Chapter 10
Isolating and Quantifying Uncertainties in the Vibration Isolation Round-Robin Challenge

Rileigh Bandy, Rachel Washington, Rebecca Morrison, and Teresa Portone

Abstract In the field of structural dynamics, mathematical models generate predictions and inform decisions about complex systems. For example, models may define a bridge's maximum allowable load, chronicle the dynamics of a jointed structure, or describe a driving car's vertical mechanical behavior. But significant discrepancies between observations and predictions often result from model-form uncertainty (e.g., missing dependencies or simplified forces). The Uncertainty Quantification in Vibration Isolation round-robin challenge considers modeling a vibration isolation experiment and the uncertainty that results from a simplified representation of the experimental configuration. As a result of model-form uncertainty, discrepancies emerge between experimental data and model predictions from a linear two-mass oscillator. Last year, we embedded a nonlinear, stochastic enrichment operator into the two-mass oscillator to correct the active damping force. This improved the time-domain amplitude of the interior mass's acceleration, but discrepancies in its time-domain frequency remained. This year, we focus on isolating and quantifying the measurement uncertainty in the experiment, uncertainties in the initial condition, and parameter uncertainties in the linear two-mass oscillator. After Bayesian calibration, the linear two-mass oscillator accurately predicts the resonant frequency of the active vibration isolation experiment.

Keywords Measurement uncertainty · Parameter uncertainty · Model discrepancy · Bayesian calibration · Mechanical oscillations

Introduction

Mathematical models are fundamental for informing predictions and decision making in structural dynamics. For example, models can determine the maximum allowable load of a bridge [1] or express the shear stress of two metal sheets bolted together [2]. However, models do not perfectly represent reality; they include uncertainties that must be quantified before making reliable predictions. These uncertainties accumulate from many sources including uncertain initial conditions or model parameters where accurate values cannot be determined *a priori*. Model-form uncertainties occur due to simplifying assumptions or missing dependencies (e.g., approximating a nonlinear force with a linear empirical function). Additionally, measurement uncertainties are the difference between observed values and reality, and they can result from imprecise measurement sensors.

The purpose of the Uncertainty Quantification in Vibration Isolation round-robin challenge (UQVI-challenge) investigated herein is to question how uncertainties arise and can be addressed when using simple, mathematical models to describe more complex system configurations [3]. Experiments explore the movement of a mass inside a frame after the frame is struck by a modal hammer, with passive and active damping applied [3, 4]. The experimental testing environment provides high-fidelity acceleration data for several combinations of system stiffness, passive and active damping, interior mass values, and high-pass filters. We explore three combinations of varied interior mass values with constant stiffness active damping,

Rileigh Bandy · Teresa Portone
Optimization and Uncertainty Quantification Department, Sandia National Laboratories, Albuquerque, NM 87123
e-mail: rjbandy@sandia.gov; tporton@sandia.gov

Rachel Washington · Rebecca Morrison
Department of Computer Science, University of Colorado, Boulder, CO 80309
e-mail: Rachel.Washington@colorado.edu; rebeccam@colorado.edu

Roland Platz et al. (eds.), *Model Validation and Uncertainty Quantification, Vol. 3*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/97887-438-0148-1_10

passive damping, and high-pass filtering. Without quantifying uncertainties, discrepancies emerge between experimental observations and model predictions [5]. Notably, we found in [5] a linear two-mass oscillator model underestimates our quantity of interest (QoI), which is the resonant frequency and its magnitude, an important consequence of the force the hammer inflicts on the interior mass.

Given experimental observations and the model, there are many approaches to estimate, reduce, and propagate uncertainties. One common approach is adding a nonintrusive, stochastic term to the model output [6, 7]. Experimental observations calibrate the stochastic term and result in a catch-all uncertainty representation that can include measurement, initial condition, parameter, and model-form uncertainties. This complicates quantifying the model-form uncertainty because it cannot be isolated.

It is necessary to quantify uncertainties separately and at their source to accurately measure their impact on predictions and to represent their intrinsic nature (e.g., model-form uncertainty originates in the right-hand-side (RHS) of the model, so it should be addressed there too) [8]. Last year, we improved upon the linear two-mass oscillator by embedding nonlinear terms into the RHS of the model to address model-form uncertainty. However, we did not consider parameter and measurement uncertainties or their impact on our representation of model-form uncertainty.

Now, we take a step back to explore if discrepancies can be clearly attributed to uncertain measurements, initial conditions, or model parameters. To address uncertainty in isolation, we focus on first quantifying measurement uncertainties in the experiment and initial condition and parameter uncertainties in the linear two-mass oscillator. We represent uncertainties using the experimental data from the middle interior mass value, and we extrapolate to scenarios with smaller and larger masses. Ultimately, we would like to determine if the linear two-mass oscillator is a sufficient representation of experimental resonant frequencies and corresponding magnitudes.

The rest of the paper is organized as follows. First, three possible distributions for the measurement uncertainty are defined. Next, we describe the linear two-mass oscillator model and plausible distributions for the initial condition and model parameter uncertainties. We detail the methodology how to generate observations, model outputs, and predictions, and Bayesian calibration details. Finally, we present our numerical results and give some concluding remarks.

Measurement Uncertainty

Before introducing a mathematical model, we define uncertainties in the experimental data. After hitting the frame with a modal hammer, the acceleration of the interior mass over time is measured using an accelerometer. These measurements are uncertain due to sensor precision and variability in the modal hammer hit (e.g., hammer force and strike location). Then, a high-pass filter of 10 Hz preprocesses the acceleration data and removes low-frequency noise [3]. The resulting experimental acceleration data collected from a single hit is

$$\boldsymbol{a}_{\text{hit}} = a_{\text{hit}}(t) \quad \forall t \in [0, 0.0002, \dots, 12.8], \tag{1}$$

where $\boldsymbol{a}_{\text{hit}}$ is a vector of the accelerations (m/s^2) from $t = 0$ seconds when the modal hammer hits the frame to the final time measured $t = 12.8$ seconds. Acceleration data is averaged over five sequential hits

$$\bar{\boldsymbol{a}} = \bar{a}(t) = \frac{1}{5}\sum_{i=1}^{5} a_{\text{hit}}^{(i)}(t) \quad \forall t \in [0, 0.0002, \dots, 12.8], \tag{2}$$

which mitigates hit uncertainty, but since five is a relatively small number of samples, hit uncertainty is not entirely removed.

The uncertainties in sensor precision and hits accumulate into the overall measurement uncertainty that we quantify with a distribution informed by the experimental data. To inform our experimental uncertainty model, we have access to $\bar{a}(t)$ and the final hit of five hits used to compute the mean. There are several possible formulations to represent the experimental uncertainty. We investigate three options: additive Gaussian, additive Gaussian with a smaller variance, and a mixture of multiplicative and additive Gaussian distributions.

Additive Gaussian Measurement Uncertainty: Additive Gaussian distributions are commonly used to model measurement uncertainty and are commonly referred to as additive white Gaussian noise [9]. As the name implies, measurement uncertainty or noise is added to the measurement. For a single hit, we assume the measured acceleration of the interior mass at time t is independent. Assuming independence imposes less structure than dependence, and it would be challenging to elicit correlations in time from the current data. (However, assuming measurements are independent in time could overvalue the data in forthcoming Bayesian calibration.) Therefore, for a single hit, we assume the measurement uncertainty is additive Gaussian as such,

$$a_{\text{hit}}(t) = \bar{a}(t) + \epsilon \sim \mathcal{N}\left(\bar{a}(t), \sigma_\epsilon^2\right). \tag{3}$$

We explore two different values for σ_ϵ^2: $\sigma_\epsilon^2 = 0.05$ (about 1% of the maximum acceleration observed in the experimental data) or $\sigma_\epsilon^2 = 0.009^2 = 8 \times 10^{-5}$ (a much smaller variance defined later to capture sensor uncertainty). The additive Gaussian distribution is not based on uncertainty information from the experiment. Therefore, we define an alternative distribution using more information.

