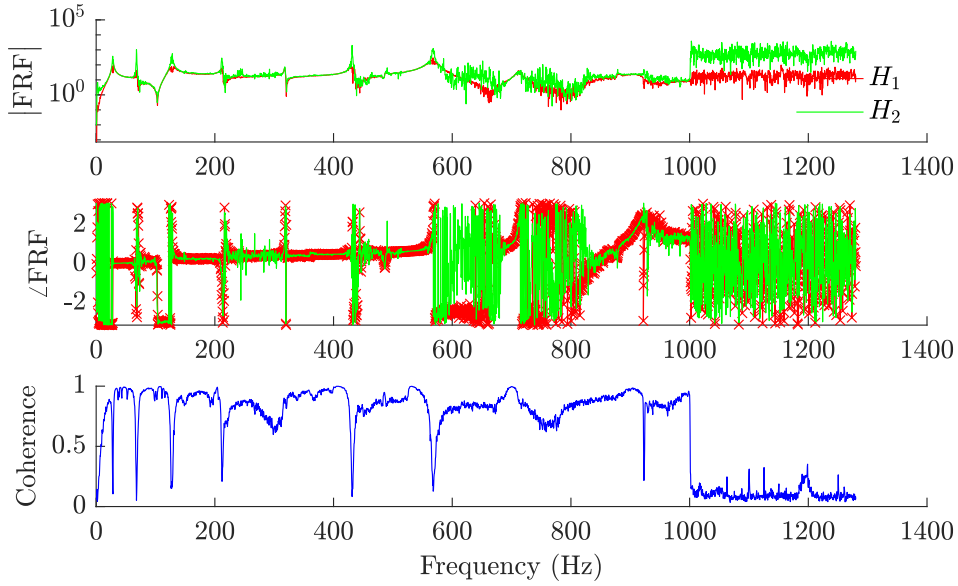


Figure B.8: FRFs extracted for the diagonal strut at a crack depth of 0mm and a static load of 68.49N

