



Foundations of population-based SHM, Part I: Homogeneous populations and forms



L.A. Bull^{a,*}, P.A. Gardner^a, J. Gosliga^a, T.J. Rogers^a, N. Dervilis^a, E.J. Cross^a, E. Papatheou^b, A.E. Maguire^c, C. Campos^c, K. Worden^a

^a Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

^b College of Engineering Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

^c Vattenfall Research and Development, New Renewables, The Tun Building, Holyrood Road, Edinburgh EH8 8AE, UK

ARTICLE INFO

Article history:

Received 19 December 2019

Received in revised form 17 April 2020

Accepted 15 July 2020

Available online 14 August 2020

Keywords:

Population-based structural health monitoring

Pattern recognition

Wind turbine monitoring

ABSTRACT

In Structural Health Monitoring (SHM), measured data that correspond to an extensive set of operational and damage conditions (for a given structure) are rarely available. One potential solution considers that information might be transferred, in some sense, between *similar* systems. A *population-based* approach to SHM looks to both model and transfer this missing information, by considering data collected from groups of similar structures. Specifically, in this work, a framework is proposed to model a population of *nominally-identical* systems, such that (complete) datasets are only available from a subset of members. The SHM strategy defines a general model, referred to as the *population form*, which is used to monitor a homogeneous group of systems. First, the framework is demonstrated through applications to a simulated population, with one experimental (test-rig) member; the form is then adapted and applied to signals recorded from an operational wind farm.

© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Population-based SHM

This paper is the first in a series of three, which will lay the foundations for a theory of *Population-Based Structural Health Monitoring* (PBSHM), it is followed by [1,2].

In conventional Structural Health Monitoring (SHM), a predictive model is learnt using the data recorded from a single structure (or system), and the model is expected to generalise to future measurements from that structure. However, for a given structure, the measured signals are usually limited, corresponding to only a fraction of the potential operational, environmental and damage conditions. Additionally, the label set (which describes what the measurements represent) might be unavailable or incomplete. If a framework can transfer information from one structure within a population to another *similar* structure for which data are missing, in order to allow diagnostic inferences on the second structure, it should bring significant advantages to practical applications.

In consequence, the *population-based* approach to SHM [1–9] aims to facilitate the transfer of valuable knowledge between groups of *similar* systems, i.e. *populations*. Intuitively, the type of population is important, as this will define what knowledge can be transferred, and by what method. This paper focusses on the simplest population case, in which all members can be (informally) considered to be *nominally-identical* [5]; for example, wind turbines within a wind farm. These

* Corresponding author.

E-mail address: l.a.bull@sheffield.ac.uk (L.A. Bull).

populations are referred to as *homogeneous*, and in certain cases, a general model (which will be referred to as the population *form*) can be used to represent the behaviour of the group, and infer the presence of damage between members.

While this work focusses on nominally-identical systems, the sister papers [1,2] address alternative technologies, considering the *heterogeneous* case, in which the population can contain more disparate members, e.g. different designs of suspension bridge. While these structures are still similar in some sense, they will not be referred to as nominally identical.

The layout of this paper is as follows. Section 2 looks to define homogeneous populations, including the *strongly-homogeneous* case; the *form* is also introduced. Section 3 extends a simulated case study (introduced in [5]), to demonstrate the form applied to an idealised strongly-homogeneous case. Through applications to measurements from an operational wind farm, Section 4 introduces a novel mixture-model algorithm [10], to approximate a form of a practical homogeneous population, representing more significant and realistic variation between individuals. Section 5 discusses the limitations of the forms approximated in this work and suggests alternative technologies to model more extended or complex population variation.

2. Homogeneous populations and forms

A convenient language to define homogeneous (as well as heterogeneous) populations borrows terminology from graph theory. The work in [1,2] will explain how structures can be represented as graphs; however, for context, a general introduction is provided here.

A graph can be used to define a simplified representation (or irreducible element model [1,4]) of a given structure: to demonstrate, examples are provided in Fig. 1. Despite their simplicity, these graphs can inform a convenient measure of *similarity* between structures [1,6], and this measure can be used to inform the level of inference that can (potentially) be made between systems within a population.

