Towards the Validation of Dynamical Models in Regions where there is no Data

P.L.Green[a], E.Chodora[b], Z.Zhu[a], S.Atamturktur[c]

[a]*Institute for Risk and Uncertainty, School of Engineering, University of Liverpool, Liverpool L69 7ZF, United Kingdom*
[b]*Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA*
[c]*Architectural Engineering, The Pennsylvania State University, University Park, PA 16801, USA*

Abstract

The creation of computer models is often driven by the need to make predictions in regions where there is no data (i.e. *extrapolations*). This makes validation challenging as it is difficult to ensure that a model will be suitable when it is applied in a region where there are no observations of the system of interest. The current paper proposes a method that can reveal flaws in a model which may be difficult to identify using traditional approaches for model calibration and validation. The method specifically targets the situation where one is attempting to model a dynamical system that is believed to possess time-invariant calibration parameters. The proposed approach allows these parameters to vary with time, even though it is believed that they are time-invariant. The of such an analysis is to identify key discrepancies - indications that a model has inherent flaws and, as a result, should not be used to influence decisions in regions where there is no data. The proposed method isn't necessarily a predictor of extrapolation performance, rather, it is a stringent test that, the authors believe, should be applied before extrapolation is attempted. The approach could therefore form a useful part of wider validation frameworks in the future.

Keywords: Extrapolative Predictions, Model Validation, Bayesian Inference, Structural Dynamics

1. Introduction

Computational models are crucial to many fields of science and engineering. Before being used to influence decision-making, such models must be validated; it must be established that the model can represent the real world with a level of accuracy sufficient for its intended purpose. Model validation is, as a result, an active research area.

In a broad sense, the current paper considers the situation where:

1. There may be several competing model structures that can, potentially, be employed by a user to simulate a quantity of interest.

2. Each model structure features a set of parameters that will need to be calibrated during the validation process.

Email address: `p.l.green@liverpool.ac.uk` (P.L.Green)

3. A set of observations, from the system of interest, are available to the user (such that this data can be used to help infer parameter estimates and/or facilitate model selection).

While such scenarios are considered by much of the literature, the current paper differs from most in that it explicitly considers the issues associated with *extrapolation*. Here, 'extrapolation' refers to a situation where a model will be used to make predictions in a region that, according to a form of distance metric, is 'very far' from current experimental observations. Using the terminology given in [1], the current work considers the situation where a model must be applied outside of its 'domain of applicability'.

Validating models that will be used to extrapolate is often confounded by the fact that, in the regions where measurements are available, calibration can mask model error. In other words, a model's parameters can often be tuned such that the model appears to be sufficiently valid compared to the available data, even if the model will be a poor extrapolator (this is illustrated in Section 4.2 of the current paper). The current work presents an approach that can, in such a scenario, illustrate the flaws in a model; providing a warning that the model should not be used for extrapolative predictions.

The proposed approach focuses on situations where it is believed that the model of interest's parameters are time-invariant. A calibration procedure is then employed that *deliberately treats the model's parameters as being time-varying, even though this is not believed to be the case*. It is by observing that a model's parameters must vary with time if it is to replicate a set of measurements with a high level of fidelity that one can identify key discrepancies in the model that may be hidden by other calibration procedures. The approach is demonstrated on applications from the field of structural dynamics.

2. Literature Review

The idea that statistical models can be used to emulate *model discrepancy* (see [2] and[3] for example) has influenced a large body of work. Such an approach accepts that there will always be a discrepancy between a model that is formulated on the understanding of the system of interest (a model that is based on physical laws, for example) and the system's true response. A purely data-based model is then trained and used to realise predictions of this discrepancy. Subsequent to calibration, then, predictions are made using a combination of the model that is based on an understanding of the system and the data-based discrepancy model[1].

One of the best-known implementations of this approach is that of Kennedy and O'Hagan [2], where Gaussian Processes (GPs) were used to create statistical models of model discrepancy (as well as emulators,

---

[1]This type of approach is sometimes called grey-box modelling, as it utilises a combination of a 'white box model' (model based on an understanding of the system) and a 'black box model' (a model that is purely data-based).

for the cases where simulations of the system of interest are computationally expensive). This idea has contributed to many works in the field of model validation, including generalised validation frameworks [4][5][6], model validation metrics that consider model discrepancy [1], metrics to establish when sufficient experimental data has been gathered in the model validation process [7] and approaches for resource allocation (between code development and experimental testing [8] or the development of substructures within a model [9]).

While undoubtedly a useful contribution, employing data-based models of discrepancy cannot help improve extrapolations that are far from the training data (GPs, for example, will typically return very quickly to predictions based only on prior knowledge of the discrepancy if extrapolated to a region where there is no data). This issue is highlighted by a number of authors. In [10] it is stated that 'the primary use of engineering models is to extrapolate to a system with new nominals when there are no new field data' but that extrapolation, using the approach described in [2], can only be performed if there is a 'modest' change in the inputs. The authors of [11] note that 'such a statistically adjusted engineering model lacks predictive power in the sense that it can be used only in the experimental conditions identical or very similar to those under which the statistical model was fitted' and in [12] it is stated that '...the discrepancy representation ... is highly dependent on calibration against observables and, hence, should not be used in situations in which it cannot be trained and tested'.

It is important to note that the above discussion isn't a direct criticism of [2] - the authors of [2] did not claim that the method will provide reliable extrapolations. A discussion of the Kennedy and O'Hagan approach is given by BrynjarsdÃşttir and O'Hagan in [13], which considers how the use of GP emulators of model discrepancy can (1) aid the identification of a system's 'true' parameter values and (2) aid extrapolations. It is concluded that true parameter values can only be uncovered if realistic prior information about the model discrepancy is available and that, even if realistic prior information is available, extrapolation is not advisable. [13] highlights these two issues using a simple case study.

The current paper considers the situation where there is no data in the region one wishes to make predictions. This separates it from [1] and [7] which (respectively) describe a 'predictive maturity index' and a 'forecasting metric', as these two approaches utilise hold-out experiments (experimental observations, in the application domain, that are deliberately removed from the training data so that they can be used to aid validation at a later stage). The proposed method is designed to reveal additional *engineering knowledge* about a particular system and, as a result, form part of the frameworks described in [11] and [12], both of which specifically consider extrapolation.

3. Methodology

3.1. General Framework

The current work specifically concerns the analysis of dynamical systems. Using $\boldsymbol{x}_i$ to represent a system's state at time $i\Delta t$ then, generally, it is assumed that the system's state evolves according to

$$\boldsymbol{x}_i = \boldsymbol{f}(\boldsymbol{x}_{i-1}, \boldsymbol{u}_i) + \boldsymbol{v}_{i-1}, \qquad \boldsymbol{v}_{i-1} \sim \mathcal{N}(\boldsymbol{v}_{i-1}; \boldsymbol{0}, \boldsymbol{Q}_{i-1}) \tag{1}$$

where Δt is a fixed time increment, $\boldsymbol{u}_i$ is a system input, $\boldsymbol{f}$ is a model that predicts the next system state given $(\boldsymbol{x}_{i-1}, \boldsymbol{u}_i)$ and $\boldsymbol{v}_{i-1}$ is a noise term that reflects the uncertainties involved in predicting the system's next state (arising because $\boldsymbol{f}$ is an approximation of the true system). Equation (1) is typically referred to as the 'prediction equation'.

Likewise, using $\boldsymbol{z}_i$ to represent an observation of some aspect of the system at time $i\Delta t$, it is also supposed that measurements are made according to

$$\boldsymbol{z}_i = \boldsymbol{h}_i(\boldsymbol{x}_i) + \boldsymbol{n}_i, \qquad \boldsymbol{n}_i \sim \mathcal{N}(\boldsymbol{n}_i; \boldsymbol{0}, \boldsymbol{R}_i) \tag{2}$$

where $\boldsymbol{h}$ is a (potentially nonlinear) function of the system's state and $\boldsymbol{n}_i$ is used to capture uncertainty in the observation process (measurement noise). Equation (2) is typically referred to as the 'observation equation'. It is noted that, in the literature, the prediction and observation equations are often written as potentially nonlinear functions of their noise terms ($\boldsymbol{v}_{i-1}$ and $\boldsymbol{n}_i$ respectively) - this has not been done here for the sake of notational simplicity later in the paper.

3.2. Extended Kalman Filter

The extended Kalman filter is one of a collection of methods that target situations where the aim is to probabilistically 'track' how the state of the system of interest changes, as more observations arrive over time. If, in the prediction and observation equations, $\boldsymbol{f}$ and $\boldsymbol{h}$ are linear functions of $\boldsymbol{x}_{i-1}$ and $\boldsymbol{x}_i$ respectively then it is possible to realise closed-form solutions for this problem. These solutions form the well-known Kalman filter. If $\boldsymbol{f}$ and $\boldsymbol{h}$ are nonlinear then, generally, such closed-form solutions are unavailable. One way to address this is to approximate $\boldsymbol{f}$ and $\boldsymbol{h}$ using linear functions (via first-order Taylor series expansions), before applying the 'standard' Kalman filter equations to the new, linearised relationships. This linearised approach is called the extended Kalman filter (EKF)[2].

---

[2]There are many other ways in which this problem could be tackled. One of the most well-known is the *particle filter* - a numerical method that is fundamentally based on importance sampling (see [14] for a tutorial). Here, however, it was found that an extended Kalman filter performed acceptably well and that the prediction and measurement equations were not sufficiently nonlinear to warrant use of a particle filter.

