Partially Structured Gaussian Processes for Grey-Box Learning in SHM

M. R. JONES, D. J. PITCHFORTH and E. J. CROSS

ABSTRACT

Combining physics with data-driven learning seeks to harness the benefits of the expressivity of machine learners whilst having known physical laws that the model must adhere to. Building these types of models have the advantage of being able to represent phenomena more complex than can be described with simplified analytical equations, whilst simultaneously allowing for greater prediction confidence and interpretability, as well as providing a mechanistic basis that predictions can revert to when data becomes scarce. With operators more frequently carrying out monitoring campaigns, and decades of engineering knowledge to call upon, it is a natural step for health monitoring practitioners to adopt physics-informed learning strategies for monitoring structures.

In many engineering scenarios, it is often the case that one will not have complete physical knowledge of a structure or system of interest, but *partial knowledge*, for which it is desirable to encode in a machine learning model. In this paper, we will explore how this type of knowledge can be embedded into Gaussian process models for processes that naturally exist as products of constituent functions, such as for processes that have both spatial and temporal dependencies. Focus will be placed on how the properties of the covariance function that has embedded physical structure dictates the predictive regime that the model may operate in within the partial knowledge setting. Application of the developed models includes learning the decoupled response of a beam under random loading into the constituent mode shapes and temporal response.

INTRODUCTION

As sensing data becomes increasingly available, it has become commonplace in many research fields to call upon data-driven learners to model complex physical processes. One example of such a field is structural health monitoring (SHM) [1], where machine learning is now regularly employed to make inferences over the health state of a structure, with objectives ranging from detecting the presence of damage [2, 3], ultimately up to predicting the remaining useful life of a structure. Whilst data-driven learn-

---

Matthew R. Jones, Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield, S6 3JD, UK. Email: matthew.r.jones@sheffield.ac.uk

ers offer a high degree of flexibility, their predictive performance is heavily conditioned on the availability of a training set that is sufficiently representative of the prediction space of interest. What this means from an SHM perspective is that there should exist data from each possible environmental and operational condition a structure may be subject to, requiring large scale investment from asset owners in terms of time spent acquiring data, and the associated financial cost. For problems that rely on spatial mapping such as learning ultrasonic feature mappings [4, 5] this requirement extends to acquiring a suitably dense training set, which, for large scale industrial structures, can be a limiting factor.

In an attempt to lessen the reliance of machine learning tools on training data, much recent work has been dedicated to incorporating known physical knowledge. In the general machine learning community, this is often referred to as physics-informed machine learning, or learning physically-consistent inductive biases. Detailed reviews on this area of research are provided in [6–8]. In SHM, terminology also extends to grey-box models, derived from one considering a physics-informed (grey-box) model as the combination of a purely physics-based (white-box) model and a purely data-driven (black-box) model [9]. The advantage here is that one can retain the flexibility provided by data-driven learners, whilst simultaneously incorporating known mechanistic laws that govern the dynamic process one is interested in capturing. Some recent examples of work in this area includes wave loading prediction [10], ultrasonic feature mapping [5, 11], capturing temperature trends during the monitoring of bridges [12], and domain adaptation for energy forecasting [13].

A natural way to handle the grey-box paradigm is through Bayesian statistics, where one begins with some prior belief. Having observed some data that can be represented as having been drawn from a distribution through the use of a likelihood function, the prior can be combined with the likelihood and a normalisation term to infer a posterior. From a grey-box perspective, the prior provides the opportunity to encode physical insight/domain awareness into the learning task. The physically encoded prior can then be combined with observational data through the likelihood to return a posterior that is a blend of data and physics-based learning, with the exact ratio of physics to data dependent on the depth of knowledge in the prior and the information provided in the data. One approach for forming Bayesian anaylsis as a learning task is to interpret the prior as a Gaussian process (GP) [14], providing the means for constructing a regression problem, whilst allowing for the inclusion of physics through embedding structure in the GP.