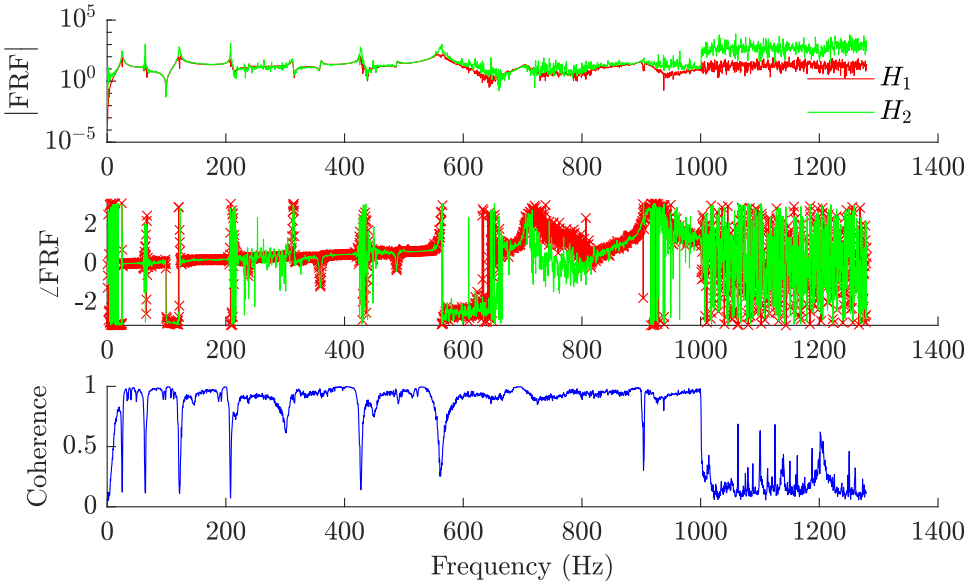


Figure B.9: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of -14.36N

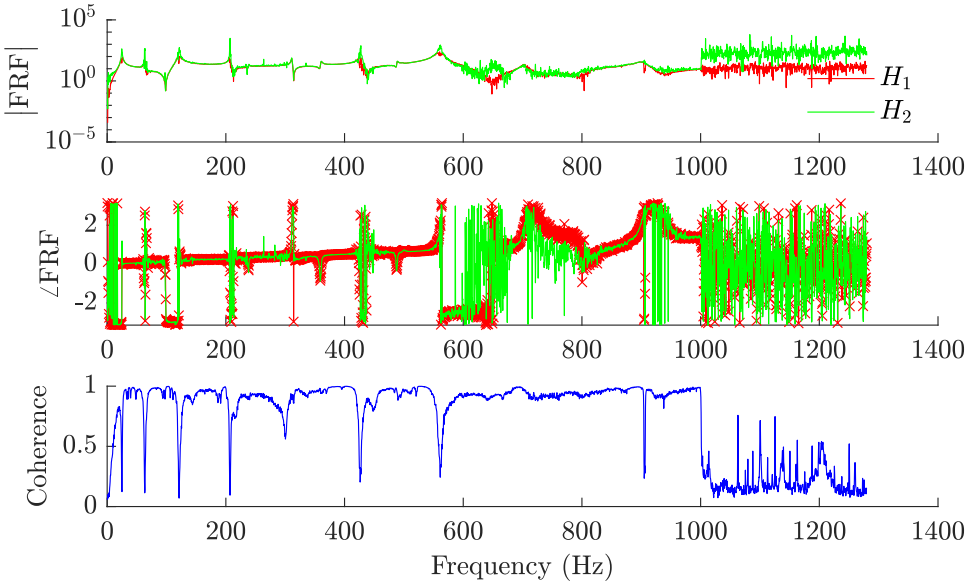


Figure B.10: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of -31.02N

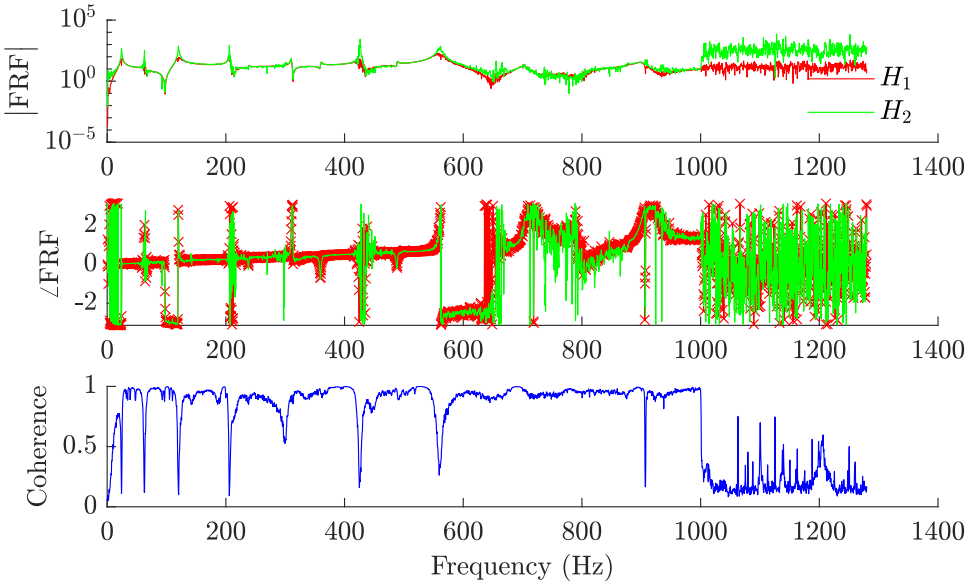


Figure B.11: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of -47.87N

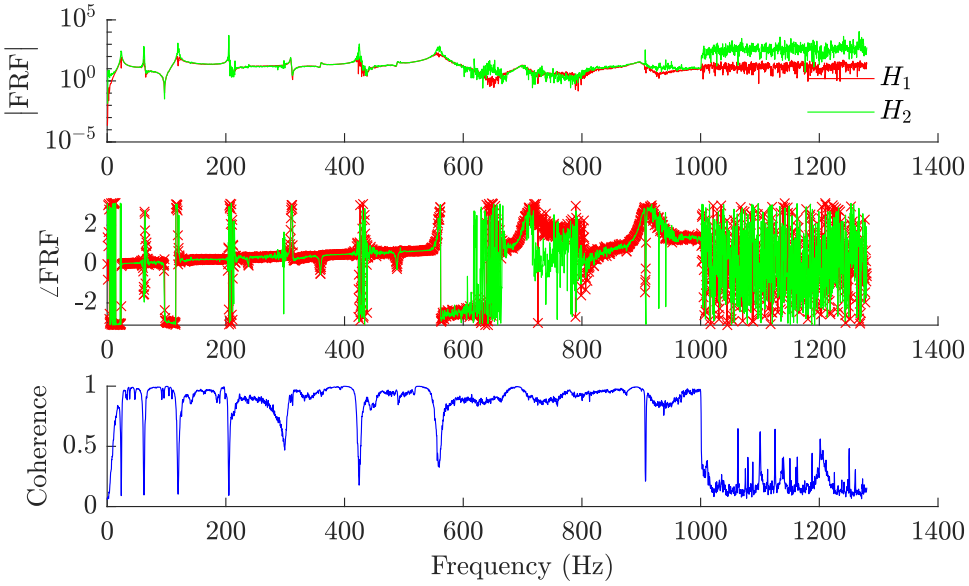


Figure B.12: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of -64.59N

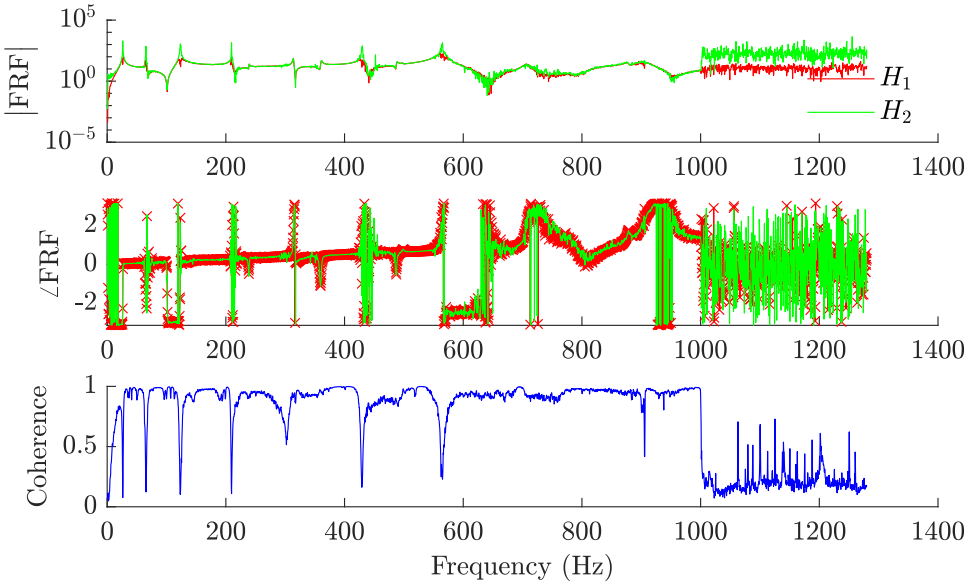


Figure B.13: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of 17.55N

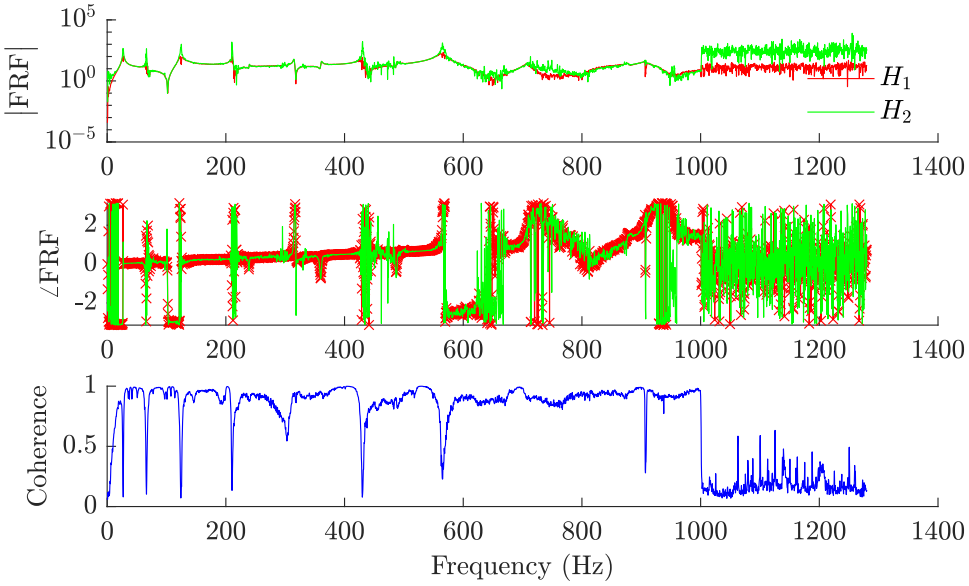


Figure B.14: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of 34.34N

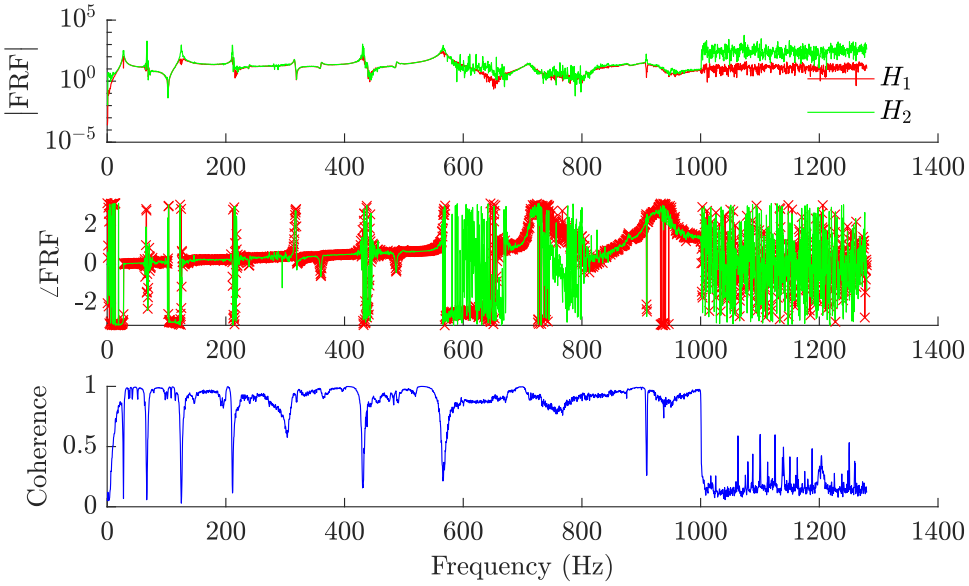


Figure B.15: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of 50.76N

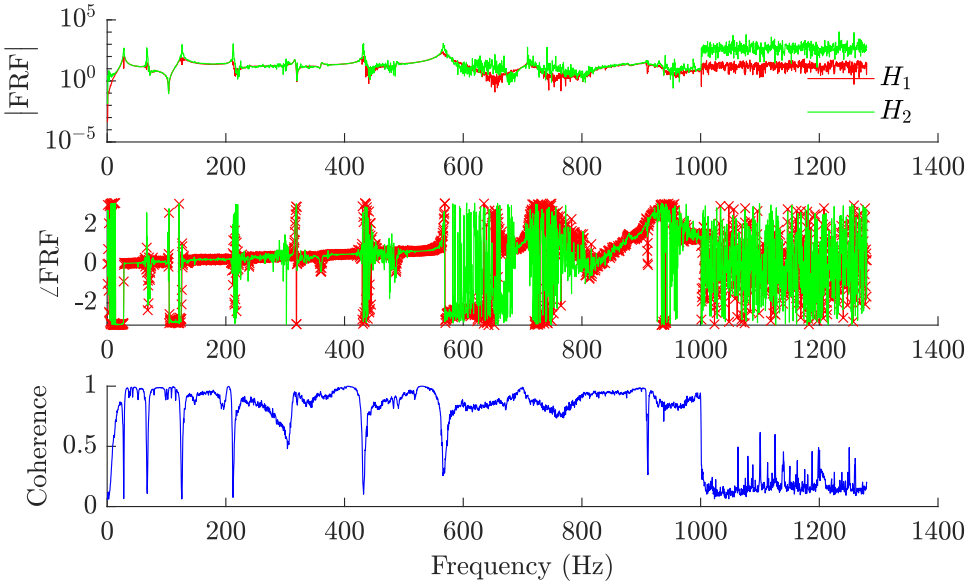


Figure B.16: FRFs extracted for the diagonal strut at a crack depth of 4mm and a static load of 66.65N