The following text gives a brief description of an extended Kalman filter, to establish notation used later in the paper. Considering the situation where

$$p(\boldsymbol{x}_{i-1} \,|\, \boldsymbol{z}_{1:i-1}) = \mathcal{N}(\boldsymbol{x}_{i-1}; \boldsymbol{m}_{i-1|i-1}, \boldsymbol{P}_{i-1|i-1}) \tag{3}$$

(such that the probability density function (PDF) of $\boldsymbol{x}_{i-1}$ given $\boldsymbol{z}_{1:i-1}$ is Gaussian with mean $\boldsymbol{m}_{i-1|i-1}$ and covariance matrix $\boldsymbol{P}_{i-1|i-1}$) then, with an EKF, the linearised prediction equation is used to predict the next system state. This yields the PDF

$$p(\boldsymbol{x}_i \,|\, \boldsymbol{z}_{1:i-1}) = \mathcal{N}(\boldsymbol{x}_i; \boldsymbol{m}_{i|i-1}, \boldsymbol{P}_{i|i-1}) \tag{4}$$

Once a new observation arrives ($\boldsymbol{z}_i$) then, through the observation equation, the current estimate of the system's state can be updated to obtain:

$$p(\boldsymbol{x}_i \,|\, \boldsymbol{z}_{1:i}) = \mathcal{N}(\boldsymbol{x}_i; \boldsymbol{m}_{i|i}, \boldsymbol{P}_{i|i}) \tag{5}$$

where the parameters of the PDFs in equations (3), (4) and (5) are as follows:

$$\boldsymbol{m}_{i|i-1} = \boldsymbol{f}(\boldsymbol{m}_{i-1|i-1}) \tag{6}$$

$$\boldsymbol{P}_{i|i-1} = \boldsymbol{Q}_{i-1} + \boldsymbol{F}_{\boldsymbol{m}_{i-1|i-1}} \boldsymbol{P}_{i-1|i-1} \boldsymbol{F}^T_{\boldsymbol{m}_{i-1|i-1}} \tag{7}$$

$$\boldsymbol{m}_{i|i} = \boldsymbol{m}_{i|i-1} + \boldsymbol{K}_i(\boldsymbol{z}_i - \boldsymbol{h}(\boldsymbol{m}_{i|i-1})) \tag{8}$$

$$\boldsymbol{P}_{i|i} = \boldsymbol{P}_{i|i-1} - \boldsymbol{K}_i \boldsymbol{H}_{\boldsymbol{m}_{i|i-1}} \boldsymbol{P}_{i|i-1} \tag{9}$$

and the following definitions have been employed:

$$\boldsymbol{K}_i = \boldsymbol{P}_{i|i-1} \boldsymbol{H}^T_{\boldsymbol{m}_{i|i-1}} \boldsymbol{S}_i^{-1} \tag{10}$$

$$\boldsymbol{S}_i = \boldsymbol{H}_{\boldsymbol{m}_{i|i-1}} \boldsymbol{P}_{i|i-1} \boldsymbol{H}^T_{\boldsymbol{m}_{i|i-1}} + \boldsymbol{R}_i \tag{11}$$

$$\boldsymbol{F}_{\hat{\boldsymbol{x}}} \equiv \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \right]_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \tag{12}$$

$$\boldsymbol{H}_{\hat{\boldsymbol{x}}} \equiv \left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} \right]_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \tag{13}$$

3.3. Application to the 4th order Runge-Kutta integration scheme

The following examples consider dynamical systems that may be subjected to a time history of known excitations: $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, ...$ Throughout the paper, the state of these systems is calculated recursively using the 4th order Runge-Kutta numerical integration scheme. To help relate this scheme back to the general prediction equation (equation (1)), some additional notation is required. Specifically, the following definitions are employed:

$$u_i = \begin{pmatrix} \boldsymbol{\delta}_{i-1} \\ \boldsymbol{\delta}_i \end{pmatrix} \tag{14}$$

$$\dot{\boldsymbol{x}}(\boldsymbol{x}_i, \boldsymbol{\delta}_i) \equiv \left[\frac{\partial \boldsymbol{x}}{\partial t} \right]_{\boldsymbol{x}=\boldsymbol{x}_i, \boldsymbol{\delta}=\boldsymbol{\delta}_i} \tag{15}$$

$$\boldsymbol{k}_1 = \dot{\boldsymbol{x}}(\boldsymbol{x}_{i-1}, \boldsymbol{\delta}_{i-1}) \tag{16}$$

$$\boldsymbol{k}_2 = \dot{\boldsymbol{x}}\left(\boldsymbol{x}_{i-1} + \frac{\Delta t}{2} \boldsymbol{k}_1, \boldsymbol{\delta}_{\text{int}} \right) \tag{17}$$

$$\boldsymbol{k}_3 = \dot{\boldsymbol{x}}\left(\boldsymbol{x}_{i-1} + \frac{\Delta t}{2} \boldsymbol{k}_2, \boldsymbol{\delta}_{\text{int}} \right) \tag{18}$$

$$\boldsymbol{k}_4 = \dot{\boldsymbol{x}}\left(\boldsymbol{x}_{i-1} + \Delta t \, \boldsymbol{k}_3, \boldsymbol{\delta}_i \right) \tag{19}$$

where Δt is the time-step of the integration scheme and $\boldsymbol{\delta}_{\text{int}} = \frac{1}{2}(\boldsymbol{\delta}_{i-1} + \boldsymbol{\delta}_i)$. Using this notation, the function $\boldsymbol{f}$ that maps $\boldsymbol{x}_{i-1}$ to $\boldsymbol{x}_i$ can be written as

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} + \frac{\Delta t}{6} \left(\boldsymbol{k}_1 + 2\,\boldsymbol{k}_2 + 2\,\boldsymbol{k}_3 + \boldsymbol{k}_4 \right) \tag{20}$$

To find an expression for $\boldsymbol{F}_{\hat{\boldsymbol{x}}}$ (equation (12)) it is noted that

$$\boldsymbol{F}_{\hat{\boldsymbol{x}}} = \boldsymbol{I} + \frac{\Delta t}{6} \left(\underbrace{\left[\frac{\partial \boldsymbol{k}_1}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}}_{\text{Term 1}} + 2 \underbrace{\left[\frac{\partial \boldsymbol{k}_2}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}}_{\text{Term 2}} + 2 \underbrace{\left[\frac{\partial \boldsymbol{k}_3}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}}_{\text{Term 3}} + \underbrace{\left[\frac{\partial \boldsymbol{k}_4}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}}_{\text{Term 4}} \right) \tag{21}$$

Evaluation of terms 1-4 in equation (21) essentially requires the application of some straightforward vector calculus. For the sake of completeness, the required derivatives are shown in detail in Appendix Appendix B. Ultimately, it can be shown that

$$\left[\frac{\partial \boldsymbol{k}_1}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} = \tilde{\boldsymbol{F}}(\hat{\boldsymbol{x}}, \boldsymbol{\delta}_{i-1}) \tag{22}$$

$$\left[\frac{\partial \boldsymbol{k}_2}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} = \tilde{\boldsymbol{F}}\left(\hat{\boldsymbol{x}} + \frac{\Delta t}{2} \boldsymbol{k}_1(\hat{\boldsymbol{x}}, \boldsymbol{\delta}_{i-1}), \boldsymbol{\delta}_{\text{int}} \right) \left(\boldsymbol{I} + \frac{\Delta t}{2} \left[\frac{\partial \boldsymbol{k}_1}{\partial \boldsymbol{x}_{i-1}} \right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} \right) \tag{23}$$

6

$$\left[\frac{\partial\,\boldsymbol{k}_3}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} = \tilde{\boldsymbol{F}}\left(\hat{\boldsymbol{x}} + \frac{\Delta t}{2}\,\boldsymbol{k}_2(\hat{\boldsymbol{x}},\boldsymbol{\delta}_{\mathrm{int}}),\boldsymbol{\delta}_{\mathrm{int}}\right)\left(\boldsymbol{I} + \frac{\Delta t}{2}\left[\frac{\partial\,\boldsymbol{k}_2}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}\right) \tag{24}$$

$$\left[\frac{\partial\,\boldsymbol{k}_4}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} = \tilde{\boldsymbol{F}}\left(\hat{\boldsymbol{x}} + \Delta t\,\boldsymbol{k}_3(\hat{\boldsymbol{x}},\boldsymbol{\delta}_{\mathrm{int}}),\boldsymbol{\delta}_i\right)\left(\boldsymbol{I} + \Delta t\left[\frac{\partial\,\boldsymbol{k}_3}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}}\right) \tag{25}$$

where $\tilde{\boldsymbol{F}}(\boldsymbol{x},\boldsymbol{\delta})$ represents the Jacobian of $\dot{\boldsymbol{x}}$, with respect to $\boldsymbol{x}_{i-1}$, evaluated at the points $\boldsymbol{x}$ and $\boldsymbol{\delta}$.

4. Example 1: Numerical Case Study

4.1. Description

Example 1 considers modelling a single degree of freedom nonlinear dynamic system whose equation of motion is

$$\ddot{x} + c\dot{x} + kx + k^*x^3 = \delta(t) \tag{26}$$

In equation (26) x is displacement, c is a linear damping coefficient, k is a linear stiffness coefficient and k^* is a nonlinear stiffness coefficient that controls the strength of the system's hardening stiffness nonlinearity. $\delta(t)$ is a random excitation that, in this case, consists of samples drawn from a zero-mean unit-variance Gaussian.

