Proceedings Paper:
Pitchforth, D.J., Jones, M.R. and Cross, E.J. (2024) Physically-informed change-point kernels for variable levels of physical knowledge inclusion in Gaussian processes. In: Proceedings of the 10th European Workshop on Structural Health Monitoring (EWSHM 2024). 11th 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. ISSN: 1435-4934. EISSN: 1435-4934.

https://doi.org/10.58286/29750

EWSHM 2024
11th European Workshop on Structural Health Monitoring

Physically-informed change-point kernels for variable levels of physical knowledge inclusion in Gaussian processes

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

University of Sheffield, Department of Mechanical Engineering, Mappin Street, Sheffield, UK, S1 3JD, d.j.pitchforth@sheffield.ac.uk

Abstract. The relative balance between physics and data within any physics-informed machine learner is an important modelling consideration to ensure that the benefits of both physics and data-based approaches are maximised. An over reliance on physical knowledge can be detrimental, particularly when the physics-based component of a model may not accurately represent the true modelled system. An under utilisation of physical knowledge potentially wastes a valuable resource, along with benefits in model interpretability and reduced demand for expensive data collection. Although adjusting the relative levels of physics and data reliance within a model is possible through the adaptation of the model structure, in practice, this can be challenging, with the relative balance produced by new model structures not always clear before they are implemented. This paper presents a means of being able to tune the balance of physics and data reliance within a model through the development of physically-informed change-point kernels for Gaussian processes. These combine more structured physical kernels, capable of enforcing physically derived behaviours, with flexible, general purpose kernels, and provide means to dynamically change the relative levels of reliance on physics and data within a model.

Keywords: Physics-informed kernel design · Change-point kernels · Gaussian process.

Introduction

Physics-Informed Machine Learning (PIML) aims to exploit the benefits of both physics and data-based modelling approaches; insight, structure and enhanced ability to extrapolate are provided through physical knowledge, whilst a data-based component increases flexibility and allows for the capture of complex relationships directly from data. In many implementations within the literature [1–3], physics is used to represent an aspect of behaviour that is better understood (e.g. a linear behaviours, general trends), whilst the data-based component tackles friction, turbulence, non-linearities or other more challenging effects. When working effectively, PIML models often outperform the individual physics and data-based models from which they are constructed.

Three of the major decisions within a PIML model are the type of physics to be included, the selection of a data-based component, and how they are integrated together to create a single model. The latter of these decisions is the focus of this paper and is key factor effecting the relative balance between physics and data within the model. Earlier work of the authors [4] discusses the relationship between how the structure of Gaussian process models affects this balance and how it may change across modelling tasks.

This work looks toward how one might be able to tune the relative reliance on physics and data within physics-informed Gaussian Process (GP) models. A GP is a flexible, non-parametric, Bayesian technique, adept within a variety of engineering regression problems, including crack growth [5, 6], tool wear [7, 8] and modelling of wind turbine power curves [9, 10]. For conciseness, an introduction to GP regression theory is not presented here, with interested readers encouraged to consult [11]. The covariance function (kernel) of a Gaussian process is responsible for the family of functions from which predictions may be drawn, with commonly used kernels enforcing properties such as smoothly varying functions, periodicity and localised behaviours [11, 12]. Through careful design of the kernel, it is possible to mimic physically desirable behaviours within drawn functions; such examples include the representation of a physical process [13, 14] or the enforcement of axial and rotational symmetries [15–18]. Here, a combination of physics-informed kernels and flexible, more general purpose, kernels are used in combination. The constructed physics-informed change-point kernels aim to dynamically switch the relative reliance between kernel components, allowing for variation of the physics-data balance within the model.

To highlight an engineering scenario with a changing reliance on a physical model, a case study is presented of wind loading of the Tamar bridge. A section of a dataset measured on the structure is used to investigate lift forces produced by high speed winds. Importantly, these lift forces only occur when winds blow across the bridge (perpendicular to the bridge length), causing a dependency of this relationship on wind direction. The implementation of physics-informed change-point kernels is used to capture the dynamic reliance on lift force, outperforming a purely data-based approach, and providing insight in to how this relationship changes.