Mixture Model Measurement Uncertainty: Given the mean acceleration data and the acceleration data from the final hit, we investigate the deviation of $a_{\text{hit}}(t)$ from $\bar{a}(t)$, shown in Figure 1. Initially the mean deviations vary depending on time, but they quickly become stable. To investigate where this switch occurs, we zoom in on the first 0.5 seconds in Figure 2a. Mean deviations consistently fall within roughly $[-0.05, 0.05]$ after $t = 0.15$. We also investigate the mean acceleration data for the first 0.5 seconds in Figure 2b, and the mean acceleration data appears to reach a stationary mode around $t = 0.25$. Therefore, we select $t_\epsilon = 0.25$ as the time when measurement uncertainty switches from time-varying to constant. We believe this is where sensor uncertainty dominates and hit uncertainty becomes less influential for the overall measurement uncertainty.

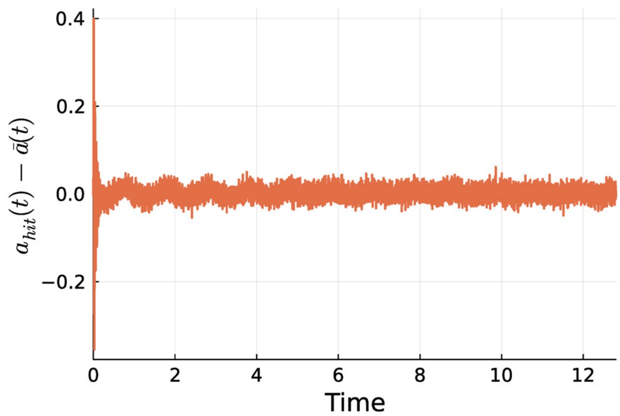


Fig. 1 Experimental acceleration mean deviations from the full time-domain.

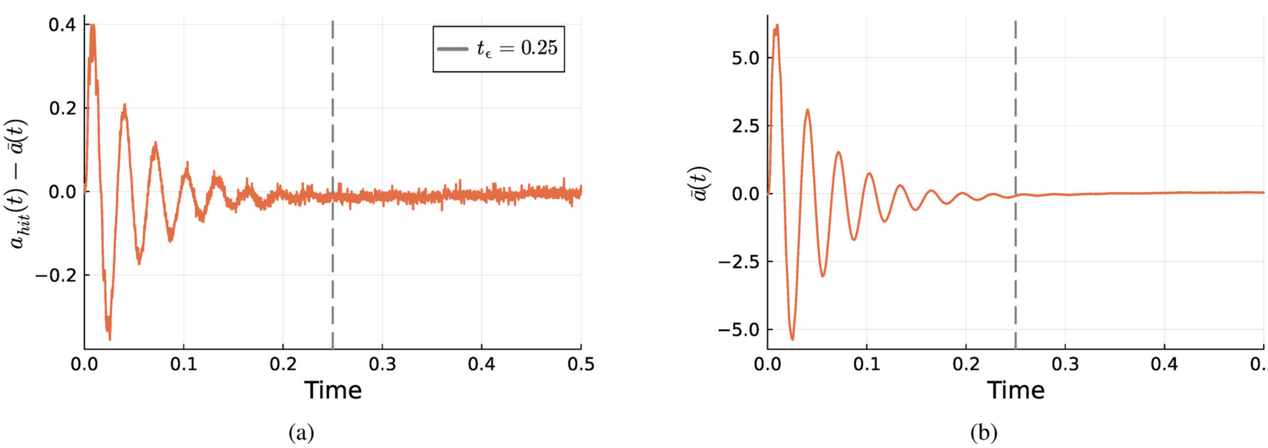


Fig. 2 Selection of the time where measurement uncertainty goes from time-varying to constant t_ϵ based on (a) experimental acceleration mean deviations and (b) mean experimental acceleration.

Given the shift from time-varying to constant mean deviations at t_ϵ, we define the measurement uncertainty in two parts, with a distribution for hit uncertainty and a distribution for sensor uncertainty. To quantify sensor uncertainty, we investigate the relationship between mean acceleration and mean deviations for times after t_ϵ in Figure 3. The mean deviations are independent of mean acceleration; they appear to be Gaussian distributed with a mean of zero and a sample standard deviation of 0.009. We assume the final hit falls within a 50% confidence interval of our sensor distribution making the distribution's standard deviation twice as large. Therefore, we characterize sensor uncertainty as

$$\epsilon_{\text{sensor}} \sim \mathcal{N}\left(0, 0.018^2\right). \tag{4}$$

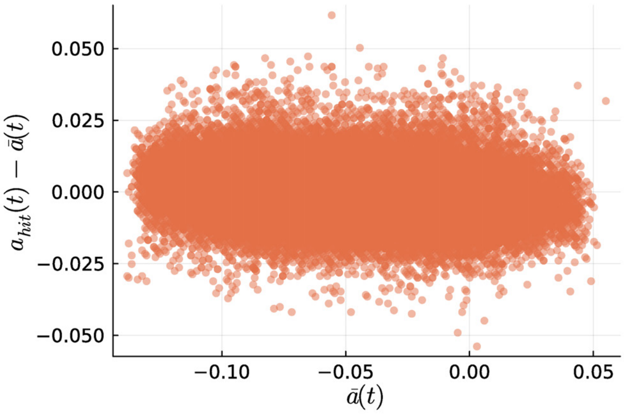


Fig. 3 Experimental acceleration mean deviations after t_ϵ.

To quantify hit uncertainty, we investigate the relationship between mean acceleration and the final hit for times before t_ϵ in Figure 4. Final hit acceleration is linearly dependent on mean acceleration. That dependence is characterized with the linear function $1.06\bar{a}(t)$, where the slope is estimated using the maximum and minimum final hit accelerations and their associated mean accelerations. Similar to sensor uncertainty, we approximate hit uncertainty standard deviation with twice the sample standard deviation of acceleration mean deviations, which is 0.1.

Therefore, we assume hit uncertainty is multiplicative, Gaussian and characterize it as

$$\epsilon_{\text{hit}} \sim \mathcal{N}\left(1.06, 0.2^2\right). \tag{5}$$

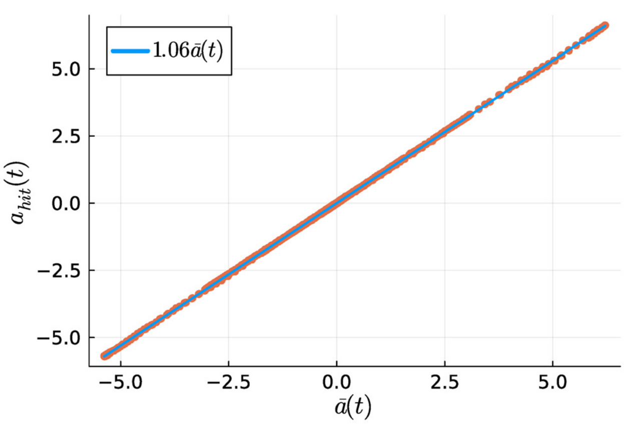


Fig. 4 Relationship between mean experimental acceleration and final hit experimental acceleration before t_ϵ.

After characterizing sensor and hit uncertainties, we bring them together to form our measurement uncertainty mixture model:

$$a_{\text{hit}}(t) = \epsilon_{\text{hit}}\bar{a}(t) + \epsilon_{\text{sensor}} \sim \mathcal{N}\left(1.06\bar{a}(t), 0.2^2\bar{a}(t)^2 + 0.018^2\right). \tag{6}$$

Measurement Uncertainty Pushforward: Could the acceleration data from the final hit plausibly be drawn from any of the three distributions of measurement uncertainty defined above? To answer this question, we use mean experimental acceleration data and Equations (3) and (6) to generate 5000 synthetic hits and compare them to acceleration data from the final hit. In Figure 5 the resulting 50% and 95% confidence intervals from the synthetic hits are depicted as the dark and lighter blue shaded regions.