To generate a set of observation data, the nonlinear system was simulated using a time step of $\Delta t = 0.01$s. Observations were created by taking the resulting displacement time history and corrupting it with zero-mean Gaussian noise (variance equal to 1×10^{-5}). The model's parameters were $k = 10$ N/m, $c = 0.1$ Ns/m and $k^* = 50$ N/m^3. Observation data was generated over a time period of 200 seconds. The full set of observation data is shown in Figure 1 while a zoomed-in portion is shown in Figure 2.
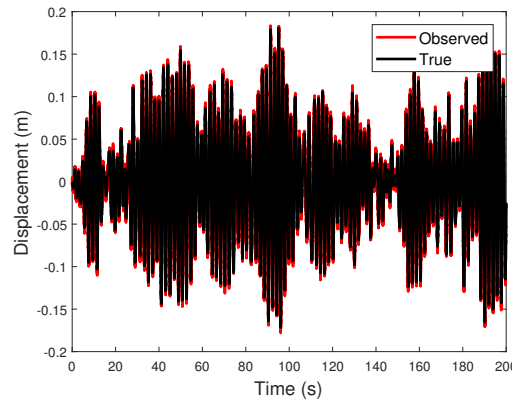

Figure 1: 200 seconds of displacement time history observations from the numerical case study. Black shows the true response of the system while red shows noisy observations.
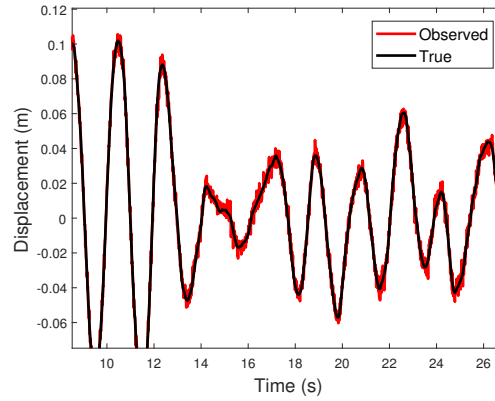
Figure 2: Numerical case study. A close-up view of the data shown in Figure 1.

The following considers the case where one is attempting to fit a linear model to data that is obtained from the nonlinear system. It is also considered that, after calibration, the model will be used to predict the response of the system to a much greater amplitude excitation than that used to generate the data in Figure 1. It is noted that the excitation used to generate the observation data is relatively low, such that the nonlinearity is difficult to detect.

4.2. Highlighting the Problem

This section illustrates that, by tuning its parameters, it is possible for the linear model to represent observations of the nonlinear system fairly accurately in the low amplitude regime where there is data, potentially giving false confidence in the model's ability to extrapolate. Specifically, a common Bayesian approach (that assumed time-invariant parameters) is adapted to the calibration of the model. The approach begins with Bayes' theorem:

$$p(\boldsymbol{\theta} \,|\mathcal{D}) \propto p(\mathcal{D}| \,\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{27}$$

where $\boldsymbol{\theta}$ is a vector of parameters to be calibrated and $\mathcal{D}$ represents training data (in this case $\mathcal{D}$ is a time history of the excitation and the observed displacement response of the system). The likelihood, $p(\mathcal{D}|\boldsymbol{\theta})$, is defined based on the assumption that each observation is equal to the response of the model, corrupted by the addition of Gaussian noise. This noise is assumed to be zero-mean, have variance R and be independent of the noise that corrupted previous observations. In the following, R is treated as a time-invariant parameter that also requires calibration such that $\boldsymbol{\theta} = (k, c, R)^{T}$.

Defining the likelihood in this way treats any differences between observations and model response (i.e. errors that arise because of both measurement noise and model discrepancy) as independent Gaussian noise[3]. Moreover, by allowing R to be calibrated, the variance of this noise can be tuned at the same time as

---

[3]This assumption is often justified by the principle of maximum entropy [15], although [16] highlights examples where prediction errors can be spatially and/or temporally correlated.

the model's stiffness and damping parameters. This framework means that the maximum likelihood set of model parameters are those that make the discrepancy between the model and the observations most closely approximate a Gaussian distribution with variance R. This way of treating uncertainty in the calibration procedure has been applied by many (including the author's of the current manuscript) and is referred to in [13] as a 'traditional approach'.

For the academic example discussed in the current section, prior distributions over the calibration parameters were chosen to be

$$p(k) = \mathcal{N}(k; 10, 3^2), \quad p(c) = \mathcal{N}(c; 0.1, 0.1^2), \quad p(R) = \mathrm{Gamma}(R; 1, 100) \tag{28}$$

such that the prior over R is a Gamma distribution whose shape and rate parameters are equal to 1 and 100, respectively. Using the first 50 seconds of data for training, the Markov chain Monte Carlo (MCMC) algorithm detailed in [17] was used to generate 1000 samples from the posterior, $p(\boldsymbol{\theta} \,|\, \mathcal{D})$.

Histograms of the resulting samples are shown 'close up' in Figure 3 while, in Figure 4, histograms are plotted alongside the true parameter values. These figures highlight two observations. Firstly the posterior distribution over k is concentrated over values that are larger than the true stiffness. This illustrates how k has been calibrated to overcome the model discrepancy (the lack of the nonlinear term). Secondly, the distribution over R is also concentrated on values that are larger than the true measurement noise variance, highlighting that this particular approach has attempted to describe the model discrepancy present in the problem as additional noise on the observations (in other words, the errors present in the model have increased the most probable value of R, given the available training data).
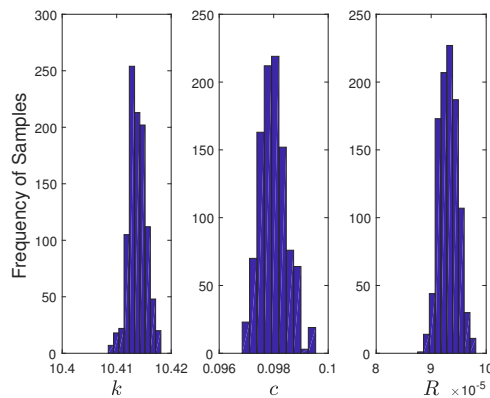

Figure 3: Histograms of samples from the posterior parameter distribution of the linear model in the numerical case study. All parameters in SI units.
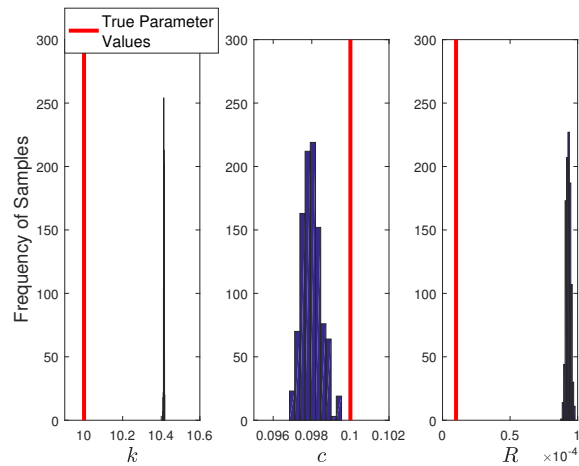
Figure 4: Histograms of samples from the posterior parameter distribution of the linear model in the numerical case study. True parameter values are marked by red lines. All parameters in SI units.

An illustration of the problem that motivates the current paper comes when the MCMC samples are used to propagate these parameter uncertainties into future model predictions. Recalling that there is only training data in a relatively low amplitude regime, Figure 5 shows the model's predictions relative to (1) the training data and (2) data that was 'held back' for validation purposes. Figure 6 shows a close-up of the model's ability to replicate the validation data. These results illustrate that, for low amplitude excitations, the model is able to closely replicate data that was not used in training; the model has been tuned to work effectively in the region where we have data, despite lacking the nonlinear stiffness term. Subsequently, the model was used to predict the response of the system to a higher amplitude excitation, whereby $\delta(t)$ consisted of samples drawn from a zero-mean Gaussian whose standard deviation was equal to 2. Figure 7 illustrates that, for this higher amplitude excitation, the model is not able to accurately replicate the response of the system (or accurately quantify the uncertainties involved in the predictions). This numerical example highlights how, without the data from the high amplitude case, it can be difficult to infer that the model would be a poor extrapolator.
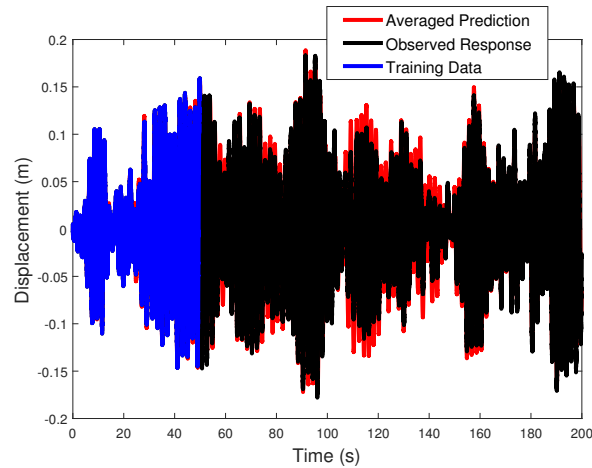
Figure 5: Propagating parameter uncertainties of the linear model into predictions (numerical case study). Blue represents training data, black represents previously unseen data and red represents the average model predictions. Bounds on model predictions are 3 standard deviations from the mean.
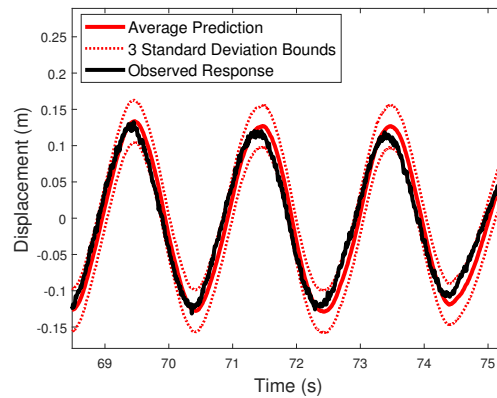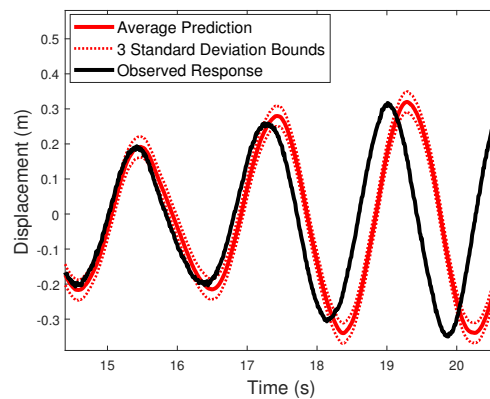

Figure 6: Propagating parameter uncertainties of the linear model into predictions of previously unseen, 'low amplitude' data (numerical case study). Bounds on model predictions are 3 standard deviations from the mean.