1 Change-point kernels

The selection of a kernel within a Gaussian process is an important modelling decision, determining the type of functions used to make predictions and, as a result, the model's performance. However, there are many instances where the selection of a single 'best' kernel for a modelling task might be challenging, for example, a system under changing conditions or the introduction of new behaviours over time. In these circumstances one might wish to utilise the behaviour of multiple kernels, with an appropriate method of switching between them. Such a method could borrow from the many existing kinds of switching model, including Mixtures of Experts (MOEs) [19, 20], Markov-switching models [21], Treed Gaussian Processes (TGPs) [22] and Regime-switching cointegration models [23]. In this work, we build upon the change-point kernel [24, 25], which controls the switching between two kernels $K_1(X, X')$ and $K_2(X, X')$ through the use of the sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-a(x-x_0)}} \tag{1}$$

where a is a gradient term, responsible for how quickly the function switches from 0 to 1, and x_0 is the switching location. The use of a sigmoid function allows the gradual transition between multiple kernels without the introduction of a discontinuity. A useful property of sigmoid functions is that $\sigma(x) + \sigma(-x) = 1$ allowing a pair of opposing sigmoids to phase between the use of two kernels. A sigmoid with a negative gradient term may be used to phase out the use of a particular kernel, whilst a sigmoid with the equivalent positive gradient may be used to phase in the use of a new kernel. The covariance function for a sigmoid is expressed:

$$K_\sigma(X, X') = \sigma(X)\sigma(X')^T \tag{2}$$

Given that products and sums of kernels are also valid kernels, a pair of opposing sigmoid kernels may be used to switch between the use of two kernels $K_1(X, X')$ and $K_2(X, X')$. This is referred to as the change-point kernel [24]:

$$K(X, X') = K_\sigma(X, X')K_1(X, X') + K_\sigma(-X, -X')K_2(X, X') \tag{3}$$

where additionally to the hyperparameters of the kernels $K_1(X, X')$ and $K_2(X, X')$, a gradient term, a is introduced to control the direction and speed of the switch and the location, x_0 determines where the switch happens.

To highlight the operation of a change-point kernel, consider a case where one may wish to transition from a process that varies quickly, to one varying more slowly. This can be achieved through a combination of short and long lengthscale Squared Exponential (SE) kernels. Draws from a SE change-point kernel are shown in Figure 1 for illustration.

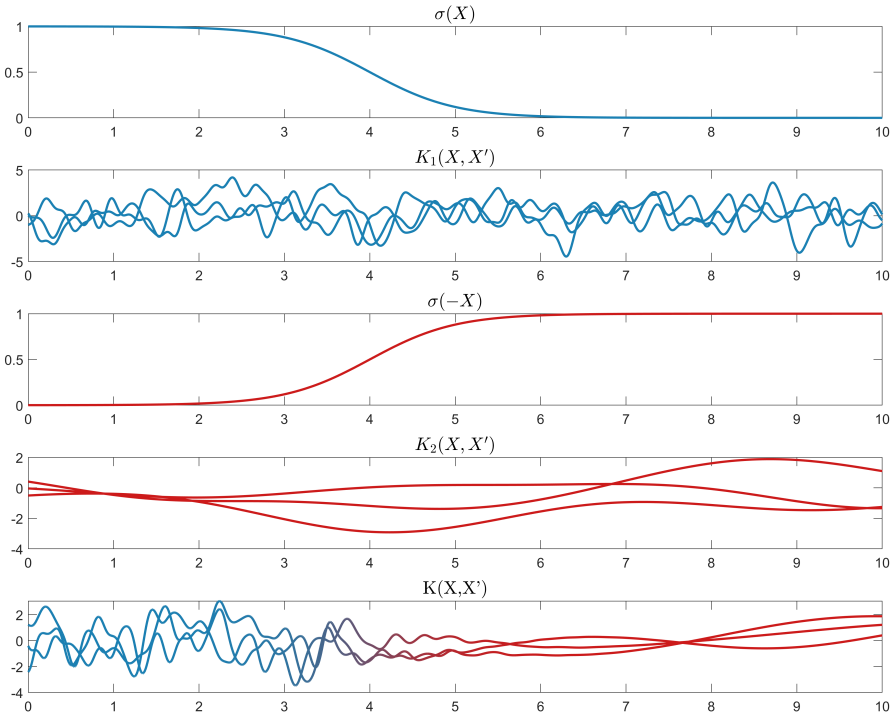

Fig. 1. Construction of a change-point kernel with $a = 2$ and $x_0 = 4$, transitioning between Squared Exponential kernels with short and long lengthscales. The sigmoid functions, $\sigma(X)$ and $\sigma(-X)$, draws from each component kernel, $K_1(X, X')$ and $K_2(X, X')$, along with draws from the combined kernel $K(X, X') = K_\sigma(X, X')K_1(X, X') + K_\sigma(-X, -X')K_2(X, X')$ are shown. The colour gradient of the plot reflects the relative weighting of each kernel; blue for $K_1(X, X')$, and red for $K_2(X, X')$.