When constructing a physics-informed model, it is imperative that the knowledge embedded is representative, at least in some way, of the actual underlying phenomena. If not abided by, the included knowledge will generally not improve model performance, and likely even hinder it. As such, it is often the case that only *partial* knowledge is available, in which only *some* insight regarding the driving physical mechanism is embedded in the model. For instance, one may possess some linear approximation of the physics, but lack an understanding of more complex behaviour that fully encapsulates the entirety of the underlying phenomena. Where this is the case, some authors have adopted a residual modelling approach, where a physical model of the more simple physics is summed with a flexible, data-driven component that captures any remanent behaviour not represented by the approximation [10]. Whilst effective, it may sometimes be more appropriate to construct a model that considers other linear operations over multiple functions.

In this paper, focus will be on the product operation. In SHM, there are many scenarios where products of multiple constituent functions are of interest. One example is the combination of spatial and temporal phenomena, which frequently arise when modelling dynamic systems, particularly under environmental and operational variations. Systems with time-dependent parameters may also be modelled as being the product of two underlying processes [15].

Whilst physics-informed machine learning is gaining a fair amount of traction, there has been much less attention on how physics may be fused into processes that are the product of multiple lower level processes in a partial knowledge setting, where one may have good insight into the governing structure of only one of the composite functions. Take the vibration response of a linear structure as an example, which can be decomposed into a product of the mode shape (spatial component) and the temporal response. In this setting, structure relating to the temporal response may be known and subsequently embedded through use of the covariance function of a linear dynamic system under random loading [16], with the mode shapes treated as unknown and need to be learnt. Even in cases where the geometry of the structure has a corresponding analytical form, such as a cantilever beam, use of such expression will enforce many assumptions that are often simplifications of a real-world system, including uniform mass, perfect boundary conditions, and homogeneous material properties, as discussed in [17].

In this paper, we consider how one may attempt to exploit partial physical/domain knowledge for processes that are formed as a product operation of multiple constituent functions. The analysis will be entirely through the lens of kernel machines [18] - specifically Gaussian processes [14] - and thus how kernel design can be exploited in this setting. It will be considered how the properties of the physical kernel, such as whether it is stationary or non-stationary, impacts the predictive capabilities of the overall model. In particular, whether interpolation or extrapolation should be performed, dependent upon how the physics held translates into a kernel representation.

The paper proceeds as follows; Section 2 will introduce the Gaussian process machinery necessary, discuss how kernel selection can be used to can embed structured physical insight, as well as the influence that particular kernel properties have on prediction capabilities in both interpolation and extrapolation. Section 3 will outline the examples used in the paper and the specific construction of the models considered, with Section 4 presenting the corresponding results and discussion. The paper concludes in Section 5.

PARTIAL PHYSICS KNOWLEDGE IN GAUSSIAN PROCESSES MODELS

A convenient model form to embed partial structure when considering products of functions are Gaussian processes; a class of Bayesian non-parametric regression techniques. The construction of this model type naturally lends itself to modelling products of functions by considering products over kernels, and will be discussed in more detail later in the section. First, let us briefly introduce the basics before progressing onto the mechanism for constructing products-of-function models.

A brief introduction to Gaussian process regression

When modelling a variable as a Gaussian process, the assumption is made that the variable is drawn from a Normal distribution, with any collection of draws made from the Gaussian process to also be jointly Normal. This property is extremely useful in a regression setting, where a GP can be used to learn a nonlinear mapping $f : X \rightarrow \mathbf{y}$, where X are inputs and $\mathbf{y}$ the corresponding targets[1]. To define a Gaussian process prior over f relies on specifying a mean and covariance function. More formally,

$$f \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot')) \tag{1}$$

where f is a GP prior with mean function $m(\cdot)$ and covariance function $k(\cdot, \cdot')$. To make predictions at unseen input locations $X_\star$, the prior may be conditioned on a set of training inputs and outputs, D, subsequently forming the joint distribution between the training and test data. This results in the following predictive equations,

$$p(y_\star) = \mathcal{N}(\mathbb{E}[y_\star], \mathbb{V}[y_\star]) \tag{2}$$

where,

$$\mathbb{E}[y_\star] = K(X_\star, X)[K(X, X) + \sigma_n^2 \mathbb{I}]^{-1}\mathbf{y} \tag{3a}$$

$$\mathbb{V}[y_\star] = K(X_\star, X_\star) - K(X_\star, X)[K(X, X) + \sigma_n^2 \mathbb{I}]^{-1} K(X, X_\star) \tag{3b}$$

