Proceedings Paper:

Jones, M.R., Pitchforth, D.J. and Cross, E.J. (2024) Gaussian process kernels for partial physical insight. In: Proceedings of the 10th European Workshop on Structural Health Monitoring (EWSHM 2024). 10th European Workshop on Structural Health Monitoring (EWSHM 2024), 10-13 Jun 2024, Potsdam, Germany. e-Journal of Nondestructive Testing, 29 (7). NDT.net GmbH & Co. KG

https://doi.org/10.58286/29859

Gaussian process priors for partial physical insight

Matthew R. JONES[1], Daniel J. PITCHFORTH[1] and Elizabeth J. CROSS[1]

University of Sheffield, Department of Mechanical Engineering, Mappin Street, Sheffield, UK, S1 3JD, matthew.r.jones@sheffield.ac.uk

Abstract. Fusing known physics into data-driven learners allows modelling practitioners to com-bine the expressive power of traditional machine learning with known mechanistic laws, where the objective is to enhance predictive performance, interpretability, and model generalisation. A core consideration that must be made when implementing a physics-informed learning architecture is how relevant knowledge will be embedded into the model structure, which, generally, is informed by the type of physics that is available. Frequently this knowledge may not be complete, with only a partial understanding of the governing physics available. In the partial knowledge setting, an important decision when constructing a model is how the interaction between known and unknown behaviours is modelled. Current practice will routinely impose assumptions on this interaction without consid-eration as to whether these assumptions are actually reflective of the system under investigation.

In this work, we will investigate the influence of the use of models that directly account for the interaction between known and unknown behaviour, following an explicit derivation path with re-spect to the partial knowledge available. It will be demonstrated how one may pursue differing but mathematically equivalent derivation paths, which are then compared to examples of what can be considered common practice when physics is embedded into data-driven learning, where the true manner in which the partial knowledge interacts with the unknown phenomena is often disregarded.

Keywords: Partial knowledge · Physics-informed learning · Derived kernels.

1 Introduction

Over the last several years, physics-informed machine learning has been increasingly adopted in an attempt to overcome the shortcomings of data-driven modelling [1,2,3]. Structural health monitoring [4] is one field that has benefited from the fusion of machine learning and domain knowledge, offering potential solutions to some of the most significant challenges in the practical deployment of intelligent asset management [3,5,6,7,8]. Some examples include lessening demands for extensive training data sets [9], as well as providing a guarantee that any predictions made adhere to underlying mechanistic laws, and thus are physically plausible [10,11], increasing the trust of operators.

When attempting to embed known physics into a statistical model, the likely scenario is that the knowledge held will not be complete, only representing *some* of the total behaviour that one is interested in modelling. Domain insight of this form may be categorised as *partial knowledge*. For instance, one may hold a linearised approximation of the system, but be unable to specify more complex, unknown physics. Alternatively, where behaviour arises across multiple physical axes, such as for spatial-temporal phenomena, it is often the case that only insight relating to either the spatial or temporal contribution is held, and not the combination of the two. Phrased in a more general sense, it is known that the

NDTnet Media and Publishing Partner
https://doi.org/10.58286/29859

overall behaviour of interest is a combination of several components, but only a subset of these individual components can be described by understood physics. Such a scenario then poses the question; how should one appropriately embed this partial knowledge into their model? In the absence of a compelling argument otherwise, a common strategy is to form a model that is the summation of the known and unknown components [3], where an expression that represents understood domain knowledge accounts for the known physics, with a flexible, data-driven component capturing any remaining structure,

$$Z(\mathbf{x},\mathbf{y}) = \underbrace{A(\mathbf{x})}_{known} + \underbrace{B(\mathbf{y})}_{unknown} \tag{1}$$

Whilst there are many examples of physics-informed models constructed as equation (1) [12,3,13], there are scenarios where partial knowledge is held in which the interaction between the known and unknown physics is not a simple summation between two components, whether it be some other type of operation, or it is unknown. In such a case, a model of the form of equation (1) will not be consistent with a model that is derived from the form of the interaction between the known and unknown physics, and so will impose structure that is not consistent with the actual driving physics. However, what is unanswered is the degree to which this assumption impacts the predictive performance of the model, particularly in the case where the assumption is incorrect.