1.1 Physics-informed change-point kernels

The framework of change-point kernels provides an effective means to vary the reliance between multiple kernels within a GP. To exploit this within a PIML setting, the integration of physics-informed kernels and flexible 'data-based' kernels is proposed alongside the option to switch on an input variable seperate from those used in the main GP. Here, a pair of sigmoid kernels $K_\sigma(Z, Z')$ and $K_\sigma(-Z, -Z')$ are used to control the explanatory power of the physics-informed kernel within the model. This leads to a kernel structure of:

$$K(Z, Z', X, X') = \underbrace{K_\sigma(Z, Z')K_{Phy}(X, X')}_{\text{Physics-informed kernel}} + \underbrace{K_\sigma(-Z, -Z')K_{Data}(X, X')}_{\text{Flexible kernel}} \tag{4}$$

where the newly introduced Z is the input to the sigmoid kernel, used to control the switching between kernels $K_{Phy}(X, X')$ and $K_{Data}(X, X')$. An effective choice of Z relates closely to how a physical relationship might change with a variable. For example, temperature, humidity, excitation level, or measures of turbulence can all effect the extent to which one might want to rely on a given piece of physical knowledge.

There are many reasons why one might want to change the relative reliance on physical knowledge within a model, an important one of which is the changing validity of a physics-based model. With any physical model, and particularly so with simple ones, assumptions must be made in order to represent the system of interest. The extent to which these assumptions hold effects the performance of the constructed model and care should be taken not to trust the results of models constructed upon invalid assumptions. Assigning a fixed degree of trust within a physical model component may be challenging when a model is required to operate over a range of conditions and allowing this to vary is therefore highly desirable.

The presence of regime-switching and localised behaviours provide alternative motivation to vary the reliance on physical knowledge. If a phenomena is known to occur in specific conditions, for example the dependency of vortex shedding on flow speed [26], one might wish to phase its occurrence in and out of a model. To investigate a use-case of this form, a case study of wind loading on the Tamar bridge will now be presented.