Gaussian processes in product spaces

As seen in equation (1), a GP is defined by a mean and kernel function. If we reduce to the zero prior mean case, then all that remains is the specification of the kernel, which essentially *weights* how a target prediction should relate to outputs already observed in the training data set. Kernels cannot, however, be any arbitrary function of a pair of inputs, and are required to satisfy two conditions; namely *positive-semi definiteness* and *symmetry* with respect to the inputs [14]. Given that linear operations acting on a kernel also produces valid kernels, then it is possible to construct a kernel that is the product of two individual kernels, thus naturally handling functions of such nature.

Kernels may also exhibit other properties, such as whether they are *stationary* or *non-stationary*, *isotropic* or *anisotropic* and order of *differentiability* to name a few. In the context of partial knowledge, *stationarity* of the kernel is perhaps the most fundamental property, ultimately governing whether the kernel allows for extrapolation or not. If a kernel is said to be stationary, then the function is only dependent upon the distance between a set of inputs $\mathrm{x} - \mathrm{x}'$, rather than the absolute value, x, and so is invariant to shifts in the input space. As an example, take the squared exponential kernel function,

$$k_{SE} = \sigma_{f,SE}^2 \exp\left(\frac{||\mathbf{x} - \mathbf{x}'||^2}{l_{SE}^2}\right) \tag{4}$$

which due to containing $\mathbf{x} - \mathbf{x}'$, serves as a stationary kernel in terms of $\mathbf{x}$. However, the linear kernel,

$$k_{lin} = (c_{lin} + \sigma_{f,lin}^2 \mathbf{x}\mathbf{x}') \tag{5}$$

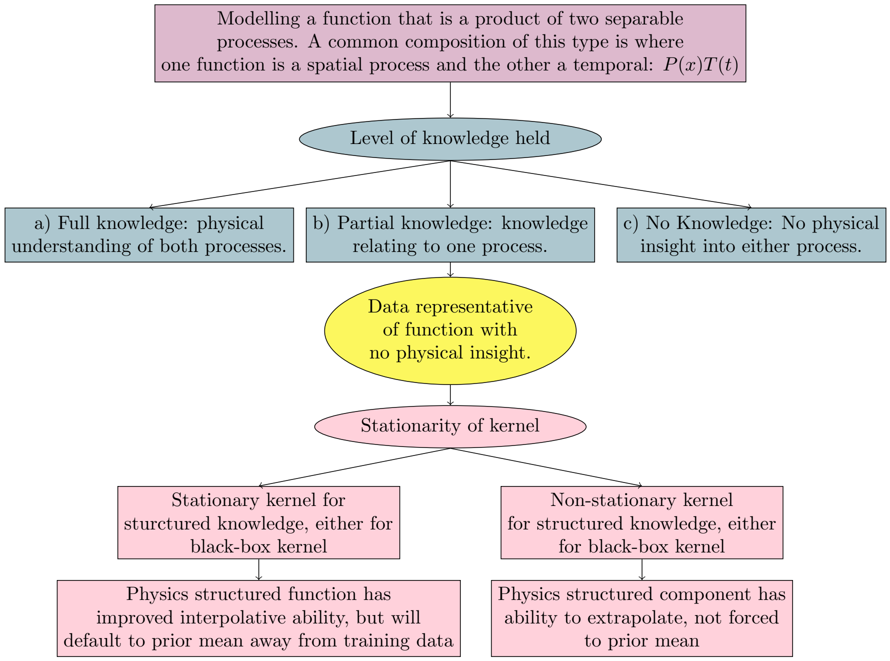


Figure 1: Workflow for embedding knowledge in products of functions as modelled by kernel functions.