In this paper, we will investigate the influence that following derivation paths for modelling known and unknown physics has on predictive performance, and comparing this to what the authors view as current, common practice for embedding domain structure into models. To construct all models in this paper, we utilise a Gaussian process (GP) framework, which enables the formulation of models that follow an explicit derivation path for the interaction between known and unknown behaviours. With GPs fully defined by a mean and covariance function, it is through the specification of these two functions that a route for appropriately representing the nature of the partial knowledge is available. Taking the general definition of both of these quantities,

$$\mu_f(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \tag{2}$$

$$k_f(\mathbf{x},\mathbf{x}') = \mathbb{E}[f(\mathbf{x})f(\mathbf{x}')] - \mathbb{E}[f(\mathbf{x})]\mathbb{E}[f(\mathbf{x}')] \tag{3}$$

these two expressions can be utilised to construct a GP that follows a derivation path from the nature of the partial knowledge available.

The first case study presented considers the influence of taking different, but equivalent derivation paths for a partially understood system. The performance of the derivation paths will be compared to one another, as well as to the model that represents what would likely be used as a 'first attempt', given some partial knowledge. The second case study will investigate the effects of incorrectly capturing the interaction between the known and the unknown components, and how this compares to the derived path. All results will be compared to a purely data-driven, black-box model.

2 Model derivation paths for partial knowledge

A convenient framework in which models can be derived to represent known physics are Gaussian processes, a type of Bayesian non-parametric regression. As introduced in the previous section, to fully define a GP requires the specification of a mean and covariance function, through which we can reflect domain knowledge. This paper assumes a basic knowledge of GPs - [14] can be consulted should the reader require a more detailed introduction.

To motivate the problem of partial knowledge, consider a function Z that is some combination of two constituent functions A and B,

$$Z(\mathbf{x},\mathbf{y}) = f\{A(\mathbf{x}), B(\mathbf{y})\} \tag{4}$$

A common mechanism for partial knowledge is for prior insight of A to be available, whilst B is completely unknown. The assumption is then made that A and B are independent Gaussian processes, where $A \sim \mathcal{GP}(\mu_A, K_A)$ and $B \sim \mathcal{GP}(\mu_B, K_B)$. In the absence of knowledge of B, we will set $\mu_B = 0$ (and so consequently $K_B = \phi_B$), and make use of a standard, flexible kernel, which here, is the squared-exponential (SE) function ([14]).

Depending on the exact form of the expression for Z, an appropriate derivation path for a GP over Z can be specified. The remainder of this section will be spent deriving an associated GP over Z such that $Z \sim \mathcal{GP}(\mu_Z, K_Z)$, depending upon the way in which the constituent functions are combined with one another, or in other words, the operator which acts between the two.

Additive The first case we will consider is an additive relationship between a known and unknown component,

$$Z = A + B \tag{5}$$

An interaction of this form is perhaps the simplest way in which insight can be built into a machine learner, with many examples arising in structural health monitoring. For instance, if attempting the model the displacement of a bridge, then it is likely there will be several trends such as wind loading, traffic loading and temperature all additively combining to give a total deck displacement. For a summation, with $\mu_B = 0$, the mean of Z is simply,

$$\mathbb{E}[Z] = \mathbb{E}[A] + \mathbb{E}[B] = \mu_A \tag{6}$$

For the covariance, one first has to make use of the fact that the autocorrelation of Z, $\phi_{ZZ'}$, can be expressed as,

$$\phi_{ZZ'} = \mathbb{E}[ZZ'] = \phi_{AA'} + \phi_{AB'} + \phi_{BA'} + \phi_{BB'} \tag{7}$$

which in combination with equation (3) returns,

$$K_{ZZ'} = K_{AA'} + \phi_{BB'} \tag{8}$$

which simply states that the autocovariance of Z is the sum of the individual autocovariance contributions (noting this is a direct consequence of the independence assumption placed between A and B). This leads to a final derived model of the form,

$$Z \sim \mathcal{GP}(\mu_A, K_{AA'} + \phi_{BB'}) \tag{9}$$

An alternative derivation path can be sought by making the prior assumption that $\mathbb{E}[Z] = 0$, which returns the covariance,

$$K_{ZZ'} = K_{AA'} + \phi_{BB'} \tag{10}$$

and an alternative derived model,

$$Z \sim \mathcal{GP}(0, K_{AA'} + \phi_{BB'}) \tag{11}$$

The above expressions can also be used in the case where A is treated deterministically, where equation (9) transforms to [15],

$$Z \sim \mathcal{GP}(A, \phi_{BB'}) \tag{12}$$

Or following the route of equation (11), a model expressed purely through the covariance,

$$Z \sim \mathcal{GP}(0, AA' + \phi_{BB'}) \tag{13}$$

This provides four different derivation paths where the interaction between known and unknown physics is additive - equations (9) (prior knowledge is stochastic) and (12) (prior knowledge is deterministic) to account for prior knowledge in the mean function, and equations (11) (stochastic) and (13) (deterministic) to account for prior knowledge in the covariance. These models are summarised in Table 1.