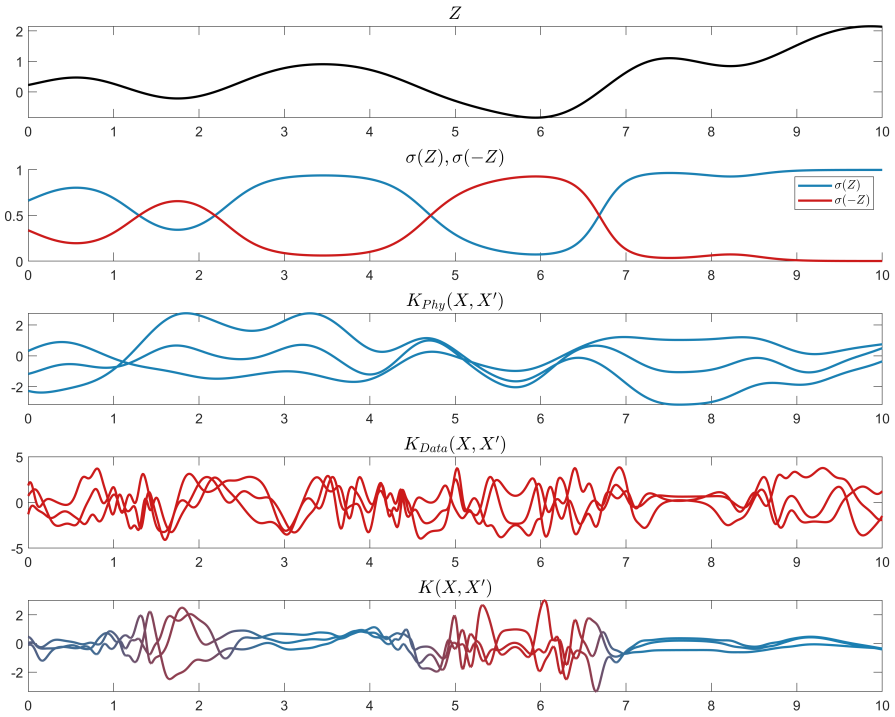

Fig. 2. Construction of a physics-informed change-point kernel, where the switching between a physics-informed kernel K_{Phy} and a flexible kernel K_{Data} is controlled by an external variable Z. Here $K_{Phy}(X, X')$ is a linear kernel, acting on an input space of $X = [\boldsymbol{x1}, \boldsymbol{x2}]$, representing knowledge of a linear process. This linear relationship is assumed not to hold for lower values of Z. The colour gradient of the plot reflects the relative weighting of each kernel; blue for $K_{Phy}(X, X')$, and red for $K_{Data}(X, X')$.

2 Directional wind loading of the Tamar bridge

The Tamar Suspension Bridge is located in the south west of the UK and was the feature of a monitoring campaign led by the Vibration Engineering Section (VES) at the University of Sheffield. The bridge is 643m long, with two towers, each 73m in height; it also forms part of a major connection to the city of Plymouth. For additional details on both the bridge and monitoring campaign, the reader is directed towards [27] and [28].

Available data from the bridge spans a three year period from 2007 to 2011, with access to measurements from accelerometers, strain gauges, annemometers, temperatures, humidity and traffic levels. Here, only a small section of this dataset (2500 points) is used from the summer of 2008, focussing on measurements of wind speed, wind direction and vertical deck acceleration. The aim of this case study is to investigate the presence of changing physical behaviours from a full scale engineering structure. The example chosen here is the dependency of response to wind load on wind direction.

Orientated in the east-west direction, the bridge is subject to higher deck accelerations when winds blow perpendicular to this at high speeds [27]. This is shown in plots of wind speed vs deck acceleration, seperated by wind

direction in Figure 3. The cause of this relationship is may be explained as lift force, incited by a pressure difference as wind flows over the bridge deck [29]. The expression for lift force is:

$$L = \frac{1}{2}\rho C_L A U^2 \tag{5}$$

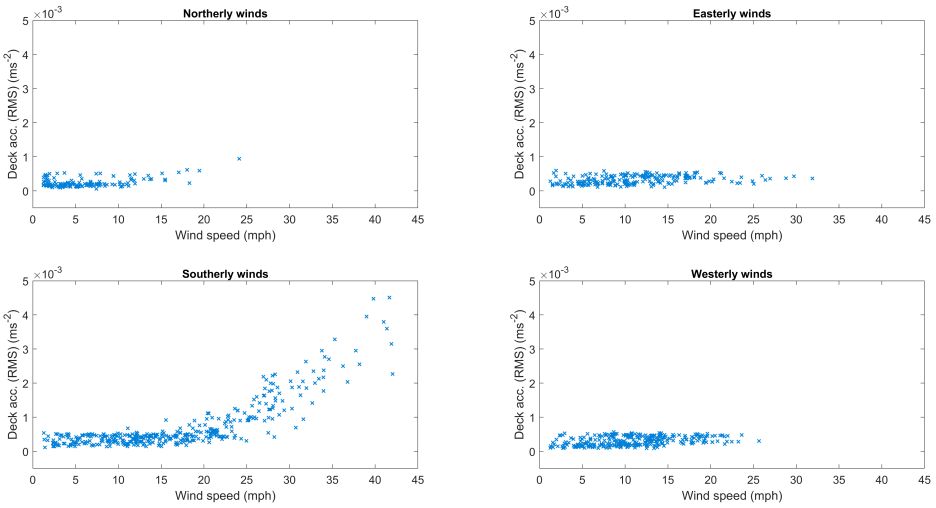

Fig. 3. Wind speed vs vertical deck acceleration for a section of the Tamar bridge dataset. Winds are separated by direction, into northerly winds ($0 \pm 20^\circ$), easterly winds ($90 \pm 20^\circ$), southerly winds ($180 \pm 20^\circ$) and westerly winds ($270 \pm 20^\circ$). The northerly and southerly winds blow sideways across the bridge and induce much higher deck accelerations, particulaly at high wind speeds.

