White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/212849/

Version: Published Version

Proceedings Paper:
Dardeno, T.A., Bull, L.A., Mills, R.S. et al. (2023) Hierarchical Bayesian modelling of a family of FRFs. In: Farhangdoust, S., Guemes, A. and Chang, F-K., (eds.) Fourteenth International Workshop on Structural Health Monitoring. Fourteenth International Workshop on Structural Health Monitoring, 12-14 Sep 2023, Stanford University, California. DEStech Publications, Inc., pp. 2897-2905. ISBN: 9781605956930.

https://doi.org/10.12783/shm2023/37065

© 2024 DEStech Publishing Inc. Reproduced in accordance with the publisher's self-archiving policy. Reprinted from the Fourteenth International Workshop on Structural Health Monitoring, 2023. Lancaster, PA: DEStech Publications, Inc.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Hierarchical Bayesian Modelling of a Family of FRFs

TINA A. DARDENO, LAWRENCE A. BULL, ROBIN S. MILLS, NIKOLAOS DERVILIS and KEITH WORDEN

ABSTRACT

Population-based structural health monitoring (PBSHM) aims to share valuable information among members of a population, such as normal- and damage-condition data, to improve inferences regarding the health states of the members. Even when the population is comprised of nominally-identical structures, benign variations among the members will exist as a result of slight differences in material properties, geometry, boundary conditions, or environmental effects (e.g., temperature changes). These discrepancies can affect modal properties and present as changes in the characteristics of the resonance peaks of the frequency response function (FRF). The hierarchical Bayesian approach provides a useful modelling structure for PBSHM, as population- and domain-level distributions are learnt simultaneously to bolster statistical strength among the parameters, and reduce variance among the parameter estimates. This paper provides an overview of current work, where hierarchical Bayesian models are developed for a small population of nominally-identical helicopter blades, using FRF data. These models account for benign variations that present as differences in the underlying dynamics across the input space, while also considering (and utilising) the similarities among the blades.

INTRODUCTION

The current work is focussed on applying population-based structural health monitoring (PBSHM) to a group of *nominally-identical* structures (i.e., a *homogeneous* population [1, 2]), with consideration for the effects of temperature changes, to improve understanding of the variability in the health states of the population, as well as the individual members. (In contrast, a *heterogeneous* population [3, 4] will have greater discrepancies among the members, such as different suspension bridge designs, and may require further processing such as domain adaptation [5–8]). However, even among nominally-identical structures, variations caused by manufacturing differences, ageing parts, and changes in testing conditions can introduce uncertainty in the underlying dy-

T.A. Dardeno[1], L.A. Bull[2], R.S. Mills[1], N. Dervilis[1], K. Worden[1], [1]Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, S1 3JD, UK.; [2]Computational Statistics and Machine Learning Group, Engineering Department, University of Cambridge, CB2 1PZ, UK

namics. For example, increased temperature may reduce stiffness, which may decrease natural frequency (a common feature of interest for many SHM systems). Likewise, damping can also be affected by temperature, and might increase or decrease depending on the material or temperature range [9]. Accounting for these benign fluctuations is important for the practical implementation and generalisation of SHM technologies, as features commonly used for damage identification are often sensitive to harmless changes as well as damage [10, 11].

Another challenge for SHM systems that rely on machine learning is data *sparsity*. Sensing networks are prone to data loss, because of sensor failure caused by harsh environmental conditions or insufficient maintenance. Transmission issues make wireless sensing networks particularly susceptible to loss, and can be caused by large transmission distances between the sensors and base station [12], software/hardware problems [13], and other issues such as weather changes, interference from nearby devices, or installation difficulties [14]. In addition, modern systems that produce large amounts of high-resolution data can suffer losses resulting from a data transfer bottleneck [15]. Significant losses higher than 30% have been reported [13, 14, 16], and a 0.38% data loss was found to have similar effects on power spectral density (PSD) as 5% additive noise [17]. Differences in sample rate among sensors could have a similar presentation, with the data captured at a lower rate seemingly missing data relative to that captured at a higher rate. Again, PBSHM addresses these sparsity issues via knowledge transfer among similar structures, so that data-rich members can support those with more limited information.

