Chapter 1
On Model Validation and Bifurcating Systems: An Experimental Case Study

Keith Worden, David J. Wagg, and Malcolm Scott

Abstract This chapter demonstrates some of the problems that can arise when validating models of nonlinear bifurcating systems and the approaches that can avoid them. Validation is the process of determining the extent to which a model accurately represents the structure or system of interest. Additional care needs to be taken when attempting to validate models of nonlinear systems because of bifurcations that may occur. These phenomena present a difficulty for validation because if a model does not precisely capture the bifurcation points, then the model's predictions could be very inaccurate, even if the model is (parametrically) very close to the real system. This situation could lead to a good model being dismissed if data generated close to a bifurcation point were used to validate it. In this chapter, experimental data were gathered from a three-storey shear building structure with a harsh nonlinearity between the top two floors, and bifurcations were observed in the structural response. Two models are constructed here, with parameters estimated using Bayesian system identification: a linear model and a nonlinear model. Selected features and metrics were then used to compare the model predictions to the test data. The results show that an appropriate model could be rejected if an inappropriate validation strategy is employed, purely as a result of slightly misplaced bifurcations. It is demonstrated that discrimination can be improved by taking modelling uncertainties into account as part of the validation process.

Keywords Validation and verification · Nonlinear systems · Bifurcations · Bayesian system identification

1.1 Introduction

The purpose of this chapter is to discuss some issues associated with the validation of computational models of nonlinear systems: in particular, systems that *bifurcate*. It is a very late follow-up to the paper [1], which discussed some of the issues in the context of a single-degree-of-freedom (SDOF) system. In this chapter, results will be presented for more realistic (MDOF) nonlinear systems and will address some of the issues involved in identifying and validating such systems. In this chapter, the definition of *model validation* is taken from [2]: "Validation is 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." The process of validation—as a baseline—typically involves comparing model simulations to experimental data; this can be done directly, or features can be extracted from the data and compared. Model uncertainties must also be taken into account during validation. A key point in the definition above is that relating to "intended use" of the model. While complete validation might be theoretically desirable, i.e., one establishes that the model provides a perfect description of physical reality, this is likely to be impossible. In practice, one would typically be satisfied by showing that the model is effective over a given set of circumstances and in making a restricted set of possible predictions over that set.

The framework for model validation for linear systems is arguably in a fairly healthy state; this is partly because of the "smoothness" of linear systems; if the parameters of a linear system model are perturbed slightly, the perturbed model predictions will only depart slightly from the predictions of the original model; there is *continuity* in the response, as a function of parameters. This state of affairs will also hold for *weakly nonlinear* systems; however, it will fail for nonlinear systems that *bifurcate* [3]. Such systems may change their output responses drastically, given even small changes in their parameters. Bifurcations can occur in many different forms, but one of the most important types in structural dynamics are the *jump phenomena* observed during experimental tests when a sine-wave excitation is varied in its frequency. The situation

K. Worden (✉) · D. J. Wagg · M. Scott
Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Sheffield, UK
e-mail: k.worden@sheffield.ac.uk; d.j.wagg@sheffield.ac.uk; m.s.scott@sheffield.ac.uk

Z. Mao (eds.), *Model Validation and Uncertainty Quantification, Volume 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.1007/978-3-031-04090-0_1

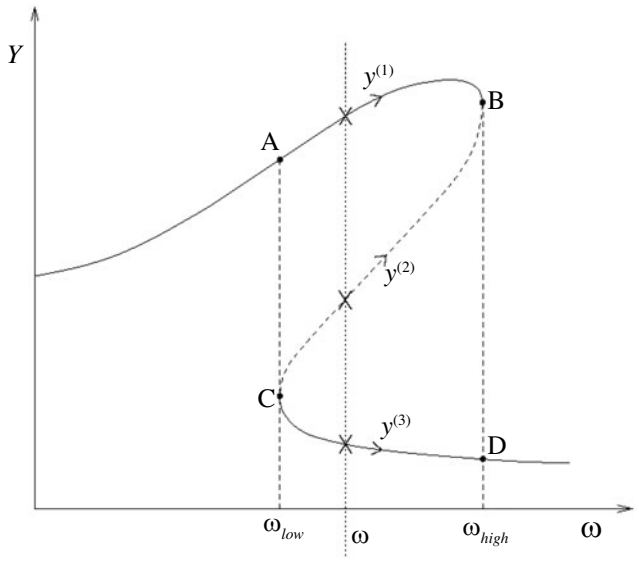


Fig. 1.1 Frequency response of the harmonically forced Duffing oscillator

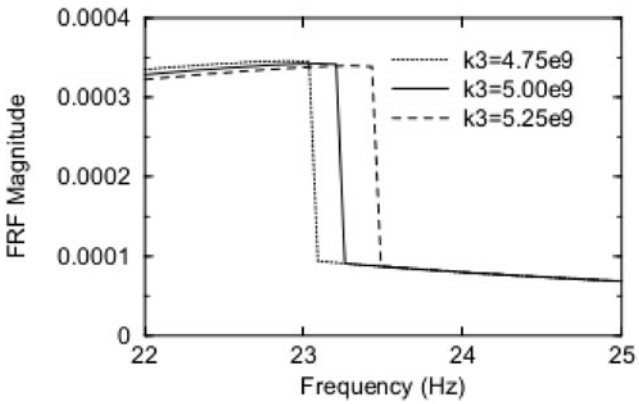


Fig. 1.2 FRFs for Duffing systems with slightly different k_3 parameters

is summed up quite nicely in Fig. 1.1, which expresses the amplitude response of an SDOF Duffing oscillator system [3],

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t). \tag{1.1}$$

As the excitation $x(t) = X\cos(\omega t)$ experiences a smooth change in the frequency ω, there is a smooth increase in the response amplitude. If X is small, such that the system has a nominally linear behaviour, the response will reach a peak (resonance) and then decline smoothly with further frequency increase. If, however, X is large, at some frequency ω_{high}, the amplitude will drop sharply from its high value to a low value (point B in the figure drops to point D). If the frequency is swept down from high values, a jump *up* will occur at some frequency ω_{low}.

A common approach to model validation is to give a quantified measure of how accurate the predictions made by the model will be, taking into account all possible sources of uncertainty [4]. Such sources include uncertainties in the excitation parameters or in the parameters of the model, if it has been identified from data. Bifurcations are clearly an issue here. Consider Fig. 1.2; this shows the range of jump frequencies obtained if the nonlinear stiffness, k_3 in Duffing's equation, is varied from 4.75×10^9 to 5.25×10^9. Clearly, a small error in this parameter, if identified from data, could make the difference between a small amplitude prediction and a large one.

The issue for validation is this: suppose that one has identified a very good model from data. Furthermore, suppose that one is attempting to validate the model by comparing its predictions from measured data, using an error metric like the normalised mean-square error (NMSE), defined by

$$J(\underline{\theta}) = \frac{100}{N\sigma_y^2}\sum_{i=1}^{N}(\ddot{y}_i - \hat{\ddot{y}}_i(\underline{\theta}))^2, \tag{1.2}$$

where the $\ddot{y}_i$ are measured samples from an acceleration time series, and $\hat{\ddot{y}}_i(\underline{\theta})$ are the corresponding predictions from the model; $\underline{\theta}$ are the model parameters and N response points are sampled and have variance σ_y^2. (With the normalisation here, an NMSE value of 5% can be considered to indicate a good model, while a value of 1% indicates excellence [3]). If the model parameters are an arbitrarily small value below the "true" parameters, a sine excitation at the "wrong" frequency will elicit a very large difference in the responses, and one might be inclined to reject a good model.