Additive Gaussian measurement uncertainty with $\sigma_\epsilon^2 = 0.05$ is pushed forward in Figures 5a and 5b, where Figure 5a shows trajectories until 0.25 seconds, and Figure 5b shows trajectories from 0.25 seconds to 12.8 seconds. This measurement uncertainty distribution encompasses the final hit's trajectory throughout time, but it could be overestimating uncertainties as acceleration dampens and uncertainty remains constant. Since $a_{\text{hit}}(t)$ is encompassed within the confidence region, the experimental acceleration data could be plausibly drawn from this distribution of measurement uncertainty.

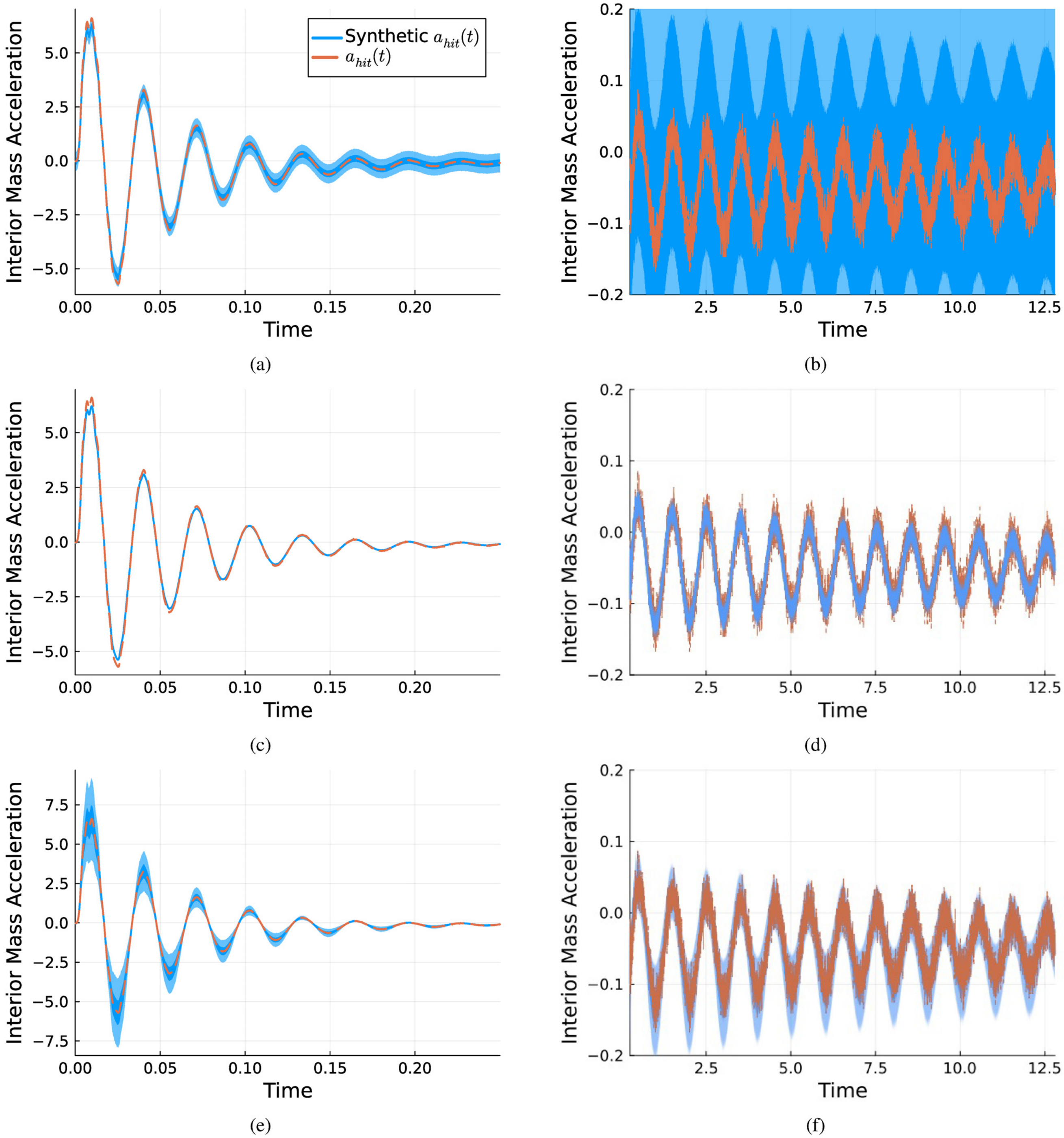


Fig. 5 Comparison of experimental acceleration data to the 50% and 95% confidence intervals (dark and lighter blue shaded regions, respectively) from synthetic data generated from three measurement uncertainty distribution: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b), additive Gaussian with $\sigma_\epsilon^2 = 8 \times 10^{-5}$ in (c, d), and the mixture model in (e, f). Subfigures (a,c,e) depict times before t_ϵ and (b,d,f) depict times after t_ϵ.

Additive Gaussian measurement uncertainty with $\sigma_\epsilon^2 = 8 \times 10^{-5}$ is pushed forward in Figures 5c and 5d. This measurement uncertainty distribution noticeably underpredicts the first three peaks and troughs compared to the final hit's acceleration, and it seems to slightly underpredict the uncertainties as acceleration dampens. Therefore, experimental acceleration data could not be plausibly drawn from this distribution of measurement uncertainty.

The mixture model measurement uncertainty is pushed forward in Figures 5e and 5f. In Figure 5e the 50% confidence interval encompasses experimental acceleration data with the most spread at the initial peaks and troughs. The spread decreases as time progresses and acceleration dampens. In Figure 5f the 95% confidence interval closely matches and encompasses most of the experimental acceleration data. Therefore, the experimental acceleration data could be plausibly drawn from this distribution of measurement uncertainty, and it is the most plausible candidate of the three measurement uncertainties. To show the impact of plausible and implausible measurement uncertainty on a mathematical model's predictions, we propagate uncertainties using all three distributions.

Parameter and Initial Condition Uncertainties

Now that we have defined three possible measurement uncertainty distributions, we can introduce a mathematical model and start quantifying its uncertainties. We start with a linear two-mass model of the experimental setup shown in Figure 6, where one mass represents the interior mass and one mass approximates the frame itself. The displacement of the interior mass is represented by $z(t)$, the displacement of the frame is $z_f(t)$, and both are measured in meters.

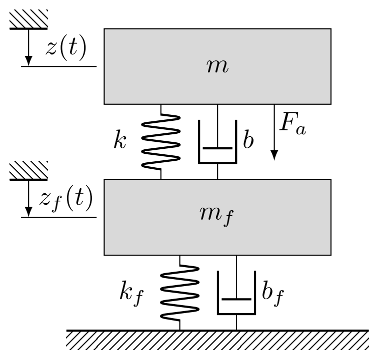


Fig. 6 Mass-spring-damper diagram of the two-mass oscillator.

The accelerations of the linear two-mass oscillator are defined as

$$\ddot{z}(t) = \frac{1}{m}\left\{-b\left(\dot{z}(t) - \dot{z}_f(t)\right) - g\dot{z}(t) - k\left(z(t) - z_f(t)\right)\right\} \tag{7}$$

$$\ddot{z}_f(t) = \frac{1}{m_f}\left\{b\left(\dot{z}(t) - \dot{z}_f(t)\right) + k\left(z(t) - z_f(t)\right) - b_f\dot{z}_f - k_f z_f\right\}. \tag{8}$$

The acceleration of the interior mass (m) is defined by $\ddot{z}(t)$, which depends on the passive damping coefficient b, the active damping coefficient g, and the spring constant k. The acceleration of the frame mass (m_f) is defined by $\ddot{z}_f(t)$, which also depends on the frame's damping coefficient b_f and spring constant k_f. The passive damping coefficients are

$$b = 2D\sqrt{\frac{k}{m}} \tag{9}$$

$$b_f = 2D_f\sqrt{\frac{k_f}{m_f}}, \tag{10}$$

where D and D_f are the passive damping ratios.