is a function of $\mathbf{x} \cdot \mathbf{x}'$, and so is sensitive to translations in the input space, and thus, non-stationary. For stationary kernels, in an extrapolation regime, the distance between training and testing inputs grows, thus forcing $\mathbf{x} - \mathbf{x}'$ to become large. Stationary kernels should therefore only be employed in settings where there is ample data *coverage* for the process one is interested in learning. If data coverage is limited, but there exists physical insight into the process being modelled, then this structure can be built into the kernel. Where a non-stationary kernel can be used to capture this structure, then it is possible to extrapolate, with the predictions not forced to a zero mean prior away from training data. If not possible to represent as a non-stationary kernel, without a mean function, predictions should only be attempted in regions of the input space that are within the range of a single lengthscale(or equivalent), which will largely be interpolation. Often, inclusion of physical structure will offer an improved interpolative performance, particularly where the training data is *sparsely* sampled[2], such as an under-sampled time signal. In this scenario, with a model structure that is representative of the dynamic behaviour of the oscillating system, then it is possible to recover the signal with less frequently sampled training data, and allows one to up-sample the resulting time signal. Figure 1 summarises the above discussion, and illustrates the workflow for embedding partial knowledge (the branch marked b) into processes described by the product of two constituent functions. Note that there are two other levels of knowledge that may be held, as indicated on the diagram with a) and c) headings. These cases will not be considered

[1] Although notation implies a 1-dimensional regression, it is possible to extend to the M-dimensional case.

[2] Note the difference between coverage and sparsity - coverage refers to how much of the total input space is covered by the possible regions of the total input space, whilst sparsity relates to how dense the training samples are from a particular region of the input space.

in this paper.

CASE STUDIES

To investigate inclusion of partial knowledge, two case studies will be considered in this paper; the first, a simulated example consisting of two lower level functions, where for one function, structure is available but with limited data coverage, whilst for the other, no domain knowledge, but good data coverage. The second example is learning the decoupled response of a vibrating beam subject to external loading such that unknown mode shape functions can be recovered.

For the first example, we consider the mapping of f_t, defined as the following,

$$f_t = f_1 f_2, \tag{6}$$

where,

$$f_1 = 5\cos^2(\mathbf{x}_1), \tag{7}$$
$$f_2 = 2\mathbf{x}_2^2. \tag{8}$$

The mapping of the combined function is plotted in Figure 2. We will imagine the scenario where one has knowledge of the process f_2 in that it is a polynomial process of order two. To model f_2, a GP with a polynomial kernel of order 2 can therefore be selected, which has the form,

$$k_{poly} = (c_{poly} + \sigma^2_{f,poly}\mathbf{x}\mathbf{x}')^2. \tag{9}$$

f_1 will be unknown and so will be modelled by a black box kernel, which in this example is a squared exponential (equation (4)). The resulting final model for f_t is thus assumed to be,

$$f_t \sim \mathcal{GP}(0, k_{SE} \cdot k_{poly}) + \sigma^2_{n,t}, \tag{10}$$

with σ_f and l denoting the respective signal variance and lengthscale hyperparameters, with σ_n the variance of the noise model.

For the second case study, the interest is in predicting the response of a randomly loaded beam. From linear modal analysis, it is possible to construct the response of a linear system δ as the sum of the contribution of all individual modes δ_r, which for a discrete number of modes R can be written as,

$$\delta(x,t) = \sum_{r=1}^{R} \delta_r(x,t). \tag{11}$$

Given that contribution of each mode can be decomposed the mode shape function Φ_r and the temporal response, T_r, we arrive at an alternative formulation of equation (11),

$$\delta(x,t) = \sum_{r=1}^{R} \delta_r(x,t) = \sum_{r=1}^{R} \Phi_r(x) T_r(t). \tag{12}$$