The original paper [1] made a number of observations on the effects of bifurcation on model validation and suggested a number of comparison features that might be less sensitive to small errors in model parameters. However, the discussion was limited to SDOF systems and did not make contact with reality in a number of respects. The objective of the current paper is to take the discussion a little further by considering MDOF systems and actually comparing identified model predictions with reality, using a real experimental structure.

1.2 Experimental Case Study

In order to discuss some of the issues referred to earlier, data were acquired from a laboratory structure; this was a model of a three-storey shear building, with substantial masses representing the floors. The mass of the columns could be neglected in comparison to the floors, so the structure had effectively three DOFs. In fact, the structure was heavily influenced by the structure discussed in [5], which was developed at Los Alamos National Laboratories. That structure was affectionately known as the "bookshelf" structure and will be referred to as such here. The main difference between the structure here and the LANL original is that the original was mounted on linear bearings and excited at the base, whereas here the base of the structure is fixed to a testing table and forcing is applied at the first storey using an electrodynamic shaker. A schematic of the new bookshelf is given in Fig. 1.3.

Each "shelf" of the bookshelf rig is made from a solid rectilinear aluminium block of dimensions 350 × 255 × 25 mm and has a mass of approximately 6.4 kg. At each corner, the shelves are connected to the upright beams using a block and four bolts. A column is bolted to the underside of the top floor, and a motion-limiting constraint, referred to here as the

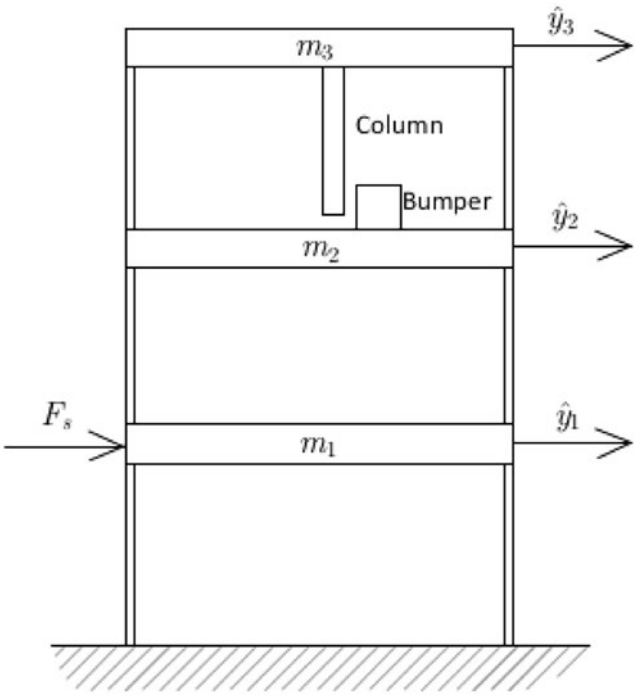


Fig. 1.3 Schematic of "bookshelf" rig

Table 1.1 Excitation frequencies and amplitudes of data

Excitation frequency (Hz)	Amplitude (N)	Excitation frequency (Hz)	Amplitude (N)	Excitation frequency (Hz)	Amplitude (N)
17	156	28	229	53	94
17.5	154	32	311	55	128
18	138	36	203	57	195
18.5	120	40	208	60	155
19	82	44	151	66	257
19.5	54	46	126	67	258
19.7	96	47	102	70	249
20	150	47.5	82	72	203
20.5	168	48	80	73	221
21	170	48.5	77	74	189
21.5	172	51.5	82	75	168
24	269	52	89		

"bumper," is bolted to the top of the second floor. The position of the bumper can be moved in order to adjust the gap d between the column and the bumper. The system responses were measured using accelerometers attached to each floor/shelf. Both data acquisition and shaker control were accomplished using an LMS SCADAS-3 interface connected to a PC running LMS Test.Lab software.

1.2.1 Test Sequence

Two groups of data were gathered, each consisting of 35 data sets generated using periodic forcing between 17 and 75 Hz; this frequency range contains the first two natural frequencies of the structure. Higher frequencies were not considered because the bumper–column impacts caused the results to become very noisy. The test frequencies were concentrated more in regions where preliminary tests showed more nonlinear phenomena. Some additional measurements were made to give a more accurate picture of the frequency response and to locate bifurcation (jump) frequencies. The first group of data was taken with the gap between the bumper and column too far apart for them to engage and is referred to here as the *linear data*. Linearity of the system in this case was confirmed using reciprocity checks and noting the absence of any response harmonics. The second group of data was gathered with the gap d set to approximately 0.5 mm and is referred to here as the *nonlinear data*. The excitation amplitudes were adjusted so that the response amplitudes for each data set were at a similar level; this was done for three reasons: (1) to ensure that the bumper engaged for the nonlinear data; (2) to ensure that the measurement noise was small compared to the measured response; and (3) to ensure that the response amplitudes were not big enough to compromise the linearity of system when the bumper did not engage. The forcing amplitudes for the nonlinear data are shown in Table 1.1; similar amplitudes were used for the linear data.

Sinusoidal excitation was chosen here as it is known to generally produce the strongest manifestation of nonlinear phenomena. A good approximation to sinusoidal forcing was fairly simple to achieve for the structure in its linear state. However, when the bumper engaged, the harsh nonlinearity made controlling the input force rather difficult. The nonlinear experimental results presented below resulted from periodic forcing of a similar nature to that shown in Fig. 1.4. The force signal contains harmonics at multiples of the forcing frequency and the natural frequencies of the system. This pollution does not pose a problem from a model validation point of view because the measured forcing from the test data could be used as an input to a model when generating features. Furthermore, despite the non-sinusoidal forcing, multiple bifurcations were still observed in the test data.

1.2.2 Response Data

For many of the excitation frequencies, the time histories of the linear and nonlinear data look very different; an example of this can be seen in the plot in Fig. 1.5 that shows the second floor response when excited at 44 Hz. The linear system

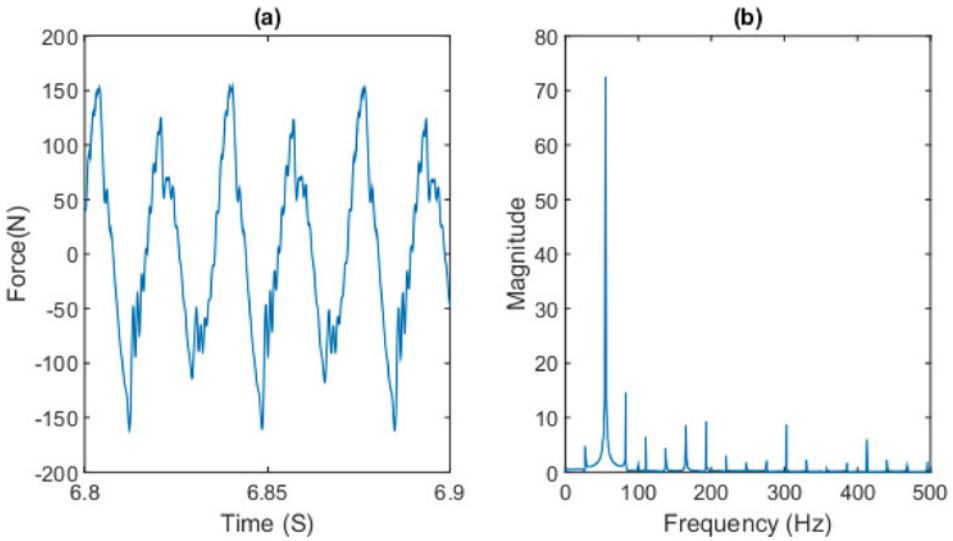


Fig. 1.4 Example of force signal applied to the bookshelf structure rig: (**a**) time history; (**b**) spectral density