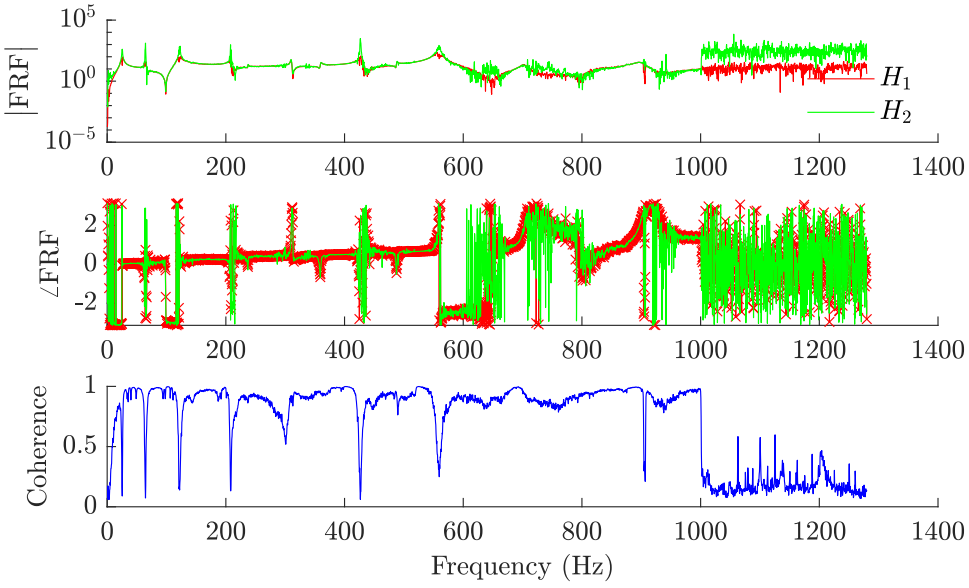


Figure B.17: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of -14.58N

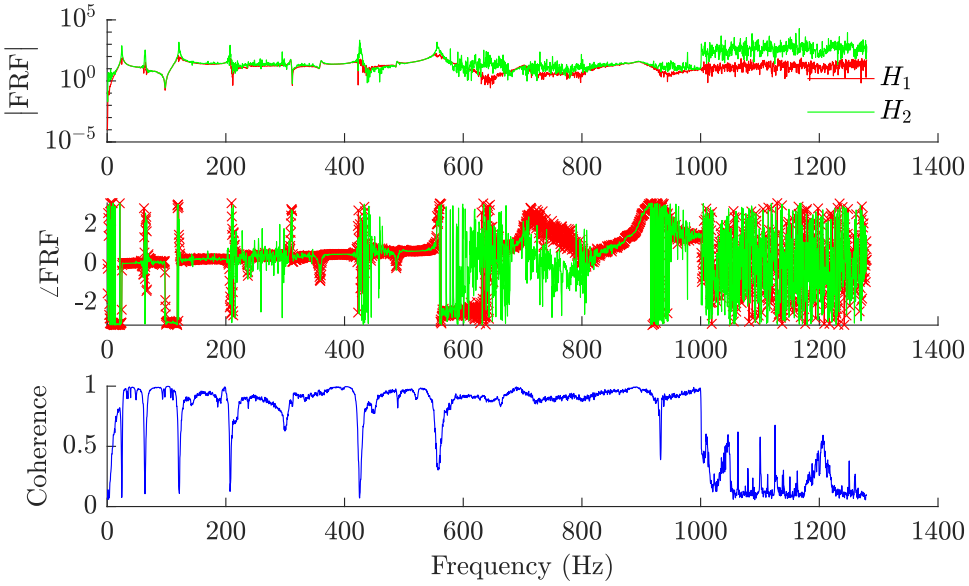


Figure B.18: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of -32.03N

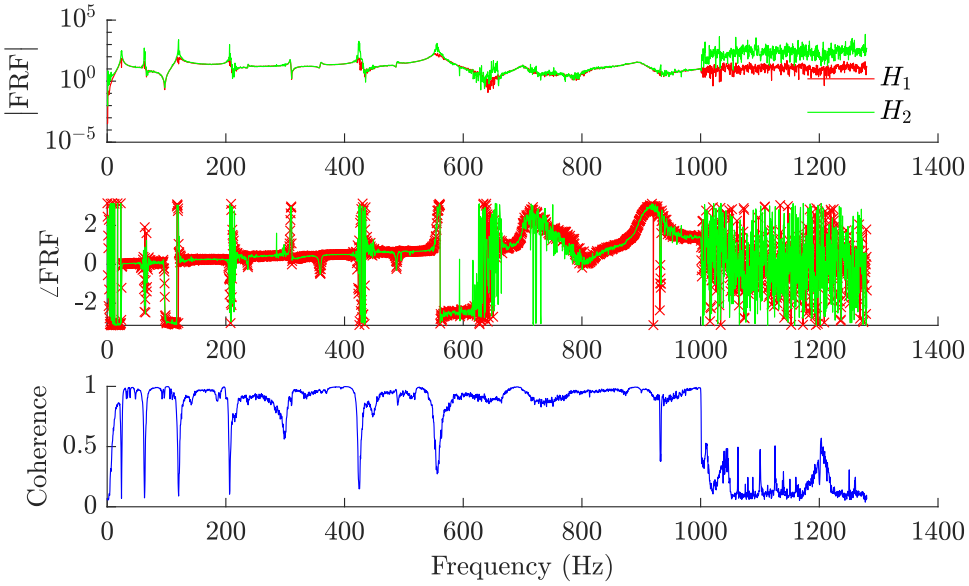


Figure B.19: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of -49.02N

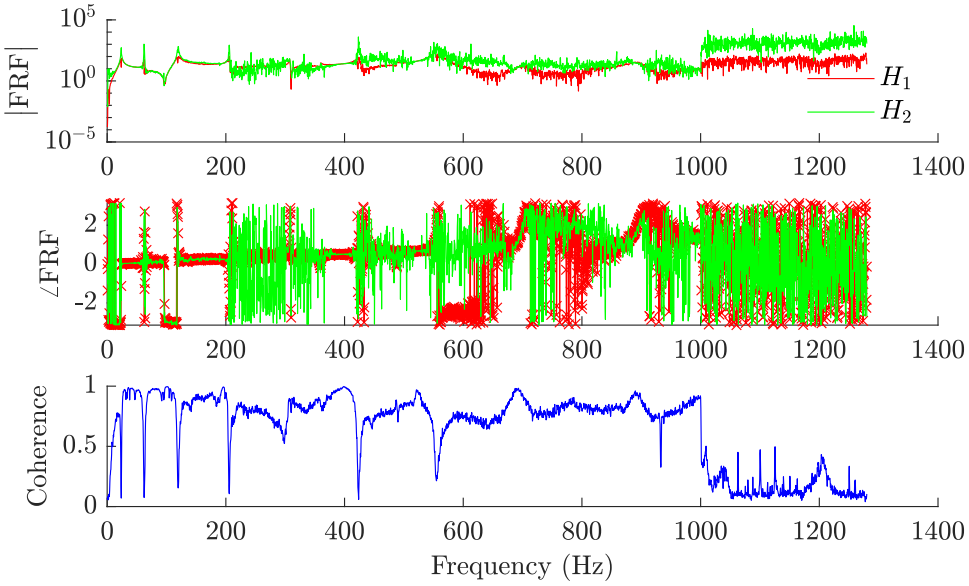


Figure B.20: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of -65.83N

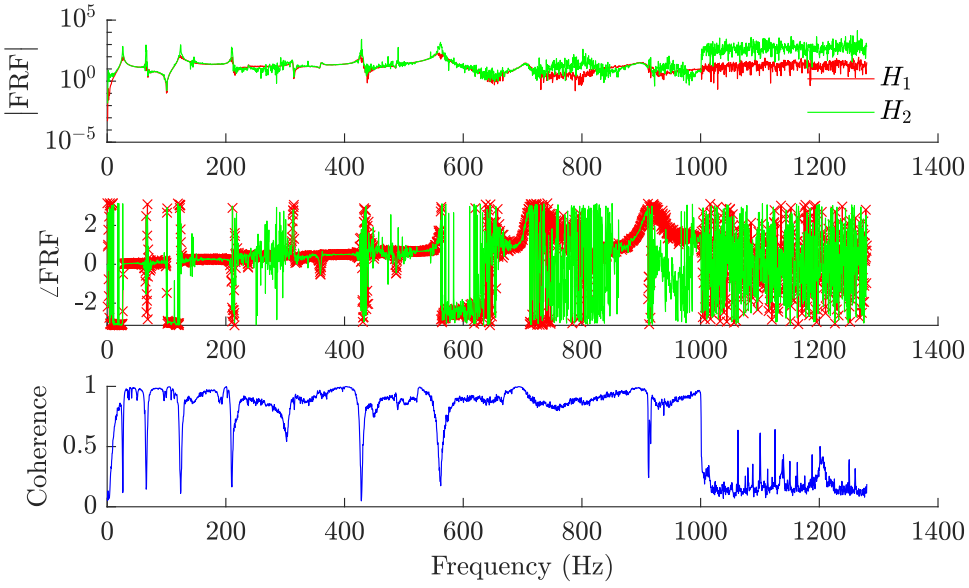


Figure B.21: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of 18.66N

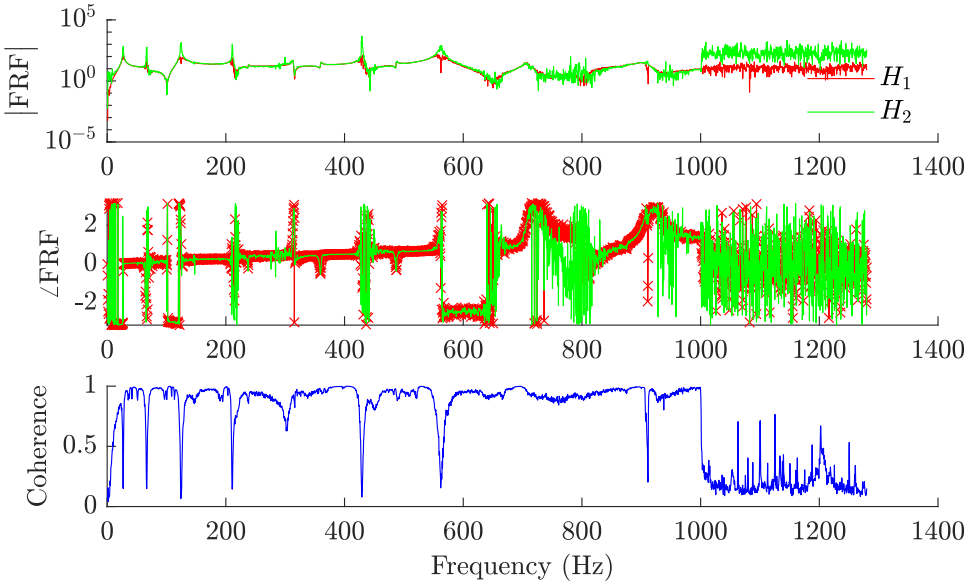


Figure B.22: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of 35.10N

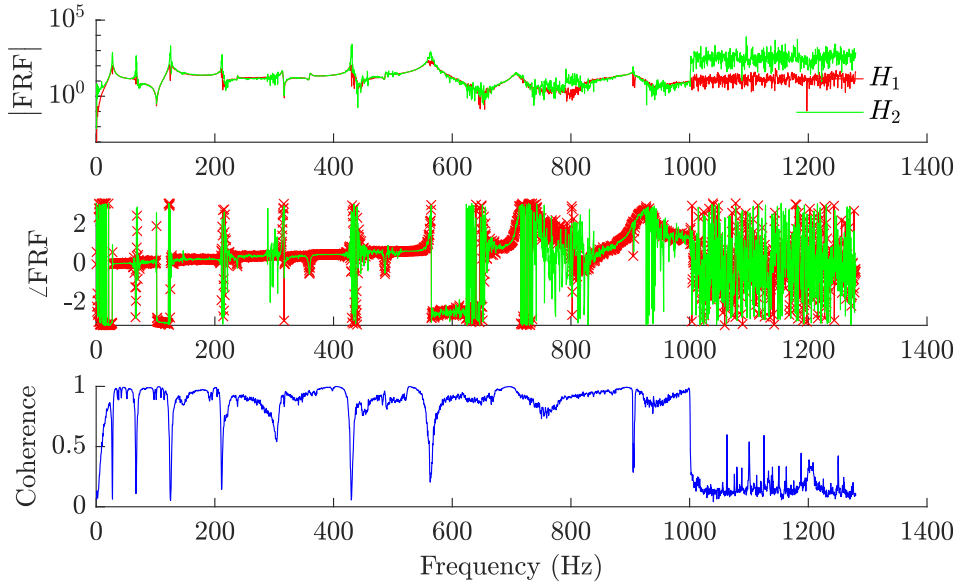


Figure B.23: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of 51.39N

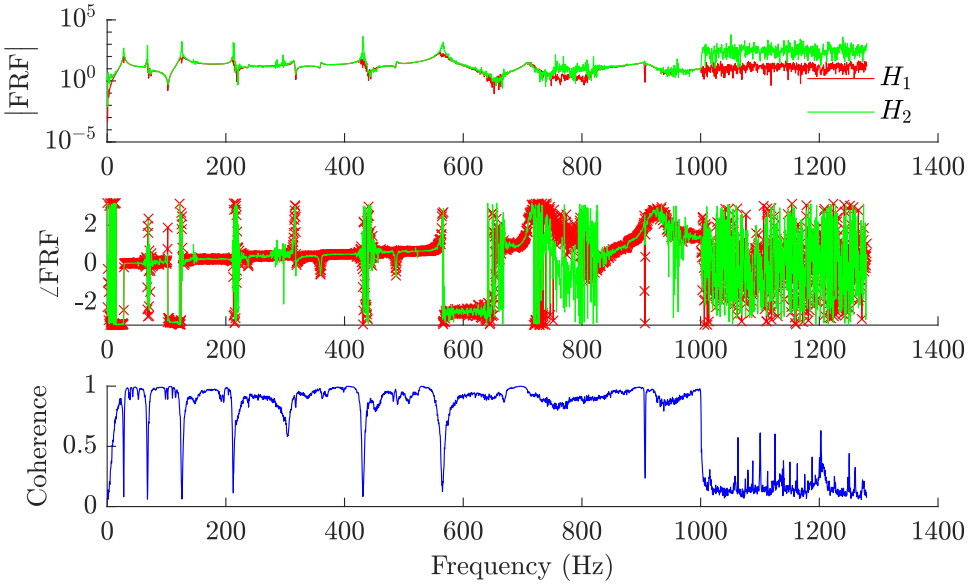


Figure B.24: FRFs extracted for the diagonal strut at a crack depth of 8mm and a static load of 67.57N