Homogenous populations can be represented using a general model, called a population *form*. The *form* is a model that attempts to capture the 'essential nature' of the population, as well as the variations encountered by the structure in operation (e.g., temperature changes) and slight differences among nominally-identical members [1]. The concept of the population form was first introduced in [1, 2], where a conventional or single Gaussian process (GP) was applied to frequency response functions (FRFs), to develop a representation for a nominally-identical population of eight degree-of-freedom (DOF) systems. To account for greater differences among the nominally-identical systems, an overlapping mixture of Gaussian processes (OMGP) [18], was used in [1, 19], to infer multivalued wind-turbine power-curve data, with unsupervised categorisation of the data. The OMGP approach [18] was again used in [20] to develop a population form for real and imaginary FRFs, obtained from four nominally-identical, full-scale, composite helicopter blades. In [21, 22], hierarchical Bayesian models were used to improve the predictive capability of simulated [21], and in-service [22], truck fleet hazard models, and wind turbine power curves [22]. It was shown that when populations of structures were allowed to share correlated information, model uncertainty was reduced [21, 22]. In addition, domains with incomplete data were able to borrow statistical strength from data-rich groups [21, 22].

Current work has involved extending the FRF-based population form to include a hierarchical Bayesian modelling structure, where a combined probabilistic FRF model is developed using vibration data from the small population of nominally-identical helicopter blades introduced in [20]. This paper presents a brief overview of the current work, including two case studies. The first case study utilised FRFs collected from all four blades, at ambient laboratory temperature, with variations among the blades resulting from manufacturing differences (e.g., small discrepancies in material properties and

geometry) and boundary conditions. Limited training data that did not fully characterise the resonance peaks were taken from two of the FRFs, while sufficient training data were taken from the remaining two FRFs, so that information could be shared with the data-poor domains via shared distributions over the parameters. Independent models were generated for comparison, to visualise the variance reduction from the combined model. The second case study considered vibration data from one of the helicopter blades collected at various temperatures in an environmental chamber. A probabilistic FRF model was again developed using a hierarchical approach. Functional relationships between temperature and the modal parameters were incorporated into the model. The learned functions were then used to extrapolate to temperatures not used to train the model, and model accuracy was evaluated by comparing the results to FRFs computed via measured vibration data.

HIERARCHICAL BAYESIAN MODELLING OF THE FRF

Hierarchical models can be used to make combined inferences, whereby domains can be treated as separate; but, at the same time, it is assumed that each domain is a realisation from a common process. This modelling structure involves *partial pooling*, and is beneficial in that population-level distributions are informed by the full data set comprised of multiple domains. In partial pooling, certain parameters are permitted to vary between domains (i.e., *varying parameters*); which are correlated by conditioning on parent variables at the population level. In this example, the natural frequency would be a varying parameter. Other parameters can be considered *shared* among members of a population (e.g., additive noise) and are learnt at the population level (these shared variables can still be sampled from parent distributions, which are also learnt at the population level). In contrast, a complete-pooling approach would consider all population data as having originated from a single source, while a no-pooling approach would involve fitting a single domain independently from the other domains.

Hierarchical models with partial pooling are particularly useful for PBSHM. Because parameters are allowed to vary at the domain level (as opposed to complete pooling), this approach can accommodate benign variations within a population. In addition, population-level variables are informed by the full data set, rather than data from a single domain (no pooling). This increase in statistical power is especially important in situations where one or more domains have limited data [21–23]. In such cases, parameters from the data-poor domains exhibit shrinkage towards the population mean (therefore borrowing information from the other domains), which tightens the parameter variance [21–23].