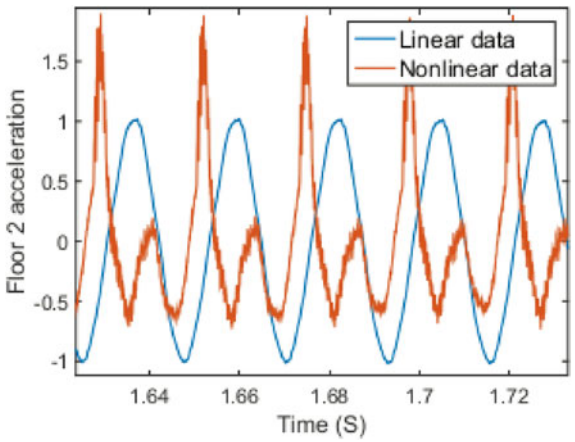


Fig. 1.5 Acceleration time histories at 44 Hz, linear and nonlinear data

response is close to sinusoidal, whereas the nonlinear response is very asymmetric, showing a strong second harmonic. Plots of frequency response values (acceleration amplitude divided by forcing amplitude) for the linear and nonlinear data are shown in Fig. 1.6. For non-harmonic responses, the amplitude here is taken as half of the peak to peak distance. When in its linear state, the structure has its first two natural frequencies at around 19.3 and 49.5 Hz; in its nonlinear state, the maximum response values were seen at 19.5 and 51.5 Hz. At the second resonance, the nonlinear data shows a "bending" of the peak that is typical of dynamic systems with a hardening stiffness [3]. Figure 1.6 also shows a sharp drop after the second peak of the nonlinear response; this is a result of a saddle-node "jump" bifurcation as discussed in the introduction here. Further tests were made to pinpoint bifurcation frequencies; as expected, these were sensitive to forcing amplitude. At forcing levels of 115, 130, and 150 N, the response amplitude jumps took place at 53.558, 54.557, and 55.065 Hz, respectively. A nice example of a jump can be seen in Fig. 1.7, which shows the acceleration time histories at 55.065 and 55.066 Hz at 150 N forcing. If the rig is excited very close to one of these bifurcation frequencies, its response has two possible stable amplitudes. By perturbing the system, it was possible to switch the response between them. Briefly resting a hand on the top floor caused it to move from the high-to-low-amplitude response; it could be returned to the high-amplitude response by giving the top floor a sharp tap.

At periodic excitations between 65 and 75 Hz, the nonlinear data shows subharmonic oscillations, as in Fig. 1.8, which shows the acceleration time history when excited at 68 Hz. It is likely that these are as a result of period-doubling bifurcations, but constructing bifurcation diagrams for the test data is problematic because of the measurement noise and stiffness coupling effects.

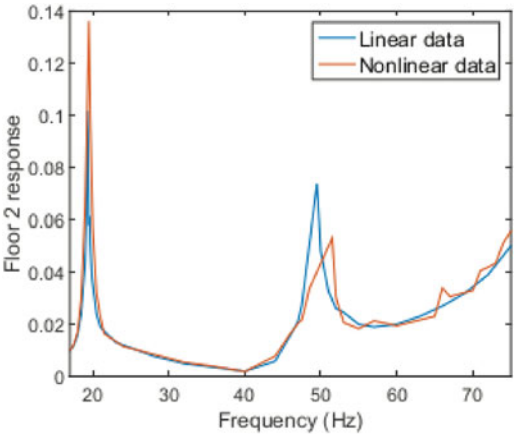


Fig. 1.6 Plot showing response of the second floor to periodic forcing, linear and nonlinear data

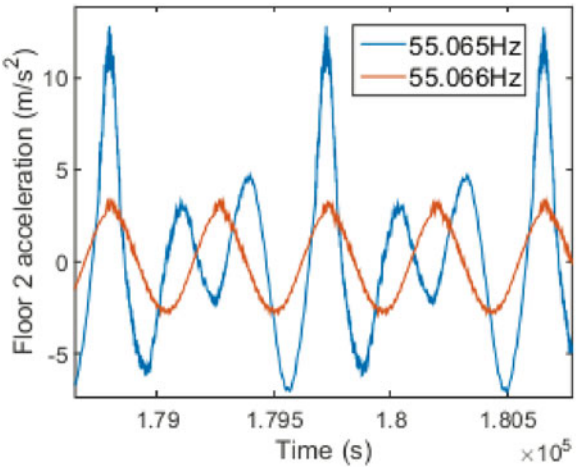


Fig. 1.7 Second floor acceleration time histories, when excited at 55.065 and 55.066 Hz

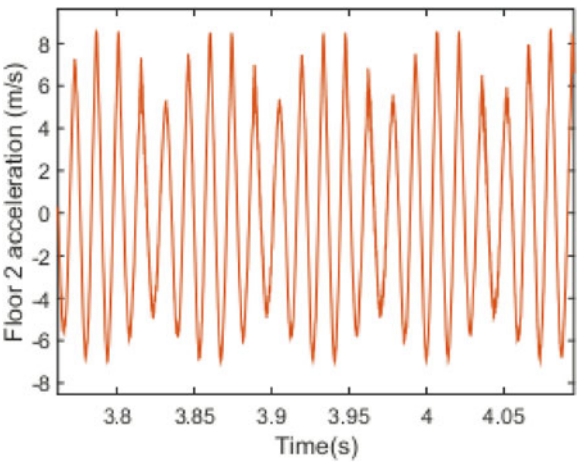


Fig. 1.8 Second floor acceleration when excited at 68 Hz

1.3 Model Development and Parameter Estimation

1.3.1 Equations of Motion

For the purposes of illustrating validation strategies, two models of the bookshelf rig have been created here: a linear model and a nonlinear model. These differ only in that the nonlinear model accounts for the bumper mechanism. In each case, the rig is modelled as a lumped-mass, three-degree-of-freedom system. For the linear model, it is assumed that there is a linear stiffness and damping between each floor, giving equations of motion,

$$\begin{aligned}
\ddot{y}_1 &= (F_s - k_1 y_1 - k_2(y_1 - y_2) - c_1\dot{y}_1 - c_2(\dot{y}_1 - \dot{y}_2))/m_1 \\
\ddot{y}_2 &= (-k_2(y_2 - y_3) - k_3(y_2 - y_3) - c_2(\dot{y}_1 - \dot{y}_2) - c_3(\dot{y}_2 - \dot{y}_3))/m_2 \\
\ddot{y}_3 &= (-k_3(y_3 - y_2) - c_3(\dot{y}_2 - \dot{y}_3))/m_3,
\end{aligned} \tag{1.3}$$

where y_i, k_i, c_i, and m_i are, respectively, the floor displacements, stiffness coefficients, damping coefficients, and floor masses for $i = \{1, 2, 3\}$ (Fig. 1.3); overdots denote differentiation with respect to time. For the nonlinear model, it is assumed that there is an additional linear stiffness between the top two floors, when the bumper is in contact; the equations of motion become

$$\begin{aligned}
\ddot{y}_1 &= (F_s - k_1 y_1 - k_2(y_1 - y_2) - c_1\dot{y}_1 - c_2(\dot{y}_1 - \dot{y}_2))/m_1 \\
\ddot{y}_2 &= (-k_2(y_2 - y_3) - k_3(y_2 - y_3) - F_c - c_2(\dot{y}_1 - \dot{y}_2) - c_3(\dot{y}_2 - \dot{y}_3))/m_2 \\
\ddot{y}_3 &= (-k_3(y_3 - y_2) - c_3(\dot{y}_2 - \dot{y}_3))/m_3,
\end{aligned} \tag{1.4}$$

where

$$F_c = \begin{cases} 0 & \text{if } (y_3 - y_2) < d \\ k_c(y_2 - y_3 + d) & \text{if } (y_3 - y_2) \geq d \end{cases} \tag{1.5}$$

with k_c the additional linear stiffness from the bumper and d the clearance between the bumper and column when the structure is at rest. Clearly, the effects of the bumper–column interaction are considerably more complicated than a simple bilinear stiffness; however, these are difficult to model with a simple lumped-mass system. The bilinear model was chosen as a baseline, as such systems are still capable of producing the significant nonlinear phenomena observed in the test data: amplitude jumps, superharmonics, and subharmonics [3].