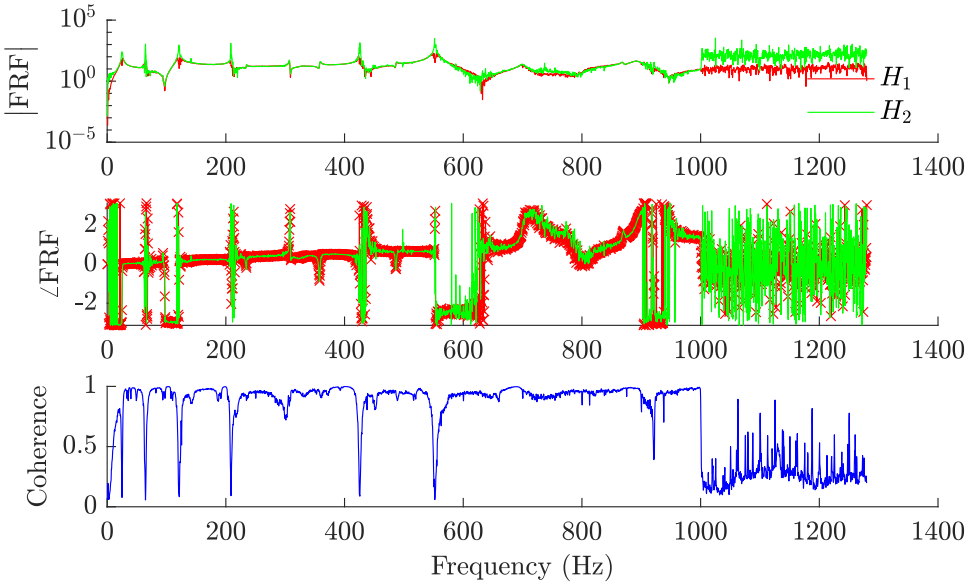


Figure B.25: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of -14.76N

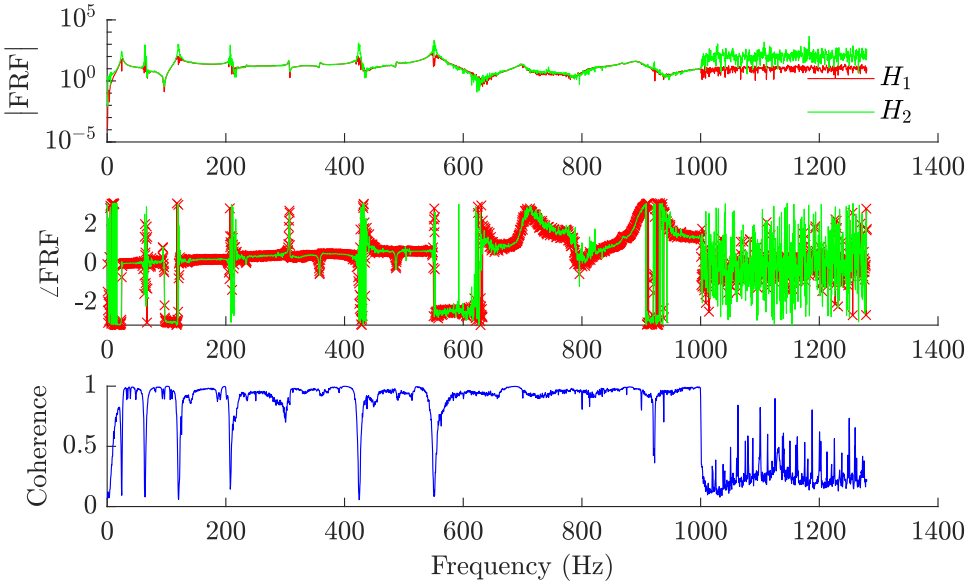


Figure B.26: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of -31.77N

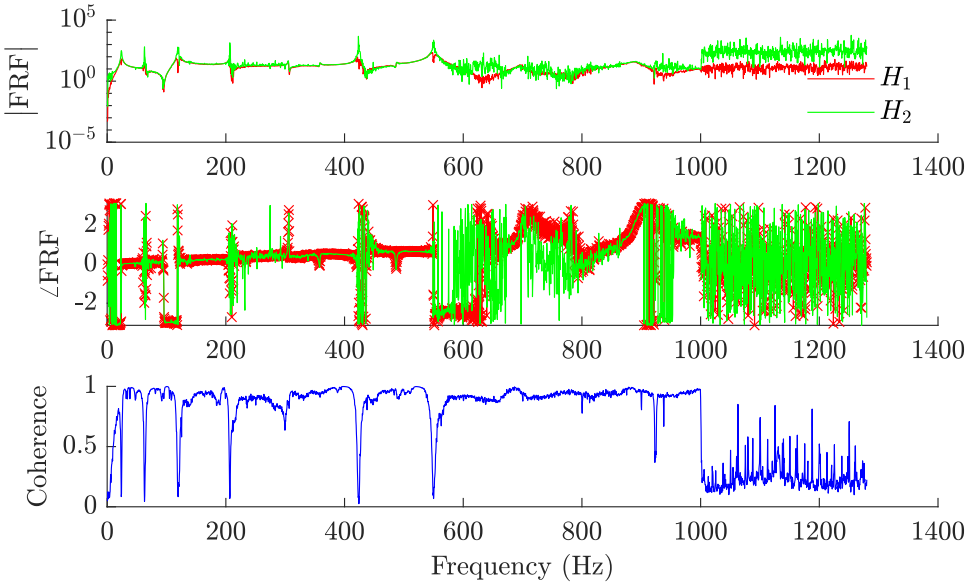


Figure B.27: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of -48.55N

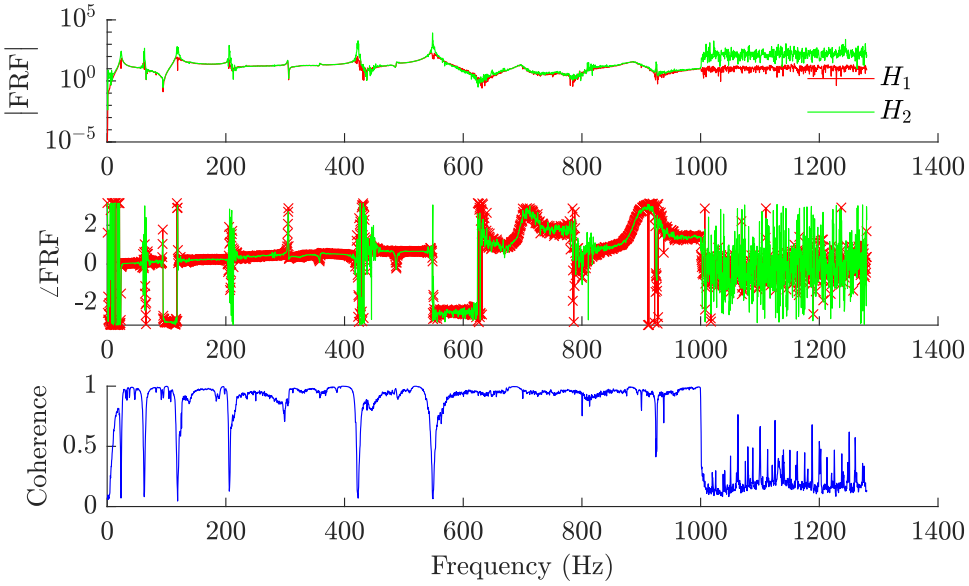


Figure B.28: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of -64.94N

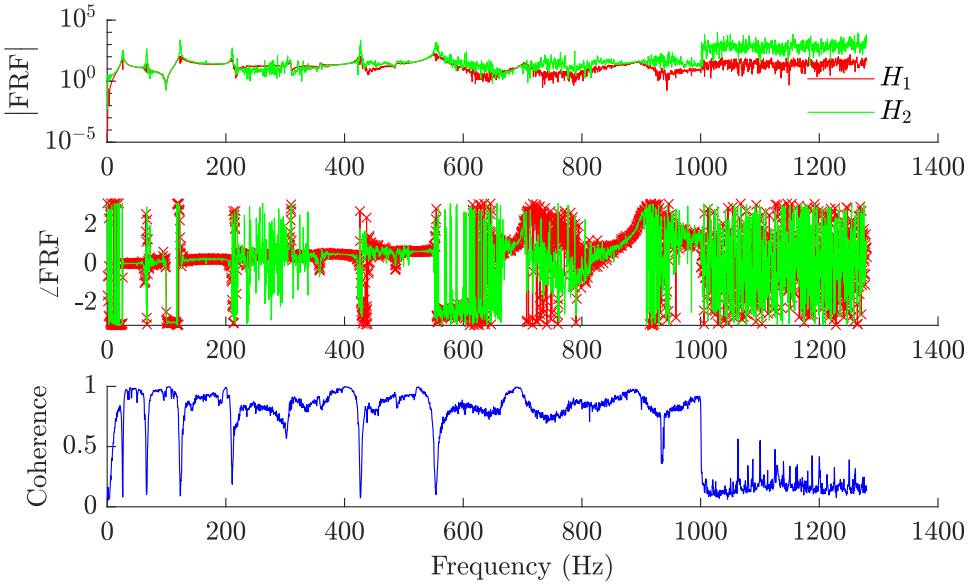


Figure B.29: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of 18.39N

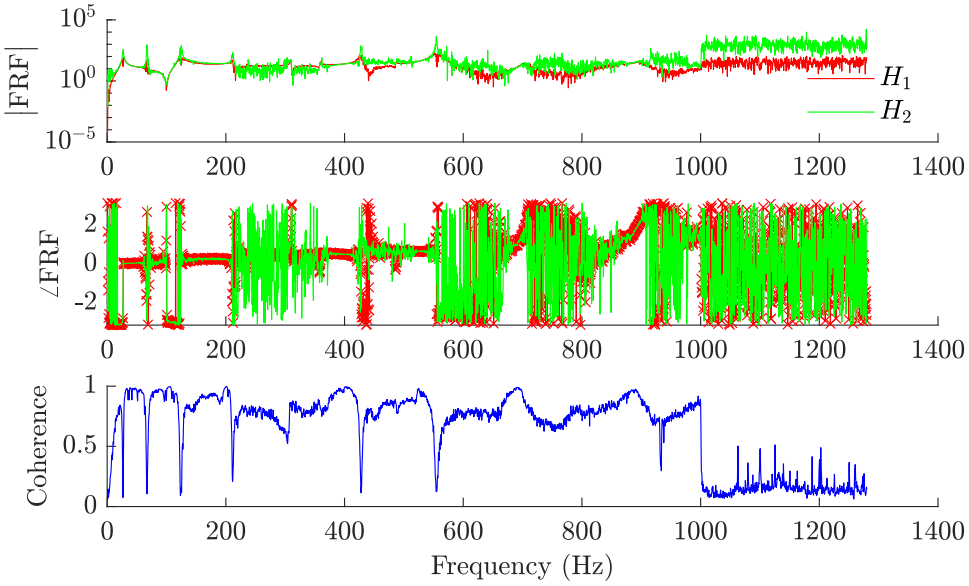


Figure B.30: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of 33.42N

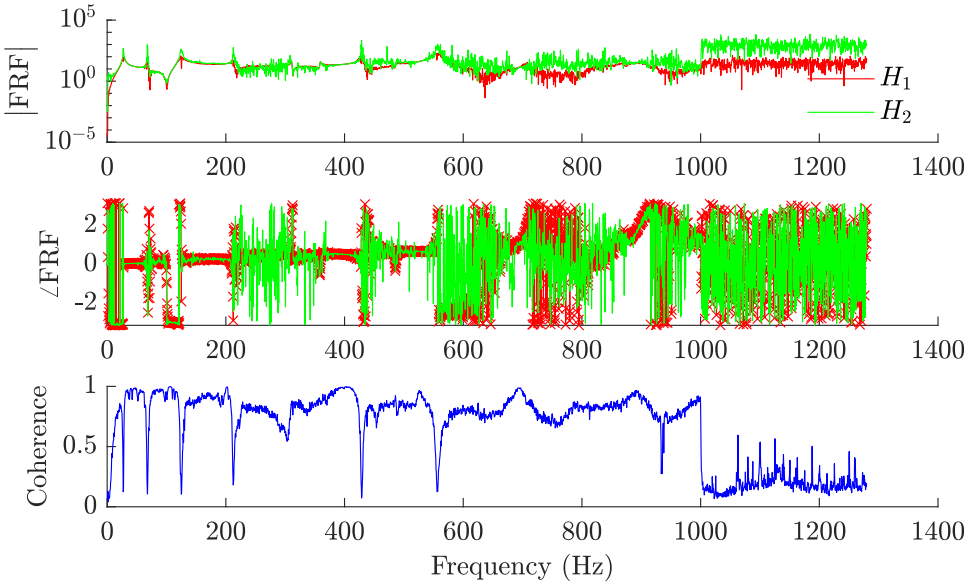


Figure B.31: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of 51.27N

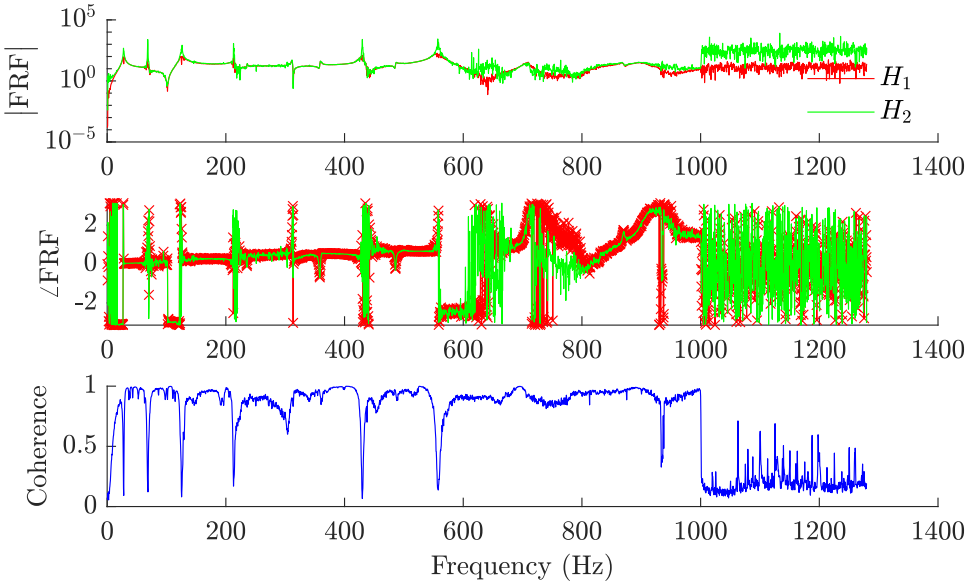


Figure B.32: FRFs extracted for the diagonal strut at a crack depth of 12mm and a static load of 66.59N