For the current work, probabilistic FRF models for the helicopter blades were developed using a hierarchical approach with partial-pooling. Note that only the real part of the FRF was considered, to simplify the analysis, although the imaginary part could be learnt using the same methods or inferred (at least in part), by exploiting the causal relationship between the real and imaginary components of the FRF [24]. Models were developed using the probabilistic programming language `Stan`. Analyses were performed using MCMC, via the no U-turn (NUTS) implementation of Hamiltonian Monte Carlo (HMC) [25, 26]. HMC uses approaches based in differential geometry to generate transitions that span the full marginal variance, so it is not susceptible to the random walk behaviour that can occur with other samplers [26].

For domains $k \in \{1, \ldots, K\}$, given frequency inputs, $\boldsymbol{\omega}_k$, and accelerance outputs, $\mathbf{H}_k$, the population data can be denoted as,

$$\{\boldsymbol{\omega}_k, \mathbf{H}_k\}_{k=1}^{K} = \left\{ \{\omega_{ik}, H_{ik}\}_{i=1}^{N_k} \right\}_{k=1}^{K} \tag{1}$$

where $\{\omega_{ik}, H_{ik}\}$ are the i^{th} pair of observations in domain k. Each domain (which are each comprised of data from a single FRF), is comprised of N_k observations, giving a total of $\sum_{k=1}^{K} N_k$ observations. Then, considering only the real component of the FRF, the objective is to learn a set of K predictors (one for each domain), related to the regression task, where the tasks satisfy,

$$\{real\left[H_{ik}\right] = f_k\left(\omega_{ik}\right) + \epsilon_{ik}\}_{k=1}^{K} \tag{2}$$

where for each observation i, the output is determined by evaluating one of K latent functions, $f_k\left(\omega_{i,k}\right)$, plus additive noise, $\epsilon_{i,k}$. While each of the k groups can be learned independently, a combined inference can be used to take advantage of the full $\sum_{k=1}^{K} N_k$ population data set. The (real) FRFs were modelled probabilistically with an assumed Gaussian distribution, with likelihood,

$$real\left[\mathbf{H}_k\left(\boldsymbol{\omega}_k\right)\right] \sim \mathcal{N}\left(\mathbf{f}_k\left(\boldsymbol{\omega}_k\right), \sigma_H^2\right) \tag{3}$$

where $\mathbf{f}_k\left(\boldsymbol{\omega}_k\right)$ is equal to the real component of the accelerance FRF, calculated via an estimation based on modal parameters,

$$real\left[\mathbf{H}_k\left(\boldsymbol{\omega}_k\right)\right] = -\boldsymbol{\omega}_k^2 \sum_{m=1}^{M} \frac{\mathrm{A}_m \left(\omega_{nat}^{2(k,m)} - \boldsymbol{\omega}_k^2\right)}{\omega_{nat}^{4(k,m)} + \boldsymbol{\omega}_k^4 + 2\boldsymbol{\omega}_k^2 \omega_{nat}^{2(k,m)} \left(2\zeta_{(k,m)}^2 - 1\right)} \tag{4}$$

where A_m is the residue for mode m, defined as the product of the mass-normalised mode shapes [24]. Note that the residue is not indexed by k; this is because it was treated as a shared variable among the different FRFs for both case studies. The natural frequency associated with the kth FRF for mode m is denoted as $\omega_{nat}^{k,m}$, and the modal damping associated with the kth FRF for mode m is $\zeta_{k,m}$ [24].

A shared hierarchy of prior distributions was placed over the modal parameters, in line with a Bayesian framework. To allow information to flow between domains, parent nodes were learnt at the population level. Some parameters were assumed shared among domains, and were therefore also learnt at the population level (e.g., noise variance). Extension of hierarchical Bayesian modelling to FRFs for the purpose of PBSHM has been explored via two cases studies. A brief overview of these cases is provided below.