The responses here were simulated using a fourth-order Runge–Kutta scheme with a timestep of 4×10^{-5}. The short sampling interval was necessary because of the discontinuous nature of the nonlinearity. The simulation allowed the forcing to be taken from actual measurements on the rig, as well as using sine-wave and random excitation. Typically, the initial 150 forcing cycles were discarded to remove transient effects.

Some preliminary results on model validation for this structure were presented in [6].

1.3.2 Parameter Estimation

The parameter estimation problem here is quite demanding; there are 11 parameters involved, and one—the clearance d—enters the equations in a nonlinear fashion. Two algorithms were applied here in order to produce the final estimates. In the first step, the *self-adaptive differential evolution* (SADE) algorithm was used [7]. SADE has proved to be a powerful choice in previous work for models that are nonlinear in the parameters and can operate directly on the measured accelerations without the need to integrate to velocity and displacement [8]. Although SADE can estimate confidence intervals for parameters [9], it assumes they have Gaussian distributions, which may be misleading for nonlinear systems. To better characterise the uncertainty in the parameters, a second identification step was applied—a Bayesian approach [10], with the SADE algorithm providing the initial parameter estimates. In this case, a vanilla Metropolis–Hastings (MH) algorithm was used. The Bayesian approach allows sampling from the actual parameter distributions, so that histograms or density estimates can be computed and displayed. Both of these algorithms, SADE and Bayesian ID, have been applied effectively on variants of the bookshelf structure in the past [11, 12].

In the first exercise here, a linear baseline model was identified using data acquired with the structure in its *nonlinear* state. As both SADE and the MH algorithm are stochastic algorithms, ten runs were carried out for each algorithm, and the results with lowest cost were selected. The training data were composed of ten sets of response data generated by excitation frequencies between 18 and 60 Hz with 150 cycles per frequency; in each case, the NMSE and parameters were averaged to give the final result. In the best run, SADE gave an acceptable NMSE of 5.4%. The MH algorithm was then run using the SADE parameter estimates as initial values; a Gaussian proposal distribution was used with variances set at 2% of the initial estimates. The priors on the parameters were taken as uniform, with bounds one order of magnitude below and above the SADE estimates. A burn-in period of 5000 iterations was used, after which 30,000 iterations were required for the algorithm to converge on a stationary distribution. With the MH step, the average NMSE between the linear model and nonlinear training data improved slightly to 5.2%. The results of the parameter estimation, of both the SADE and MH steps, are reported in Table 1.2.

Table 1.2 Linear model parameter values found using SADE and the MH algorithm

Parameter	SADE estimate	MH—mean	MH—standard deviation
m_1	6.4 kg	6.89 kg	0.402 kg
m_2	6.64 kg	6.67 kg	0.39 kg
m_3	6.26 kg	6.08 kg	0.355 kg
k_1	4.32×10^5 N/m	4.29×10^5 N/m	0.229×10^5 N/m
k_2	6.39×10^5 N/m	6.28×10^5 N/m	0.357×10^5 N/m
k_3	3.99×10^5 N/m	3.57×10^5 N/m	0.145×10^5 N/m
c_1	53.0 Ns/m	76.8 Ns/m	8.30 Ns/m
c_2	99.2 Ns/m	118 Ns/m	8.76 Ns/m
c_3	6.2 Ns/m	5.19 Ns/m	0.429 Ns/m

Table 1.3 Nonlinear model parameter values found using SADE and the MH algorithm

Parameter	SADE estimate	MH—mean	MH—standard deviation
m_1	6.69 kg	6.76 kg	0.412 kg
m_2	6.65 kg	6.63 kg	0.401 kg
m_3	6.26 kg	6.17 kg	0.395 kg
k_1	4.36×10^5 N/m	4.28×10^5 N/m	0.260×10^5 N/m
k_2	6.34×10^5 N/m	6.38×10^5 N/m	0.346×10^5 N/m
k_3	3.56×10^5 N/m	3.58×10^5 N/m	0.157×10^5 N/m
c_1	78.0 Ns/m	79.0 Ns/m	7.37 Ns/m
c_2	110.2 Ns/m	119.1 Ns/m	8.62 Ns/m
c_3	6.9 Ns/m	5.12 Ns/m	0.463 Ns/m
k_c	5.06×10^5 N/m	4.83×10^5 N/m	0.384×10^5 N/m
d (mm)	0.480	0.491	0.0358

The second exercise here was to estimate parameters for the nonlinear model from the *nonlinear structure*. In this case, a three-stage process was used because the MH algorithm failed to work without good initial estimates, but SADE struggled to produce good estimates when all 11 parameters were included. In the first stage, parameters were estimated for the linear model with the structure in its *linear state*. In the second stage, the nonlinear model was estimated with the *nonlinear* data, with the mass and stiffness estimates held at the values from the previous stage; this produced estimates of k_c and d and refined the damping estimates. With this procedure, the first stage gave an NMSE of 3.0% for the linear model and 4.3%, which were considered acceptable. Using the SADE estimates as initial estimates, as before, the MH algorithm was then run, with the same settings as used previously. The MH run reduced the NMSE to 3.7%, averaged over the three floors. The parameters from these algorithms are given in Table 1.3. Apart from the low NMSE values, it is reassuring that the mass estimates are all close to the calculated mass of the floor blocks (6.4 kg), with m_3 slightly lower as it has less column mass entrained. The parameter histograms from MH are presented in Fig. 1.9; note that some of these appear quite non-Gaussian.

1.4 Validation Features and Metrics for Nonlinear Systems

In structural dynamics applications, a *feature* is a subset of data extracted from the raw data of a dynamic response, which can be used to compare experimental data to a model output. A *metric* represents a quantitative evaluation of the similarity between recorded and predicted features. Described below are a selection of features and the metrics used in the following sections. The features are time histories, FRFs, and bifurcation points, and the metrics are NMSEs and t-statistics. Only features appropriate to periodically excited systems are considered here. A good discussion of features for validating models of randomly excited systems is given in [5].