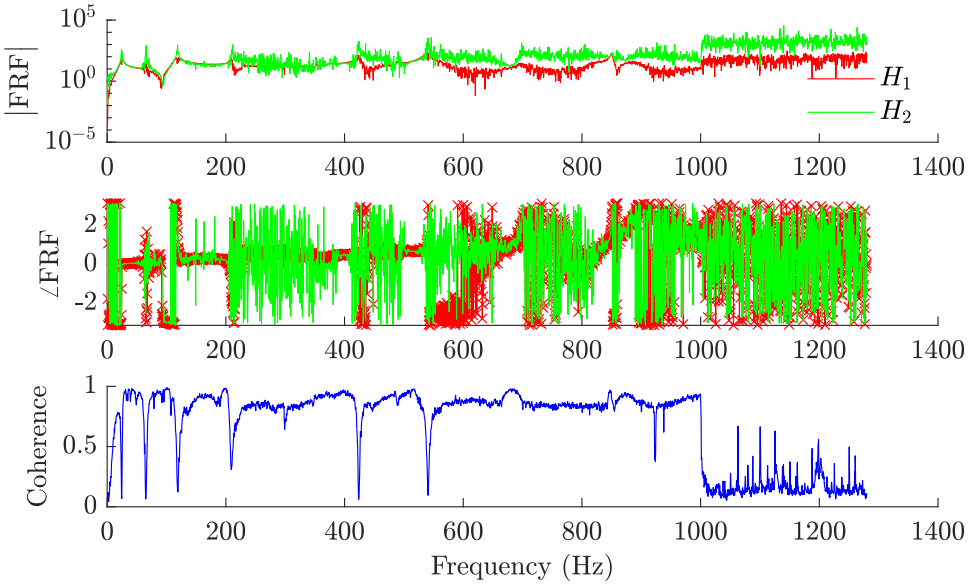


Figure B.33: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of -14.94N

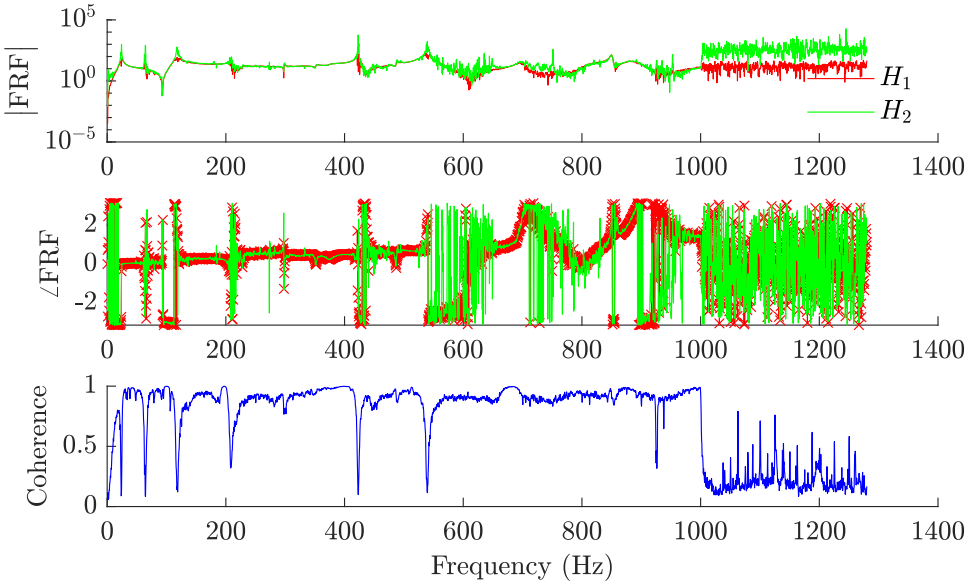


Figure B.34: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of -31.93N

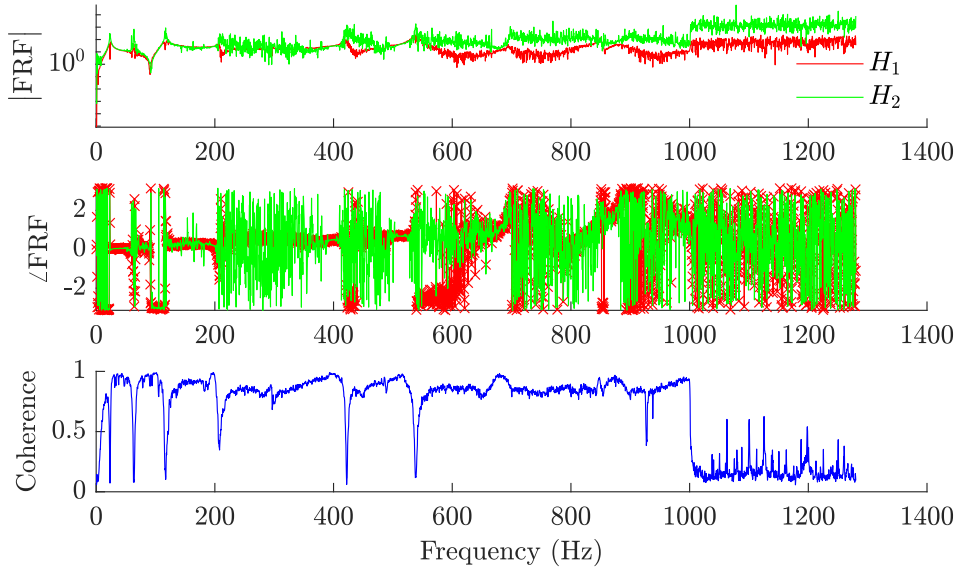


Figure B.35: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of -48.32N

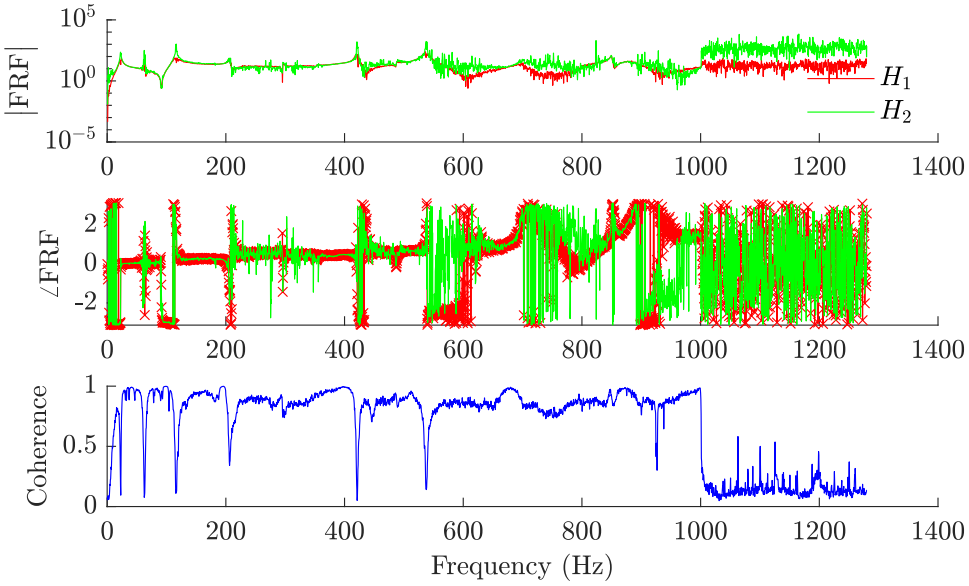


Figure B.36: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of -64.91N

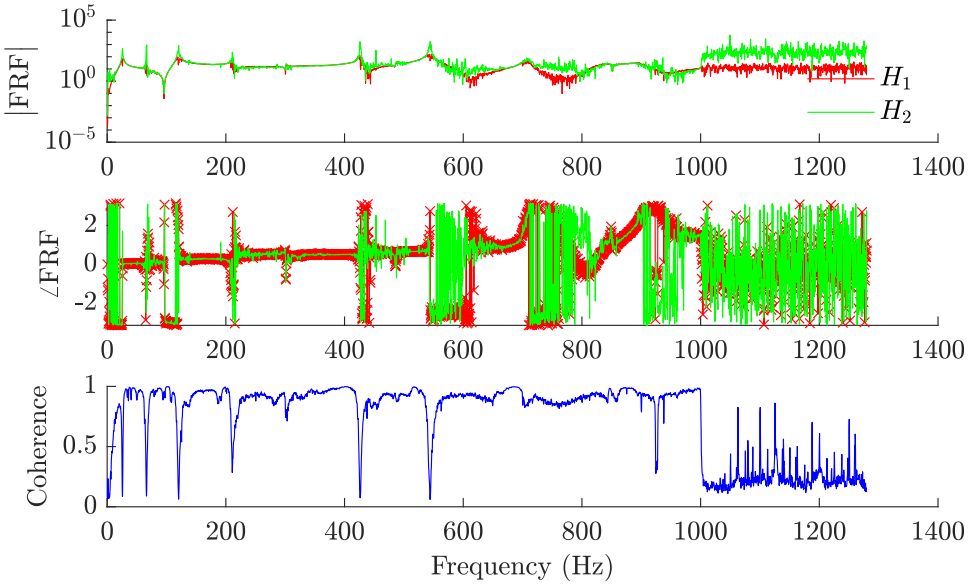


Figure B.37: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of 18.79N

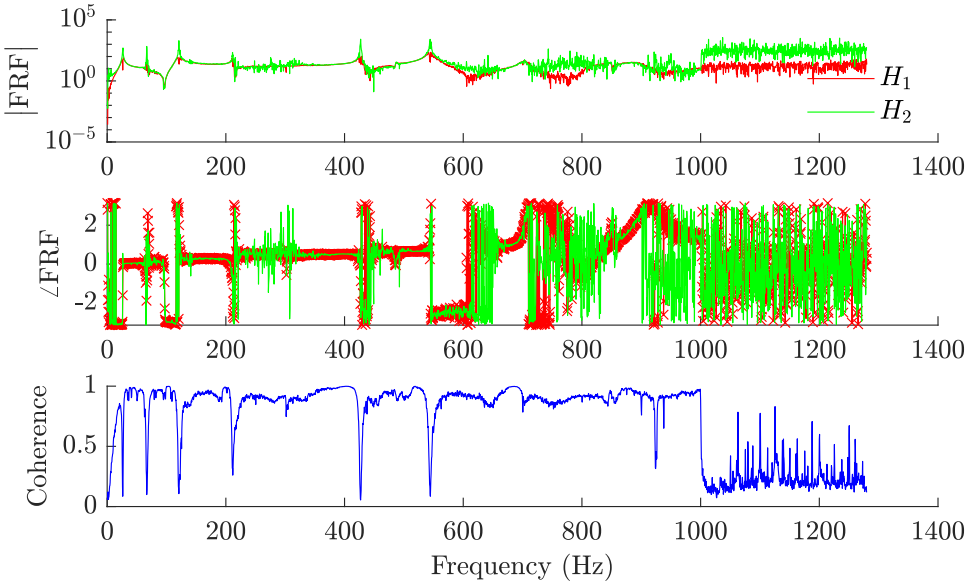


Figure B.38: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of 36.70N

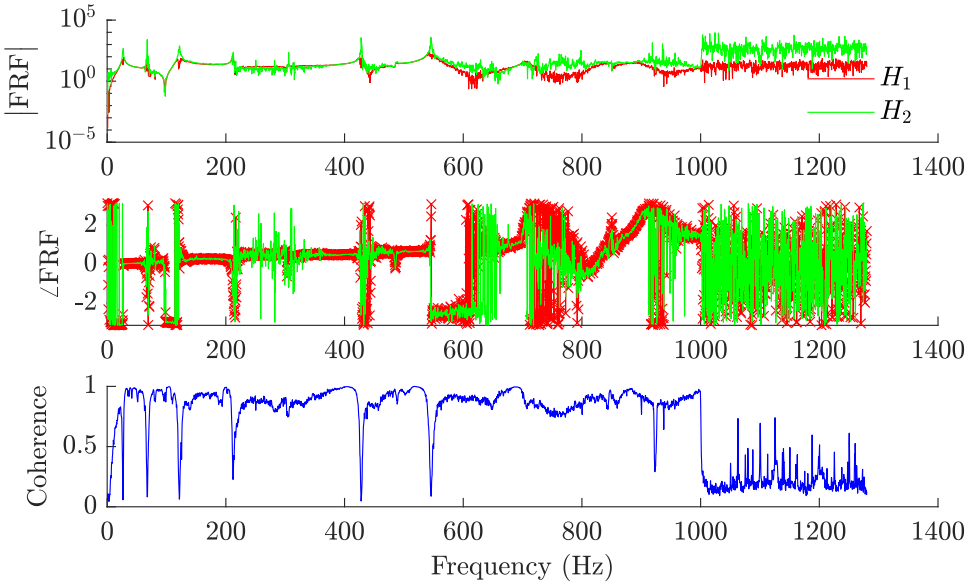


Figure B.39: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of 50.33N

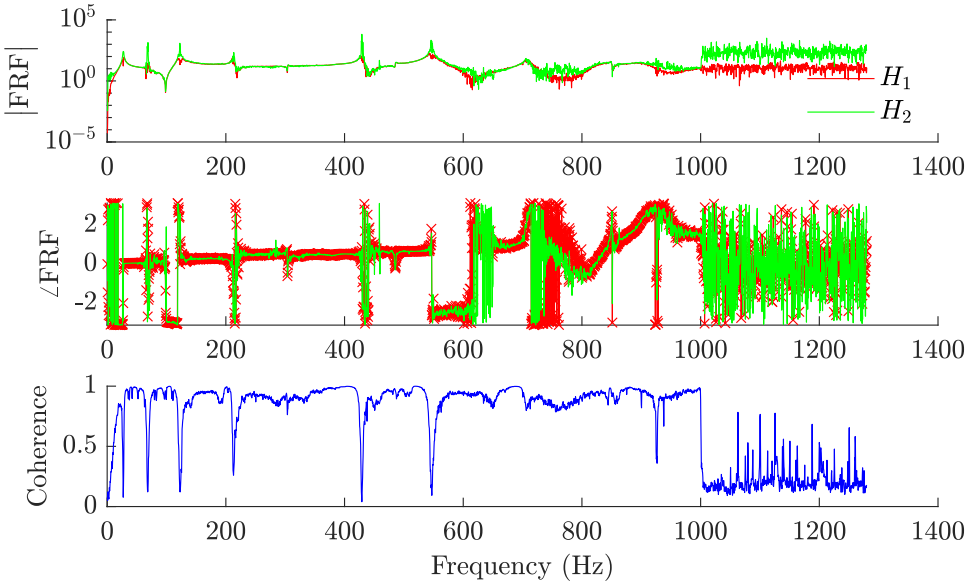


Figure B.40: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of 66.77N

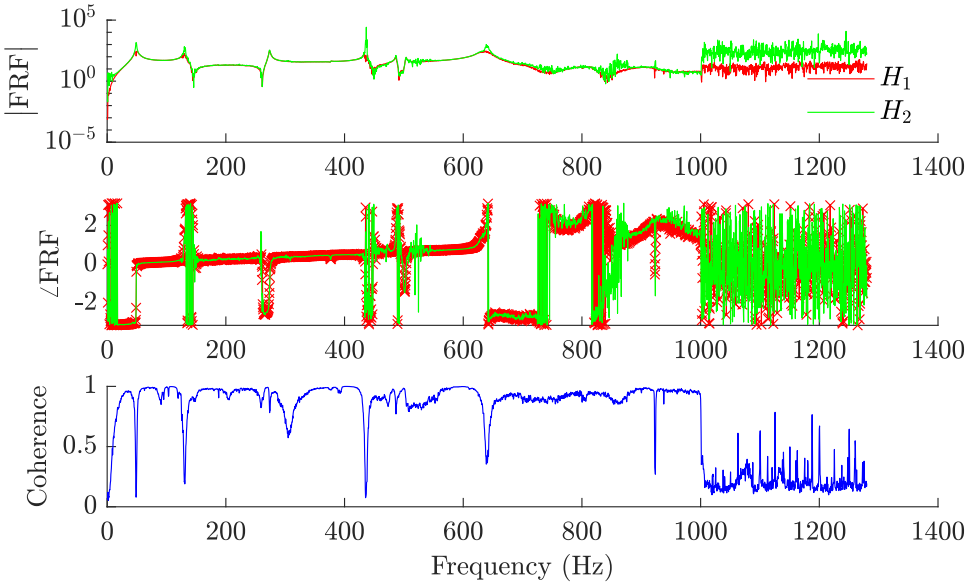


Figure B.41: FRFs extracted for the vertical strut at a crack depth of 0mm and a static load of -30.98N

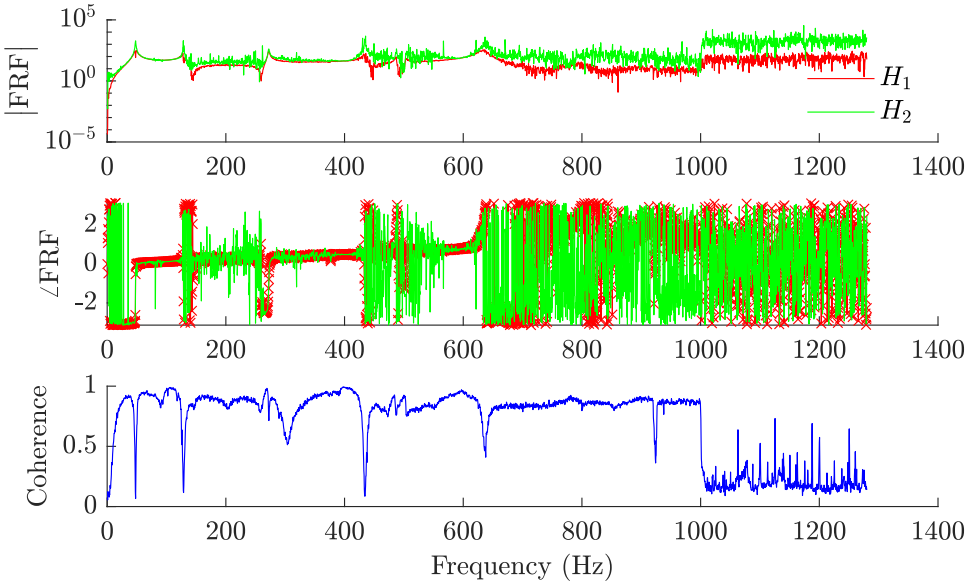


Figure B.42: FRFs extracted for the vertical strut at a crack depth of 0mm and a static load of -63.84N

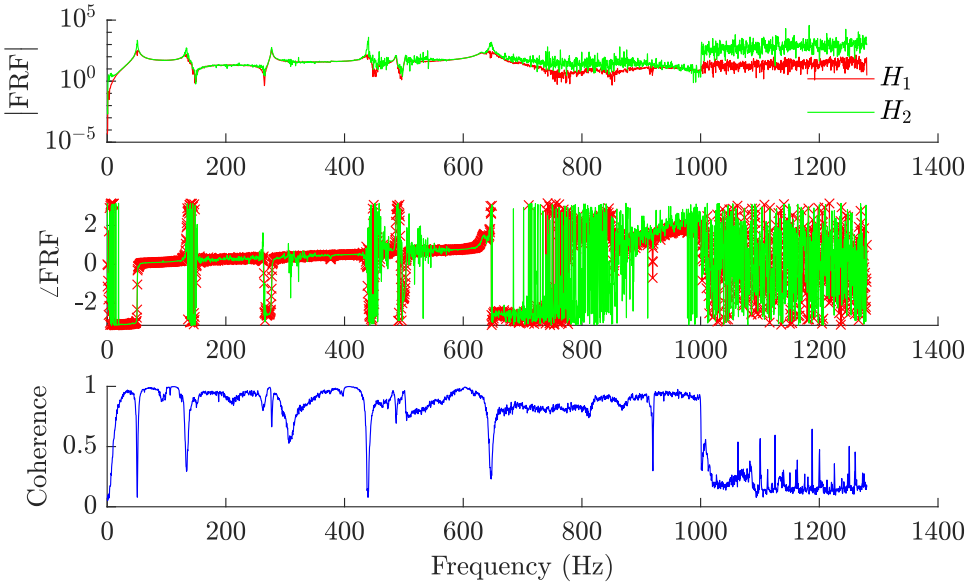


Figure B.43: FRFs extracted for the vertical strut at a crack depth of 0mm and a static load of 32.60N

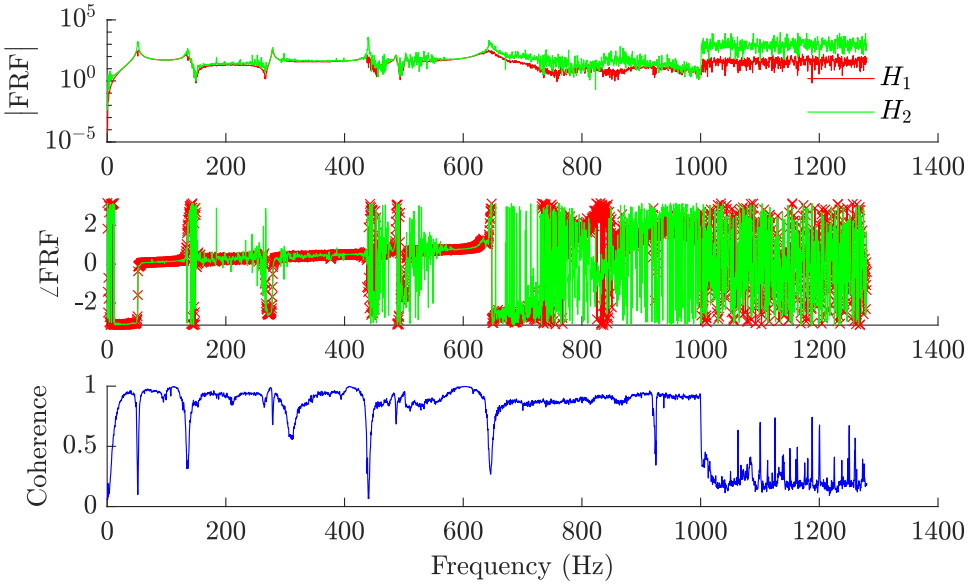


Figure B.44: FRFs extracted for the vertical strut at a crack depth of 0mm and a static load of 65.15N