The linear two-mass oscillator only considers damping and spring forces, so there is not an analogous impulse force to the modal hammer hit from the experiment. Instead, the linear two-mass oscillator is integrated forward in time starting after the initial impulse force subsides, which we estimate is at $T_i = 0.01$ seconds. Therefore, the initial conditions for the linear two-mass oscillator $\boldsymbol{Z} = (z(T_i), z_f(T_i), \dot{z}(T_i), \dot{z}_f(T_i))$ are uncertain, and we must approximate them with a distribution using prior information. We define the prior distribution on the initial conditions as

$$p(\boldsymbol{Z}) = \prod_{i=1}^{4} p(Z_i) = \prod_{i=1}^{4} \mathcal{U}\left(-0.5, 0.5\right). \tag{11}$$

The initial condition priors are maximum entropy priors, where the distribution with the highest entropy that fits a given set of constraints is selected [10]. Here, the maximum entropy priors are uniform, since we expect the initial displacements and velocities to be small and between roughly -0.5 and 0.5 in order for the initial interior mass acceleration to be around $5\,\mathrm{m/s^2}$.

The model parameters of the linear two-mass oscillator $\boldsymbol{\theta} = (k, D, g, k_f, D_f)$ are also uncertain. The nominal model parameters $\boldsymbol{\theta}^0 = \left(k^0, D^0, g^0, k_f^0, D_f^0\right)$ are derived from the experimental apparatus before gathering data, and they are shown in Table 1. For example, the interior mass's spring constant is derived from the assumed bending stiffness EI and length l, $k^0 \propto \frac{EI}{l^3}$. While the nominal values give us good initial estimates, they are uncertain and could be improved using experimental data. Therefore, we approximate them with a distribution using prior information. We define the prior distribution on the model parameters as

$$p(\boldsymbol{\theta}) = \prod_{i=1}^{5} p(\theta_i) = \prod_{i=1}^{5} \mathcal{N}\left(\theta_i^0, 0.2\theta_i^0\right). \tag{12}$$

The model parameters assume maximum entropy priors centered at their nominal values with standard deviations of 20% of those nominal values.

Table 1 Nominal parameter values derived from the experimental apparatus before gathering data.

Parameter		
k^0	32583	N/m
D^0	4.85	%
g^0	26	Ns/m
k_f^0	722	N/m
D_f^0	0.1	%

Methodology

After isolating and quantifying the uncertainties in measurements, initial conditions, and parameters, we must determine whether the linear two-mass oscillator adequately represents the experiment. To address this, we perform Bayesian calibration using post-processed observations from the experiment.

Observations: During the round-robin experiments, five sequential hits were conducted, and the results were filtered using a high-pass analogue filter with a cutoff frequency of 10 Hz and then averaged [4]. Using this averaged and filtered data, we interpolated the time series, $\bar{\boldsymbol{a}}$, to obtain discrete data at uniform time intervals. The resulting observations are

$$\boldsymbol{D}(\boldsymbol{X}) = \text{interpolate}\left(\bar{a}(t; \boldsymbol{X})\right), \quad \forall t \in (T_i, T_i + tstep, \ldots, 12.8), \tag{13}$$

where the experiment is initialized with masses

$$\boldsymbol{X} = (m, m_f). \tag{14}$$

The time after the initial impulse force is $T_i = 0.01$ seconds and the timestep between discrete observations is $tstep = 0.001$ seconds. We subsample calibration data from one set of masses for a short time-window to create calibration observations, defined as

$$\begin{aligned} \boldsymbol{D}_c &\subset \boldsymbol{D}\left(\boldsymbol{X}_c\right) = \boldsymbol{D}\left(m = 0.9249\,\text{kg}, m_f = 9.33\,\text{kg}\right) \\ \boldsymbol{D}_c &= \text{interpolate}\left(\bar{a}(t; \boldsymbol{X}_c)\right), \quad \forall t \in (T_i, T_i + tstep, \ldots, 0.072). \end{aligned} \tag{15}$$

This short time-window is selected to include the first two wavelengths after the initial impulse force subsides.

We subsample validation data to include an interior mass both larger and smaller than the calibration scenario, while the frame mass remains the same:

$$\boldsymbol{D}_v = \left\{\boldsymbol{D}\left(\boldsymbol{X}_{v_1}\right), \boldsymbol{D}\left(\boldsymbol{X}_{v_2}\right)\right\} = \left\{\boldsymbol{D}\left(m = 0.7853\,\text{kg}, m_f = 9.33\,\text{kg}\right), \boldsymbol{D}\left(m = 1.0613\,\text{kg}, m_f = 9.33\,\text{kg}\right)\right\}. \tag{16}$$

Model Output: The linear two-mass oscillator simulates the interior mass's movement; we then sample the trajectory of the interior mass's acceleration at discrete intervals in time. This results in the following output:

$$\boldsymbol{Y}(\boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{X}) = \ddot{z}(t; \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{X}), \quad \forall t = (T_i, T_i + tstep, \ldots, 12.8), \tag{17}$$

where $\boldsymbol{\theta}$ are model parameters and $\boldsymbol{Z}$ are initial conditions. Outputs for the calibration scenario are defined as

$$\begin{aligned}&\boldsymbol{Y}_c(\boldsymbol{\theta},\boldsymbol{Z}) \subset \boldsymbol{Y}(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_c)\\&\boldsymbol{Y}_c(\boldsymbol{\theta},\boldsymbol{Z}) = \ddot{z}(t;\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_c), \quad \forall t \in (T_i, T_i + tstep, \ldots, 0.072),\end{aligned} \tag{18}$$

and outputs for the validation scenarios are defined as

$$\boldsymbol{Y}_v(\boldsymbol{\theta},\boldsymbol{Z}) = \left\{\boldsymbol{Y}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_1}\right), \boldsymbol{Y}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_2}\right)\right\}. \tag{19}$$

Model Predictions of the QoI: Acceleration of the interior mass can also be analyzed in the frequency domain. We transform observations and model outputs using the discrete Fourier transform (DFT):

$$\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}\right) = \left|\mathrm{DFT}\left(\boldsymbol{Y}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}\right)\right)\right|, \tag{20}$$

where the frequencies above the Nyquist frequency are removed [11]. The time domain (in seconds) is transformed to the frequency domain (in Hz), and the interior mass's acceleration (in m/s^2) is transformed to the absolute magnitude (in m/s^2). We use the time-domain from T_i until 12.8 seconds to improve the resolution of the frequency-domain data. Since the linear two-mass oscillator model has two degrees of freedom, there are two modes of vibration and two frequencies where the modes occur (resonant frequencies). The lower resonant frequency ($\approx 2\pi$ Hz) corresponds to the frame, and the higher resonant frequency ($\approx 10^{1.5}$ Hz) corresponds to the interior mass [4]. We omit the frame's resonant frequency, and our QoI is the interior mass's resonant frequency and associated magnitude:

$$\mathrm{QoI} = (\alpha, \lambda^*) = \left(\max_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X};\lambda\right)\right\}, \operatorname*{argmax}_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X};\lambda\right)\right\}\right), \tag{21}$$

where $\Lambda = (10, 100)$ Hz. This QoI is important for quantifying the interior mass's response to forcing.