For homogeneous populations, the concept of *structural equivalence* is important; this implies that the graphs used to represent structures are *topologically equivalent* [1], with ground-nodes occurring in the same location. Examples of graphs that represent structurally-equivalent wind turbines are shown in Fig. 1a and b; intuitively, the associated graphs are identical (drawn differently to demonstrate the non-uniqueness of layout of graphical representations). On the other hand, Fig. 1c shows a graph that is topologically equivalent to those in Fig. 1a and b, but *not* structurally equivalent – due to an inconsistent ground node. Finally, a graphical representation of a three-span bridge is shown in Fig. 1d; this is topologically and structurally *inequivalent* to all the other graphs in Fig. 1 – clearly, the graphs are dissimilar.

Another useful feature in graphical representations of structures are node and edge *attributes* (see Part II [1] for details). In this case, the attributes are parameter sets associated with the graphical elements, used to describe aspects relevant in fixing the dynamic behaviour of the structure – such as the topology, materials, and geometry [1].

Having introduced these concepts, a homogeneous population can be defined:

Definition 1. A population of M structures $\{S_i\}_{i=1}^M$ is homogeneous if individual members are pair-wise structurally equivalent, with material, geometric, and physical parameters Θ_i (i.e. graph attributes) that can be considered to be random draws from an underlying base-distribution $p(\Theta_i)$.

As a result, the distribution $p(\Theta_i)$ describes population variation in terms of the parameters Θ_i . If any pair of members are not structurally equivalent, the population is *heterogeneous*.

2.1. Strongly-homogeneous populations

A further restriction considers the *strongly-homogeneous* case. In this scenario, the densities $p(\Theta_i)$ associated with the parameters of the population are considered to be unimodal with low dispersion. In the strictest sense, the associated densities would be Dirac functions over the parameter values, such that each member is an identical system. A more realistic example, considers a population of components of the same model, but subject to manufacturing tolerances. In this setting, methods from conventional SHM [11] can potentially be applied to population data, as demonstrated in the first case study, in Section 3.

Importantly, the strongly-homogeneous case represents an idealised population. Returning to the wind farm example, while the population is homogeneous in terms of structural equivalence, it is unlikely that the group will be *strongly-homogeneous*. For example, consider the complex distributions that would be associated with the parameters that describe the boundary conditions at the seabed, or local interactions with the wind, as well as differences relating to operational practice. In consequence, the case study in Section 4 demonstrates one method to move beyond the restriction to the strongly-homogeneous case.

2.2. The population form

The concept of the *Form* for population-based SHM was motivated by the work of Plato. Plato's theory of Forms (also sometimes called the theory of Ideas) is referred to across Plato's dialogues; although, it is first anticipated in the Meno [12] and considered in greatest depth in The Republic [13]. Forms were considered to be the essence of things and existed

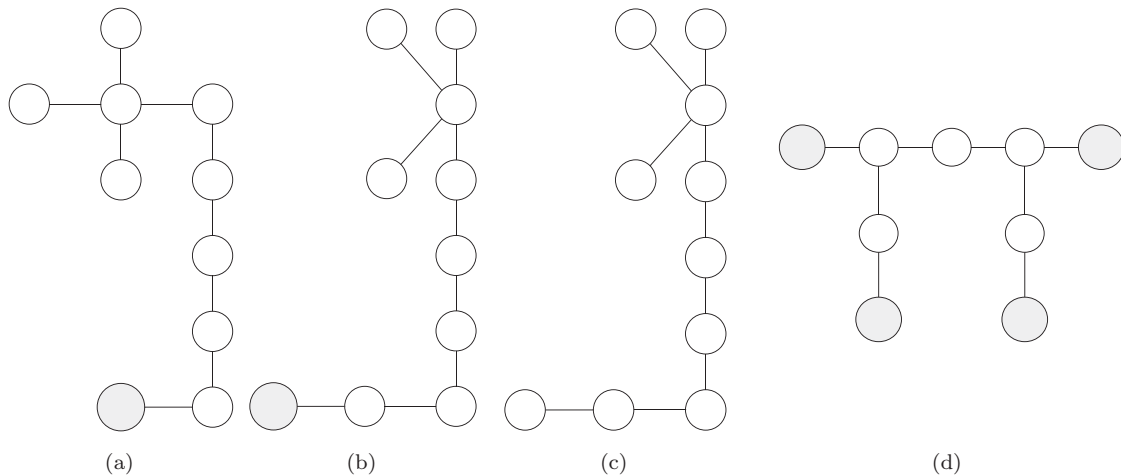


Fig. 1. Graphical representations of structures. Circles are nodes and lines are edges, ground nodes are shaded. (a) and (b): two topologically-equivalent graphs, used to represent *structurally-equivalent* wind turbines. (c): a graph that is topologically equivalent to (a) and (b) but not structurally, due to inconsistency in ground nodes. (d) a graph used to represent a three-span bridge, that is topologically and structurally *inequivalent* to (a), (b) and (c).