Table 1. Summary of derived models for additive combination of known and unknown behaviour.

Model	Assumptions made	Zero mean	Model form
1A	stochastic known and unknown	n	$\mathcal{GP}(\mu_A, K_{AA'} + \phi_{BB'})$
1B	stochastic known and unknown	y	$\mathcal{GP}(0, K_{AA'} + \phi_{BB'})$
1C	deterministic known, stochastic unknown	n	$\mathcal{GP}(A, \phi_{BB'})$
1D	deterministic known, stochastic unknown	y	$\mathcal{GP}(0, AA' + \phi_{BB'})$

Products Another common type of interaction that may be encountered between known and unknown components is a product, and defines the second case of interaction that will be considered,

$$Z = A \cdot B \tag{14}$$

There are many examples of this interaction in engineering, such as the dynamic response of a structure, where a spatial component multiplied by a temporal response gives the total response of a vibrating system. Under the earlier assumption of independence between A and B, the mean of Z is,

$$\mathbb{E}[Z] = \mathbb{E}[A] \cdot \mathbb{E}[B] = 0 \tag{15}$$

due to setting the prior mean over μ_B to be zero. To be able to derive the autocovariance for Z, one can make use of the general expression for the covariance of a product of two functions [16],

$$K_{ZZ'} = \mu_A\mu_{A'}K_{BB'} + \mu_A\mu_{B'}K_{BA'} + \mu_B\mu_{A'}K_{AB'} + \mu_B\mu_{B'}K_{AA'} + K_{AA'}K_{BB'} + K_{AB'}K_{BA'} \tag{16}$$

Accounting for independence and only retaining the means that are non-negative simplifies to,

$$K_{ZZ'} = \mu_A\mu_{A'}\phi_{BB'} + K_{AA'}\phi_{BB'} \tag{17}$$

The final derived model for Z arising in the product setting is therefore,

$$Z \sim \mathcal{GP}(0, \mu_A\mu_{A'}\phi_{BB'} + K_{AA'}\phi_{BB'}) \tag{18}$$

If the knowledge held about A is that it is deterministic, then the model reduces to,

$$Z \sim \mathcal{GP}(0, AA'\phi_{BB'}) \tag{19}$$

Table 2. Summary of derived models for multiplicative combination of known and unknown behaviour.

Model	Assumptions made	Zero mean	Model form
2A	stochastic known and unknown	y	$\mathcal{GP}(0, \mu_A\mu_{A'}\phi_{BB'} + K_{AA'}\phi_{BB'})$
2B	deterministic known, stochastic unknown	y	$\mathcal{GP}(0, AA'\phi_{BB'})$

3 Do derivation paths matter?

Having established the necessary mathematical machinery for deriving models to account for the interaction between our known and unknown physics, we can now advance to considering several studies which will allow us to investigate the efficacy of the various forms of models. Two case studies will be considered here. The first will investigate the impact of following differing, but mathematically equivalent, derivation paths. The second will consider the impact of wrongly assuming the interaction between the known and unknown physics, and comparing with the correct derivation path. For each case study,

$$Z(\mathbf{x}) = f\{A(\mathbf{x}), B(\mathbf{x})\} \tag{20}$$

therefore restricting equation (4) to a univariate case. It will be assumed that the form of A is known, with B unknown, with each component given by,

$$\begin{aligned} A &= \alpha \cos(15\mathbf{x})^3 & \alpha &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2) \\ B &= \mathbf{x}^2(\cos(10\mathbf{x})^2 + \sin(5\mathbf{x})) + \epsilon & \epsilon &\sim \mathcal{N}(0, \sigma_\epsilon^2) \end{aligned} \tag{21}$$