The QoI prediction for the calibration scenario is defined as

$$\mathrm{QoI}_c = (\alpha_c, \lambda_c^*) = \left(\max_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_c;\lambda\right)\right\}, \operatorname*{argmax}_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_c;\lambda\right)\right\}\right), \tag{22}$$

and QoIs for the validation scenarios are defined as

$$\begin{aligned}\mathrm{QoI}_{v_1} &= \left(\alpha_{v_1}, \lambda_{v_1}^*\right) = \left(\max_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_1};\lambda\right)\right\}, \operatorname*{argmax}_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_1};\lambda\right)\right\}\right)\\\mathrm{QoI}_{v_2} &= \left(\alpha_{v_2}, \lambda_{v_2}^*\right) = \left(\max_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_2};\lambda\right)\right\}, \operatorname*{argmax}_{\lambda\in\Lambda}\left\{\boldsymbol{F}\left(\boldsymbol{\theta},\boldsymbol{Z},\boldsymbol{X}_{v_2};\lambda\right)\right\}\right).\end{aligned} \tag{23}$$

Bayesian Calibration: We propagate our uncertainties to the model output using Bayesian calibration. The posterior distribution of the model parameters and initial conditions is sampled using the Adaptive Scaling within Adaptive Metropolis (ASWAM) method in Julia [12, 13]. The posterior is

$$\pi_{post} = p(\boldsymbol{\theta},\boldsymbol{Z} \mid \boldsymbol{D}_c) \propto \pi_{like}\pi_{pri}. \tag{24}$$

The prior distribution is the product of the initial condition and model parameter priors:

$$\pi_{pri} = p(\boldsymbol{\theta})p(\boldsymbol{Z}). \tag{25}$$

The likelihood distribution measures how well the model explains experimental data given the measurement uncertainty. Therefore, we explore three possible likelihood distributions:

$$\pi_{like} = p(\boldsymbol{D}_c \mid \boldsymbol{\theta},\boldsymbol{Z}) = \begin{cases}\mathcal{N}\left(\boldsymbol{Y}_c(\boldsymbol{\theta},\boldsymbol{Z}); \boldsymbol{D}_c, 0.05\boldsymbol{I}\right) & \text{(additive Gaussian)}\\\mathcal{N}\left(\boldsymbol{Y}_c(\boldsymbol{\theta},\boldsymbol{Z}); \boldsymbol{D}_c, 8\times10^{-5}\boldsymbol{I}\right) & \text{(small additive Gaussian)}\\\mathcal{N}\left(\boldsymbol{Y}_c(\boldsymbol{\theta},\boldsymbol{Z}); 1.06\boldsymbol{D}_c, \mathrm{diag}\left(0.2^2\boldsymbol{D}_c^2\right) + 0.018^2\boldsymbol{I}\right) & \text{(mixture model)}\end{cases} \tag{26}$$

following the measurement uncertainties defined in Equations (3) and (6).

Calibration took $300,000$ Markov Chain Monte Carlo (MCMC) steps with five parallel tempering levels. In adaptive parallel tempering, each level had an independent random walk sampler, and the first level was randomly swapped with an auxiliary level (levels two through five) in order to reach an average switching probability of 0.234 [12]. The first $50,000$ MCMC steps were thrown out as burnin, and the lag was 50 MCMC steps. Each calibration took under ten minutes on a single Apple M3 Max CPU core, and data was stored on an SSD. We assessed convergence using trace plots, where each initial condition and parameter value is plotted against filtered MCMC steps. We examined the trace plots for visual patterns, where no major trends or jumps occurred indicating that the chains likely reached a stationary distribution [14].

While calibration with all three likelihoods appears to have converged, calibration using the small additive Gaussian likelihood was the most challenging. A narrower range in initial conditions and model parameters was sampled, and sometimes that range was limited by the priors. For example, with small additive Gaussian likelihood, samples of the active damping coefficient g ranged from 36 to 40, which were roughly two to three standard deviations larger than the prior mean of 26. With the mixture model likelihood, samples of g ranged from 15 to 30.

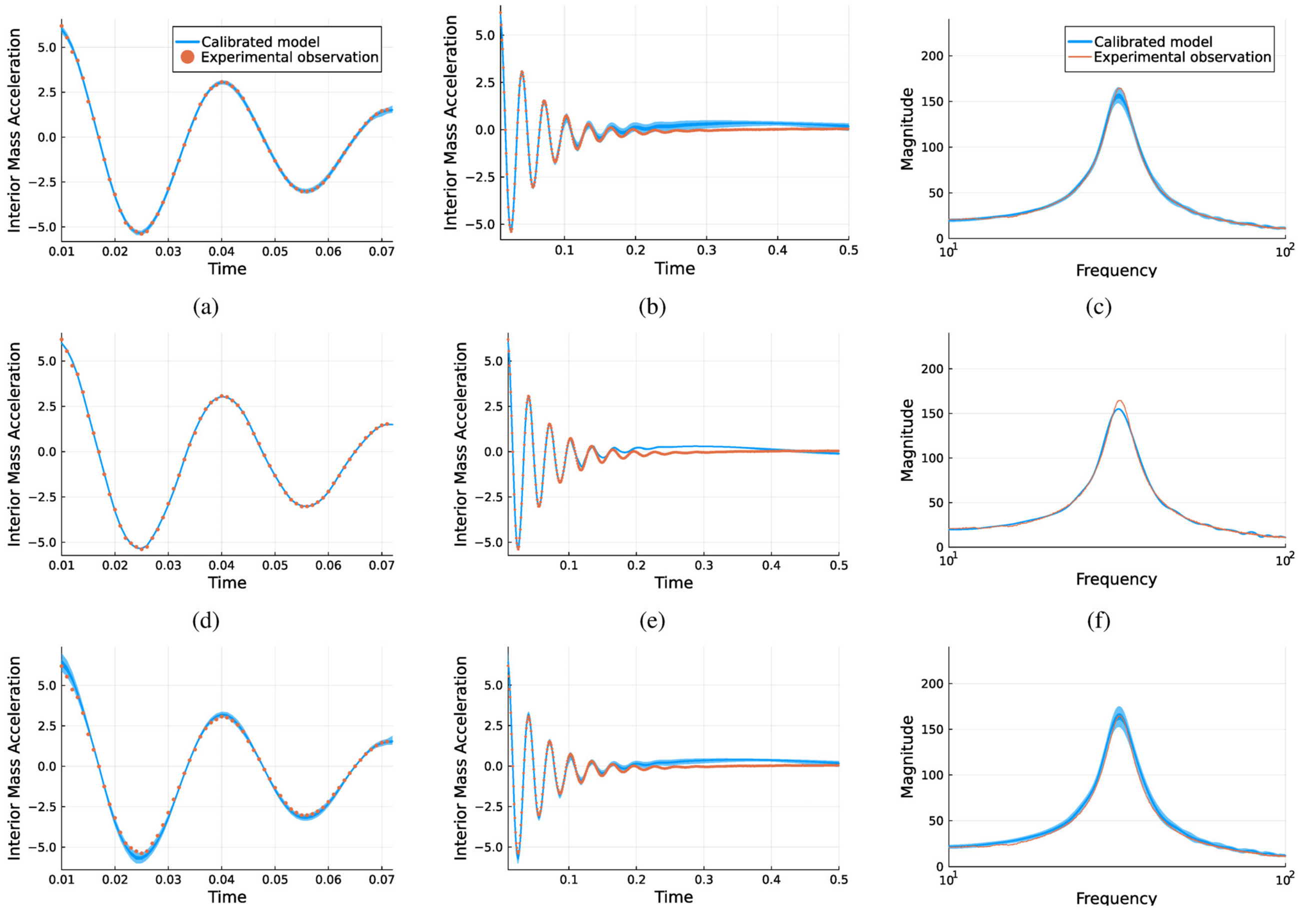
Numerical Results

After calibrating the linear two-mass oscillator, uncertainties propagate to model outputs, transformed outputs in the frequency-domain, and QoI predictions. Numerical results from the calibrated model with the three likelihood formulations are compared to experimental observations for the calibration scenario and two validation scenarios. For the validation scenarios, we update the interior mass, but do not recalibrate the initial conditions or model parameters. We omit the frame mass's data during calibration and validation because we anticipate it introducing additionally, unresolved model-form uncertainties [5].