as abstract entities; they were considered eternal, immutable, and to represent the highest level of reality, independent of ordinary objects in the human world. Ordinary objects derive their nature and properties by 'participating' in the Forms; e.g. all cats in the world are recognisable as such because they participate in the Form of *cat*.

The idea of the Form is extended in the current work to capture, not only the essence of things, but the extent of variations in their participants: the extended Form of cat would not only constitute Plato's original Form, but a summary of the expected variations in participants encountered in human reality. This extension is needed in order to capture the needs of population-based SHM; a Form here is thus the essential object of interest, together with a measure of the expected variations from the ideal across the population. For example, consider a particular model of a wind turbine; one could argue that the essential nature of the turbine is captured in the complete design specification; in reality, a population of real turbines will differ from the ideal because of manufacturing tolerances, and perhaps other inconsistencies.

In this work, from this point forward, the term *form* will be used in a mathematical sense to denote a model (in some feature space) of the object of interest which attempts to capture the two ingredients of the extended Form above¹: the *essential nature* of the object and the *variations encountered* when the object is embodied in the real world. Hence, the term *form* will herein refer to a model which represents, or approximates closely, the extended Form defined above.

The 'object of interest' need not be a structure itself, but a specific feature or measurement vector which represents the structure for SHM purposes; thus, the feature space associated with a population of structures is part of the description of the form. Considering the above, the model should be validated within the population and chosen feature space, such that it is capable of representing the variation between individuals. Variation between members is less significant when considering the strongly-homogeneous case, however, when moving to more practical examples, the form will become increasingly complex – as demonstrated in the case studies.

In this work, a statistical model of the Form is defined using benchmark normal-condition data. Importantly, the data are incomplete (or absent) from several members within the population. Through comparisons between future measurements and the form, the model is used for novelty detection across the entire population, to monitor the condition of each member.

2.2.1. Gaussian process regression of the form

Any suitable combination of model and/or feature space can be used to represent the population form. If the features of interest are functional (e.g. wind turbine power curves, or frequency response functions), a Gaussian process (GP) regression is an excellent candidate to describe a form: the predictive mean can be regarded as capturing the essence of the feature, while the predictive variance represents variations across the population. If the features of interest are pointwise and multivariate, a candidate representation of the form might be a Gaussian Mixture Model (GMM); although subtleties may arise in extracting the essence from a multi-modal GMM – this subtlety then propagates into the capture of variations.

As *functional features* are the focus of this work, Gaussian Process (GP) regression is applied in each case study². Gaussian processes can be used to solve regression problems through a Bayesian machine learning approach [14,15]. The GPs exhibit a number of desirable properties for this application: they are nonparametric, automatically return confidence intervals, and they are capable of modelling data with a low signal-to-noise ratio. A brief review of the theory is provided here; details can be found in [14,15].

¹ To distinguish between abstract entities and their concrete model representations, the latter will be distinguished by a capital, i.e. Forms.

² In each case study, a parametric model could arguably be applied; however, the focus of this work is to introduce the general concept of the *form*; therefore, GPs are used (with appropriate parameterised mean functions) as a *general* regression tool – to provide a flexible representation of the form.