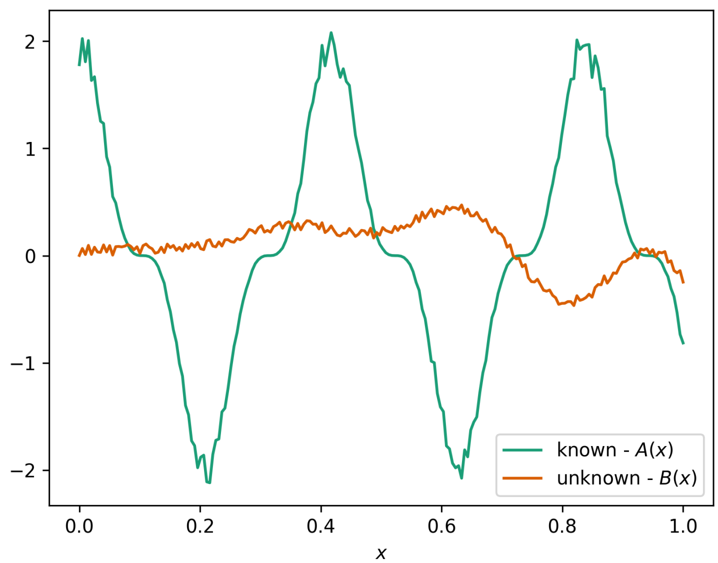

Fig. 1. Simulated functions A (assumed known) and B (assumed unknown).

The purpose of the use of these functions is to replicate the real-life scenario where it is often the case that a simpler trend is known (A), whilst other influencing behaviour, generally of a more complex nature, is unknown (B). The appearance of both of these functions is provided in Figure 1, with the function parameters set as $\mu_\alpha = 2, \sigma_\alpha = 0.1, \sigma_\epsilon = 0.1$.

3.1 Case study 1: comparing derivation paths

For the first case study, the effect of pursing differing derivation paths but for the same system will be investigated. As seen in the previous section, when the interaction between the known and unknown can be expressed as a summation, $Z(\mathbf{x}) = A(\mathbf{x}) + B(\mathbf{x})$, there arises two possible derivation paths, either through the mean function of the GP, or through the kernel function.

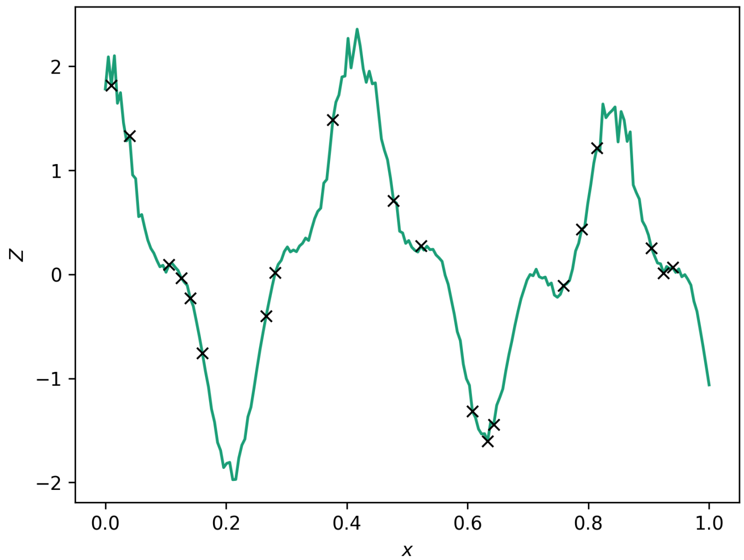

Fig. 2. Additive interaction between known and unknown components. Training points for all models given by black crosses.

As A is a stochastic process, these two derivation paths correspond to a model of the form $Z \sim \mathcal{GP}(\mu_A, K_{AA'} + \phi_{BB'})$ (1A, Table 1) and $Z \sim \mathcal{GP}(0, K_{AA'} + \phi_{BB'})$ (1B, Table 1). A third model is also formed, with the known component represented in the mean function, whilst a single, flexible, data-driven kernel accounts for the remainder of the behaviour - this corresponds to a model of the form $Z \sim \mathcal{GP}(A, \phi_{BB'})$ (1C, Table 1). This is likely the structure that a modelling practitioner would choose as a 'first-attempt' at embedding partial knowledge into a system without considering a specific derivation path. As such, it serves as a benchmark for the inclusion of physics in an ad-hoc manner, as is standard procedure for much physics-informed machine learning. Finally, a standard zero mean, squared exponential kernel is also included, serving as the example of a 'black box' model, $Z \sim \mathcal{GP}(0, K_{SE})$.