Figure 7: Propagating parameter uncertainties of the linear model into predictions of previously unseen, 'high amplitude' data (numerical case study). Bounds on model predictions are 3 standard deviations from the mean.

To further illustrate the problem, model selection is now considered. Specifically, two potential models

are now considered: model 1 ($\mathcal{M}_1$) is the linear model that was used in the previous analysis while model 2 ($\mathcal{M}_2$) includes the nonlinear stiffness term. Model 2 has the same structure as equation (26) and therefore represents a scenario where there is no model error. The following analysis illustrates how, even if a model whose structure was exactly correct is included in the analysis, well-known model selection criteria, based on the available data, may still lead to the selection of the incorrect model.

The calibration parameters of $\mathcal{M}_1$ and $\mathcal{M}_2$ are now written as $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ respectively, such that

$$\boldsymbol{\theta}_1 = (k, c, R)^T, \qquad \boldsymbol{\theta}_2 = (k, c, k^*, R)^T \tag{29}$$

Here, for illustrative purposes, *Bayes factor* is used as a model selection indicator.[4] Evaluating Bayes factor involves computing the relative probability of the two model structures, given the available measurement data:

$$\frac{\Pr(\mathcal{M}_1|\mathcal{D})}{\Pr(\mathcal{M}_2|\mathcal{D})} = \frac{p(\mathcal{D}|\mathcal{M}_1)\Pr(\mathcal{M}_1)}{p(\mathcal{D}|\mathcal{M}_2)\Pr(\mathcal{M}_2)} \tag{30}$$

where

$$p(\mathcal{D}|\mathcal{M}_i) = \int p(\mathcal{D}|\boldsymbol{\theta}_i, \mathcal{M}_i)p(\boldsymbol{\theta}_i|\mathcal{M}_i)d\boldsymbol{\theta}_i, \qquad i = 1, 2 \tag{31}$$

In this case, equal prior probability mass was assigned to each model such that $\Pr(\mathcal{M}_1) = \Pr(\mathcal{M}_2)$.

The integrals shown in equation (31) can be estimated using a variety of numerical methods ([19][20][21] for example). Here, Bayes factor was estimated using Laplace's method, which employs Gaussian approximations of the posterior parameter PDFs of the two models (see appendix Appendix A for details). It was found that:

$$\log\left(\frac{\Pr(\mathcal{M}_1|\mathcal{D})}{\Pr(\mathcal{M}_2|\mathcal{D})} \approx 5622\right) \tag{32}$$

Consequently, according to this particular model selection criterion, the incorrect linear model is viewed as preferable relative to the nonlinear model. This is despite the fact that the nonlinear model is actually a perfect representation of the system of interest. The results reflect that, in this low amplitude region where the effects of the nonlinearity is relatively small, the chosen model selection criterion views the nonlinear model as overly complex (indeed, in the region where data is available, the linear model does appear to perform well).

---

[4]The authors acknowledge that many different model selection criteria exist and that the question of how model selection criteria should be defined is still widely debated (as illustrated in [18], for example). The authors do not intend to critique these criteria here, except to say that the majority of model selection criteria (1) penalise overly complex models (as these are more prone to overfitting) and (2) define 'model complexity' as a function of the number of tunable parameters that are in a model. Bayes' factor is an example of such a criterion that is used relatively frequently in the literature.

The next section (Section 4.3) illustrates the application of the proposed approach to the same situation, using the same set of observation data.

4.3. Proposed Approach

In the proposed approach, to treat a model's parameters as time-variant, the parameters are included in the definition of the system's state. This allows them to be 'tracked', alongside the system's displacement and velocity, using the EKF (see [22] for a similar approach that utilised particle filtering techniques). Here it is illustrated how the approach can be implemented for $\mathcal{M}_2$ (the model that includes the nonlinear stiffness term), where numerical integration is conducted using the RK4 integration scheme. Implementation of the approach to $\mathcal{M}_1$ is very similar and, as such, is not discussed in detail here.

The state of $\mathcal{M}_2$ at time $i\Delta t$ is now defined as $\boldsymbol{x}_i = (x_i, \dot{x}_i, k_i, c_i, k_i^*)^T$. In order to implement the EKF, one must hypothesise how the model's parameters vary through time i.e. a definition for $\boldsymbol{f}$ in equation (1) must be realised. As with [22], the chosen prediction model assumes that each system parameter is equal to its preceding value, such that

$$\dot{\boldsymbol{x}} = \begin{pmatrix} \dot{x} \\ \delta - kx - c\dot{x} - k^*x^3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{33}$$

Consequently, the Jacobian $\tilde{\boldsymbol{F}}$ is given by

$$\tilde{\boldsymbol{F}}(\boldsymbol{x}, \delta) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -(k + 3k^*x^2) & -c & -x & -\dot{x} & -x^3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{34}$$

One must also quantify the uncertainty in this predictive model (in other words, the covariance matrix $\boldsymbol{Q}_{i-1}$ in equation (1) must be defined). Here, $\boldsymbol{Q}_{i-1}$ is defined as follows

$$\boldsymbol{Q}_{i-1} = \begin{bmatrix} \boldsymbol{0}_{2\times2} & \boldsymbol{0}_{2\times3} \\ \boldsymbol{0}_{3\times2} & \boldsymbol{G}_{i-1} \end{bmatrix} \tag{35}$$

where $\boldsymbol{G}_{i-1}$ is a diagonal matrix that, for model $\mathcal{M}_2$, is of size 3×3. The approach assumes that there is zero uncertainty associated in the prediction of displacement and velocity, and that all of the uncertainty is manifested in the forward prediction of the model's time-varying parameters. This is deliberate - the intention is that, by tracking the parameter estimates as they vary with time, it will be possible to infer indications of model error. Following on from [22], the diagonal elements of $\boldsymbol{G}$ were chosen such that the coefficient of variation in the drift of each parameter is fixed. For the numerical case study shown here

13

the coefficient of variation was set equal to 0.1 % (note that the influence of this term is investigated more thoroughly in Section 5).
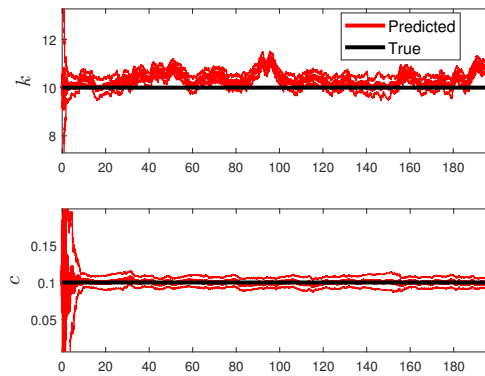

Figure 8: Tracking time varying parameters for the linear model (numerical case study). Black lines show true parameter values, red lines show estimated parameter values. Bounds on parameter estimates are 3 standard deviations from the mean.

Figure 8 shows the results that were obtained when the parameters of the linear model were treated as being time-variant, and tracked using the EKF. As one would expect, this method allows the displacement and velocity to be tracked very accurately - a comparison between the observed and predicted is not shown here because the two are essentially indistinguishable. Figure 8 shows that, relatively speaking, the system's damping value converges while the linear stiffness varies with time.
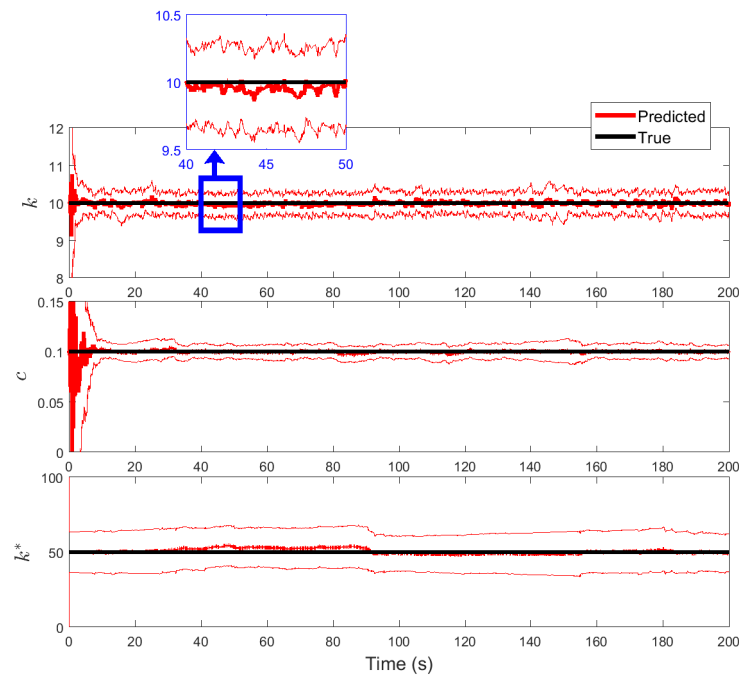
Figure 9: Numerical case study. Tracking time varying parameters for the nonlinear model. Black lines show true parameter values, red lines show estimated parameter values. Bounds on parameter estimates are 3 standard deviations from the mean. Blue axes show a 'close up' view of results relating to k, the linear stiffness.
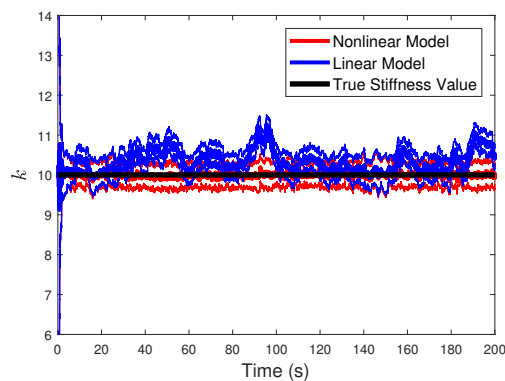

Figure 10: Numerical case study. Tracking the stiffness for the nonlinear model (red) and linear model (blue). Bounds on parameter estimates are 3 standard deviations from the mean.