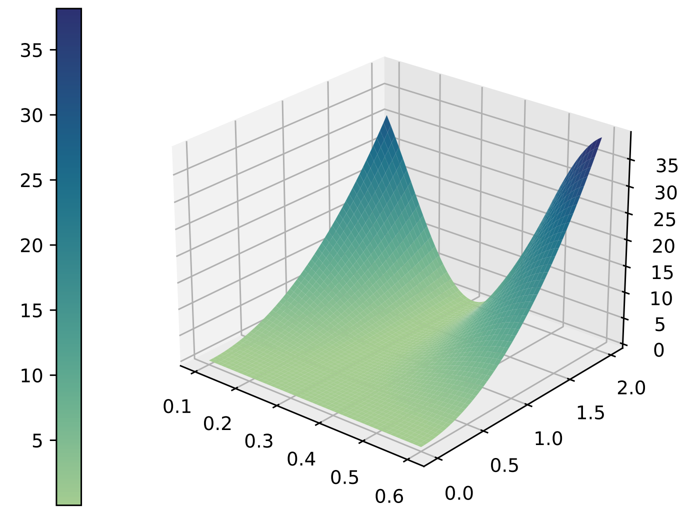


Figure 2: Mapping of equation (6). Left side axis is $\mathbf{x}_1$, right side axis is $\mathbf{x}_2$.

It can be seen that equation (12) is an example of a spatial temporal function where the spatial and temporal components are independent of one another. In this example, we consider the case where we have knowledge of the covariance of the temporal response, but not of the mode shapes, which is often the case for operational structures as discussed previously. To model the temporal response, the SDOF covariance as derived in [16] will be used, which is expressed as,

$$k_{SDOF} = \frac{\sigma_{force}}{4m^2\zeta\omega_n^3} \exp^{-\zeta\omega_n|\tau|} \left(\cos(\omega_d \tau) + \frac{\zeta\omega_n}{\omega_d} \sin(\omega_d|\tau|) \right) \tag{13}$$

m is the equivalent SDOF mass, ω_n and ω_d the natural and damped natural frequency, ζ the damping ratio, σ_{force} the variance of the forcing and τ equal to $||\mathbf{x} - \mathbf{x}'||$. For the spatial component, as the modeshapes are assumed unknown but with representative data available, a black box kernel will be used, which is chosen to be the SE kernel. This results in an assumed model for the rth mode of,

$$\delta_r(x, t) \sim \mathcal{GP}(0, k_{SE,r} \cdot k_{SDOF,r}) + \sigma_{n,r}^2 \tag{14}$$

To consider the combination of R modes, a sum can be formed over the product kernel such

$$\delta(\mathrm{x}, t) \sim \mathcal{GP}(0, \sum_{r=1}^{R} k_{SE,r} \cdot k_{SDOF,r}). \tag{15}$$

To generate data for this system, the response of a fixed-free beam under random forcing is simulated at a sampling rate of 8192Hz for one second. The forcing profile can be seen in Figure 3, with the corresponding response shown in Figure 4. Although in this case, there are closed form analytical expressions for the mode shape functions, the purpose

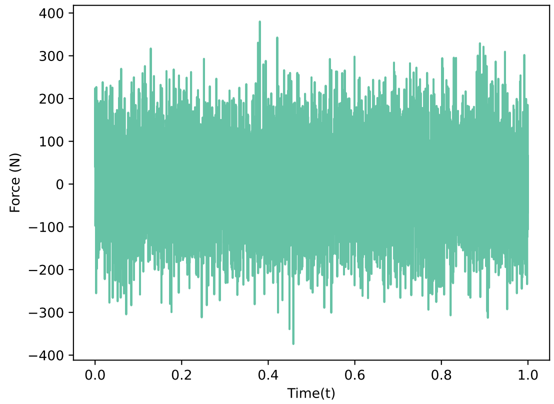


Figure 3: Input forcing to the system.

of the case study is to demonstrate how partial knowledge may be embedded into kernel structures, which subsequently can be extended for more challenging systems such as those with imperfect boundary conditions and other challenges of non-idealised systems.