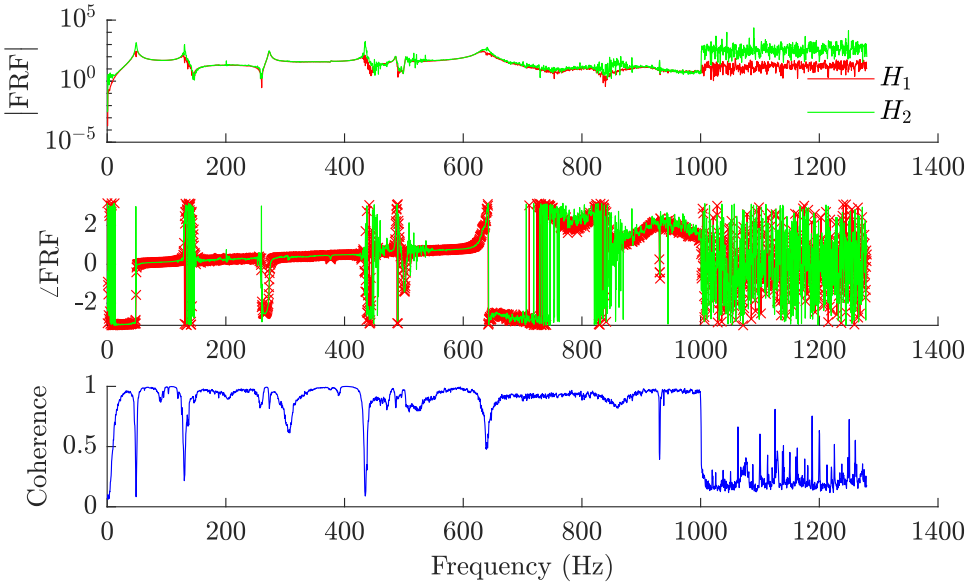


Figure B.45: FRFs extracted for the vertical strut at a crack depth of 4mm and a static load of -30.09N

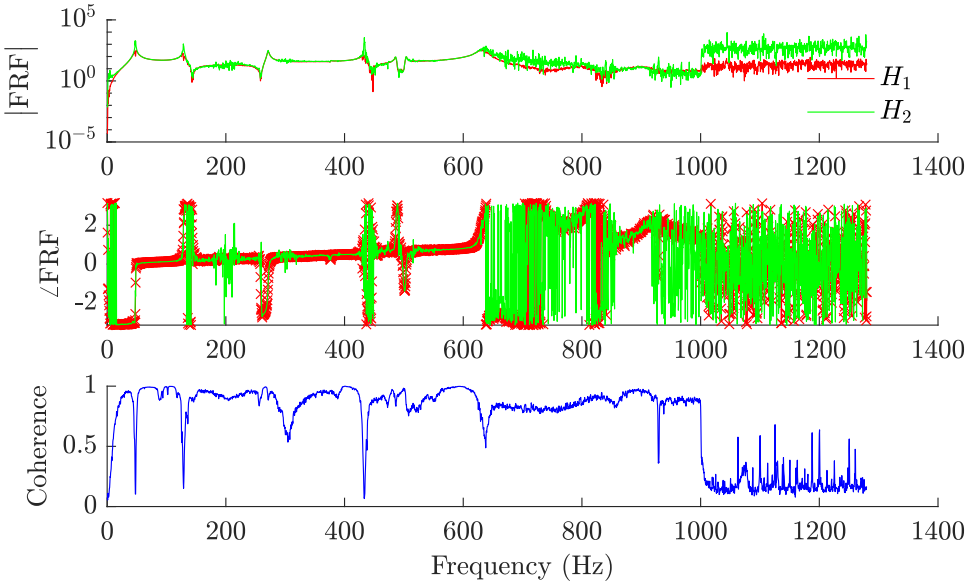


Figure B.46: FRFs extracted for the vertical strut at a crack depth of 4mm and a static load of -62.93N

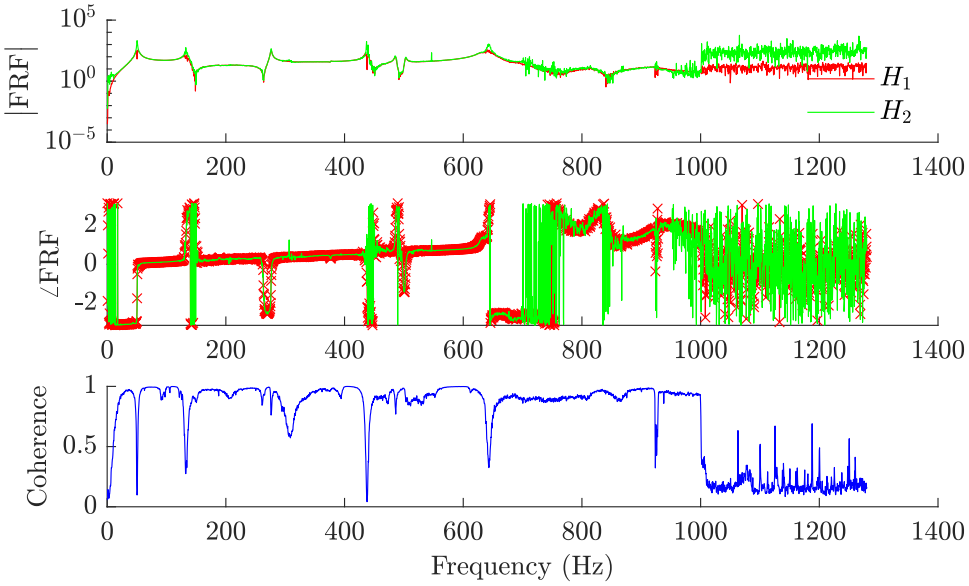


Figure B.47: FRFs extracted for the vertical strut at a crack depth of 4mm and a static load of 32.97N

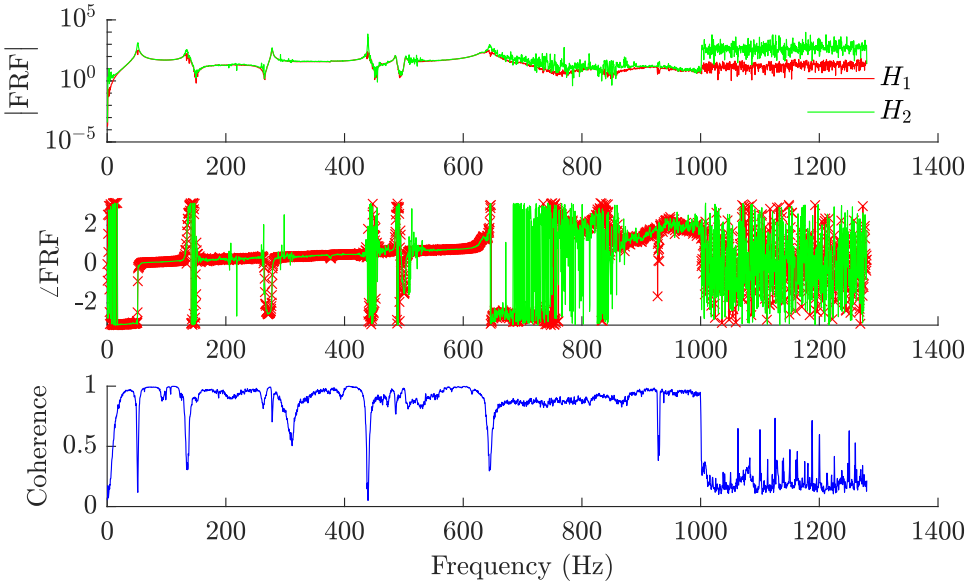


Figure B.48: FRFs extracted for the vertical strut at a crack depth of 4mm and a static load of 65.68N

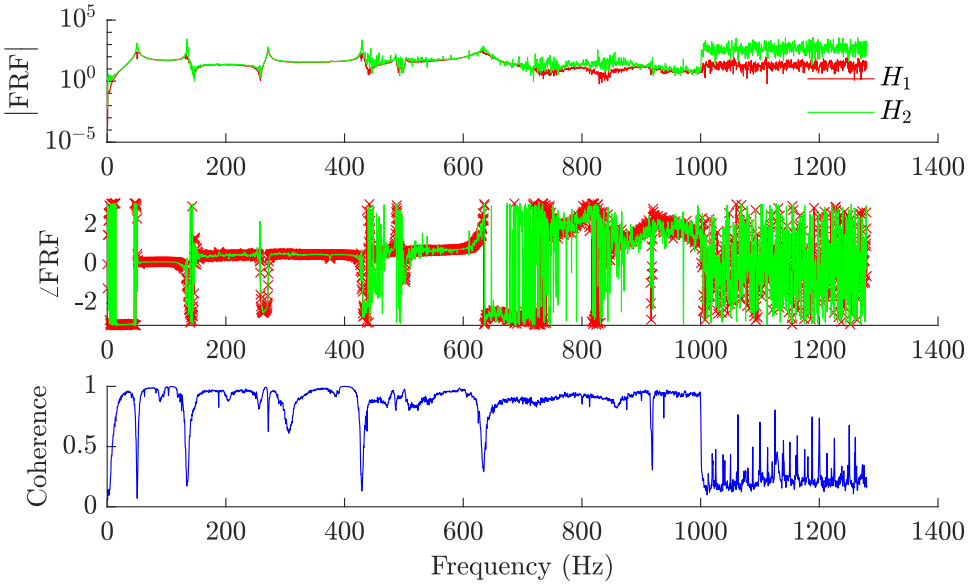


Figure B.49: FRFs extracted for the vertical strut at a crack depth of 8mm and a static load of -31.25N

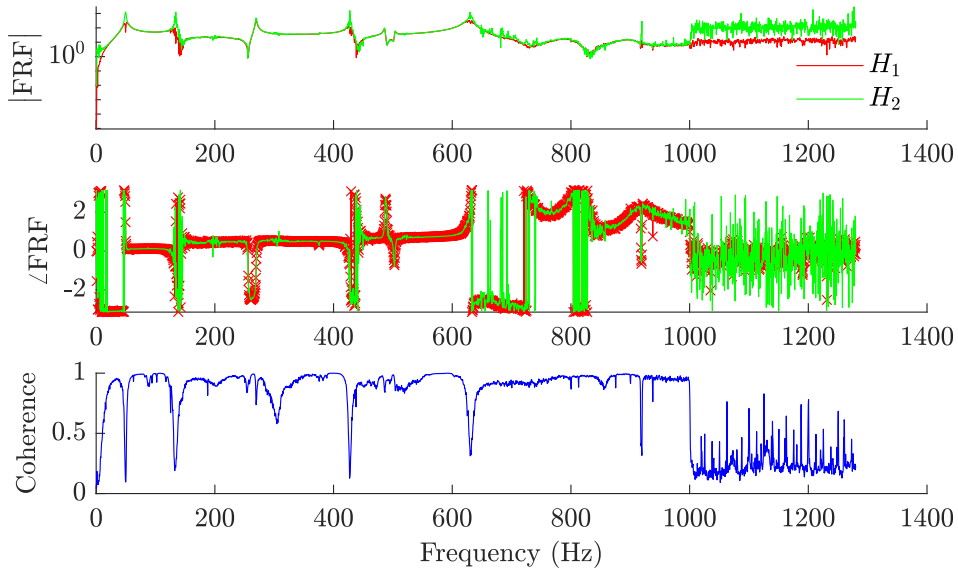


Figure B.50: FRFs extracted for the vertical strut at a crack depth of 8mm and a static load of -64.00N

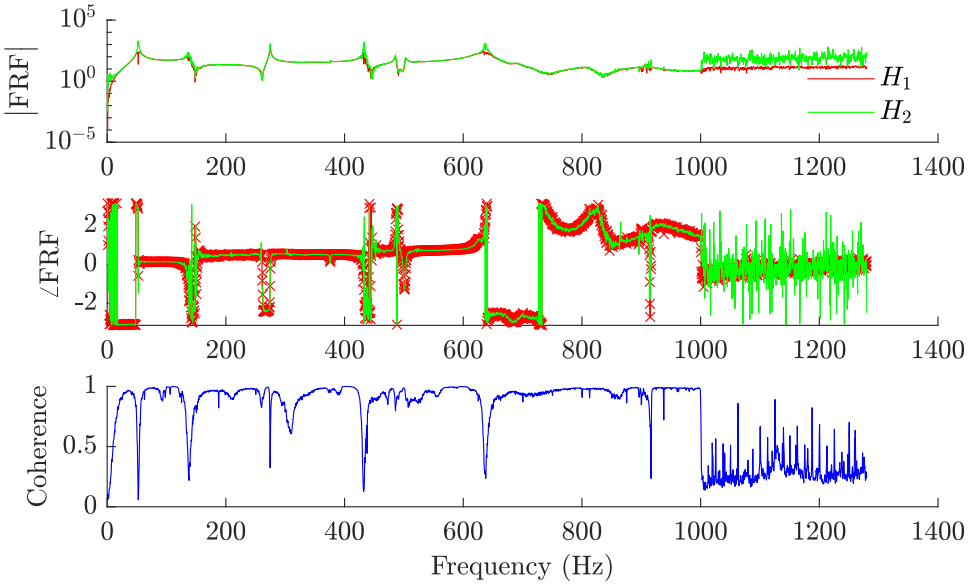


Figure B.51: FRFs extracted for the vertical strut at a crack depth of 8mm and a static load of 33.06N