To aid comparison this analysis was repeated for the nonlinear model. These results are shown in Figure 9 while Figure 10 shows a comparison between the stiffness values that were identified, as a function of time, for both the linear and nonlinear models. To achieve an excellent fit to the observation data it can be see that the stiffness of the linear model had to vary significantly more than the stiffness of the nonlinear model.

Recalling that the analysis is (deliberately) applied to systems whose parameters, it is believed, should not be varying with time the important question is: can the results in Figure 10 be used to infer that the linear model would extrapolate poorly if used at high amplitudes? The authors concede that the answer is

15

not straightforward and it is clear that these results would require a significant amount of interpretation, however, it is also argued that allowing the parameters to vary with time provides a _different type of information_ relative to the more traditional approach shown in Section 4.2. It is certainly conceivable that, by studying the results in the current section, a practitioner may lead to a different set of conclusions about the suitability of the linear model for extrapolations compared to if they had studied the results in Section 4.2 (i.e. the more traditional approach).

A more in depth discussion of how such results could be utilised in a general setting is postponed until Section 6. First, in Section 5, it is illustrated how the approach may be applied to an experimental case study.

5. Example 2: Experimental Case Study

5.1. Description

This section analyses the dynamics of the two storey aluminium structure shown in Figure 11. Testing was conducted whereby lateral vibrations were induced by a shaker at the base of the structure. Measurements were obtained from 3 triaxial accelerometers that were placed at the base, first floor and second floor of the structure. The structure was excited sinusoidally at a frequency close to its first linear natural frequency, at a variety of amplitudes. The full set of resulting observation data is shown in Figure 12.
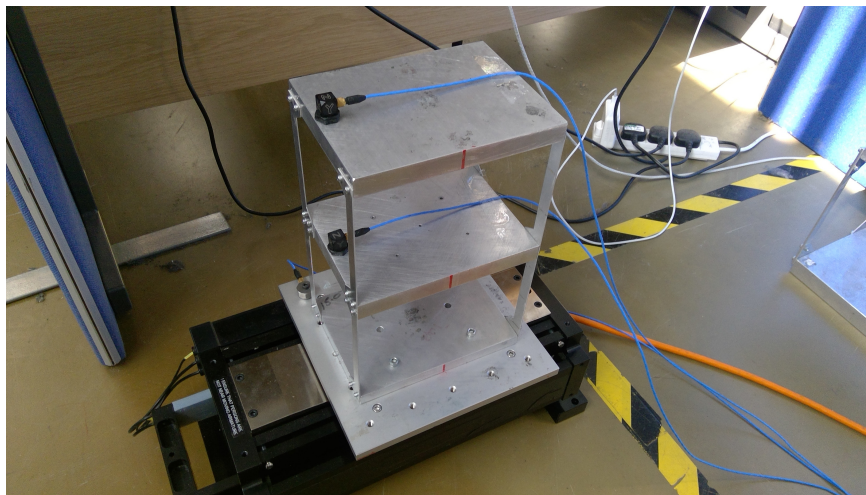

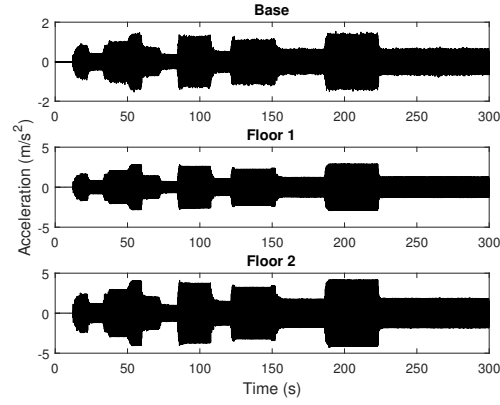Figure 11: Test rig for the experimental case study.

Figure 12: Time history of acceleration measurements for the experimental case study.

In the following, measurements of the system's dynamic response are used to infer a model of the structure. Again, the situation where the model must be used to extrapolate far beyond available data is considered (for example, the model may be needed to predict the response of the structure to excitation levels far higher than those shown in Figure 12).

A 2DOF shear-storey model of the building was developed. Defining x_1 and x_2 as the displacement of floors 1 and 2 respectively, the dynamic model is based on the following:

$$\ddot{x}_1 = \frac{1}{m_1} \left[k_2(x_2 - x_1) + c_2(\dot{x}_2 - \dot{x}_1) - c_1(\dot{x}_1 - \dot{x}_{\text{base}}) - k_1(x_1 - x_{\text{base}}) \right] \tag{36}$$

$$\ddot{x}_2 = \frac{1}{m_2} \left[-c_2(\dot{x}_2 - \dot{x}_1) - k_2(x_2 - x_1) \right] \tag{37}$$

where x_{base} is the displacement of the base, k_1 and c_1 are the stiffness and damping between the base and floor 1, k_2 and c_2 are the stiffness and damping between floor 2 and floor 1 and m_i is the mass of the ith floor. Measurements of base displacement and velocity were obtained via numerical integration of the observed acceleration time histories (after the appropriate signals had been passed through a low-pass butterworth filter). The two masses, m_1 and m_2, were measured directly (with, it was assumed, a sufficient degree of accuracy) thus leaving the parameters k_1, k_2, c_1 and c_2 to be calibrated.

In the next two sections two different approaches to the calibration and validation of this model are illustrated. The first shows the results of a Bayesian analysis where it is assumed that the parameters do not vary with time (this is the same 'traditional approach' that was demonstrated in Section 4.2 of the numerical case study). The second is the proposed approach, whereby the calibration parameters are treated as being time-variant. It is important to emphasise that, in the subsequent analyses, the aim is not necessarily to identify a model which could be relied upon for extrapolations. Rather, the intention is to illustrate how the proposed approach can reveal useful information that, after interpretation, helps to establish a model's validity outside of the region where data is available.

17

5.2. Time-invariant parameters

Here, the first 20 seconds of the data shown in Figure 12 were used for training. The observations of base excitations were treated as an input to the model while acceleration measurements from the two floors were used to infer parameter estimates. Fundamentally, the analysis proceeded in the same manner as described in Section 4.2 (the 'traditional approach'). To formulate the likelihood it was assumed that, at each floor, the acceleration observations were equal to the response of the model plus the addition of independent Gaussian noise. The variance of the noise that corrupts observations for floors 1 and 2 was defined as R_1 and R_2 respectively. In the following, the masses of the floors are assumed known while the remaining stiffness and damping terms were treated as calibration parameters alongside R_1 and R_2. Prior distributions were as follows:

$$p(k_1) = p(k_2) = \mathcal{N}(1500, 300^2), \quad p(c_1) = p(c_2) = \mathcal{N}(10, 2^2),$$

$$p(R_1) = p(R_2) = \text{Gamma}(1 \times 10^{-3}, 1) \tag{38}$$

(noting that, to help form $p(k_1)$ and $p(k_2)$, initial estimates of the inter-storey stiffnesses were realised by modelling each column as a cantilever beam). Once again, MCMC was used to generate 1000 samples from the posterior parameter distribution. Histograms of these samples are shown in Figure 13. Using the samples to propagate parameter uncertainty into model predictions, the ability of the model to replicate data that was not used for training is shown in Figure 14. It can be seen that the model appears to perform relatively well and that the observed response stays within three standard deviations of the mean prediction.
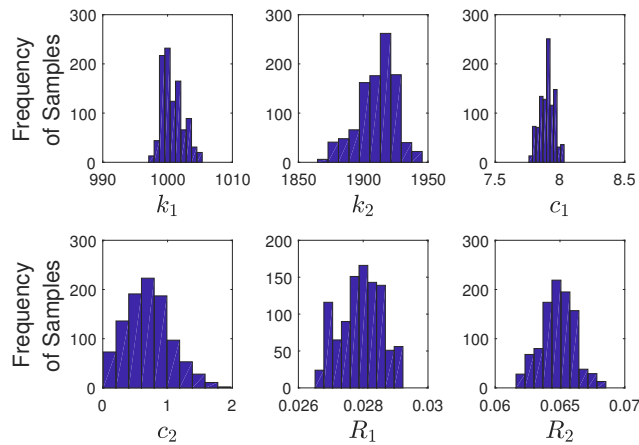

Figure 13: Histograms of samples from the posterior parameter distribution (experimental case study). All parameters in SI units.
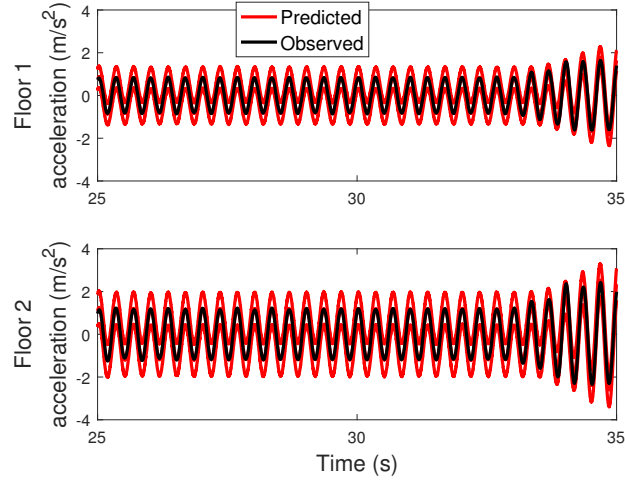
Figure 14: Propagating parameter uncertainties of the linear model into predictions (experimental case study). Black represents previously unseen data and red represents model predictions. Bounds on model predictions are 3 standard deviations from the mean.