Case 1: FRFs from Multiple Nominally-Identical Structures

The first case used FRFs collected from four full-scale, composite, nominally-identical helicopter blades, with variations among the blades resulting from manufacturing differences (e.g., small discrepancies in material properties and geometry), and boundary conditions. The FRFs were computed from vibration data collected at ambient laboratory temperature, with the blades in an approximately fixed-free boundary condition (for a more detailed discussion of these tests, see [20]). All experiments used in this analysis

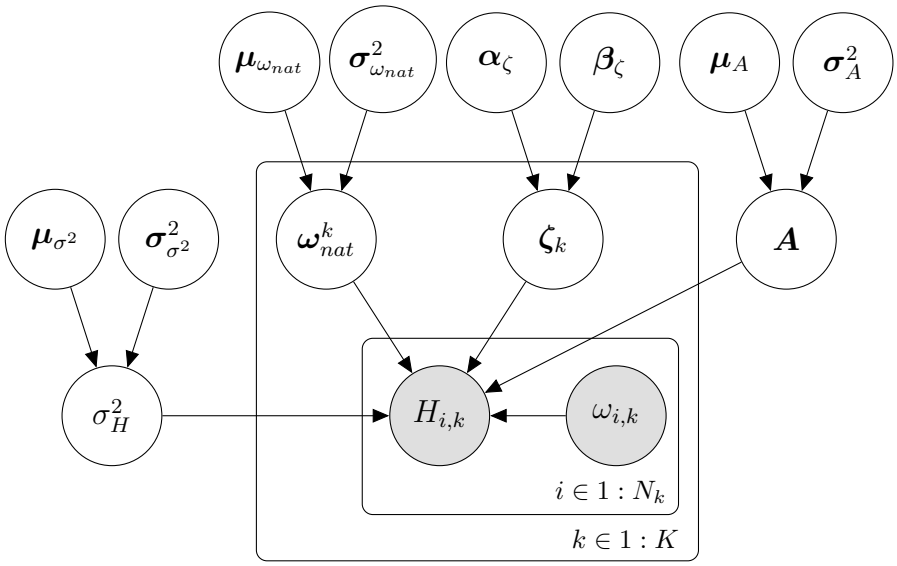

Figure 1. Graphical representation of the hierarchical FRF model developed for the first case.

were performed at the Laboratory for Verification and Validation (LVV) in Sheffield, UK, using Siemens PLM LMS SCADAS hardware and software. To highlight the benefits of hierarchical modelling (specifically, variance shrinkage towards the population mean, particularly when data are sparse), limited training data that did not fully characterise the resonance peaks were taken from two of the FRFs, while sufficient training data were taken from the remaining two FRFs. This situation is representative of incomplete data in the time domain, which would reduce the number of spectral lines in the frequency domain.

To simplify the analysis, a narrow frequency band was selected between 24 and 61 Hz, with the fourth and fifth bending modes of the blades dominating the response in this band. A combined probabilistic FRF model was then generated according to the graphical model shown in Figure 1. Independent models were also computed to demonstrate the variance reduction achieved using the partial-pooling technique. FRFs for each blade were computed via Eq. (4), from the samples of the modal parameters. Total variance was estimated by adding the standard deviation of the FRFs to the expectation of the noise variance. Posterior predictive mean and 3-sigma deviation for the partial pooling and independent models are plotted in Figures 2a to 2d, respectively, and show that by borrowing information from data-rich domains (Figures 2a and 2b), within the population, the variance for the data-poor FRFs (Figures 2c and 2d) was reduced compared to independent modelling.