where ρ is fluid density, C_L is lift coefficient, A is projected area and U is wind speed. The quadratic term in this expression aligns with the observations in Figure 3, particularly for southerly winds. To incorporate this knowledge within our model, a second order polynomial kernel is used to represent lift force:

$$K_{Lift}(U, U') = \sigma_L^2 (UU^T + c)^2 \tag{6}$$

where σ_L^2 scales the variance and c allows the introduction of lower order terms (linears and constants). Draws from this kernel will be quadratic, enforcing desirable structure within the predictions of the model. Using this kernel alone would not be sensible as this quadratic behaviour is clearly not visible for easterly and westerly winds. Instead, it is integrated within a larger change-point kernel structure:

$$\begin{aligned} K(\theta, \theta', U, U') = & \underbrace{K_\sigma(\cos(2\theta), \cos(2\theta')) K_\sigma(U, U') K_{Lift}(U, U')}_{\text{Introduce lift force at high N/S winds...}} \\ & + \underbrace{K_\sigma(-\cos(2\theta), -\cos(2\theta')) K_\sigma(-U, -U') K_{SE}(U, U')}_{\text{... otherwise, use a Squared Exponential kernel}} \end{aligned} \tag{7}$$

where θ is wind angle, U is wind speed, $K_\sigma(\cos(2\theta), \cos(2\theta'))$ is a sigmoid kernel for wind direction, $K_\sigma(U, U')$ is a sigmoid kernel for wind speed, $K_{Lift}(U, U')$ is the polynomial lift kernel and $K_{SE}(U, U')$ is a Squared Exponential (SE) kernel. The aim of this structure is to allow the knowledge of lift force to be utilised only for higher speed northerly and southerly winds. The term $\cos(2\theta)$ is used a measure of how well a wind direction aligns with north or south.

2.1 Results

To test the performance of models, a training set of 500 randomly selected points were used from the 2500 datapoint section of the dataset. The remaining unseen data was used as the test set. Models were optimised inline with standard GP practice [11], using the Negative Log Marginal Likelihood (NLML) as a cost function.

To allow the comparison of the proposed physics-informed change-points model with a benchmark case, a purely data-based model was also investigated. This took the form of a GP with a SE kernel, $K_{SE}(U, U', \theta, \theta')$, given access to the same model inputs of wind speed and wind direction. Predictions for this model are shown in Figure 4, with predictions from the change-point model shown in Figure 5.

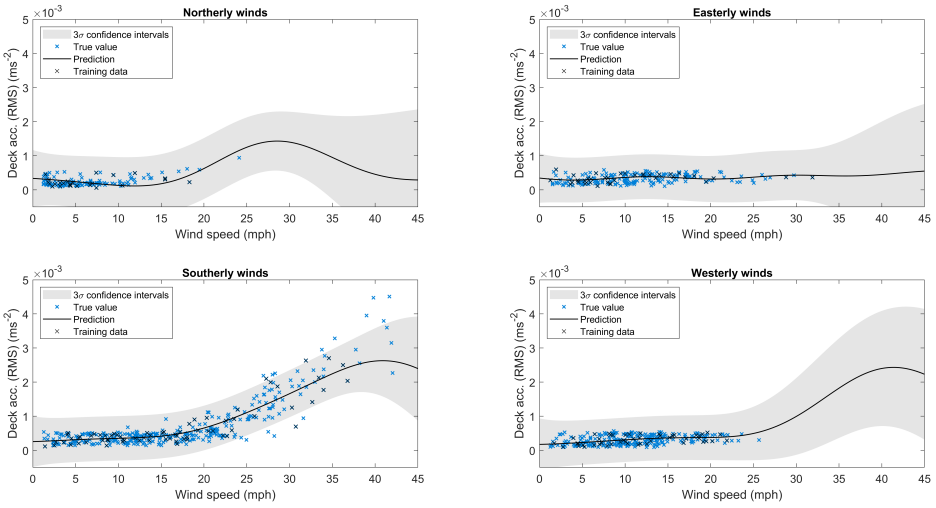

Fig. 4. Predictions of vertical deck acceleration on the Tamar bridge, separated by wind direction, using a Gaussian process with a Squared Exponential kernel $K_{SE}(U, U', \theta, \theta')$. The model was shown a random scatter of 500 datapoints from a 2500 point dataset during training.