5.3. Time-variant parameters

This section describes the application of the proposed approach to the experimental case study. To track how the model's parameters vary with time, the system's state was defined as

$$\begin{bmatrix} x_1 & x_2 & \dot{x}_1 & \dot{x}_2 & k_1 & k_2 & c_1 & c_2 \end{bmatrix}^T \tag{39}$$

while the derivative of the state with respect to time was

$$\begin{bmatrix} \dot{x}_1 & \dot{x}_2 & \ddot{x}_1 & \ddot{x}_2 & 0 & 0 & 0 & 0 \end{bmatrix}^T \tag{40}$$

where $\ddot{x}_1$ and $\ddot{x}_2$ are defined by equations (36) and (37) respectively. The matrices $\tilde{\boldsymbol{F}}$ and $\boldsymbol{H}$ that correspond to this system are shown in Appendix Appendix C.

The matrix $\boldsymbol{Q}$ was defined as described previously, such that the coefficient of variation in the drift of each parameter was set equal to 0.1%. Defining z_1 and z_2 as the observed acceleration of floors 1 and 2 respectively, the full set of parameter tracking results are shown in Figure 15. It was observed that, once again, the acceleration predictions made by the EKF were almost indistinguishable from the observed. Note that, in the following figures, uncertainty bounds are omitted for the sake of clarity.
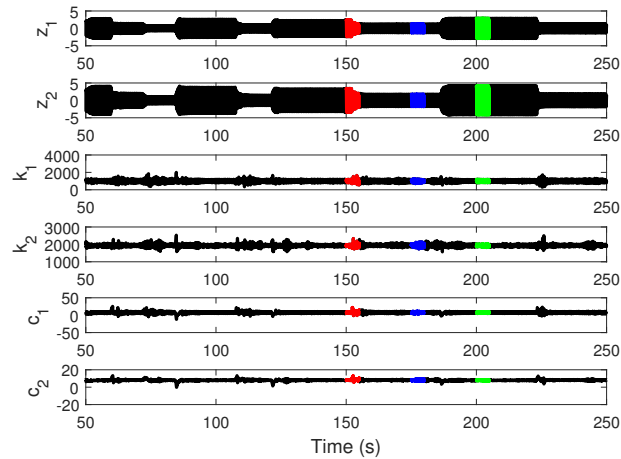
Figure 15: Experimental case study. Panels 1 and 2 show the observed acceleration of floors 1 and 2 respectively. Remaining panels show the tracked values of the model's parameters (mean values only). Red, blue and green portions are shown close-up in Figures 16, 17 and 18 respectively. SI units used throughout.

The red, blue and green portions of Figure 15 are shown close-up in Figures 16, 17 and 18 respectively.

Figure 16 shows a region where the system transitions from a relatively high amplitude response to a relatively low amplitude response. When in steady state it is observed that, for all four parameters, a periodic oscillation is required for the model to fit the training data very precisely (see results between 150 and 152 seconds, for example). This type of behaviour was observed for all steady state portions of the data. During a transitional period (see between 152 and 154 seconds, for example), as well as relatively large changes in stiffness, the magnitude of both damping terms approximately doubles. The reasons for this could be many (for example, from experience, it is known that the response of this system is very sensitive to the torque used when bolting the columns to the two floors - a possible indication that a linear damping model may be insufficient). For now this is left as speculation. The authors wish to emphasise that the current paper is not contributing an overall framework for validation but, instead, is providing a tool that could form part of such a framework in the future. It is also argued that, relative to the more traditional approach, this example highlights how treating the parameters as time-varying can lead to a richer source of information for future model development. At the very least, the results indicate that the model performs worse in transitional regions and that, before being used for extrapolation, more efforts should perhaps be diverted towards understanding the damping mechanisms present in the structure. Such conclusions would have been more difficult to reach using the results presented in Section 5.2.
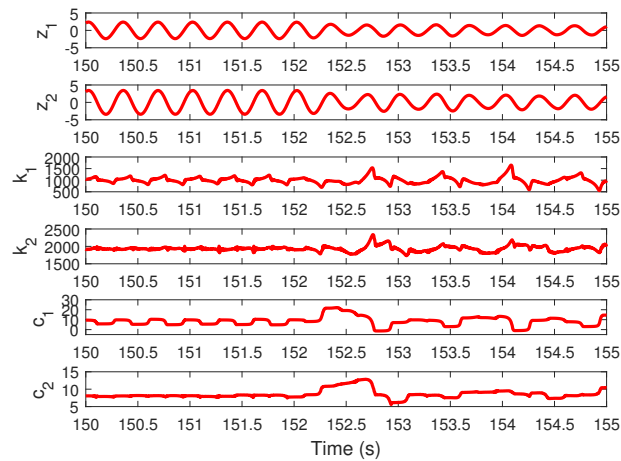
Figure 16: Close-up view of the red portion of results in Figure 15 (experimental case study).

Figures 17 and 18 show the results obtained when the system is at steady state, for a relatively low and high amplitude response respectively. It is interesting to note that, regardless of amplitude, near-periodic oscillations in all four parameters are required for the model to accurately fit the training data.
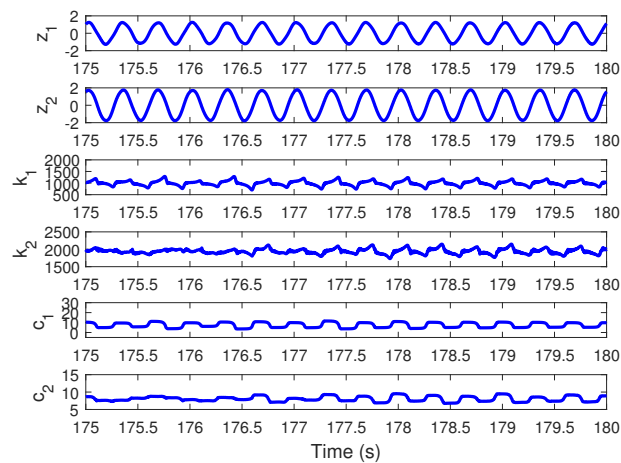

Figure 17: Experimental case study. Close-up view of the blue portion of results in Figure 15.
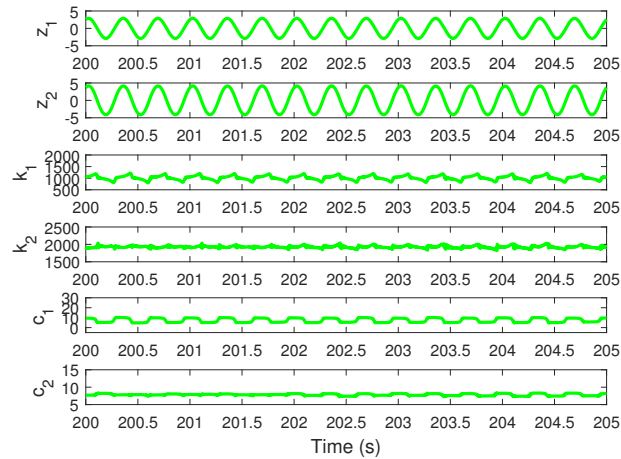
Figure 18: Experimental case study. Close-up view of the green portion of results in Figure 15.

6. Discussion and future work

One of the key parameters in the proposed approach is the 'drift term', which essentially controls the magnitude of the diagonal elements of the matrix $\boldsymbol{G}$ in equation (35). Selection of the drift term requires some thought. If the drift term is large then large variations in parameters become more probable. This can lead to over fitting, whereby the calibration parameters are varied such that the model begins to fit the measurement noise in the data. Likewise, if the drift term is small then the parameters cannot vary quickly enough for the model to fit the observation data.