The function that is being learnt in this case study is illustrated in Figure 2. The training points that each GP model is conditioned on are given by the black crosses. The predictive results for each of these models are then plotted in Figure 3. The first observation that can be made is that both of the models that follow the derivation path (1A and 1B) offer an enhanced performance over those that do not, with model 1C consistently failing to capture the peaks and troughs of the simulated function. Additionally, the derived models offer tighter uncertainty bounds across much of the function space, which is likely due to their ability to directly capture the covariance contribution of A, instead of attempting to learn this in conjunction with B as is the case for the two models not following the derivation path. Overall, it appears as through the inability of model 1C to explicitly account for the stochastic nature of the embedded physics worsens performance compared to the derived models, and highlights the benefit of pursuing the derived model route. With regards to each of the derived models, it can be seen that model 1A and model 1B both perform comparably, with almost identical predictive means and confidence bounds. This is confirmed by the mean standardised log loss (MSLL) returned by moth models being similar, which is a metric that assesses how well the predictive distribution represents the true distribution. Whilst this particular example suggests there is no strong argument for pursuing a derivation through the mean or kernel route, this is something that the authors will be investigating further. Finally, perhaps as expected, the black-box models returns the least favourable performance.

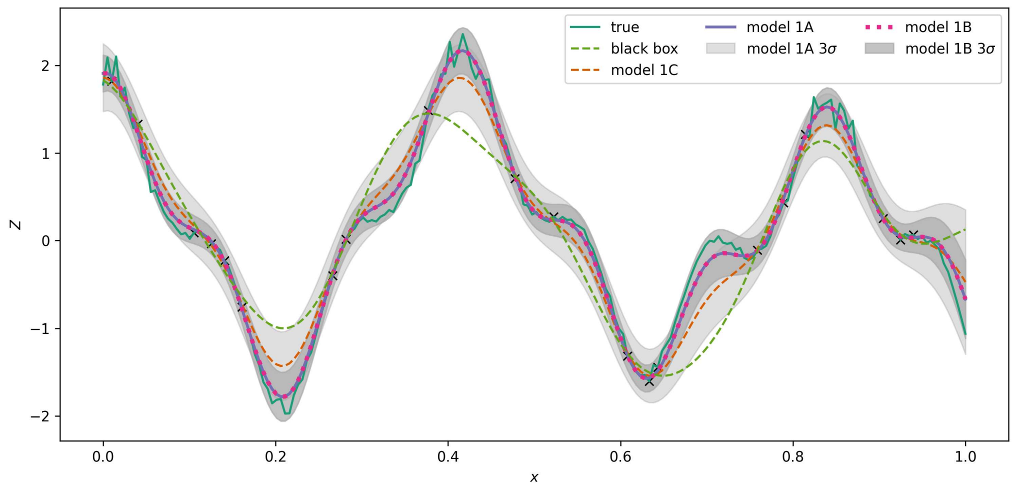

Fig. 3. Predictive performance on additive function for each form of model considered. MSLLs: 1A (-2.21), 1B (-2.27), 1C (-1.13), black box (-0.51). Note that the confidence bound for 1A are almost identical to that of 1B. The confidence bounds for 1B and the black box models have been omitted to ensure the figure is clear.

3.2 Case study 2: misspecified interaction between known and unknown behaviours

As a second case study, we will consider the result of misspecifying the interaction between the known and unknown behaviour. To motivate this example, we will consider the case where Z is a product of A and B, $Z(\mathbf{x}) = A(\mathbf{x}) \cdot B(\mathbf{x})$. In this setting, following the derivation path returns only one model option, $Z \sim \mathcal{GP}(0, \mu_A \mu_{A'} \phi_{BB'} + K_{AA'} \phi_{BB'})$ (2A, Table 2). The second model considered is that of a single product kernel, $Z \sim \mathcal{GP}(0, K_{AA'} \phi_{BB'})$, named 'product kernel', which is the approach a modeller would likely adopt given knowledge of the physics and unknown components interacting in such a way, but without taking the full derivation route.