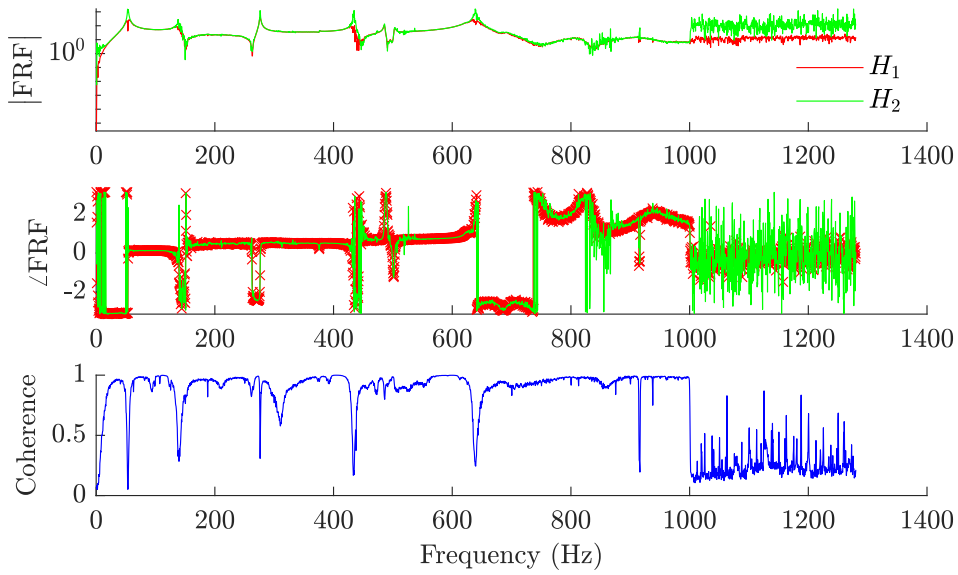


Figure B.52: FRFs extracted for the vertical strut at a crack depth of 8mm and a static load of 65.55N

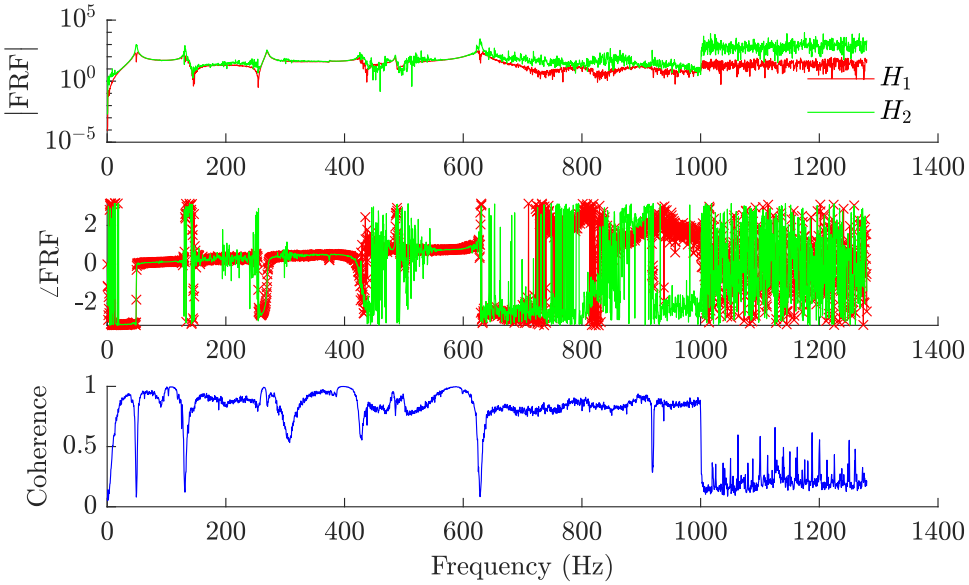


Figure B.53: FRFs extracted for the vertical strut at a crack depth of 12mm and a static load of -31.43N

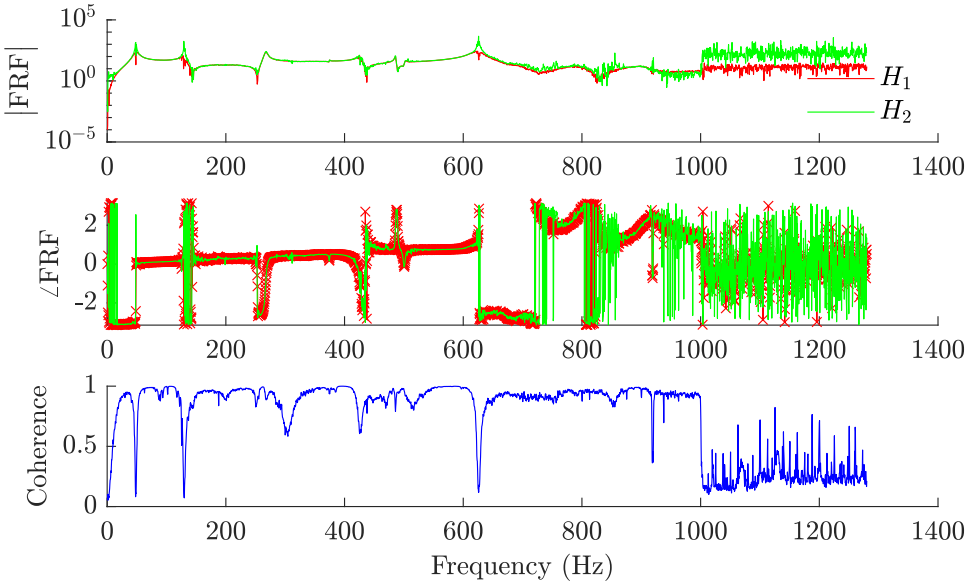


Figure B.54: FRFs extracted for the vertical strut at a crack depth of 12mm and a static load of -64.11N

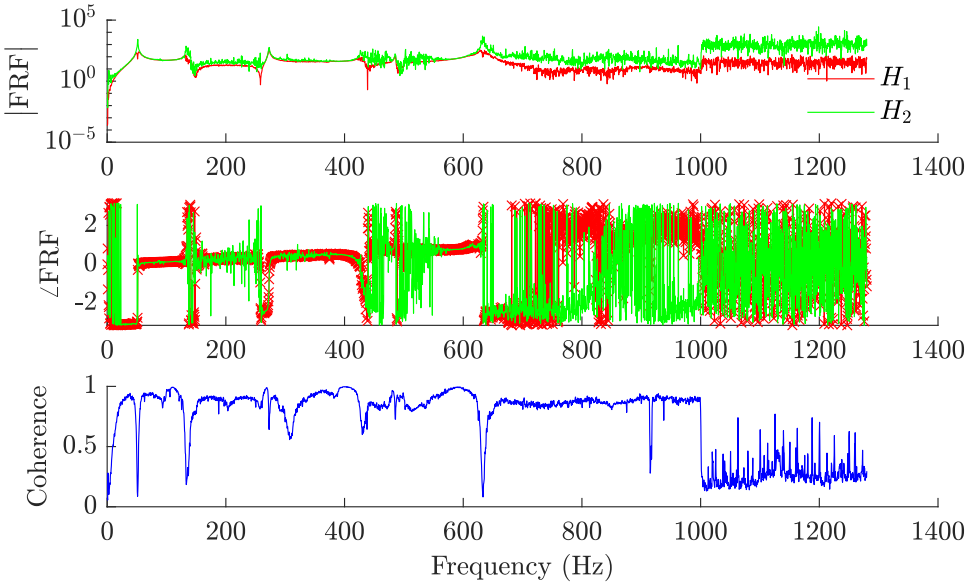


Figure B.55: FRFs extracted for the vertical strut at a crack depth of 12mm and a static load of 33.60N

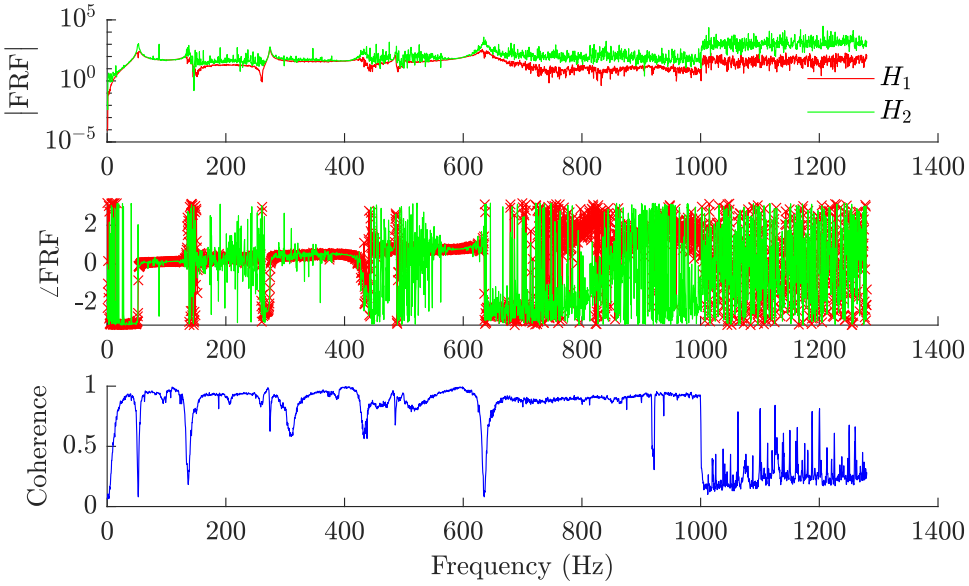


Figure B.56: FRFs extracted for the vertical strut at a crack depth of 12mm and a static load of 63.83N

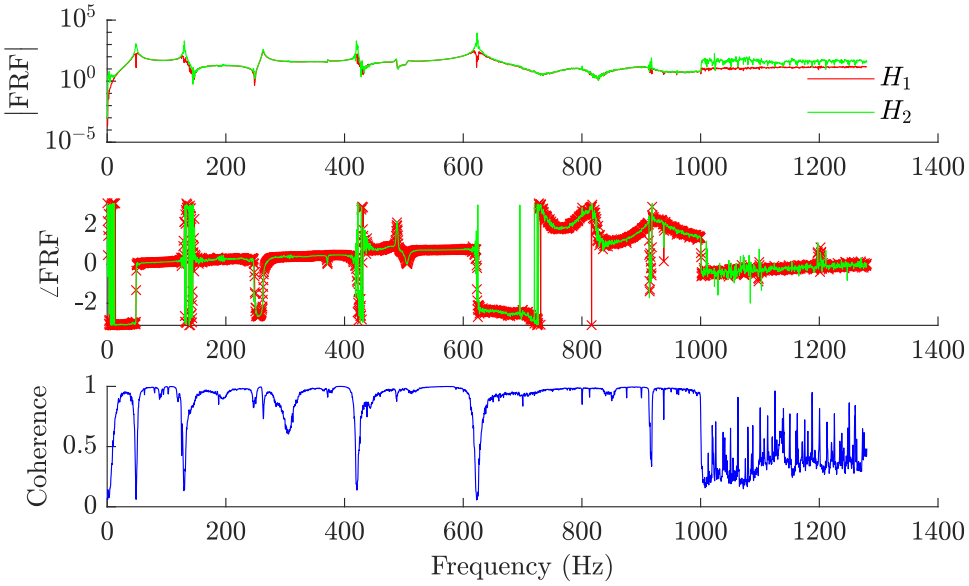


Figure B.57: FRFs extracted for the diagonal strut at a crack depth of 16mm and a static load of -30.43N

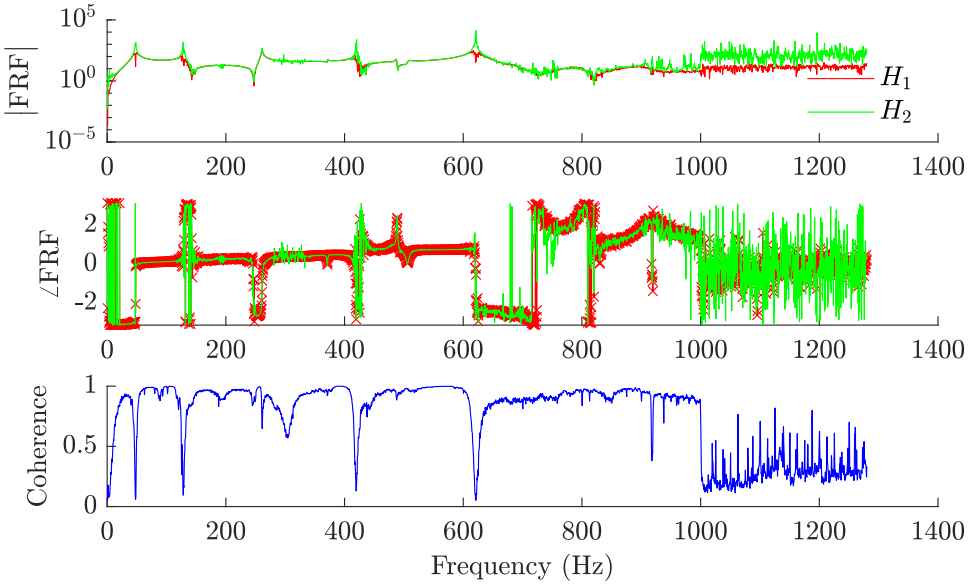


Figure B.58: FRFs extracted for the vertical strut at a crack depth of 16mm and a static load of -62.99N