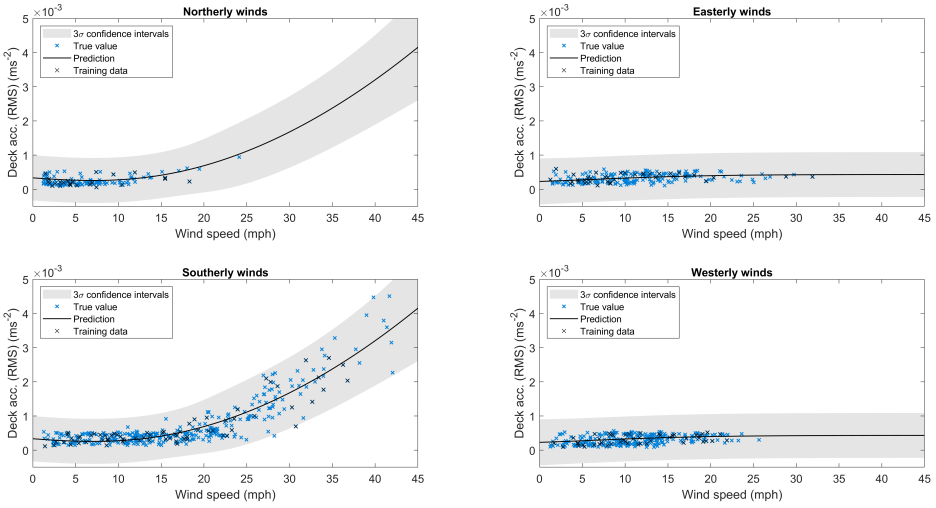

Fig. 5. Predictions of vertical deck acceleration on the Tamar bridge, separated by wind direction, using a Gaussian process with a physics-informed change-point kernel. The model was shown the same random scatter of 500 datapoints from a 2500 point dataset during training as the SE kernel in Figure 4.

The first observation from the results is that the change-point model was better able to cope with the changing behaviours present within the dataset than the purely data-based approach. Across the full time series, the change-point kernel achieved an NMSE of 24.25%, compared with the SE kernels' 29.74%. The introduction of the quadratic lift force for higher speed northerly and southery winds can be seen, whilst importantly, it is not introduced in to the easterly or westerly winds. For the SE kernel however, some difficulty can been seen in distinguishing between behaviours. For example, the prediction of deck acceleration for westery winds was pulled upwards by the presence of high speed southerly training data. Although these points do occupy different places in the input space $[U, \theta]$, the SE kernel could not effectively seperate these relationships. A suspected reason for this is an over estimation for the lengthscale acting on wind angle, preventing predictions varying quickly enough w.r.t this variable.

An advantage of many PIML methods, also seen here, is an ability to use physical knowledge to assist with extrapolation. Beyond the highest northerly and southerly wind speeds observed, the change-point model is able to rely on the non-stationary lift kernel to predict further from observed data. For the SE kernel however, predictions tended back towards the prior mean of zero. For the section of the dataset used for this work, there was a lack of high speed northerly wind data, making predictions in this region challenging. In the UK, the prevalling wind direction is typically south-westerly, causing an increased likelihood to encouter higher wind speeds in this direction [30]. This is further exagerated near the south-west coast, where the Tamar bridge is located. The rarity of higher speed northery winds does not make a models performance during their occurence any less important and an effective model should be able to cope with reduced data in this region. There are many reasons why data for particular event may be rarer to measure and the use of physical knowledge is one way to try alleviate this.

A point of discussion relevant to kernel design is the number of hyperparameters within the model. Typically, the construction of more complex kernels comes at the expense of the introduction of additional hyperparameters, therfore increasing computational demands for optimisation. Here, the number of hyperparameters grows from three for the SE kernel, to eleven for the physics-informed change-point kernel leading to an increase in computation time by a factor of 1.7. To alleviate this, several of the introduced hyperparameters were bounded during the optimisation. For example, the switching location of the wind speed sigmoid x_{w0} was bounded between 5 and 30 mph. The use of physical knowledge to reduce optimiser search space is one of the advantages of hyperparameters with physical meaning.