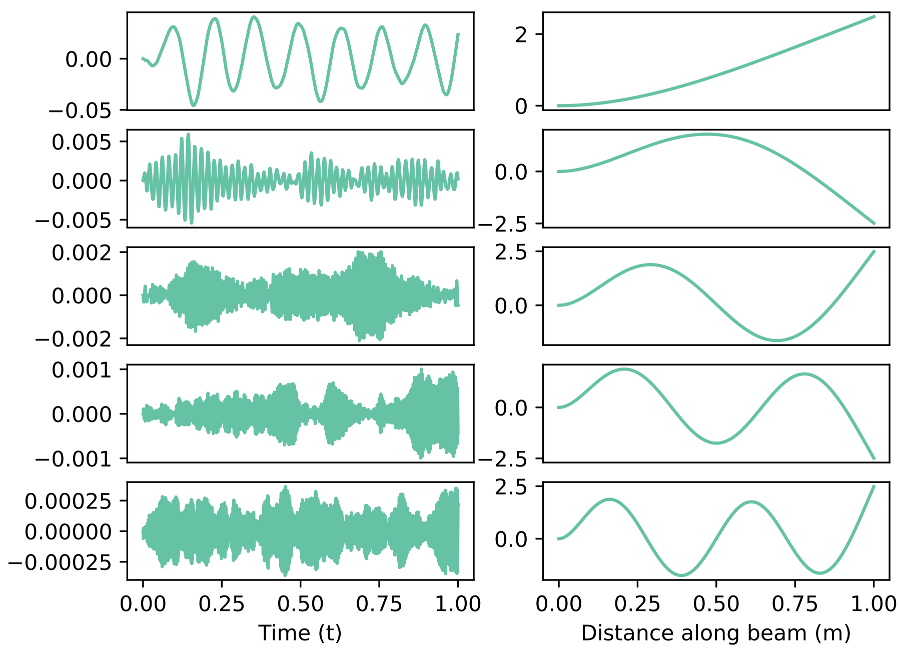
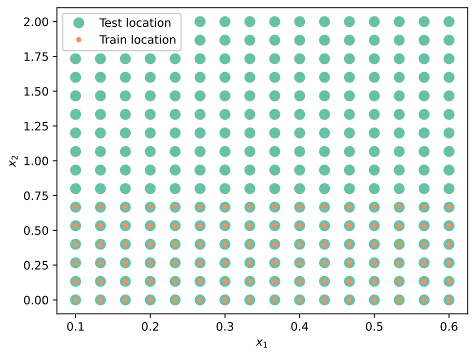
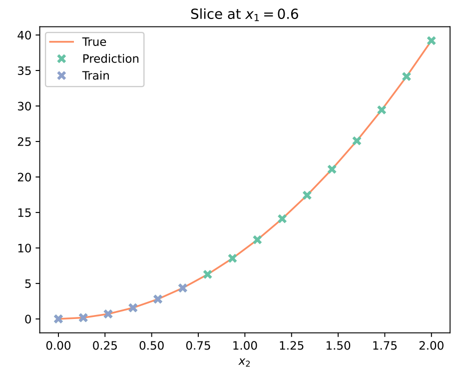


Figure 4: Decomposed temporal response of first five modeshapes (left) and the associated modeshapes (right) for the simulated beam.



(a) Training and testing data locations for case study 1.



(b) Predictions on the test set across a 1d slice at $\mathrm{x}_1 = 0.6$.

Figure 5: Training/testing regime for case study 1 and corresponding predictions at slice of x_1.

RESULTS AND DISCUSSIONS

For the first case study, a training set of half of the total input space is considered, with a testing set of the full simulated input space, as shown in Figure 5a. The training set is representative of having limited data coverage for f_2, for which physical structure is encoded through the kernel, whilst having full data for process f_1, where no domain insight is available. Plotting the predictions along a slice of the inputs at $\mathrm{x}_1 = 0.6$ in Figure 5b, it can be seen that the GP model is able to accurately predict the mapping, including in regions of extrapolation. Again, this is a result of it being possible that the physical insight possessed is translatable to a non-stationary kernel.

Having demonstrated through a simple problem how partial knowledge can be exploited in kernels in a product of functions setting, let us now move on to a task with direct engineering relevance in learning a vibration response decomposition. For the purposes of demonstration, only the first mode is considered. However, similar to the free vibration setting as demonstrated in [17], it is possible to extend the kernel to be a sum over multiple modes. Should identifiability issues arise, one strategy is to make informed selections of the temporal kernel hyperparameters, such as fixing the natural frequency or damping ratio of a given mode. In the work in this paper, w_n in the temporal kernel is set at the corresponding natural frequency, which would be possible to obtain from a variety of identification techniques such as stochastic subspace identification [19]. If one did not have confidence in fixing at a determinant value, a Bayesian treatment of the hyperparameters may be pursued, placing a prior that reflects the confidence a modeller has, over the parameter.