The first feature considered is the system response time history itself, typically either displacement, velocity, or acceleration. Acceleration is the most commonly used as it is arguably the easiest to measure. A common metric used for comparing time histories is the NMSE defined in Eq. (1.2). The advantages of comparing time series for model validation are that they are information rich; disadvantages are that they are unlikely to be the "quantity of interest" for the intended use of the model, the model might do a good job of predicting important aspects of system behaviour but still have a poor NMSE

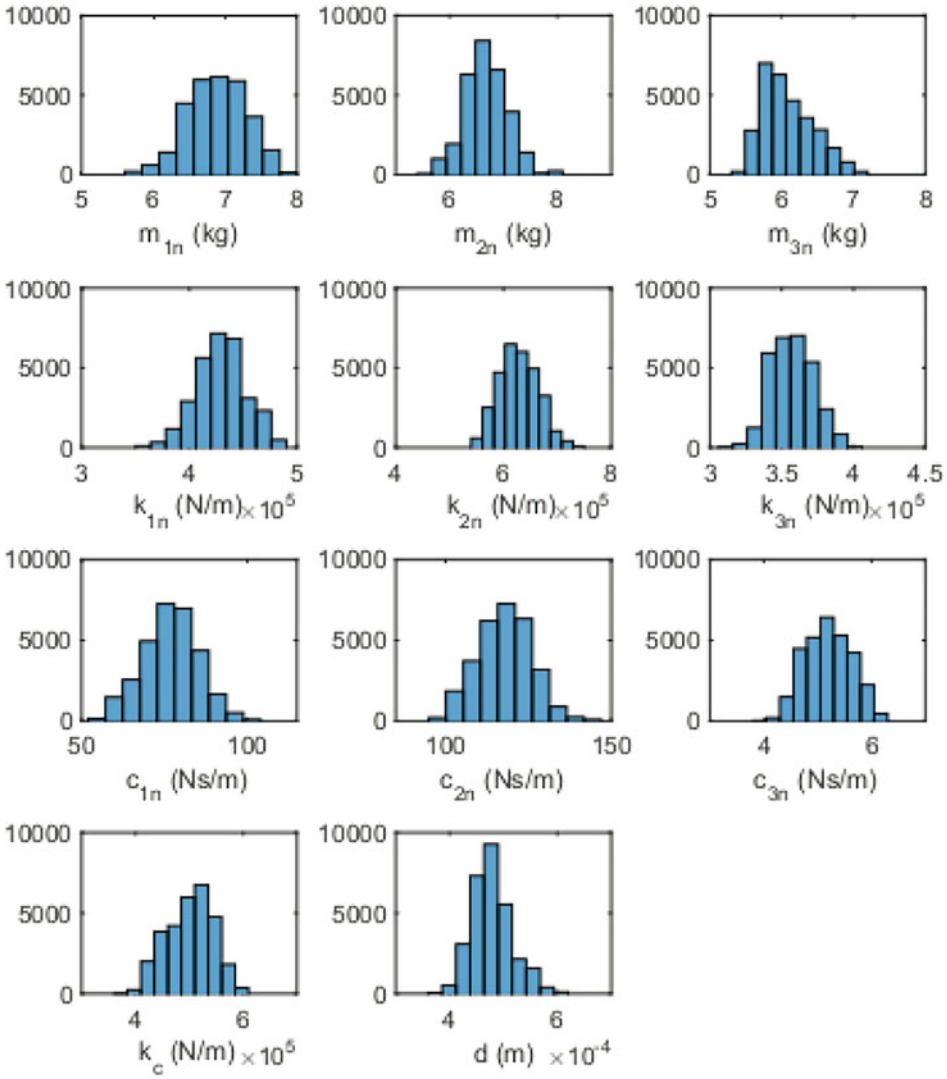


Fig. 1.9 Histograms of the nonlinear parameter values generated by the MH algorithm

between time histories, and it is difficult to take uncertainties into account. Furthermore, as discussed in [1], the NMSE has problems with systems that are sensitive to initial conditions.

Frequency response functions (FRFs) are options for validation features; these are straightforwardly estimated for random excitation or swept-sines. However, care is needed for nonlinear systems, as different excitations produce different FRF estimates, and an FRF estimated by random excitation may not represent well the behaviour of the system with harmonic forcing [3]. FRFs for swept-sine excitation of nonlinear structures can be challenging to obtain because precise control of the input waveform is needed. FRFs can be compared in the same way as time histories, using an NMSE as the metric.

While features and metrics can be used in a deterministic fashion for validation, it is much better to account for uncertainties. If parameter uncertainties can be propagated to give response value probability *distributions*, appropriate metrics exist for comparing these [5]. (As discussed in [1], FRFs and probability densities can be features that are insensitive to initial conditions.) A good discussion on statistical tests for validation metrics is given in [13]. Hypothesis tests for validation usually require an assumption about the distribution of interest; often they are taken to be Gaussian. In the Gaussian case, a simple measure of discrepancy is the t-statistic,

$$t = \frac{|r - \overline{r}|}{s}, \tag{1.6}$$

where r is the measured feature and $\bar{r}$ and s are the sample mean and standard deviation of the predicted feature distribution. If the response distribution is not Gaussian, it may be possible to make a transformation to bring it closer before computing the t-statistic. The Box–Cox transformation is a simple transformation used to remove skewness from distributions [14], as follows:

$$r_\lambda = \frac{\lambda r - 1}{\lambda}, \tag{1.7}$$

where λ is a constant that is chosen to minimise the skewness of the transformed distribution.

For nonlinear systems, a good approach is to select features that directly address that nonlinearity; in the current context, a good feature would be the jump or frequency bifurcation. It is clearly important to know whether the structure and model will jump at different frequencies.

The following section will describe the application of these features and metrics in validating the models identified for the experimental bookshelf structure.

1.5 Validation

This section presents the results of validating the linear and nonlinear models of the bookshelf structure using three different features: time histories, FRFs, and bifurcation frequencies. For each feature, deterministic and non-deterministic metrics were considered, the latter taking uncertainty into account. Twenty-five data sets, gathered using periodic forcing between 17 and 75 Hz, were used for model validation. These were separate from the training data used to estimate the parameters. In the analysis, the uncertainty in the parameter estimates from the MH algorithm was propagated into the model outputs.

The Markov chains from the algorithm generated 30,000 samples of each parameter. First of all, a Maximin Latin Hypercube design was used to select 1000 subsamples that covered the parameter space. For each parameter, the model was run for 300 forcing periods with the measured shaker forces as excitation, and the first 150 cycles were discarded to eliminate transients. For each simulation, the NMSE was calculated, giving a distribution over the runs; the frequency response amplitude was also calculated for each set of response data. These distributions were used to calculate the non-deterministic validation metrics in the following two sections.

1.5.1 Time Histories

The deterministic validation metric used here was the NMSE between modelled and measured acceleration time histories. In each model run, the measured force from the shaker was used as input. Model parameter values were taken as the mean of the distributions estimated in Sect. 1.3.2. The average NMSEs for the linear and nonlinear model were 17% and 35%, respectively. These might be considered high, but it should be noted that this is the first time the models have encountered validation data. At first sight, it appears that the linear model is performing better; however, a closer look is needed. In fact, the high average for the nonlinear models is because of three data sets generating very high errors; at 52 and 53 Hz, the nonlinear model gives errors of 200% and 500%, respectively. These high values are a result of the model misplacing the bifurcation frequency and predicting high response amplitudes when the true values are low (and *vice versa*). This highlights a disadvantage of using time-series features for models of bifurcating systems; the model is penalised for replicating the bifurcation behaviour of the system unless it does so near-perfectly. After removing the three "outliers," the NMSEs are plotted as shown in Fig. 1.10, from which it is clear that the nonlinear model generally produces better predictions.

Figure 1.11 shows the modelled and measured time histories of the second floor acceleration when the structure is excited at 21 Hz. While the nonlinear model reproduces the harmonic content seen in the real system, it actually generates an error of 6% compared to the 3% of the linear model. This effect is because of the incorrect phase of the harmonics; this is another example of the model being penalised for replicating the nonlinear behaviour of the system unless it does so extremely closely.

When the parameter uncertainty is propagated, distributions of the NMSEs result, as shown in Fig. 1.12. The plot shows the maximum and minimum values for each frequency. As before, the errors are dominated by some very high values; however, in the mean (with outliers removed), the nonlinear model performs better than the linear one. At some frequencies close to the bifurcation point, the NMSE distributions were bimodal depending on whether they correctly predicted the upper or lower branch for the amplitude.