Case 2: FRFs from a Single Structure with Temperature Variation

The second case used FRFs computed from data collected in an environmental chamber at temperatures ranging from -20 to 30°C in increments of 5°C), from a single blade in an approximately fixed-free boundary condition. Testing was again performed at the LVV in Sheffield, UK. A probabilistic FRF model was developed using a hierarchical and partial-pooling approach, to accommodate the changes resulting from temperature variations, according to the graphical model shown in Figure 3. To simplify the analysis,

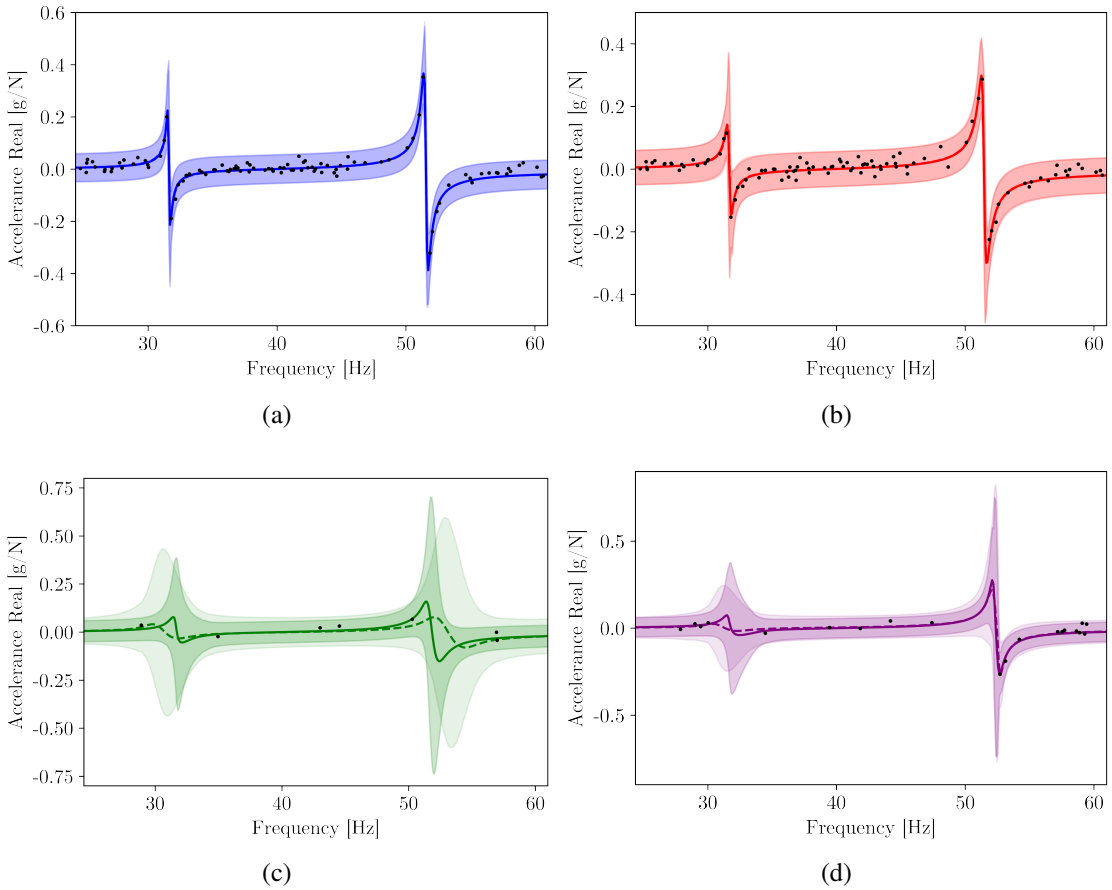

Figure 2. Posterior predictive mean and 3-sigma deviation for independent (no-pooling) and partial-pooling models, (a) Blade 1 (blue), (b) Blade 2 (red), (c) Blade 3 (green) and (d) Blade 4 (purple). Training data are shown in a black scatter plot. Posterior predictive means are shown as solid and dashed lines for the partial-pooling and independent models, respectively. The variance is represented by shaded regions, where the independent model variances are shown in a lighter color than those for the partial-pooling models.

a narrow frequency band was selected between 135 and 155 Hz, with a higher-order bending mode of the blade dominating the response in this band. The relationships between temperature and natural frequency, and between temperature and damping, were represented using Taylor series expansions (first and second order expansions for natural frequency and damping, respectively). Polynomial coefficients were determined for the temperature-varied FRFs, using a subset of the FRFs as training data.