Calibration Scenario: Numerical results are compared to experimental observations for the calibration scenario in Figures 7a, 7d and 7g, where the calibrated model closely matches experimental observations for all three likelihood distributions. However, the likelihood directly influences the calibrated model confidence intervals. For example, the small additive Gaussian likelihood in Figure 7d results in narrow confidence intervals in the output, where 50% and 95% confidence intervals are visually indistinguishable. When the mixture model likelihood is used in Figure 7g, model output is most uncertain at the first few peaks and troughs, which tracks with its measurement uncertainty pushforward in Figure 5e. When the time-frame is extended in Figures 7b, 7e and 7h, discrepancies emerge after the first couple oscillations and persist in time. Notably, with the small additive Gaussian likelihood in Figure 7e, discrepancies are larger and appear earlier than in Figures 7b and 7h. Time-domain trajectories are transformed to the frequency-domain in Figures 7c, 7f and 7i. The calibrated models with additive Gaussian likelihood and mixture model likelihood encompass most experimental observations, while the calibrated model with small additive Gaussian likelihood underpredicts the resonant frequency's magnitude. Since the resonant frequency and its magnitude are our QoI, it is important for the calibrated model to predict them well.

We explore the calibration scenario QoI predictions in Figure 8, where the marginal calibrated model distributions of the resonant frequency and its magnitude are compared to the experiment. The marginals of the calibrated models with additive Gaussian likelihood and mixture model likelihood encompass the experimentally observed resonant frequency and maximum magnitude. The marginals from the mixture model likelihood have higher density around the experiment, indicating that the experimentally observed QoI was more likely to be generated from the model with mixture model likelihood than the model with additive Gaussian likelihood. The experimentally observed maximum magnitude falls outside of the marginal distribution of the calibrated model with small additive Gaussian likelihood in Figure 8d. Therefore, we deem the linear two-mass oscillator with small additive Gaussian likelihood insufficient for predicting the QoI of the experiment. Because the linear two-mass oscillator with small additive Gaussian likelihood is not predictive in the calibration scenario, we do not extend its predictions to the validation scenarios.

We investigated extending the time-frame of the calibration scenario to incentivize model output to be closer to experimental observations, but Bayesian calibration would not converge. This could be caused by correlation of the measurement uncertainties overtime, which we assumed to be independent. Increasing the time-frame of the calibration data caused the importance of the likelihood to increase. However, the likelihoods also became more restrictive. Specifically, the likelihoods did not allow for much deviation in the time-domain frequency. Many parameter and initial condition combinations of the model resulted in a frequency shift that the likelihoods determined improbable. Therefore, our MCMC chains would not mix and oscillated between a few parameter and initial condition combinations regardless of the number of parallel levels we added, the number of steps we took, or even the MCMC sampling methods used.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 7 Comparison of experimental observations to the 50% and 95% confidence intervals (dark and lighter blue shaded regions) from the calibrated model with three likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b, c), additive Gaussian with $\sigma_\epsilon^2 = 8 \times 10^{-5}$ in (d, e, f), and the mixture model in (g, h, i) for (a, d, g) the calibration time-window, (b, e, h) extended times, and (c, f, i) the frequency-domain.

We hypothesize additional, unresolved uncertainties caused Bayesian calibration of the extended time-frame to fail. Notably, and in contrast to our work last year, we did not resolve model-form uncertainties in the linear two-mass oscillator. For example, we approximate the frame of the experiment as a mass, and we assume there are only linear spring and damping forces acting on the system. In the future, we would like to explore these uncertainties further.

Validation Scenarios: In each validation scenario, the interior mass is changed, but we do not recalibrate the linear two-mass oscillator. Instead, we propagate samples from the calibrated model's initial conditions and model parameters to model outputs in the validation scenarios. We assume the experimental test setup was not disassembled and reassembled between scenarios, and the only change in the experiment was increasing/decreasing the interior mass. Then, we test to see if the linear two-mass oscillator reliably extrapolates to the validation scenarios.

Numerical results are compared to experimental observations for validation scenario 1 in Figure 9, where the time-domain frequency of the calibrated model is slightly slower than experimental observations for the additive Gaussian likelihood and the mixture model likelihood. In the frequency-domain, predicted resonant frequencies are slightly lower than the experimentally observed resonant frequency. These discrepancies indicate potential unresolved model-form uncertainty since a perfect model of the experiment should be able to extrapolate to unobserved mass values.

We explore the QoI predictions for validation scenario 1 in Figure 10, where the marginal calibrated model distributions of the resonant frequency and maximum magnitude are compared to the experiment. The experimentally observed resonant frequency is in the right tail of both models' marginal distributions, while the experimentally observed magnitude is in a higher density region of the marginal distribution from the mixture model likelihood. Because the marginals