A third model representative of the scenario in which one naively attempts to embed the held knowledge within the mean function of the model, $Z \sim \mathcal{GP}(A, \phi_{BB'})$ (1C, Table 1) is also considered, which incorrectly assumes the form of the interaction of the partial knowledge with the rest of the system. Finally, a black-box model is also included as in the previous example, $Z \sim \mathcal{GP}(0, K_{SE})$. The function being modelled in this case study is shown in Figure 4, with the training points highlighted by black crosses. The predictions of each of these models are plotted in Figure 5.

The first result that can be seen is that by following the derivation path, instead of naively placing the known component inside the mean function, the predictive performance is much improved, particularly away from training locations. What this result demonstrates is that if one is to incorrectly assume the

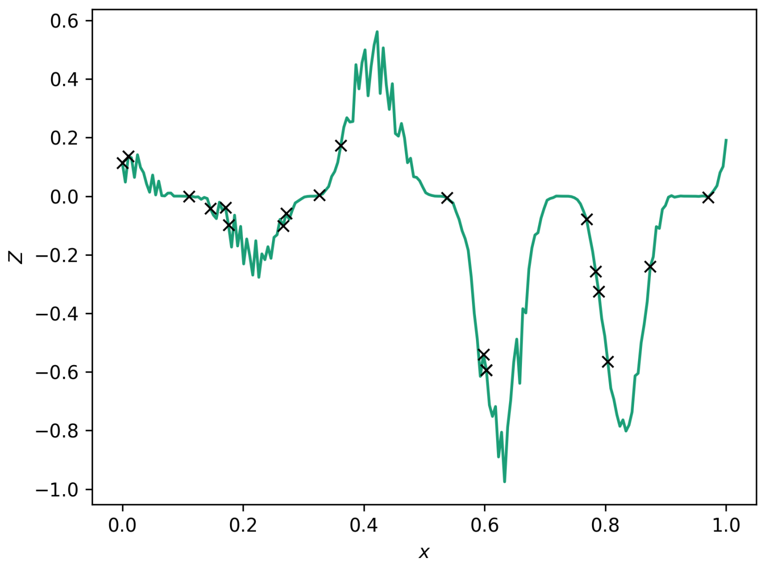

Fig. 4. Multiplicative interaction between known and unknown components. Training points for all models given by black crosses.

manner in which the partial knowledge enters the system, then this can significantly hinder the performance of the subsequent predictions. This observation is reinforced by the black-box model actually offering an improved performance over making an incorrect assumption and placing the partial knowledge in the mean function, demonstrating that wrongly including partial knowledge worsens performance than if one were to not include the physics at all. A second observation is that one does not gain an improved performance from following the derivation path versus a single product kernel. However, as with the previous examples, further investigation is needed across a range of case studies.

4 Conclusions and future work

In this paper, the framework in which partial physical knowledge of a system is embedded into a machine learner has been investigated. Comparisons were made between models derived from the form of the mathematical relationship between known and unknown behaviours and those that arise from current, standard practice in physics-informed machine learning. This comparison included how one may purse different, but equivalent derivation routes, and the subsequent effects this has on predictive performances. Additionally, the perils of incorrectly assuming both the form of the knowledge and the manner in which it interacts with the unknown behaviour were demonstrated.