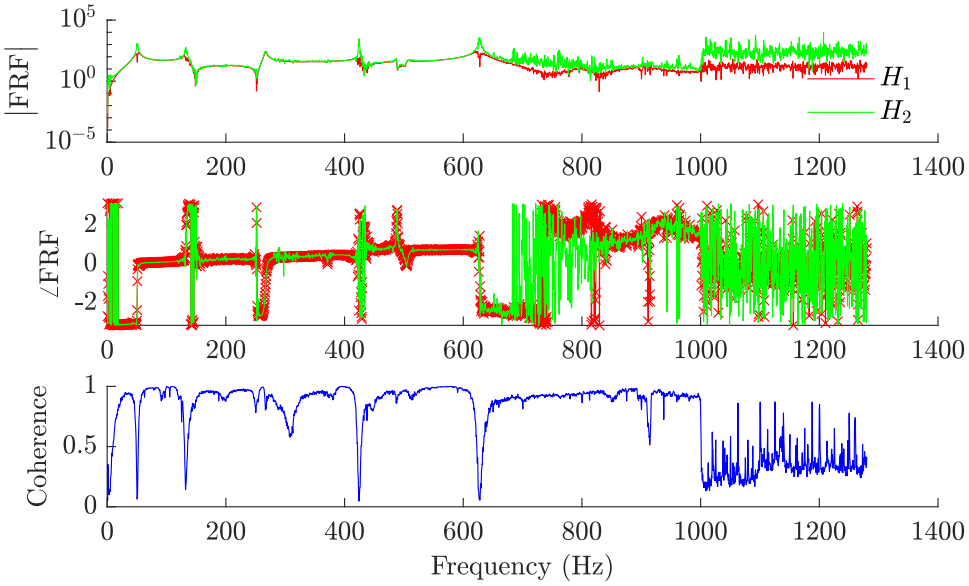


Figure B.59: FRFs extracted for the vertical strut at a crack depth of 16mm and a static load of 33.04N

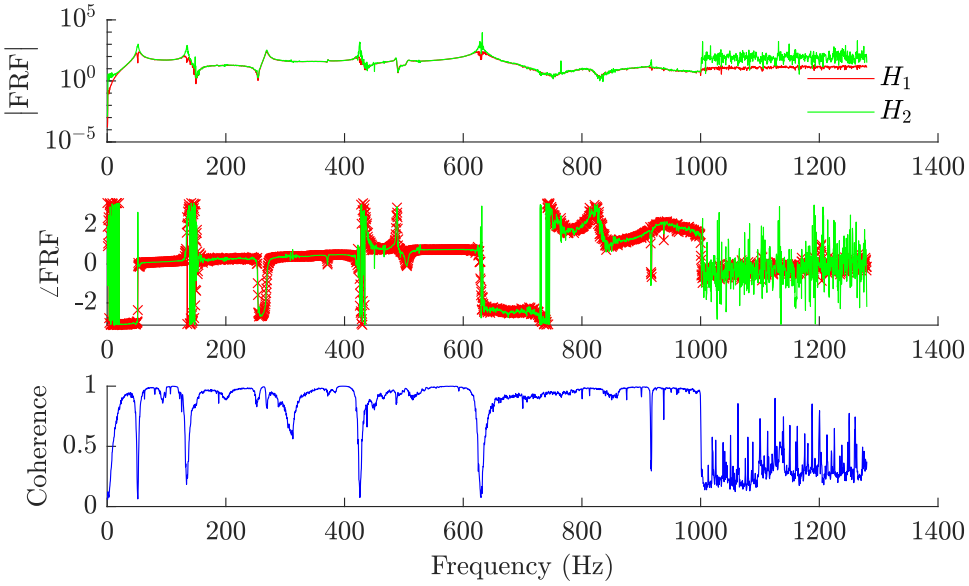


Figure B.60: FRFs extracted for the vertical strut at a crack depth of 16mm and a static load of 65.25N

B.2 Modal data

Tables of the modal analysis results are presented as follows. These were estimated using a single-DoF curvefitting algorithm since the resonance peaks were observed to be relatively well-separated. The algorithm used returned some damping ratio values that were just below zero (which is not physically possible); because the damping ratios were not to be used in this thesis these were ignored. The algorithm returned these values because it used an unconstrained optimisation process to carry out the curve fit; given that the strut was extremely lightly damped, damping ratios close to zero could be wrongly estimated in this way.

The load values given in the following tables are averaged across the true values for each damage extent. The individual load values at each test point can be found in Table 6.7.

		Natural frequencies (Hz)							Damping ratios							
Average load (N)		-3	-20	-37	-53	30	47	63	79	-3	-20	-37	-53	30	47	63	79
Mode 1	0mm	25.85	24.71	24.16	23.84	26.53	26.94	27.79	28.15	-0.001	-0.002	0.003	0.009	-0.003	-0.009	0.003	0.001
	4mm	24.66	24.77	23.63	23.82	26.17	27.03	27.61	27.77	-0.004	-0.006	0.001	-0.001	-0.003	0.004	0.001	0.004
	8mm	24.92	24.88	24.34	23.92	26.26	26.81	27.39	27.90	-0.002	-0.004	-0.010	-0.007	0.000	-0.002	0.001	-0.007
	12mm	25.02	24.06	23.44	23.46	26.22	26.88	27.63	27.74	0.007	-0.004	0.002	-0.005	0.001	0.002	0.003	-0.004
	16mm	24.36	23.57	23.34	22.86	25.50	26.50	26.94	27.05	0.001	0.013	-0.006	-0.002	0.000	-0.008	0.000	-0.006
Mode 2	0mm	63.86	63.69	62.95	61.71	65.97	66.47	66.83	67.84	0.000	-0.003	0.001	0.003	0.000	0.003	-0.001	0.000
	4mm	64.18	62.77	62.77	61.76	65.08	66.33	66.57	67.27	0.002	0.001	0.005	0.004	0.002	0.002	0.000	-0.001
	8mm	64.32	63.85	62.76	62.42	65.67	66.82	67.18	67.74	0.002	0.005	-0.002	0.000	0.000	0.000	-0.001	0.001
	12mm	64.56	63.77	63.33	62.70	66.45	66.72	68.02	68.77	0.000	-0.001	-0.002	0.000	-0.001	0.002	0.000	0.008
	16mm	64.58	64.33	63.01	62.89	66.17	66.86	67.85	68.20	0.005	-0.002	0.003	0.006	0.003	0.004	0.001	0.003
Mode 3	0mm	122.29	121.65	119.92	119.17	124.86	125.96	126.19	127.50	0.000	0.000	-0.004	-0.003	0.002	0.001	-0.001	0.005
	4mm	121.95	121.34	120.15	119.48	123.16	124.96	125.30	125.82	0.002	-0.001	0.001	0.000	0.001	0.000	-0.003	0.001
	8mm	121.93	121.23	120.03	119.28	123.79	124.54	125.16	126.48	0.002	0.002	0.000	0.003	0.004	0.003	0.000	0.000
	12mm	121.38	120.51	120.18	118.68	122.97	123.81	125.79	125.76	0.000	0.002	0.000	-0.001	0.002	0.003	0.004	0.005
	16mm	118.94	117.80	115.55	115.76	120.42	120.84	120.92	122.54	0.009	0.003	0.002	0.000	0.001	0.000	0.000	0.002
Mode 4	0mm	207.82	206.55	205.68	205.24	209.39	210.43	211.39	212.17	-0.001	0.000	0.000	0.000	0.001	0.001	0.002	0.002
	4mm	207.74	207.08	206.28	204.96	209.80	210.68	211.00	212.24	-0.001	0.000	0.001	0.000	0.001	0.001	0.001	0.000
	8mm	207.87	207.69	206.46	205.81	209.78	210.74	212.06	212.75	0.000	0.001	0.002	0.002	0.001	0.000	0.001	0.002
	12mm	208.78	207.32	206.70	205.78	210.59	211.39	212.86	213.32	0.000	0.000	0.002	0.001	0.004	0.002	0.003	0.003
	16mm	208.65	208.07	207.47	206.28	211.07	211.33	211.72	213.87	0.007	0.007	0.007	0.006	0.001	0.002	0.005	0.002
Mode 5	0mm	427.27	425.82	425.93	424.35	429.65	429.63	430.57	431.62	-0.001	0.000	0.000	0.000	0.001	0.000	-0.001	-0.001
	4mm	426.97	426.19	425.64	424.82	429.96	430.42	431.07	432.53	-0.001	-0.002	-0.002	-0.002	0.000	0.000	-0.001	-0.002
	8mm	426.74	425.65	424.60	426.52	427.89	429.01	430.71	430.79	0.000	-0.001	-0.001	0.000	-0.001	0.000	0.000	-0.001
	12mm	425.25	424.34	423.98	422.97	427.08	428.23	429.33	430.30	0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.000
	16mm	423.80	422.97	422.31	421.21	426.32	426.84	427.96	429.63	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000
Mode 6	0mm	563.47	561.63	560.57	559.42	564.50	563.71	566.45	568.21	-0.002	-0.002	-0.003	-0.003	-0.002	0.000	-0.001	-0.001
	4mm	562.47	561.87	560.11	558.64	564.98	564.63	567.55	567.45	-0.003	-0.004	-0.001	-0.003	0.000	0.000	0.000	-0.003
	8mm	559.38	558.16	555.62	562.46	569.49	563.11	564.00	564.66	-0.002	-0.004	0.000	0.000	0.000	-0.001	-0.001	-0.002
	12mm	555.03	554.42	550.00	549.76	553.93	557.00	559.45	557.12	0.000	0.001	0.000	0.001	-0.002	0.000	0.000	0.000
	16mm	558.02	559.01	517.74	563.74	540.25	538.41	546.68	543.74	0.001	0.000	0.011	0.000	-0.005	-0.002	-0.004	-0.005

Table B.1: Full modal analysis results for experimental data drawn from the diagonal strut

		Natural frequencies (Hz)				Damping ratios			
	Average load (N)	**-19**	**-52**	**45**	**77**	**-19**	**-52**	**45**	**77**
Mode 1	0mm	49.09	47.81	50.97	51.88	-0.004	−0.003	0.000	0.000
	4mm	48.65	47.61	49.97	51.76	0.000	-0.003	0.000	0.000
	8mm	50.61	49.48	52.78	53.16	0.007	0.001	0.005	0.000
	12mm	49.65	48.59	50.94	52.28	-0.002	0.003	-0.002	-0.001
	16mm	48.92	48.05	50.36	51.54	0.003	-0.007	-0.003	0.003
Mode 2	0mm	131.11	128.73	133.86	136.38	-0.007	−0.006	-0.007	-0.002
	4mm	130.20	128.66	132.47	135.35	-0.001	-0.001	-0.001	-0.008
	8mm	134.24	134.19	137.20	138.93	0.013	0.000	0.002	0.004
	12mm	131.11	129.55	134.47	136.31	-0.001	-0.005	-0.001	-0.006
	16mm	129.40	128.05	132.78	134.07	0.000	-0.001	-0.002	-0.007
Mode 3	0mm	273.80	272.35	277.03	278.73	-0.003	-0.003	-0.002	-0.003
	4mm	272.90	270.97	276.20	277.90	-0.004	−0.005	-0.002	-0.003
	8mm	271.27	269.35	274.71	276.34	-0.002	-0.002	-0.001	-0.001
	12mm	269.28	267.27	272.70	274.27	-0.005	-0.007	-0.003	-0.003
	16mm	262.98	261.49	266.61	268.62	-0.005	−0.004	-0.006	-0.005
Mode 4	0mm	436.04	433.97	439.94	440.84	0.000	0.000	0.001	0.001
	4mm	435.44	433.12	438.18	439.06	0.000	0.000	0.001	0.000
	8mm	429.16	428.01	431.69	433.63	0.001	0.001	0.001	0.002
	12mm	427.63	426.94	430.43	431.81	0.006	0.006	0.007	0.008
	16mm	422.43	419.25	424.03	425.47	-0.001	0.002	0.002	0.001
Mode 5	0mm	640.16	635.20	646.31	645.19	-0.003	-0.005	-0.001	-0.002
	4mm	639.98	635.94	642.92	644.43	-0.004	−0.007	-0.002	-0.003
	8mm	633.15	631.22	635.47	639.06	0.000	-0.003	0.000	-0.003
	12mm	629.16	626.97	633.02	635.76	-0.001	-0.002	-0.001	0.000
	16mm	623.57	624.21	629.45	631.18	0.001	0.004	-0.001	0.002

Table B.2: Full modal analysis results for experimental data drawn from the vertical strut

Appendix C

Publications

C.1 Journal papers

J. Wilson, P. Gardner, G. Manson, R. J. Barthorpe, 2023, Application of hierarchical verification and validation in a forward model-driven structural health monitoring strategy. Submitted to Structural Health Monitoring 2023

C.2 Conference papers

J. Wilson, P. Gardner, G. Manson, R. J. Barthorpe, 2023, On the selection and validation of component damage models for prediction of damage-state behaviour of a truss bridge. Proceedings of IMAC XLI International Conference on Modal Analysis, Austin, USA.

J. Wilson, P. Gardner, G. Manson, R. J. Barthorpe, 2022, The role of features in a hierarchical modelling strategy for forward model-driven structural health monitoring. Proceedings of USD2022 International Conference on Noise and Vibration Engineering, Leuven, Belgium.

J. Wilson, P. Gardner, G. Manson, R. J. Barthorpe, 2022, Hierarchical model verification and validation for structural health monitoring using dynamic substructuring. Proceedings of the European Workshop on Structural Health Monitoring 2022, Palermo, Italy.