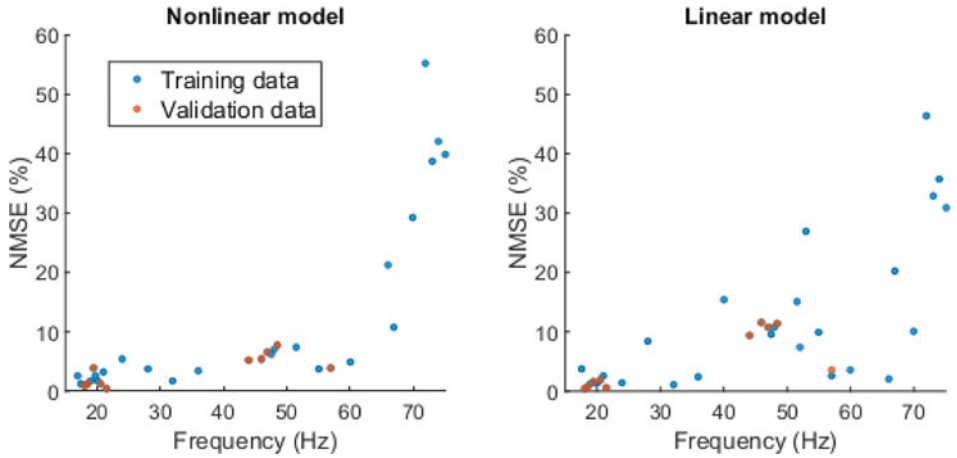


Fig. 1.10 NMSE values below 30% for nonlinear training and validation data, linear and nonlinear models

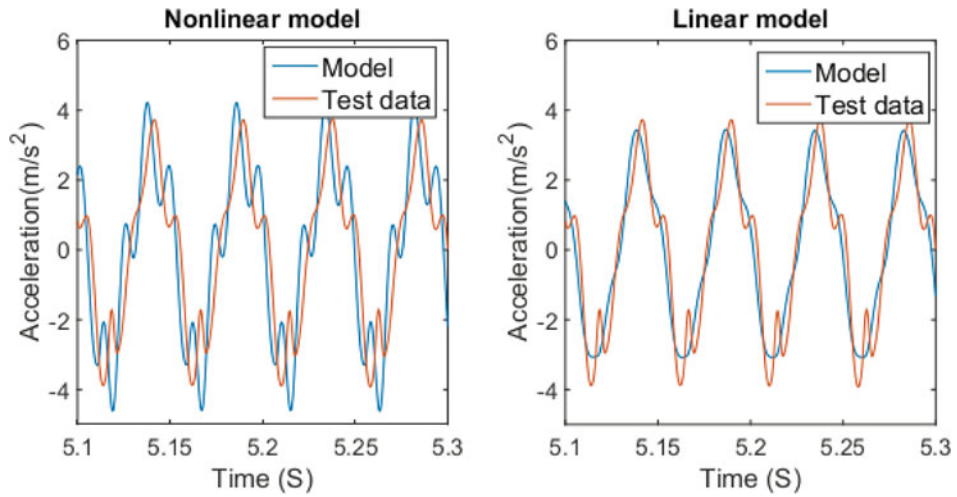


Fig. 1.11 Modelled and measured acceleration when excited at 21 Hz, linear and nonlinear models

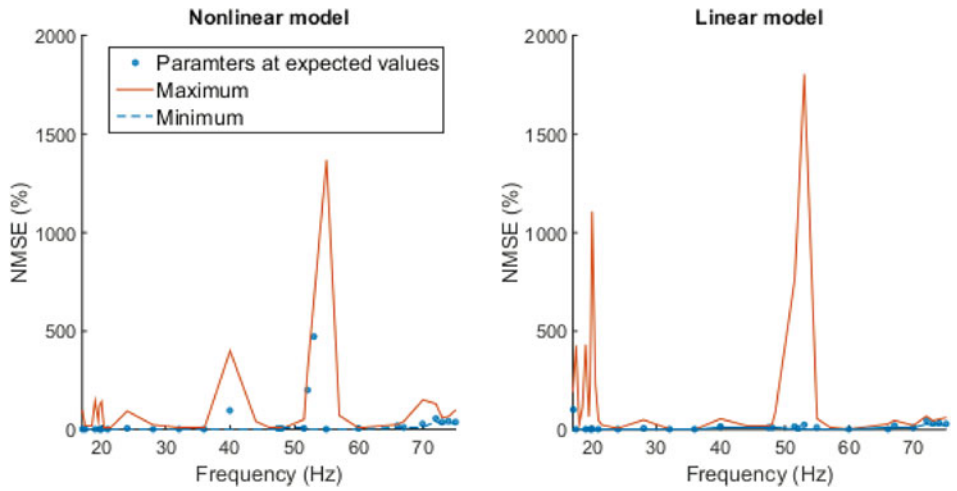


Fig. 1.12 Plot showing maximum and minimum NMSE values for each forcing frequency

1.5.2 Frequency Response

Plots of the modelled and measured frequency responses (acceleration amplitude divided by forcing amplitude) for the second floor are shown in Fig. 1.13. The model frequency responses were generated using time histories simulated with the model parameters at their expected values. In this case, the linear model appears to outperform the nonlinear one; the respective errors are 30% are 60%. The discrepancy is the result of the nonlinear model incorrectly predicting the lower amplitude beyond 52 Hz.

The model parameter uncertainty was propagated once more, to give distributions for the frequency response values. Many of these distributions looked Gaussian "by eye," generally those away from the second natural frequency. Some of the distributions showed marked skewness; an example for the 18 Hz frequency is given in Fig. 1.14. This figure also shows that the mode of the nonlinear model distribution is at the measured value from the structure, while the linear model is clearly biased. At frequencies close to the second natural frequency, the response distributions for the nonlinear model became bimodal again because the model incorrectly predicted the amplitude branch.

A plot showing the uncertainty in the modelled frequency responses of the second floor, alongside the test data, is given in Fig. 1.15. The shaded area is bounded by the maximum and minimum values of the predicted response distributions. The bimodal nature of the distributions for the nonlinear model can be seen in the figure between 51.5 and 55 Hz. For both models, the majority of the measured response values is within the bounds of the model distributions; however, the nonlinear model has a single exception at 52 Hz, while the linear model has four exceptions at 53, 55, 57, and 72 Hz.

In order to quantify the distance between the measured values and the model distributions, the response value distributions were transformed using Box–Cox transformations and the t-statistics were computed in each case (see Sect. 1.4). In almost every case, the t-statistic was lower for the nonlinear model than for the linear, with the main exceptions occurring around

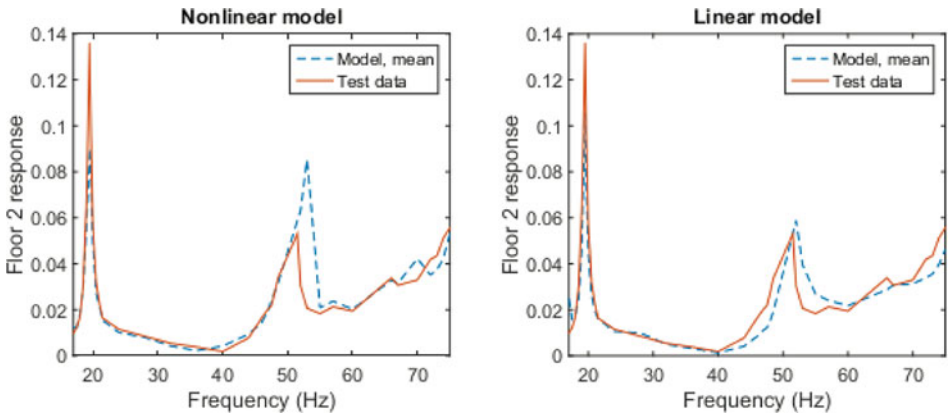


Fig. 1.13 Modelled and measured responses of Floor 2 to periodic forcing at different frequencies, nonlinear and linear models