Predictions were then made at 'unseen' temperatures (not used in training) using the expectation of population-level variables, and compared to measured FRFs to evaluate model accuracy. The extrapolated FRFs are shown in Figure 4, ranging from -20 to 30°C in increments of 5°C, with decreasing temperature from left to right (as natural frequency increases with decreasing temperature). In Figure 4, the extrapolated FRFs are shown as solid blue lines, with a shaded blue region indicating the variance bounds. The measured FRFs used to train the model (without added noise), are shown as solid red lines, and the measured FRFs used to test the model at 'unseen' temperatures are shown as dashed black lines. From the figure, it is clear that excellent agreement was achieved between the extrapolated and measured FRFs at the temperatures used to train the model, as expected.

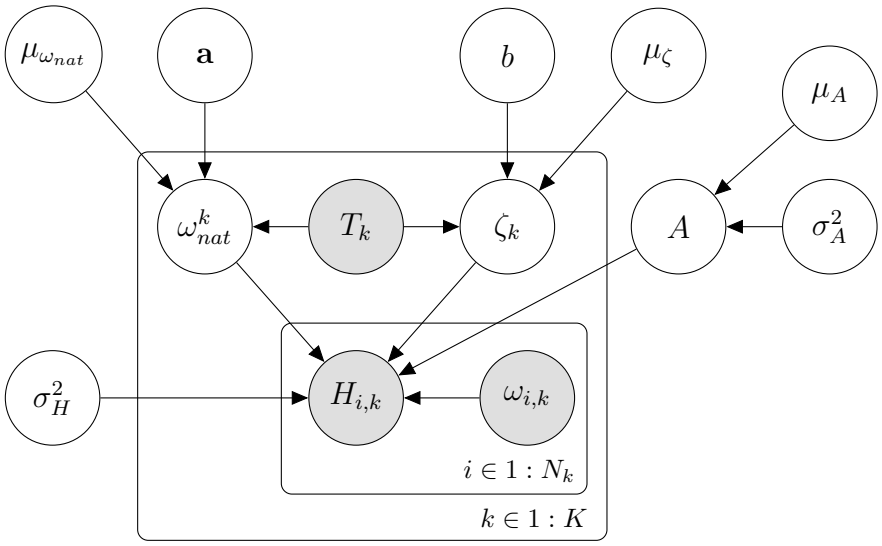

Figure 3. Graphical representation of hierarchical FRF model for the second case.

Likewise, the FRFs at temperatures between those used to train the model (e.g., at 0°C and 15°C) show excellent agreement. Notably, the FRFs at colder temperatures (further away from the training data) still show good agreement. Detailed discussions of the cases presented herein can be found in [27].

CONCLUSIONS

Ongoing work involves the development of probabilistic FRF models using a hierarchical Bayesian approach, that account for benign variations and similarities among nominally-identical structures. Brief overviews of two cases were presented. The first case demonstrated how this modelling approach can reduce variance in data-poor domains by allowing information transfer with data-rich domains, via shared population-level distributions. The second case showed that including functional relationships in the modelling structure to describe temperature variations allows prediction beyond the training data. Future work will investigate the structure of the data and hierarchical models.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC), via grant reference EP/W005816/1. This research made use of The Laboratory for Verification and Validation (LVV), which was funded by the EPSRC (via EP/J013714/1 and EP/N010884/1), the European Regional Development Fund (ERDF), and the University of Sheffield. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

The authors would like to extend special thanks to Michael Dutchman of the LVV, for helping set up the experiments, and also Domenic Di Francesco of the Alan Turing Institute, for his advice when designing the hierarchical models.