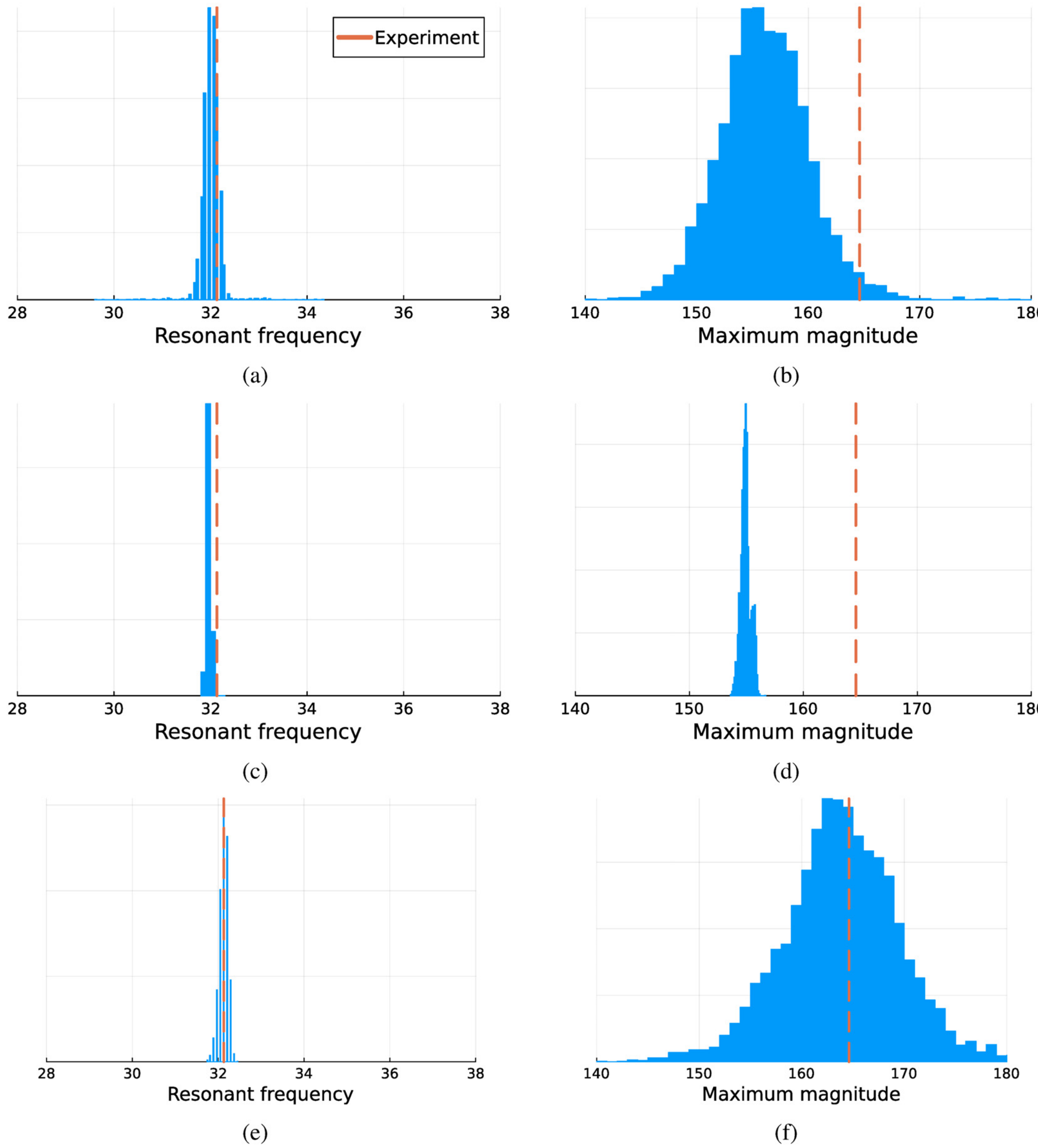


Fig. 8 Comparison of calibration scenario experimental QoI (resonant frequency and maximum magnitude) to their marginal distributions from the calibrated model with three likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b), additive Gaussian with $\sigma_\epsilon^2 = 8 \times 10^{-5}$ in (c, d), and the mixture model in (e, f).

from both likelihoods are similar, additional quantitative validation is needed to determine the best model for this validation scenario.

Numerical results are compared to experimental observations for validation scenario 2 in Figure 11, where the time-domain frequency of the calibrated models closely matches experimental observations. In Figure 11b, the first peak/trough are slightly underestimated/overestimated by the additive Gaussian likelihood, while in Figure 11c the calibrated model with a mixture model likelihood is able to encompass more observations in its 95% confidence interval.

Validation scenario 2 QoI predictions are shown in Figure 12. Again, the experimentally observed resonant frequency is in the right tail of both models' marginal distributions. The experimentally observed magnitude is in a higher density region of the marginal distribution from the additive Gaussian likelihood. Quantitative validation is needed to determine the best model for this validation scenario as well.

Model validation, which evaluates if a mathematical model is a reliable representation of reality, strongly depends on the QoI. The linear two mass oscillator with either additive Gaussian likelihood or mixture model likelihood are promising

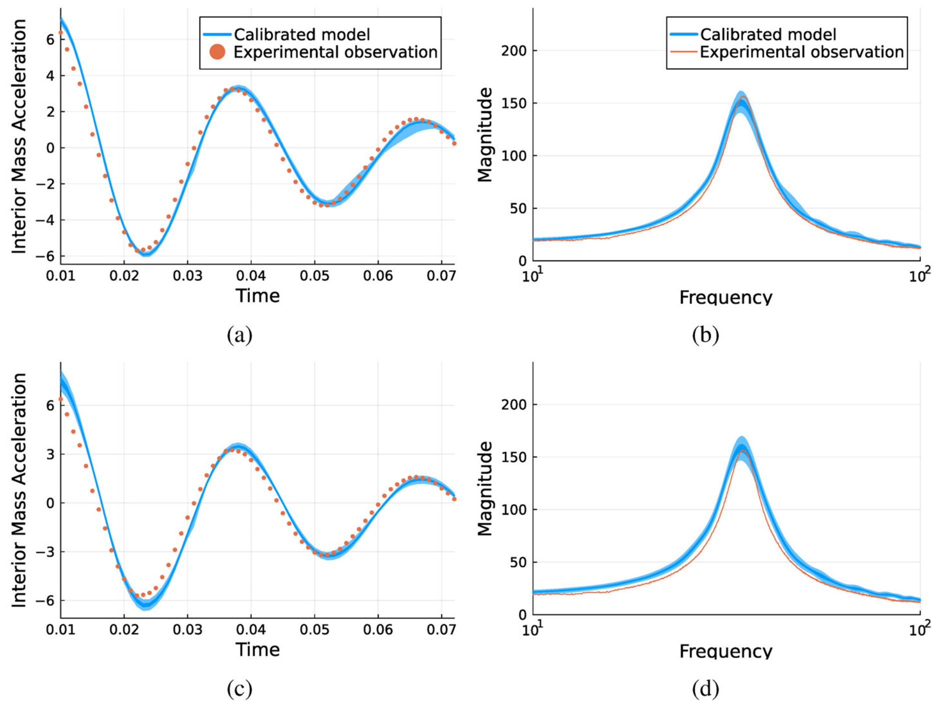


Fig. 9 Validation scenario 1 ($m = 0.7853$ kg) comparison of experimental observations to model output 50% and 95% confidence intervals (dark and lighter blue shaded regions) from the calibrated model with two likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b) and the mixture model in (c, d) for (a, c) the calibration time-window and (b, d) the frequency-domain.

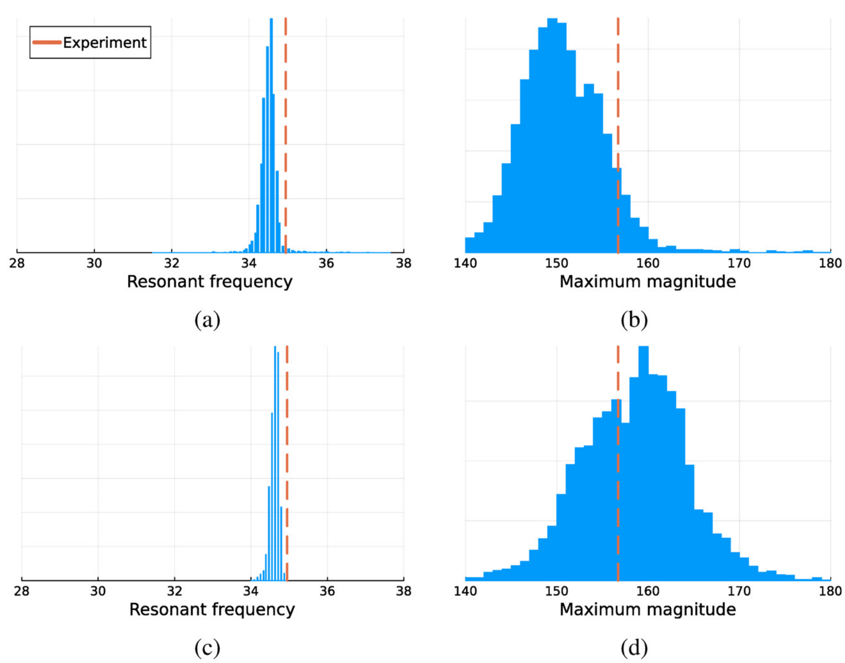


Fig. 10 Comparison of validation scenario 1 experimental QoI (resonant frequency and maximum magnitude) to their marginal distributions from the calibrated model with two likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b) and the mixture model in (c, d).

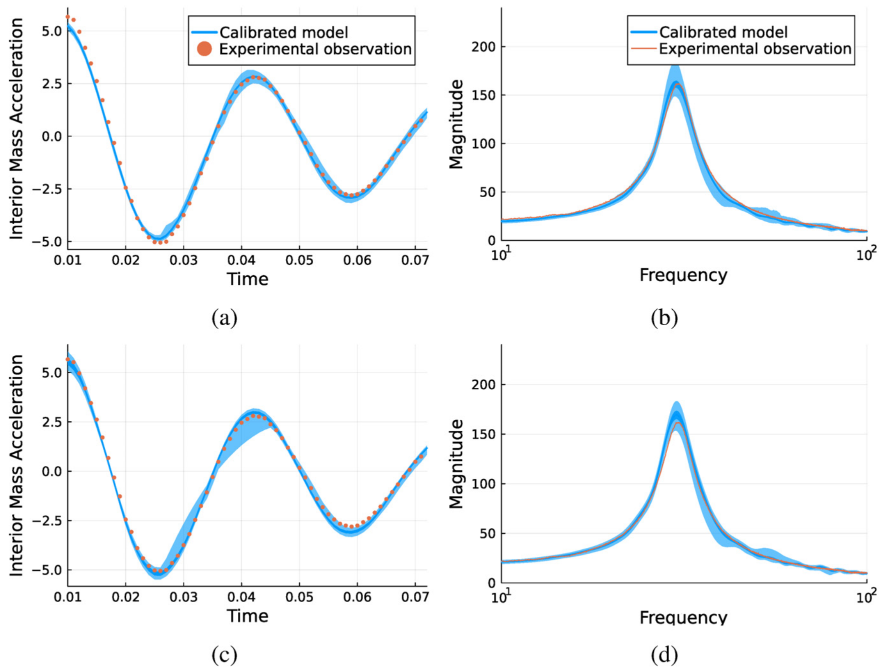


Fig. 11 Validation scenario 2 ($m = 1.0613$ kg) comparison of experimental observations to model output 50% and 95% confidence intervals (dark and lighter blue shaded regions) from the calibrated model with two likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b) and the mixture model in (c, d) for (a, c) the calibration time-window and (b, d) the frequency-domain.