Another common motivation for PIML methods is an ability to provide additional interpretability within results. This is achieved here through the physically informative hyperparameters of the sigmoid kernels. Plots displaying the sigmoids learned from the data, constructed using these hyperparameters, are shown in Figure 6. These plots are able to intuitively display, in terms of where and how quickly, how the relationships modelled within the data change. Here, the model was able to learn from the data that changing direction of wind rapidly introduced lift force as wind approached northerly and southerly directions ($\cos(2\theta) \to 1$). The dependency of the relationship on wind speed was more gradual, seen within the shallower gradient of the learned wind speed sigmoid. The yellow region of the 2d sigmoid surface highlights a region of high speed northerly and southerly winds in which the lift force kernel has the highest explanatory power within the model. Here, we had a known regime change in mind that we were trying to capture, however similar plots could be of particular help where one might be less sure about how behaviours within a dataset change.

3 Conclusions

The first development of a newly proposed kernel structure was presented in the form of the physics-informed change-point kernel. This kernel provides a means to incorporate physical knowledge within the kernel of a Gaussian process whilst being able to vary the extent to which the model relies upon this knowledge.

A dataset of directional wind loading of the Tamar bridge was used to highlight a case where physical knowledge of a phenomena may only apply in particular regimes. The lift force induced by high wind speed, represented here with a polynomial kernel, only acted when wind hit the bridge side-on. The change-point kernel was able to learn this and present results of the changing regimes in an inerpretable manner. Performance was also improved over a purely data-based approach.

Acknowledgements

The authors gratefully acknowledge the support of Safran Landing Systems, the University of Sheffield and Innovate UK through the OLLGA project grant 10040817. We would like to thank the Vibration Engineering Section for their work collecting the Tamar bridge dataset.

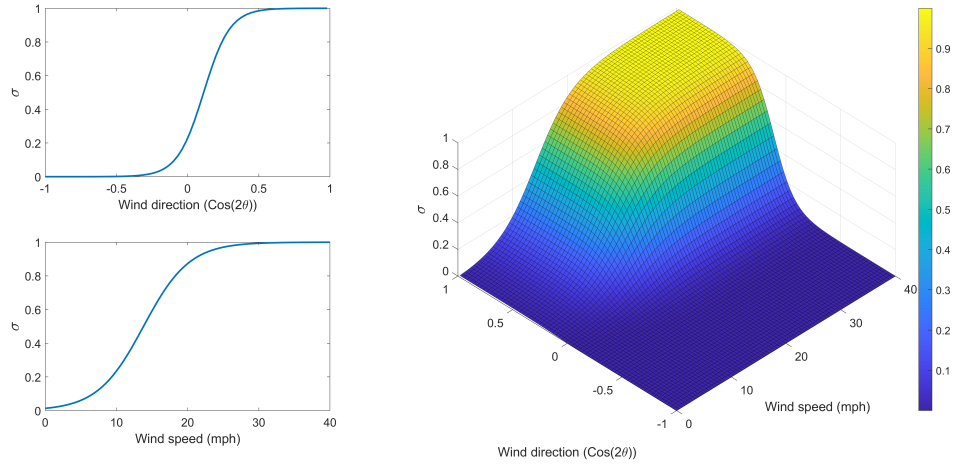

Fig. 6. Learned sigmoid functions for the wind direction (top left) and wind speed (bottom left) from the physics-informed change-point kernel. The 2D sigmoid surface (right) across both wind direction and speed highlights the region in yellow in which the physics-based kernel becomes active. This represents a region of northely/southerly winds at high speeds.