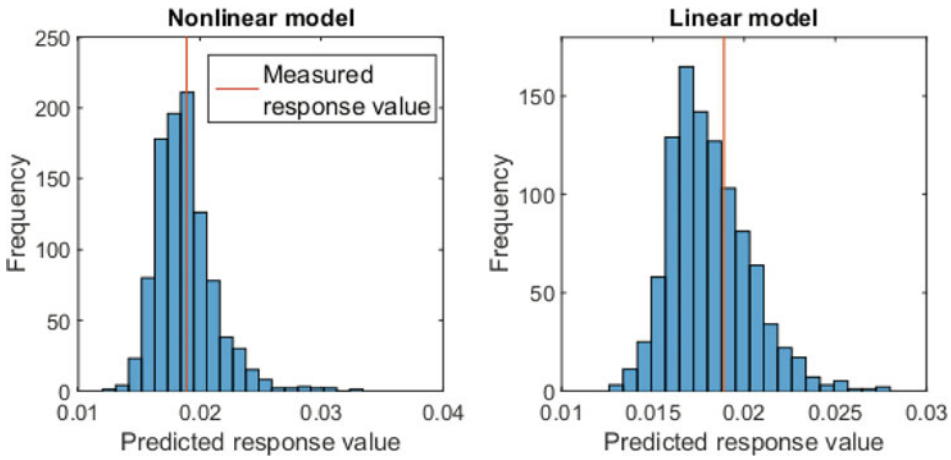


Fig. 1.14 Histogram showing predicted response value of second floor at 18 Hz, nonlinear and linear models. Predicted response value

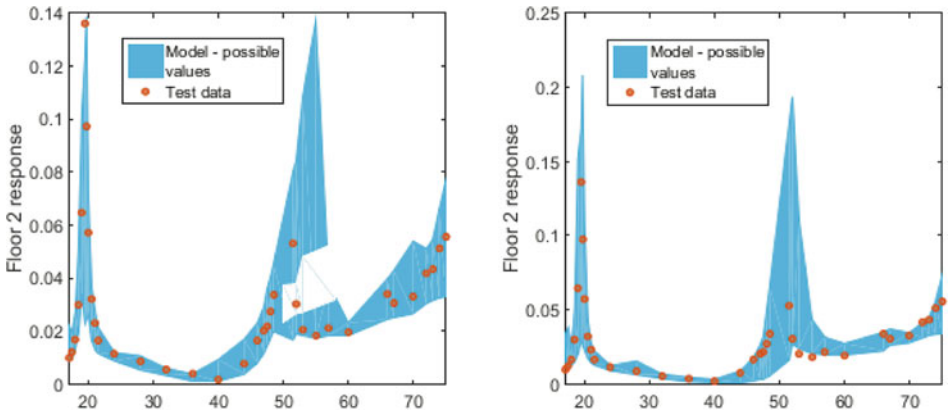


Fig. 1.15 Uncertainty of the modelled response of second floor to periodic forcing at different frequencies. Excitation frequency (Hz) t-statistic, nonlinear model t-statistic, linear model

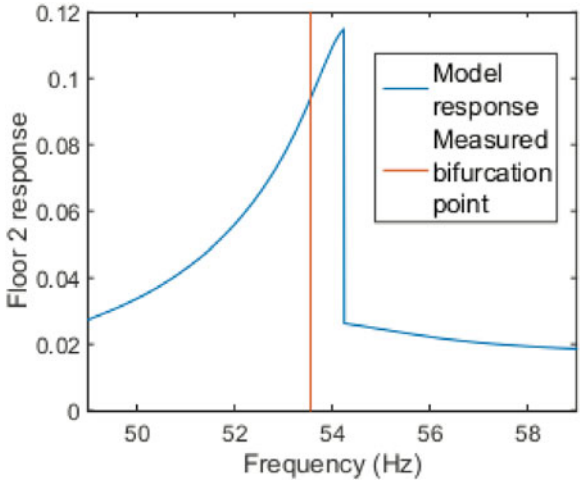


Fig. 1.16 Modelled response close to bifurcation point: excitation amplitude 115 N

the bifurcation frequency where the response distributions for the nonlinear model were bimodal. The mean values of the t-scores over all the frequencies tested were 1.30 for the nonlinear model and 1.66 for the linear model. Care must be taken with the t-statistic, as it can favour models with higher uncertainty (because of the variance in its denominator), and this is not necessarily what is needed. In this case, the parameter distributions had comparable variances, but this uncertainty propagated through to *narrower* response distributions for the nonlinear model, so the t-statistics are actually underestimating the nonlinear model performance.

1.5.3 Bifurcation Behaviour

In the tests, the bookshelf rig showed two types of bifurcations: a saddle-node bifurcation leading to a jump in amplitude just above the second natural frequency and period-doubling bifurcations leading to subharmonics between 65 and 75 Hz. Linear systems do not bifurcate, so the linear model cannot replicate this behaviour. A comparison between the bifurcation behaviours of the bookshelf rig and the nonlinear model is given below.

Data were gathered to show the frequencies at which amplitude jumps occurred at two different forcing amplitudes: 115 and 150 N. For each amplitude, a time history was recorded just above and below the bifurcation point. These time histories are referred to here as the *bifurcation data*. At 115 and 130 N, with parameters set at their expected values, the model shows amplitude jumps at similar points to those seen in the test data, as seen in Fig. 1.16. This plot replicates a stepped-sine test but uses measured forces from the rig at the bifurcation point, rather than sine waves. The frequency is varied by changing the timestep between points. At 150 N, the model did not jump but did show a discontinuity in its transition to the low-amplitude response, shown in Fig. 1.17. One should recall that this is a bilinear system rather than one of Duffing type [3].

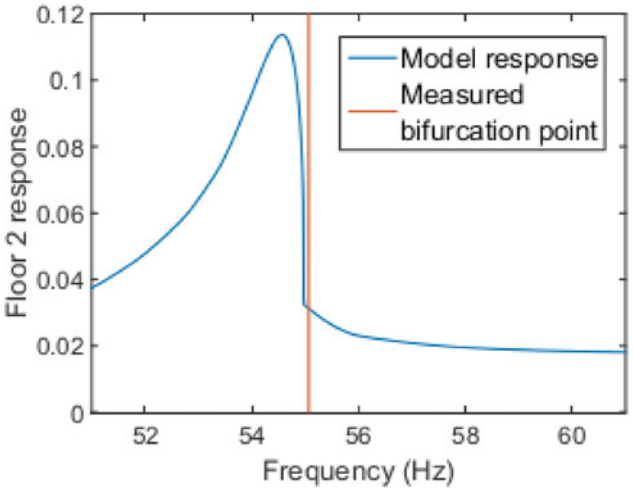


Fig. 1.17 Modelled response close to bifurcation point: excitation amplitude 150 N

Fig. 1.18 Histogram showing model uncertainty in the predicted bifurcation frequency: forcing amplitude 115 N