Training on the response of the first mode at 10 evenly spaced location at every 128th time point (every 1/64s), model performance is evaluated at every 32nd time point (1/256s) at the same spatial locations. The decomposed predictions on the test set are shown in Figure 6. It can be seen that the model has been able to correctly recover each component of the decomposition, notably at a much higher sample rate than the model was trained on.

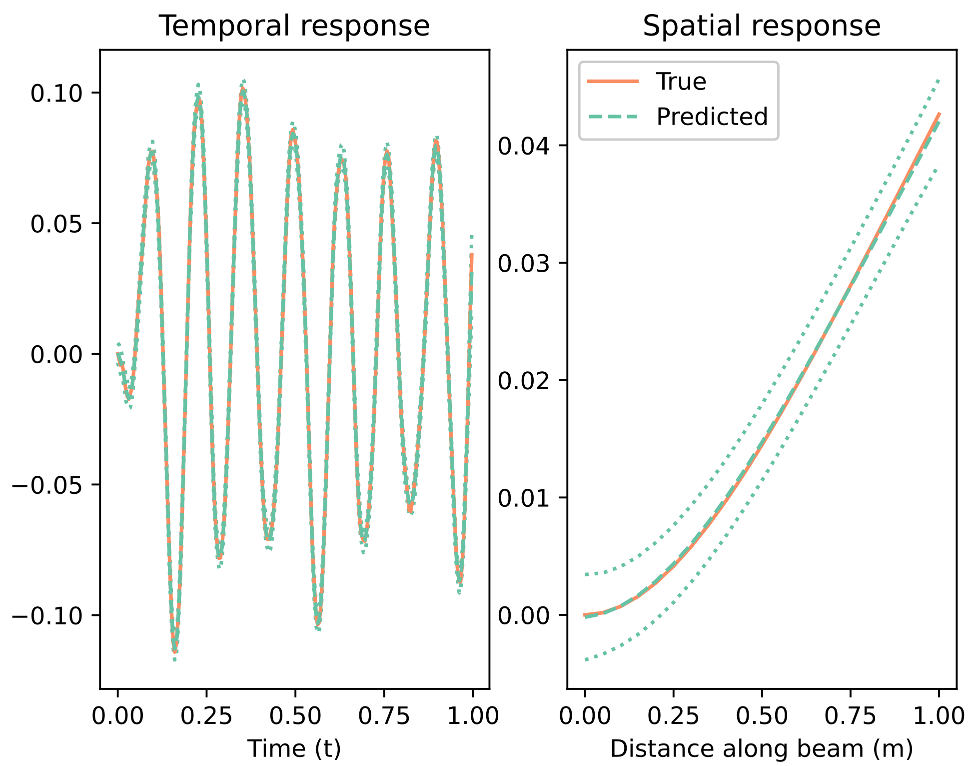


Figure 6: Decomposition of first mode contribution for cantilever beam response. Uncertainty on predictions indicated by dashed lines, quantified as 3 standard deviations from the output of the GP.

CONCLUDING REMARKS

In this paper, it has been investigated how product combinations of kernels can be utilised for embedding partial knowledge into systems that can be modelled as products of constituent functions. Consideration has been given to how the properties of the kernel with physical structure impact the predictive regimes that the model may be expected to perform well in; more specifically, if the kernel is stationary or non-stationary. It was demonstrated that when available physical insight can be represented as a non-stationary kernel, the model may be used in extrapolation, providing data is available for other constituent functions for which no domain insight is available. In the case where knowledge is captured with a stationary kernel, an improved interpolative performance can be expected, for which benefits were shown in both up-sampling a time signal, as well as for learning the decoupled response of a vibrating beam into the corresponding mode shape and and temporal response.

ACKNOWLEDGEMENTS

The authors wish to gratefully acknowledge the support of the UK Engineering & Physical Sciences Research Council, through grant EP/S001565/1.