To illustrate how the drift term may affect results, the experimental case study was revisited using coefficient of variations in the drift of each parameter set equal to 0.001%, 0.01% and 0.1% respectively. Figure 19 shows the resulting parameter estimates for these three cases, (specifically focusing on the transitional region shown in Figure 16). As one would expect, larger drift terms leads to larger variations in the parameters. Figure 19 shows the ability of the model to replicate the observed acceleration data for these three scenarios. While the closeness of the overlays makes it difficult to see the measurement data, it can be observed that the red and green lines are almost exactly the same while the blue line, which represents the lowest drift term, deviates slightly from the others. Figure 19 therefore illustrates how, if the drift term is too low, the model loses fidelity with respect to the measurement data.
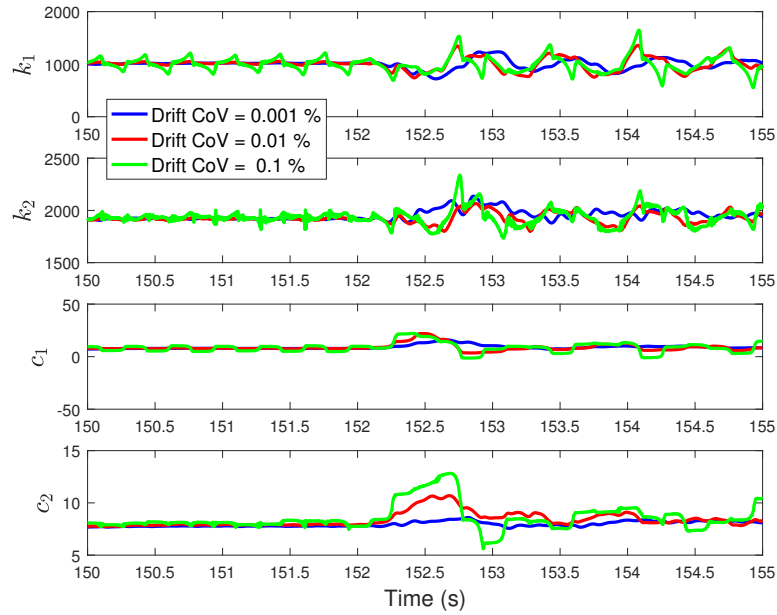
Figure 19: Experimental case study. Tracking the values of the model's parameters (mean values only) using different drift coefficients of variation. All parameters in SI units.
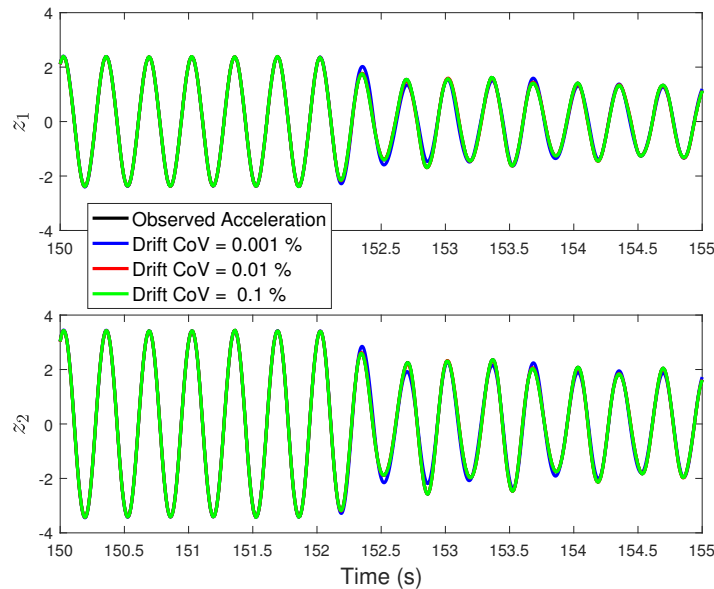

Figure 20: Experimental case study. Tracking the acceleration of floors 1 and 2 (mean values only) using different drift coefficients of variation. Acceleration in SI units.

In the simulated case study, model error was deliberately introduced to a well-defined problem. This was to illustrate how the proposed methodology could behave in such a scenario. For the experimental case study, where results indicate that the model should be improved before being relied upon for extrapolations, a similar procedure could be used to aid model development. Specifically, in a similar manner to Section 4.3, numerical experiments could be conducted whereby models with deliberately introduced discrepancies

23

could be calibrated against synthetic data until results similar to those obtained from the experiment are observed (parameters varying periodically with time, large increases in damping in transitional regions etc.)

To this end, the authors revisited the numerical case study in Section 4.3 using a sinusoidal excitation instead of a random excitation. Interestingly this did not lead to results similar to the experimental case study - the parameters did not have to vary periodically for an excellent fit to the training data to be achieved. This, perhaps, indicates that the type of model error being simulated (missing stiffness nonlinearity) is not what was witnessed in the experimental case study. Future work may investigate other types of model error (for example, the errors induced when reducing the system to a 2 degree-of-freedom model). Whether or not such model error can be 'reverse engineered' from such results is an interesting topic for future work.

Further work will also seek to combine the proposed approach with a sensitivity analysis. This would help to establish, in a more rigorous manner, whether the witnessed variations in parameter values are truly significant. Combining the approach with a sensitivity analysis of the model in the region where it is to be extrapolated could be an important part of an overall validation framework. In particular, a detailed study concerning the affect of 'extrapolation severity' (in other words, the possible consequences of extrapolating further beyond the available data) should be conducted. Ultimately, the intention is to embed the proposed approach into validation frameworks that specifically consider extrapolation (see [11] and [12] for example).

7. Conclusions

This paper focuses on the situation where a model of a dynamical system is required to perform extrapolations (predictions in regions where there is no data). It considers how such a model could be validated, given that observation data is only available in a region far from the model's application domain. Such situations are often complicated by the fact that, through tuning a model's calibration parameters, it is often possible to gain false confidence in a model that will, in reality, perform poorly when used for extrapolation. Such a scenario is demonstrated using a numerical case study.

For the situation where it is believed that a model's calibration parameters should be time-invariant, a calibration procedure that allows these parameters to vary with time (even though it is believed that, physically, this should not be the case) is proposed. Results suggest that, by following such an approach, it is possible to identify important flaws in a model that would otherwise be hidden by more traditional calibration procedures. This is illustrated by a numerical and an experimental case study.

Appendix A. Evaluating Bayes factor

Bayesian system identification is first performed on $\mathcal{M}_2$ (that is, the situation where there is no model error). This analysis is performed in a similar manner to that described at the beginning of Section 4.2, where a prior distribution over k^* is now introduced:

$$p(k^*) = \mathcal{N}(k^*; 50, 10^2) \tag{A.1}$$

The resulting MCMC samples are shown in Figure A.21. It can be observed that the true values of all parameters lie within the main probability mass of the posterior PDF. This includes the noise term, R. Such results can be expected because the analysis considers an identification problem where there is no model error.
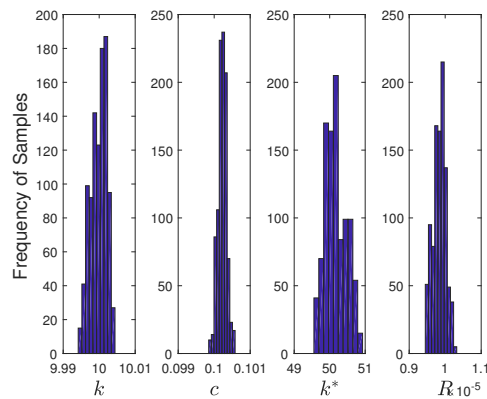

Figure A.21: Numerical case study. Histograms of samples from the posterior parameter distribution of the nonlinear model (all parameters in SI units.)

Having obtained posterior parameter samples for both model structures, Gaussian approximations of the two posterior parameter PDFs were realised. The quality of these approximations are shown in Figures A.22 and A.23.
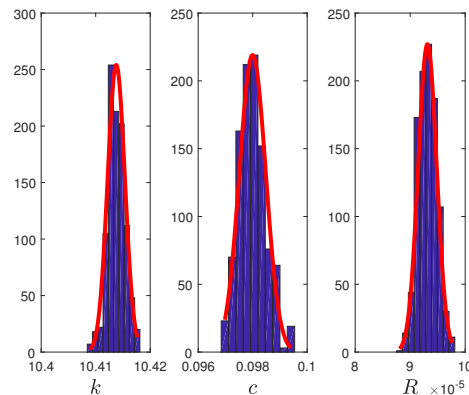

Figure A.22: Numerical case study. Gaussian fit to histograms of samples from the posterior parameter distribution of the linear model (all parameters in SI units.)
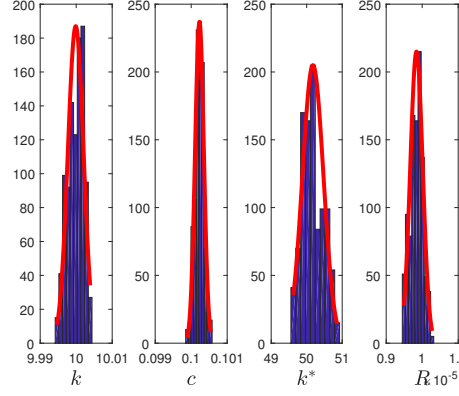
Figure A.23: Numerical case study. Gaussian fit to histograms of samples from the posterior parameter distribution of the nonlinear model (all parameters in SI units.)

Bayes factor was then estimated using Laplace's method. Firstly, the product of prior and the likelihood are written as

$$p(\mathcal{D}|\,\boldsymbol{\theta}_i,\mathcal{M}_i)p(\boldsymbol{\theta}_i\,|\mathcal{M}_i) = \exp(-J(\boldsymbol{\theta}_i)) \tag{A.2}$$

before a first-order Taylor series expansion of J is taken around the most probable parameter vector, $\boldsymbol{\theta}_i$:

$$J(\boldsymbol{\theta}_i) \approx J(\hat{\boldsymbol{\theta}}_i) + \frac{1}{2}(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^T \boldsymbol{A}(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i) \tag{A.3}$$

This allows one to write

$$p(\mathcal{D}|\,\boldsymbol{\theta}_i,\mathcal{M}_i)p(\boldsymbol{\theta}_i\,|\mathcal{M}_i) \approx \exp\left(J(\hat{\boldsymbol{\theta}}_i)\right)\mathcal{N}(\boldsymbol{\theta}_i;\hat{\boldsymbol{\theta}}_i,\boldsymbol{A}^{-1}) \tag{A.4}$$

where $\boldsymbol{A}$ is a Hessian matrix, containing second order derivatives of J with respect to $\boldsymbol{\theta}_i$. The Taylor series approximation then allows one to write

$$p(\mathcal{D}|\mathcal{M}_i) \approx \exp(J(\hat{\boldsymbol{\theta}}_i))\int \mathcal{N}(\boldsymbol{\theta}_i;\hat{\boldsymbol{\theta}}_i,\boldsymbol{A}^{-1})d\,\boldsymbol{\theta}_i \tag{A.5}$$

and

$$\log p(\mathcal{D}|\mathcal{M}_i) \approx J(\hat{\boldsymbol{\theta}}_i) + \frac{1}{2}\log\left((2\pi)^{N_\theta}|\boldsymbol{A}^{-1}|\right) \tag{A.6}$$

The elements of $\boldsymbol{A}$ can be estimated using finite difference methods, however, as samples from the posterior parameter distributions of both models were available in this case, the inverse of $\boldsymbol{A}$ was instead be realised by directly estimating the posterior covariance matrix. For both models the most probable parameter vector was chosen to be the mean of the MCMC samples.