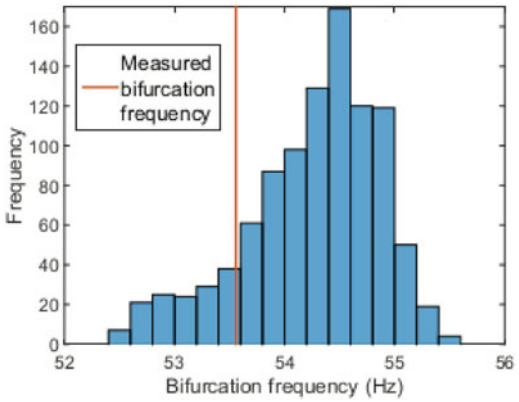
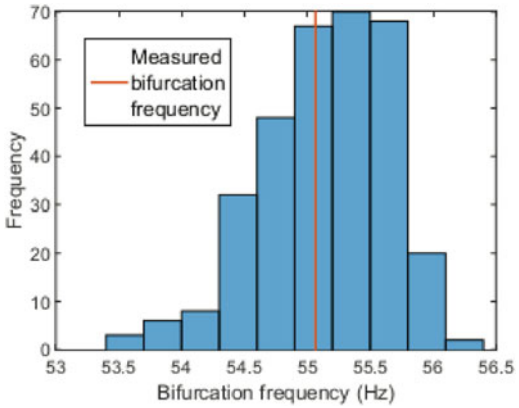


Fig. 1.19 Histogram showing model uncertainty in the predicted bifurcation frequency: forcing amplitude 150 N



As before, the uncertainty in the predicted bifurcation frequency was characterised by propagating the parameter uncertainty through the model. The same MLH design of 1000 parameter values was used as in the previous section, and the bifurcation point was found by bracketing it between converging pairs of upper/lower frequencies. The resulting distributions for amplitudes of 115 and 150 N are shown in Figs. 1.18 and 1.19.

While the model did not jump with an excitation amplitude of 150 N and the parameters at their expected values, it did jump at this forcing amplitude for around 30% of the parameter vectors in the MLH design. At 115 N, the model bifurcated for all 1000 vectors in the MLH design. An intermediate study at 130 N showed jumps at 80% of the sampled parameters. These results suggest that nonlinear model is doing quite a good job of replicating the saddle-node bifurcation behaviour of the system.

1.6 Discussion and Conclusions

In this chapter, experimental data was gathered from a three-storey experimental structure with a harsh nonlinearity. Two models were then created of the structure, one linear and one nonlinear. A selection of validation features and metrics were then applied to the models. The different approaches to validation gave quite different outcomes. In terms of the NMSEs for the time histories and frequency responses, the raw statistics made the linear model look much better than the nonlinear model; however, a closer look and proper consideration of uncertainties showed the opposite conclusion. Furthermore, the nonlinear model successfully captured the bifurcation behaviour of the bookshelf rig in qualitative terms. The results show that a model that accurately represents the physics of the true nonlinear structure can appear much worse than one that does not and could be rejected as a result of slightly misplaced bifurcation points, unless the correct validation strategy is employed. This issue still occurs if only a small proportion of the validation data is close to the bifurcation point. The results suggest that the NMSE is a poor validation metric unless an extremely accurate model is required because the model can be penalised for replicating the nonlinear behaviour of a system unless it does so nearly perfectly.

While it is known that it is important to take uncertainties into account when validating a model, the results here show that this is particularly true for bifurcating systems. The problems illustrated here for validating models of nonlinear bifurcating systems can be mitigated by ensuring that model uncertainties are appropriately accounted for. This can be done by propagating model uncertainties to give probability distributions for the model outputs and using statistical tests to compare them to the test data. Only parameter uncertainty was considered here, but ideally, when validating a model, all forms of modelling and experimental uncertainty should be taken into account. It is possible that nonlinearity may introduce difficulties in incorporating other forms of uncertainty, but this was considered to be outside the scope of this study.

It is worth reiterating two of the purposes of model validation: first, choosing whether to accept a model or to reject it and perform further model development; second to quantify the accuracy with which the model can make predictions. These two steps are often done as part of the same process, i.e., the model will be accepted when it reaches an acceptable level of predictive accuracy. However, for nonlinear systems, it may be desirable to introduce some extra criteria that must be met before the model is accepted because of the wide range of behaviours they can exhibit; bifurcations being a perfect example. Taking the above case study as an example, suppose that the quantity that the model is required to predict is the response to periodic forcing, then this would dictate the feature that is used to assess the model accuracy. If one wanted to provide further confirmation that the model had captured the system characteristics fully, then further features could be investigated. The above results suggest that comparing the bifurcation characteristics would be a more suitable way of doing this than comparing the time histories.

It should be noted that nonlinearity causes a diverse range of phenomena, only a small number of which were observed here. Care should therefore be taken in generalising these conclusions to different nonlinear systems.

Acknowledgments The authors would like to thank the UK EPSRC for funding via the Established Career Fellowship EP/R003645/1 and the Programme Grants EP/R006768/1 and EP/K003836/1. The authors would also like to thank Drs Rob Barthorpe and Daniela Tiboaca of the University of Sheffield for help in commissioning the experimental rig and for useful discussions on Bayesian inference and validation.

References

1. Worden, K.: Some thoughts on model validation for nonlinear systems. In: Proceedings of IMAC XVII – 17^{th} International Modal Analysis Conference, Orlando, FL (2001)
2. Guide for the verification and validation of computational fluid dynamics simulations. Technical Report KAIAA-G-077-1998, American Institute of Aeronautics and Astronautics, 1998
3. Worden, K., Tomlinson, G.R.: Nonlinearity in Structural Dynamics: Detection, Identification and Modelling. Institute of Physics Press (2001)
4. Thacker, B.H., Doebling, S.W., Hemez, F.M., Anderson, M.C., Pepin, J.E., Rodriguez, E.A.: Concepts of model verification and validation. Technical report, Los Alamos National Laboratory, 2004
5. Nishio, M., Hemez, F.M., Worden, K., Park, G., Takeda, N., Farrar, C.R.: Feature extraction for structural dynamics model validation. In: Conference Proceedings of the Society for Experimental Mechanics, pp. 153–163 (2011)
6. Scott, M., Tiboaca, O.D., Barthorpe, R.J., Wagg, D.J., Worden, K.: On the validation of nonlinear MDOF system models. In: Proceedings of 27^{th} International Conference on Noise & Vibration Engineering, Leuven (2016)
7. Qin, A.K., Suganthan, P.N.: Self-adaptive differential evolution algorithm for numerical optimization. In: Proceedings of 2005 IEEE Congress on Evolutionary Computation, vol. 2, pp. 1785–1791 (2005)
8. Worden, K., Manson, G.: On the identification of hysteretic systems, Part I: fitness landscapes and evolutionary identification. Mech. Syst. Signal Process. **29**, 201–212 (2012)
9. Worden, K., Becker, W.E.: On the identification of hysteretic systems, Part II: Bayesian sensitivity analysis and parameter confidence. Mech. Syst. Signal Process. **29**, 213–227 (2012)

10. Worden, K., Hensman, J.J.: Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference. Mech. Syst. Signal Process. **32**, 153–169 (2012)
11. Worden, K., Tiboaca, O.D., Antoniadou, I., Barthorpe, R.J.: System identification of an MDOF experimental structure with a view towards validation and verification. In: Proceedings of the 33^{rd} International Modal Analysis Conference, Orlando, FL (2015)
12. Tiboaca, O.D., Green, P.L., Barthorpe, R.J., Antoniadou, I., Worden, K.: Bayesian inference and RJMCMC in structural dynamics – on experimental data. In: Proceedings of the 34^{th} International Modal Analysis Conference, Orlando, FL (2016)
13. Liu, Y., Chen, W., Arendt, P., Huang, H.Z.: Towards a better understanding of validation metrics. J. Mech. Des. **133**, 071005 (2011)
14. Box, G.E.P., Cox, D.R.: An analysis of transformations. J. R. Stat. Soc. B **26**, 211–252 (1964)