REFERENCES

1. Farrar, C. R. and K. Worden. 2012. *Structural Health Monitoring: A Machine Learning Perspective*, John Wiley & Sons.
2. Park, G., A. C. Rutherford, H. Sohn, and C. R. Farrar. 2005. "An outlier analysis framework for impedance-based structural health monitoring," *Journal of Sound and Vibration*, 286(1-2):229–250.
3. Dervilis, N., M. Choi, S. Taylor, R. Barthorpe, G. Park, C. Farrar, and K. Worden. 2014. "On damage diagnosis for a wind turbine blade using pattern recognition," *Journal of sound and vibration*, 333(6):1833–1850.
4. Jones, M. R., T. J. Rogers, K. Worden, and E. J. Cross. 2022. "A Bayesian methodology for localising acoustic emission sources in complex structures," *Mechanical Systems and Signal Processing*, 163:108143.
5. Haywood-Alexander, M., N. Dervilis, K. Worden, E. J. Cross, R. S. Mills, and T. J. Rogers. 2021. "Structured machine learning tools for modelling characteristics of guided waves," *Mechanical Systems and Signal Processing*, 156:107628.
6. Willard, J., X. Jia, S. Xu, M. Steinbach, and V. Kumar. 2020. "Integrating physics-based modeling with machine learning: A survey," *arXiv preprint arXiv:2003.04919*, 1(1):1–34.
7. Karniadakis, G. E., I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. 2021. "Physics-informed machine learning," *Nature Reviews Physics*, 3(6):422–440.
8. Von Rueden, L., S. Mayer, K. Beckh, B. Georgiev, S. Giesselbach, R. Heese, B. Kirsch, J. Pfrommer, A. Pick, R. Ramamurthy, et al. 2021. "Informed Machine Learning–A taxonomy and survey of integrating prior knowledge into learning systems," *IEEE Transactions on Knowledge and Data Engineering*, 35(1):614–633.
9. Cross, E. J., S. Gibson, M. Jones, D. Pitchforth, S. Zhang, and T. Rogers. 2022. "Physics-Informed Machine Learning for Structural Health Monitoring," in *Structural Health Monitoring Based on Data Science Techniques*, Springer, pp. 347–367.
10. Pitchforth, D. J., T. J. Rogers, U. T. Tygesen, and E. J. Cross. 2021. "Grey-box models for wave loading prediction," *Mechanical Systems and Signal Processing*, 159:107741.
11. Jones, M. R., T. J. Rogers, and E. J. Cross. 2023. "Constraining Gaussian processes for physics-informed acoustic emission mapping," *Mechanical Systems and Signal Processing*, 188:109984.
12. Zhang, S., T. J. Rogers, and E. J. Cross. 2021. "Gaussian process based grey-box modelling for SHM of structures under fluctuating environmental conditions," in *European Workshop on Structural Health Monitoring: Special Collection of 2020 Papers-Volume 2*, Springer, pp. 55–66.
13. Conti, Z. X., R. Choudhary, and L. Magri. 2023. "A physics-based domain adaptation framework for modeling and forecasting building energy systems," *Data-Centric Engineering*, 4:e10.
14. Williams, C. K. and C. E. Rasmussen. 2006. *Gaussian Processes for Machine Learning*, vol. 2, MIT press Cambridge, MA.
15. Jones, M. R., T. J. Rogers, and E. J. Cross. 2022. "Physical Covariance Functions for Dynamic Systems with Time-Dependent Parameters," in *European Workshop on Structural Health Monitoring: EWSHM 2022-Volume 3*, Springer, pp. 381–391.

16. Cross, E. J. and T. J. Rogers. 2021. “Physics-derived covariance functions for machine learning in structural dynamics,” *IFAC-PapersOnLine*, 54(7):168–173.

17. Pitchforth, D., T. Rogers, U. Tygesen, and E. Cross. 2022. “Incorporation of partial physical knowledge within Gaussian processes,” in *Proceedings of the 30th international conference on noise and vibration engineering (ISMA 2022)*.

18. Hofmann, T., B. Schölkopf, and A. J. Smola. 2008. *Kernel methods in machine learning*.

19. Peeters, B. and G. De Roeck. 1999. “Reference-based stochastic subspace identification for output-only modal analysis,” *Mechanical systems and signal processing*, 13(6):855–878.