For a set of N inputs x_i , and corresponding outputs y_i , (i.e. the training data, $\mathcal{D} = \{x_i, y_i\}_{i=1}^N = \{\mathbf{x}, \mathbf{y}\}$) GP regression looks to define the predictive distribution, given a new input x_* and the available training data \mathcal{D} . It is assumed that observations can be modelled by some noiseless latent function f of the inputs, plus an independent noise term, ϵ_i [10],

$$y_i = f(x_i) + \epsilon_i \quad (1)$$

A GP prior is set over the latent functions f , with covariance function $k(x_i, x_r)$, and a Gaussian prior is placed over the noise term, such that,

$$f(x_i) \sim \mathcal{GP}(m(x_i), k(x_i, x_r)), \quad \epsilon_i \sim \mathcal{N}(0, \sigma_n^2) \quad (2)$$

This expression introduces the first hyperparameter σ_n^2 , which specifies the noise variance. Importantly, the GP prior should introduce *a priori* knowledge over the expected functions, given engineering judgement. This knowledge is included via the mean function $m(x_i)$ and the covariance function $k(x_i, x_r)$.

The covariance encodes the degree of coupling between y_i and y_r , and therefore determines important properties, including the process variance and smoothness [10]. One of the best-known covariance functions is the squared-exponential, and it is applied in this work,

$$k(x_i, x_r) = \sigma_0^2 \exp\left(-\frac{1}{2l^2}(x_i - x_r)^2\right) \quad (3)$$

Eq. (3) introduces two more hyperparameters: the length-scale l , and the process variance σ_0 . The length scale determines how fast the correlation between outputs decays across the input space, while σ_0 determines the signal variance [10].

Common practice [14,16,10] assumes a zero-mean function for the prior of the GP. To satisfy this assumption, \mathcal{D} should be normalised – usually by subtracting the sample mean and dividing by the standard deviation [14]. However, in many cases, an explicit mean can be incorporated given engineering judgement – even as an approximation. Conveniently, a parametric definition of $m(x_i)$ leads to a natural expression of such prior information, while improving the interpretability of the model. Different mean functions are appropriate in each case study, depending on the application. For now, in general notation, the mean is characterised by another set of hyperparameters α , such that,

$$m(x_i; \alpha) \quad (4)$$

The collected hyperparameters of the model are $\theta = \{l, \sigma_0, \sigma_n, \alpha\}$.

The joint distribution between the training data $\mathcal{D} = \{x_i, y_i\}_{i=1}^N$ and some previously unseen observations, $\{\mathbf{x}_*, \mathbf{y}_*\} = \{x_{*i}, y_{*i}\}_{i=1}^M$ is a multivariate Gaussian,

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(\mathbf{x}) \\ m(\mathbf{x}_*) \end{bmatrix}, \begin{bmatrix} K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N & K_{\mathbf{x}, \mathbf{x}_*} \\ K_{\mathbf{x}, \mathbf{x}} & K_{\mathbf{x}, \mathbf{x}_*} + \sigma_n^2 \mathbf{I}_M \end{bmatrix}\right). \quad (5)$$

The notation $[A]_{i'}$ refers to the i^{th} row in column i' of matrix A , while $[\mathbf{a}]_i$ is the i^{th} element of vector \mathbf{a} [10]. As such, $[K_{\mathbf{xx}}]_{i'}$ = $k(x_i, x_r)$, $[K_{\mathbf{xx}}]_{i'i'}$ = $k(x_i, x_{r'})$, and $[K_{\mathbf{x}, \mathbf{x}_*}]_{i'i'}$ = $k(x_{*i}, x_r)$. \mathbf{I}_N is an $N \times N$ identity matrix.

By conditioning the joint distribution in (5) on the observed training data in \mathcal{D} , the predictive distribution over \mathbf{y}_* can be defined [14],

$$\begin{aligned} p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D}) &= \mathcal{N}(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \\ \boldsymbol{\mu}_* &\triangleq m(\mathbf{x}_*) + K_{\mathbf{x}, \mathbf{x}}(K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N)^{-1}(\mathbf{y} - m(\mathbf{x})) \\ \boldsymbol{\Sigma}_* &\triangleq K_{\mathbf{x}, \mathbf{x}_*} - K_{\mathbf{x}, \mathbf{x}}(K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N)^{-1}K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_M \end{aligned} \quad (6)$$

i.e. a posterior predictive mean $\mathbb{E}[\mathbf{y}_*] = \boldsymbol{\mu}_*$ and variance $\mathbb{V}[\mathbf{y}_*] = \text{diag}(\boldsymbol{\Sigma}_*)$.