References

1. 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. In: Structural Health Monitoring Based on Data Science Techniques, pp. 347–367. Springer (2022)
2. Karniadakis, G., Kevrekidis, Y., Lu, L., Perdikaris, P., Wang, S., Yang, L.: Physics-informed machine learning. Nature Reviews Physics pp. 1–19 (2021)
3. von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Walczak, M., Pfrommer, J., Pick, A., Ramamurthy, R., Garcke, J., Bauckhage, C., Schuecker, J.: Informed machine learning - a taxonomy and survey of integrating prior knowledge into learning systems. IEEE Transactions on Knowledge and Data Engineering (2021)
4. 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)
5. Pfingstl, S., Zimmermann, M.: On integrating prior knowledge into Gaussian processes for prognostic health monitoring. Mechanical Systems and Signal Processing **171**, 108917 (2022)
6. An, D., Kim, N.H., Choi, J.H.: Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. Reliability Engineering & System Safety **133**, 223–236 (2015)
7. Tóth, M., Brown, A., Cross, E., Rogers, T., Sims, N.D.: Resource-efficient machining through physics-informed machine learning. Procedia CIRP **117**, 347–352 (2023), 19th CIRP Conference on Modeling of Machining Operations
8. Kong, D., Chen, Y., Li, N.: Gaussian process regression for tool wear prediction. Mechanical Systems and Signal Processing **104**, 556–574 (2018)
9. Rogers, T., Gardner, P., Dervilis, N., Worden, K., Maguire, A., Papatheou, E., Cross, E.: Probabilistic modelling of wind turbine power curves with application of heteroscedastic Gaussian process regression. Renewable Energy **148**, 1124–1136 (2020)
10. Bull, L., Gardner, P., Rogers, T., Dervilis, N., Cross, E., Papatheou, E., Maguire, A., Campos, C., Worden, K.: Bayesian modelling of multivalued power curves from an operational wind farm. Mechanical Systems and Signal Processing **169**, 108530 (2022)
11. Rasmussen, C.E., Williams, C.K.I.: Gaussian processes for Machine Learning. The MIT Press (2005)
12. Schulz, E., Speekenbrink, M., Krause, A.: A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. Journal of Mathematical Psychology **85**, 1–16 (2018)
13. Cross, E.J., Rogers, T.J.: Physics-derived covariance functions for machine learning in structural dynamics. IFAC-PapersOnLine **54**, 168–173 (2021), 19th IFAC Symposium on System Identification SYSID 2021
14. Kok, M., Solin, A.: Scalable magnetic field slam in 3d using Gaussian process maps. In: 2018 21st International Conference on Information Fusion (FUSION). pp. 1353–1360 (2018)
15. Noack, M.M., Sethian, J.A.: Advanced stationary and nonstationary kernel designs for domain-aware Gaussian processes. Communications in Applied Mathematics and Computational Science **17** (2022)
16. Haywood-Alexander, M., Dervilis, N., Worden, K., Cross, E.J., Mills, R.S., Rogers, T.J.: Structured machine learning tools for modelling characteristics of guided waves. Mechanical Systems and Signal Processing **156**, 107628 (2021)

17. Klus, S., Gelß, P., Nüske, F., Noé, F.: Symmetric and antisymmetric kernels for machine learning problems in quantum physics and chemistry. Machine Learning: Science and Technology **2** (2021)
18. Grisafi, A., Wilkins, D., Csányi, G., Ceriotti, M.: Symmetry-adapted machine learning for tensorial properties of atomistic systems. Physical Review Letters **120** (2018)
19. Baldacchino, T., Cross, E.J., Worden, K., Rowson, J.: Variational bayesian mixture of experts models and sensitivity analysis for nonlinear dynamical systems. Mechanical Systems and Signal Processing **66-67**, 178–200 (2016)
20. Baldacchino, T., Worden, K., Rowson, J.: Robust nonlinear system identification: Bayesian mixture of experts using the t-distribution. Mechanical Systems and Signal Processing **85**, 977–992 (2017)
21. Kim, C.J.: Dynamic linear models with markov-switching. Journal of Econometrics **60**, 1–22 (1994)
22. Worden, K., Cross, E.: On switching response surface models, with applications to the structural health monitoring of bridges. Mechanical Systems and Signal Processing **98**, 139–156 (2018)
23. Shi, H., Worden, K., Cross, E.J.: A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. Mechanical Systems and Signal Processing **103**, 381–397 (2018)
24. Wilson, A.G.: The change point kernel. Tech. rep. (2013)
25. Wilson, A.G.: Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes. Ph.D. thesis, University of Cambridge (2014)
26. Sarpkaya, T.S.: Wave Forces on Offshore Structures. Cambridge University Press (2010)
27. Cross, E.J.: On Structural Health Monitoring in Changing Environmental and Operational Conditions. Ph.D. thesis, University of Sheffield (2012)
28. Koo, K.Y., Brownjohn, J.M.W., List, D.I., Cole, R.: Structural health monitoring of the Tamar suspension bridge. Structural Control and Health Monitoring **20**, 609–625 (2013)
29. Kerenyi, K., Sofu, T., Guo, J.: Hydrodynamic forces on inundated bridge decks. United States Department of Transportation (2009)
30. Lapworth, A., McGregor, J.: Seasonal variation of the prevailing wind direction in britain. Weather **63**, 365–368 (2008)