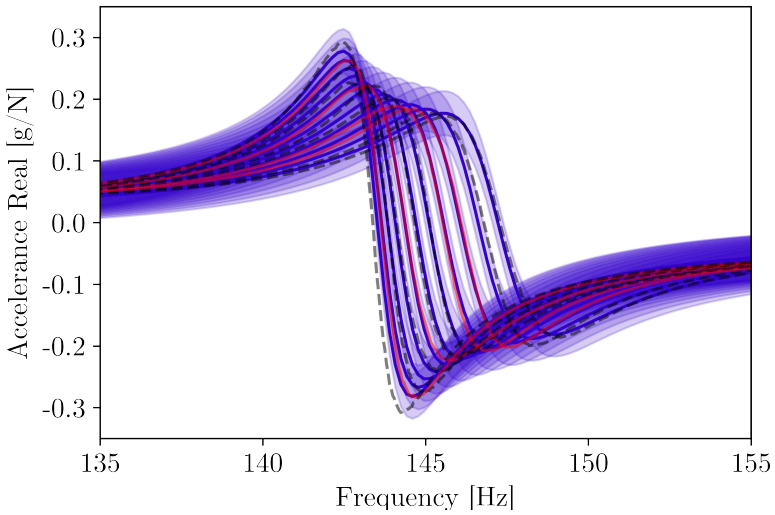

Figure 4. Extrapolated FRFs using population-level variables. The extrapolated means are shown as solid blue lines, with shaded blue regions indicating variance bounds (3-sigma deviation). Solid red lines are the FRFs used for training the model, while dashed black lines are the measured FRFs at 'unseen' temperatures.

REFERENCES

1. L.A. Bull, P.A. Gardner, J. Gosliga, T.J. Rogers, N. Dervilis, E.J. Cross, E. Papatheou, A.E. Maquire, C. Campos, and K. Worden. 2021. "Foundations of population-based SHM, Part I: Homogeneous populations and forms," *Mech. Syst. Signal Process.*, 148:107141.
2. L.A. Bull, T.J. Rogers, N. Dervilis, E.J. Cross, and K. Worden. July 9-10, 2019. "A Gaussian process form for population-based structural health monitoring," in *Proceedings of the 13th International Conference on Damage Assessment of Structures (DAMAS 2019), Porto, Portugal*, Springer.
3. J. Gosliga, P.A. Gardner, L.A. Bull, N. Dervilis, and K. Worden. 2021. "Foundations of Population-based SHM, Part II: Heterogeneous populations–Graphs, networks, and communities," *Mech. Syst. Signal Process.*, 148:107144.
4. P.A. Gardner, L.A. Bull, J. Gosliga, N. Dervilis, and K. Worden. 2021. "Foundations of population-based SHM, Part III: Heterogeneous populations–mapping and transfer," *Mech. Syst. Signal Process.*, 149:107142.
5. P. Gardner, R. Fuentes, N. Dervilis, C. Mineo, S.G. Pierce, E.J. Cross, and K. Worden. 2020. "Machine learning at the interface of structural health monitoring and non-destructive evaluation," *Philos. Trans. Royal Soc. A*, 378(2182):20190581.
6. P. Gardner, X. Liu, and K. Worden. 2020. "On the application of domain adaptation in structural health monitoring," *Mech. Syst. Signal Process.*, 138:106550.
7. P. Gardner, L.A. Bull, N. Dervilis, and K. Worden. 2022. "On the application of kernelised Bayesian transfer learning to population-based structural health monitoring," *Mech. Syst. Signal Process.*, 167:108519.
8. J. Poole, P. Gardner, N. Dervilis, L.A. Bull, and K. Worden. 2022. "On statistic alignment for domain adaptation in structural health monitoring," *Struct. Health Monit.*, 0(0):1–20.
9. M. Colakoglu. 2008. "Effect of temperature on frequency and damping properties of polymer matrix composites," *Adv. Compos. Mater.*, 17(2):111–124.
10. K. Worden, H. Sohn, and C. Farrar. 2002. "Novelty detection in a changing environment: Regression and interpolation approaches," *J. Sound Vib.*, 258:741–761.