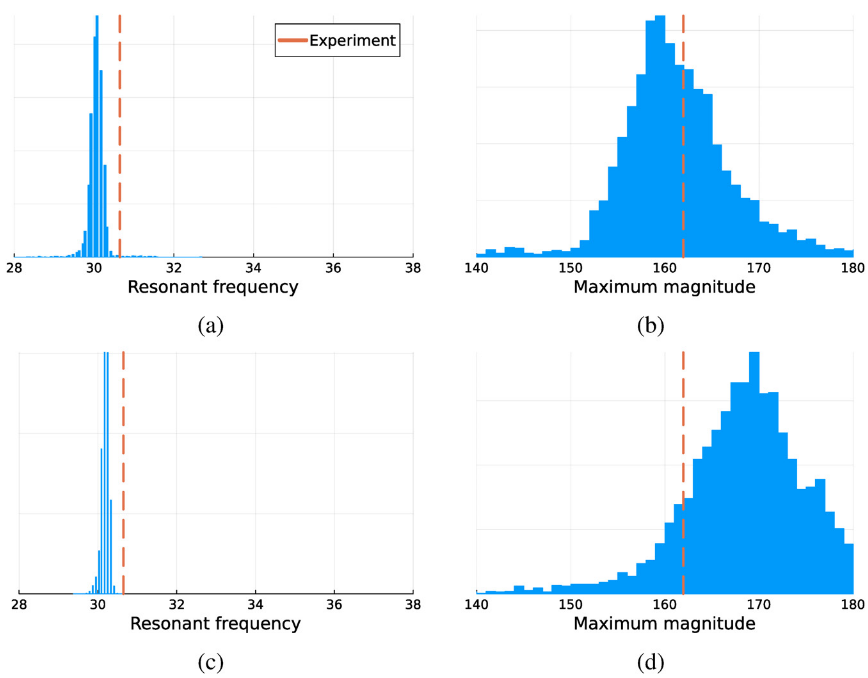


Fig. 12 Comparison of validation scenario 2 experimental QoI (resonant frequency and maximum magnitude) to their marginal distributions from the calibrated model with two likelihood distributions: additive Gaussian with $\sigma_\epsilon^2 = 0.05$ in (a, b) and the mixture model in (c, d).

models for predicting the experiment's resonant frequency and maximum magnitude. If the QoI changed to the time-domain acceleration data up to 0.5 seconds, all of these models would be invalidated.

Conclusion

In this work, we investigate and reduce discrepancies between experiments and the linear two-mass oscillator model caused by uncertainties in measurements, initial conditions and model parameters. We isolate and address each uncertainty at its source, and we inform parameter and initial condition uncertainties using experimental data. After Bayesian calibration, the linear two-mass oscillator with additive Gaussian and mixture model measurement uncertainty accurately predict the resonant frequency and associated magnitude of the active vibration isolation experiment. However, discrepancies in the time-domain could be reduced as trajectories dampen. We hypothesize that model-form uncertainties caused these discrepancies, and in the future, we plan to integrate our work from last year into this approach to quantify model-form uncertainties.

The importance of quantifying measurement uncertainty cannot be overstated and is evident in the strong impact of the likelihood in Bayesian calibration and model predictions. In the future, more work could be done to quantify a robust distribution for the measurement uncertainty. If we had access to multiple individual hits instead of one, we could more directly represent measurement uncertainties rather than modeling them as we did here. For example, we could build a covariance matrix to represent correlations in time instead of assuming independence.

Getting back to a primary motivation of the UQVI-challenge, we show that model validation depends on the QoI. The QoI should be selected based on the use case of the model, since the model may be valid for predicting one QoI but not another as we demonstrated. In order to assess diverse measures in quantifying uncertainties of various mathematical models, we need careful consideration of use cases and rigorous validation with uncertainty quantification to help determine which models are sufficient and to enable selection between multiple candidate models.

Acknowledgments This research was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan. The authors would like to thank Kyle Neal and Manuel Vega for their help.

References

1. Shahabpoor, E., Pavic, A., and Racic, V. "Identification of mass–spring–damper model of walking humans". In *Structures*, volume 5, pages 233–246. Elsevier (2016)
2. Segalman, D.J., Gregory, D.L., Starr, M.J., Resor, B.R., Jew, M.D., Lauffer, J.P., and Ames, N.M. "Handbook on dynamics of jointed structures". *Sandia National Laboratories, Albuquerque* (2009)
3. Platz, R. "Comprehensive testing environment to evaluate approaches in uncertainty quantification for passive and active vibration isolation". In *Model Validation and Uncertainty Quantification, Volume 3: Proceedings of the 40th IMAC, A Conference and Exposition on Structural Dynamics 2022*, pages 97–106. Springer (2022)
4. Platz, R. "Approach to assess basic deterministic data and model form uncertaint in passive and active vibration isolation". In *International Conference on Uncertainty in Mechanical Engineering*, pages 208–223. Springer (2021)
5. Bandy, R., Portone, T., and Morrison, R. "Stochastic model correction for the adaptive vibration isolation round-robin challenge". In *Model Validation and Uncertainty Quantification, Volume 3: Conference Proceedings of the Society for Experimental Mechanics Series*. Springer (2022)
6. Kennedy, M.C. and O'Hagan, A. "Bayesian calibration of computer models". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464 (2001)
7. Higdon, D., Gattiker, J., Williams, B., and Rightley, M. "Computer model calibration using high-dimensional output". *Journal of the American Statistical Association*, 103(482):570–583 (2008)
8. Oliver, T.A., Terejanu, G., Simmons, C.S., and Moser, R.D. "Validating predictions of unobserved quantities". *Computer Methods in Applied Mechanics and Engineering*, 283:1310–1335 (2015)
9. Liu, W. and Lin, W. "Additive white Gaussian noise level estimation in SVD domain for images". *IEEE Transactions on Image processing*, 22(3):872–883 (2012)
10. Jaynes, E.T. "On the rationale of maximum-entropy methods". *Proceedings of the IEEE*, 70(9):939–952 (1982)

11. Frigo, M. and Johnson, S.G. “The design and implementation of FFTW3”. *Proceedings of the IEEE*, 93(2):216–231 (2005) Special issue on “Program Generation, Optimization, and Platform Adaptation”.
12. Miasojedow, B., Moulines, E., and Vihola, M. “An adaptive parallel tempering algorithm”. *Journal of Computational and Graphical Statistics*, 22(3):649–664 (2013)
13. Vihola, M. “Ergonomic and reliable Bayesian inference with adaptive Markov chain Monte Carlo”. *Wiley statsRef: statistics reference online*, pages 1–12 (2020)
14. Plummer, M., Best, N., Cowles, K., Vines, K., et al. “CODA: convergence diagnosis and output analysis for MCMC”. *R news*, 6(1):7–11 (2006)