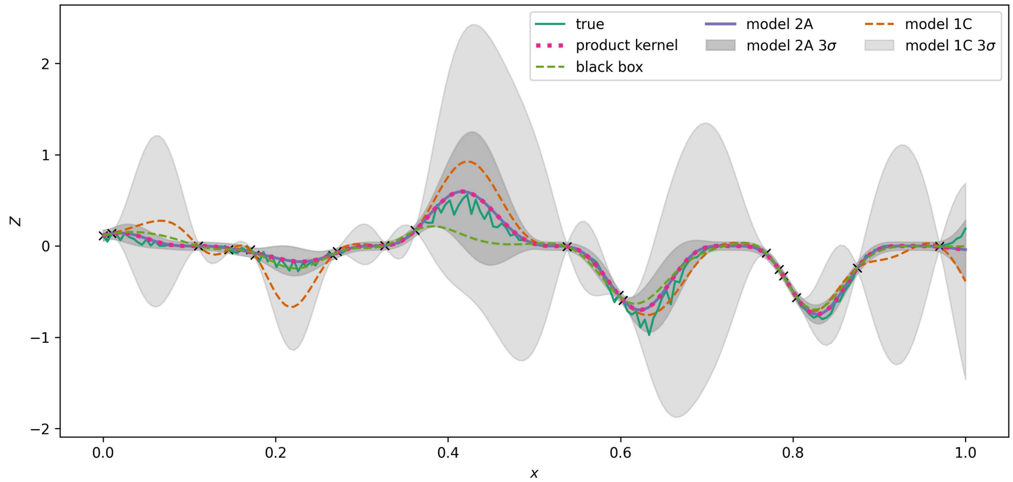

Fig. 5. Predictive performance on additive function for each form of model considered. MSLLs: 2A (-2.13), product kernel (-2.13), 1C (-0.53), black box (-1.42). Note that the confidence bounds for 2A are almost identical to that of the product kernel. The confidence bounds for the product kernel and black box models have been omitted to ensure the figure is clear.

Acknowledgements

The authors gratefully acknowledge the support of Safran Landing Systems, the University of Sheffield and Innovate UK through the OLLGA project grant 10040817.

References

[1] Willard, J., Jia, X., Xu, S., Steinbach, M., Kumar, V.: Integrating physics-based modeling with machine learning: A survey. arXiv preprint arXiv:2003.04919 **1**(1), 1–34 (2020)
[2] Von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Pfrommer, J., Pick, A., Ramamurthy, R., et al.: 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 (2021)
[3] Cross, E.J., Rogers, T.J., Pitchforth, D.J., Gibson, S.J., Zhang, S., Jones, M.R.: A spectrum of physics-informed Gaussian processes for regression in engineering. Data-Centric Engineering **5**, e8 (2024)
[4] Farrar, C.R., Worden, K.: An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **365**(1851), 303–315 (2007)
[5] Di Lorenzo, D., Champaney, V., Marzin, J., Farhat, C., Chinesta, F.: Physics informed and data-based augmented learning in structural health diagnosis. Computer Methods in Applied Mechanics and Engineering **414**, 116186 (2023)
[6] Lai, Z., Mylonas, C., Nagarajaiah, S., Chatzi, E.: Structural identification with physics-informed neural ordinary differential equations. Journal of Sound and Vibration **508**, 116196 (2021)
[7] Miele, S., Karve, P., Mahadevan, S.: Multi-fidelity physics-informed machine learning for probabilistic damage diagnosis. Reliability Engineering & System Safety **235**, 109243 (2023)

[8] Cross, E.J., Gibson, S.J., Jones, M.R., Pitchforth, D.J., Zhang, S., Rogers, T.J.: Physics-informed machine learning for structural health monitoring. Structural Health Monitoring Based on Data Science Techniques pp. 347–367 (2022)
[9] Jones, M.R., Rogers, T.J., Cross, E.J.: Constraining Gaussian processes for physics-informed acoustic emission mapping. Mechanical Systems and Signal Processing **188**, 109984 (2023)
[10] Al-Adly, A.I., Kripakaran, P.: Physics-informed neural networks for structural health monitoring: a case study for Kirchhoff–Love plates. Data-Centric Engineering **5**, e6 (2024)
[11] Mclean, J.H., Jones, M.R., O'Connell, B.J., Maguire, E., Rogers, T.J.: Physically meaningful uncertainty quantification in probabilistic wind turbine power curve models as a damage-sensitive feature. Structural Health Monitoring **22**(6), 3623–3636 (2023)
[12] Rogers, T.J., Holmes, G.R., Cross, E., Worden, K.: On a grey box modelling framework for nonlinear system identification. In: Special Topics in Structural Dynamics, Volume 6: Proceedings of the 35th IMAC, A Conference and Exposition on Structural Dynamics 2017. pp. 167–178. Springer (2017)
[13] Pitchforth, D.J., Rogers, T.J., Tygesen, U.T., Cross, E.J.: Grey-box models for wave loading prediction. Mechanical Systems and Signal Processing **159**, 107741 (2021)
[14] Williams, C.K., Rasmussen, C.E.: Gaussian Processes for Machine Learning, vol. 2. MIT press Cambridge, MA (2006)
[15] Kennedy, M.C., O'Hagan, A.: Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63**(3), 425–464 (2001)
[16] Bohrnstedt, G.W., Goldberger, A.S.: On the exact covariance of products of random variables. Journal of the American Statistical Association **64**(328), 1439–1442 (1969)