Appendix B. Evaluating terms 1-4 in equation (21)

$$\left[\frac{\partial\,\boldsymbol{k}_1}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} = \left[\frac{\partial\dot{\boldsymbol{x}}(\boldsymbol{x}_{i-1},\boldsymbol{\delta}_{i-1})}{\partial\,\boldsymbol{x}_{i-1}}\right]_{\boldsymbol{x}_{i-1}=\hat{\boldsymbol{x}}} \equiv \tilde{\boldsymbol{F}}(\hat{\boldsymbol{x}},\boldsymbol{\delta}_{i-1}) \tag{B.1}$$

26

Defining

$$s_2 = x_{i-1} + \frac{\Delta t}{2} \, k_1(x_{i-1}, \delta_{i-1}) \tag{B.2}$$

then

$$\left[\frac{\partial \, k_2}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} = \left[\frac{\partial \dot{x}(s_2, \delta_{\text{int}})}{\partial \, s_2} \frac{\partial \, s_2}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \tag{B.3}$$

$$= \left[\frac{\partial \dot{x}(s_2, \delta_{\text{int}})}{\partial \, s_2} \right]_{s_2=\hat{x}+\frac{\Delta t}{2} \, k_1(\hat{x},\delta_{i-1})} \left(I + \frac{\Delta t}{2} \left[\frac{\partial \, k_1}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.4}$$

$$= \tilde{F} \left(\hat{x} + \frac{\Delta t}{2} \, k_1(\hat{x}, \delta_{i-1}), \delta_{\text{int}} \right) \left(I + \frac{\Delta t}{2} \left[\frac{\partial \, k_1}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.5}$$

Likewise, defining

$$s_3 = x_{i-1} + \frac{\Delta t}{2} \, k_2(x_{i-1}, \delta_{\text{int}}) \tag{B.6}$$

allows one to write

$$\left[\frac{\partial \, k_3}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} = \left[\frac{\partial \dot{x}(s_3, \delta_{\text{int}})}{\partial \, s_3} \frac{\partial \, s_3}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \tag{B.7}$$

$$= \left[\frac{\partial \dot{x}(s_3, \delta_{\text{int}})}{\partial \, s_3} \right]_{s_3=\hat{x}+\frac{\Delta t}{2} \, k_2(\hat{x},\delta_{\text{int}})} \left(I + \frac{\Delta t}{2} \left[\frac{\partial \, k_2}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.8}$$

$$= \tilde{F} \left(\hat{x} + \frac{\Delta t}{2} \, k_2(\hat{x}, \delta_{\text{int}}), \delta_{\text{int}} \right) \left(I + \frac{\Delta t}{2} \left[\frac{\partial \, k_2}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.9}$$

Finally, writing

$$s_4 = x_{i-1} + \Delta t \, k_3(x_{i-1}, \delta_{\text{int}}) \tag{B.10}$$

gives

$$\left[\frac{\partial \, k_4}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} = \left[\frac{\partial \dot{x}(s_4, \delta_i)}{\partial \, s_4} \frac{\partial \, s_4}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \tag{B.11}$$

$$= \left[\frac{\partial \dot{x}(s_4, \delta_i)}{\partial \, s_4} \right]_{s_4=\hat{x}+\Delta t \, k_3(\hat{x},\delta_{\text{int}})} \left(I + \Delta t \left[\frac{\partial \, k_3}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.12}$$

$$= \tilde{F} \left(\hat{x} + \Delta t \, k_3(\hat{x}, \delta_{\text{int}}), \delta_i \right) \left(I + \Delta t \left[\frac{\partial \, k_3}{\partial \, x_{i-1}} \right]_{x_{i-1}=\hat{x}} \right) \tag{B.13}$$

Appendix C. Matrices $\boldsymbol{F}$ and $\boldsymbol{H}$ for experimental case study

$$
\tilde{\boldsymbol{F}} =
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\frac{1}{m_1}(-k_2-k_1) & \frac{k_2}{m_1} & \frac{1}{m_1}(-c_2-c_1) & \frac{c_2}{m_1} & \frac{1}{m_1}(x_{\text{base}}-x_1) & \frac{1}{m_1}(x_2-x_1) & \frac{1}{m_1}(\dot{x}_{\text{base}}-\dot{x}_1) & \frac{1}{m_1}(\dot{x}_2-\dot{x}_1) \\
\frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{c_2}{m_2} & -\frac{c_2}{m_2} & 0 & \frac{x_1-x_2}{m_2} & 0 & \frac{\dot{x}_1-\dot{x}_2}{m_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\tag{C.1}
$$

Noting that acceleration measurements are begin utilised and that acceleration is related to the system's state by equations (36) and (37), it can be show that

$$
\boldsymbol{H} =
\begin{bmatrix}
\frac{1}{m_1}(-k_2-k_1) & \frac{k_2}{m_1} & \frac{1}{m_1}(-c_2-c_1) & \frac{c_2}{m_1} & \frac{1}{m_1}(x_{\text{base}}-x_1) & \frac{1}{m_1}(x_2-x_1) & \frac{1}{m_1}(\dot{x}_{\text{base}}-\dot{x}_1) & \frac{1}{m_1}(\dot{x}_2-\dot{x}_1) \\
\frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{c_2}{m_2} & -\frac{c_2}{m_2} & 0 & \frac{x_1-x_2}{m_2} & 0 & \frac{\dot{x}_1-\dot{x}_2}{m_2}
\end{bmatrix}
\tag{C.2}
$$

(in fact, these are just the 3rd and 4th rows of $\tilde{\boldsymbol{F}}$).

[1] F. Hemez, H.S. Atamturktur, and C. Unal. Defining predictive maturity for validated numerical simulations. *Computers & structures*, 88(7-8):497–505, 2010.

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

[3] M. Goldstein and J. Rougier. Probabilistic formulations for transferring inferences from mathematical models to physical systems. *SIAM journal on scientific computing*, 26(2):467–487, 2004.

[4] M.J. Bayarri, J.O. Berger, R. Paulo, J. Sacks, J.A. Cafeo, J. Cavendish, C.H. Lin, and J. Tu. A framework for validation of computer models. *Technometrics*, 49(2):138–154, 2007.

[5] M.J. Bayarri, J.O. Berger, M.C. Kennedy, A. Kottas, R. Paulo, J. Sacks, J.A. Cafeo, C.H. Lin, and J. Tu. Predicting vehicle crashworthiness: Validation of computer models for functional and hierarchical data. *Journal of the American Statistical Association*, 104(487):929–943, 2009.

[6] C. Unal, B. Williams, F. Hemez, S.H. Atamturktur, and P. McClure. Improved best estimate plus uncertainty methodology, including advanced validation concepts, to license evolving nuclear reactors. *Nuclear Engineering and Design*, 241(5):1813–1833, 2011.

[7] S. Atamturktur, F. Hemez, B. Williams, C. Tome, and C. Unal. A forecasting metric for predictive modeling. *Computers & Structures*, 89(23-24):2377–2387, 2011.

[8] S. Atamturktur, J. Hegenderfer, B. Williams, M. Egeberg, R.A. Lebensohn, and C. Unal. A resource allocation framework for experiment-based validation of numerical models. *Mechanics of Advanced Materials and Structures*, 22(8):641–654, 2015.

[9] J. Hegenderfer and S. Atamturktur. Prioritization of code development efforts in partitioned analysis. *Computer-Aided Civil and Infrastructure Engineering*, 28(4):289–306, 2013.

[10] M.J. Bayarri, J.O. Berger, J. Cafeo, G. Garcia-Donato, F. Liu, J. Palomo, R.J. Parthasarathy, R. Paulo, J. Sacks, and D. Walsh. Computer model validation with functional output. *The Annals of Statistics*, pages 1874–1906, 2007.

[11] V.R. Joseph and H. Yan. Engineering-driven statistical adjustment and calibration. *Technometrics*, 57(2):257–267, 2015.

[12] T.A. Oliver, G. Terejanu, C.S. Simmons, and R.D. Moser. Validating predictions of unobserved quantities. *Computer Methods in Applied Mechanics and Engineering*, 283:1310–1335, 2015.

[13] J. Brynjarsdóttir and A. OÊijHagan. Learning about physical parameters: The importance of model discrepancy. *Inverse Problems*, 30(11):114007, 2014.

[14] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002.

[15] E.T. Jaynes. *Probability theory: the logic of science*. Cambridge university press, 2003.

[16] E. Simoen, C. Papadimitriou, and G. Lombaert. On prediction error correlation in Bayesian model updating. *Journal of Sound and Vibration*, 332(18):4136–4152, 2013.

[17] P.L. Green. Bayesian system identification of dynamical systems using large sets of training data: A MCMC solution. *Probabilistic Engineering Mechanics*, 42:54–63, 2015.

[18] J.B. Kadane and N.A. Lazar. Methods and criteria for model selection. *Journal of the American statistical Association*, 99(465):279–290, 2004.

[19] P.J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[20] J. Skilling et al. Nested sampling for general Bayesian computation. *Bayesian analysis*, 1(4):833–859, 2006.

[21] J. Ching and Y.C. Chen. Transitional Markov chain Monte carlo method for Bayesian model updating, model class selection, and model averaging. *Journal of engineering mechanics*, 133(7):816–832, 2007.

[22] J. Ching, J.L. Beck, and K.A. Porter. Bayesian state and parameter estimation of uncertain dynamical systems. *Probabilistic engineering mechanics*, 21(1):81–96, 2006.