11. H. Sohn. 2007. “Effects of environmental and operational variability on structural health monitoring,” *Philos. Trans. A: Math. Phys. Eng. Sci.*, 365:539–560.
12. J. Pei, C. Kapoor, T. Graves-Abe, Y. Sugeng, and J. Lynch. 2005. “Critical Design Parameters and Operating Conditions of Wireless Sensor Units for Structural Health Monitoring,” in *Proceedings of the 23rd International Modal Analysis Conference (IMAC)*.
13. J. Meyer, R. Bischoff, G. Feltrin, and M. Motavalli. 2010. “Wireless sensor networks for long-term structural health monitoring,” *Smart Struct. Syst.*, 6.
14. Y. Bao, Y., H. Li, X. Sun, Y. Yu, and J. Ou. 2013. “Compressive sampling–based data loss recovery for wireless sensor networks used in civil structural health monitoring,” *Struct. Health Monit.*, 12(1):78–95.
15. A. Malekloo, E. Ozer, M. AlHamaydeh, and M. Girolami. 2022. “Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights,” *Struct. Health Monit.*, 21(4):1906–1955.
16. J. Lynch and K. Loh. 2006. “A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring,” *Shock Vibr. Dig.*, 38:91–128.
17. T. Nagayama, S. Sim, Y. Miyamori, and B.F. Spencer Jr. 2007. “Issues in structural health monitoring employing smart sensors,” *Smart Struct. Syst.*, 3(3):299–320.
18. M. Lázaro-Gredilla, S. Van Vaerenbergh, and N.D. Lawrence. 2012. “Overlapping mixtures of Gaussian processes for the data association problem,” *Pattern Recognit.*, 45(4):1386–1395.
19. L.A. Bull, P.A. Gardner, T.J. Rogers, N. Dervilis, E.J. Cross, E. Papatheou, A.E. Maquire, C. Campos, and K. Worden. 2021. “Bayesian modelling of multivalued power curves from an operational wind farm,” *Mech. Syst. Signal Process.*:108530.
20. T.A. Dardeno, L.A. Bull, R.S. Mills, N. Dervilis, and K. Worden. 2022. “Modelling variability in vibration-based PBSHM via a generalised population form,” *J. Sound Vib.*, 538:117227.
21. L.A. Bull, M. Dhada, O. Steinert, T. Lindgren, A. Kumar, and M. Girolami. July 4-7, 2022. “Population-Level Modelling for Truck Fleet Survival Analysis,” in *Proceedings of the 10th European Workshop on Structural Health Monitoring (EWSHM), Palermo, Italy*, Springer.
22. L.A. Bull, D. Di Francesco, M. Dhada, O. Steinert, T. Lindgren, A.K. Parlikad, A.B. Duncan, and M. Girolami. “Hierarchical Bayesian modeling for knowledge transfer across engineering fleets via multitask learning,” *Comput.-Aided Civ. Inf.*, 00(n/a):1– 28.
23. A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. 2013. *Bayesian Data Analysis*, CRC Press, Boca Raton, FL.
24. K. Worden and G.R. Tomlinson. 2001. *Nonlinearity in Structural Dynamics: Detection, Identification and Modelling*, Institute of Physics Publishing, Bristol, UK.
25. M.D. Hoffman and A. Gelman. 2014. “The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.*, 15(1):1593–1623.
26. M. Betancourt and M. Girolami. 2015. “Hamiltonian Monte Carlo for hierarchical models,” *Current trends in Bayesian methodology with applications*, 79(30):2–4.
27. T.A. Dardeno, R.S. Mills, N. Dervilis, K. Worden, and L.A. Bull. 2023. “On the hierarchical Bayesian modelling of frequency response functions,” *arXiv preprint arXiv:2307.06263*.