In order to learn the hyperparameters, $\theta = \{l, \sigma_0, \sigma_n, \alpha\}$, a Type-II maximum likelihood approach is taken here [14]; this is equivalent to empirical Bayes [16]. As such, the *marginal likelihood* of the model $p(\mathbf{y} | \mathbf{x}, \theta)$ is maximised; this utilises the Bayesian Occam's razor [17,14], to find the minimally-complex model given the observed training data \mathcal{D} . This optimisation is normally performed as a minimisation over the negative log-marginal-likelihood, for convenience and numerical stability; thus, the hyperparameters are estimated through the following optimisation [16],

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \{-\log p(\mathbf{y} | \mathbf{x}, \theta)\}, \quad (7)$$

where,

$$\begin{aligned} -\log p(\mathbf{y} | \mathbf{x}, \theta) &= -\log \mathcal{N}(\mathbf{y} | m(\mathbf{x}), K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N) = \frac{N}{2} \log(2\pi) + \frac{1}{2} \log |K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N| \dots \\ &+ \frac{1}{2} [(\mathbf{y} - m(\mathbf{x}))^\top (K_{\mathbf{xx}} + \sigma_n^2 \mathbf{I}_N)^{-1} (\mathbf{y} - m(\mathbf{x}))] \end{aligned} \quad (8)$$

3. Case study I: Strongly-homogeneous population

To demonstrate the form applied to a *strongly-homogeneous* population, 19 structurally-equivalent systems are simulated, leading to a population $\{S_i\}_{i=1}^{19}$. As in previous work [5], each member S_i is a model realisation of an experimental rig, designed at the Los Alamos National laboratory [11]. The test-rig itself acts as the 20th member in the population, such that the total group is $\{S_i\}_{i=1}^{20}$. A schematic of the 8-DOF system is shown in Fig. 2; $u_i(t)$ is the system input (forcing) on mass i at time t , and $y_i(t)$ is the system response (output) of mass i at time t .

3.1. Modal analysis

In order to simulate signals that resemble the experimental data, the system parameters – mass m_j , stiffness k_j , and damping c_j – are identified by minimising a sum-of-squares error between the analytical and experimental frequency response functions (FRFs). Note, the j index indicates the j^{th} degree of freedom, such that $j = \{1, \dots, 8\}$.

The experimental FRF can be calculated from the forcing $u_1(t)$ and acceleration $z_8(t)$ time-series measurements, such that,

$$H_{18}(\omega) = \frac{S_{zu}(\omega)}{S_{uu}(\omega)} \quad (9a)$$

$$\begin{aligned} S_{zu}(\omega) &\triangleq \mathbb{E}[Z_8(\omega)\bar{U}_1(\omega)] \\ S_{uu}(\omega) &\triangleq \mathbb{E}[U_1(\omega)\bar{U}_1(\omega)] \end{aligned} \quad (9b)$$

$$\begin{aligned} U_1(\omega) &\triangleq \mathcal{F}[u_1(t)] \\ Z_8(\omega) &\triangleq \mathcal{F}[z_8(t)] \end{aligned} \quad (9c)$$

where ω denotes frequency; S_{zu} and S_{uu} are the standard definitions of the auto-spectra and the cross-spectra [18]; bar notation denotes the complex conjugate; and \mathcal{F} is the fast Fourier transform. The analytical FRF for an 8DOF system is defined under linear assumptions with proportional damping [18],

$$\tilde{H}_{18}(\omega) = \omega^2 \sum_{j=1}^8 \frac{\Psi_{1k} \Psi_{8k}}{\hat{m}_j \omega^2 - i \hat{c}_j \omega - \hat{k}_j} \quad (10)$$

$$\begin{aligned} \Psi &\triangleq \text{eig}\{M^{-1}K\} \\ \hat{\mathbf{m}} &\triangleq \text{diag}\{\Psi^T M \Psi\} \\ \hat{\mathbf{k}} &\triangleq \text{diag}\{\Psi^T K \Psi\} \\ \hat{\mathbf{c}} &\triangleq \text{diag}\{C\} \end{aligned}$$

where the expression $\text{eig}(A)$ returns a matrix V , whose columns are eigenvectors, such that $AV = V\Lambda$, and Λ is the diagonal matrix of corresponding eigenvalues. The mass, stiffness and damping matrices, M, K and C , are defined by standard practice from the set of physical parameters, herein denoted $\Theta = \{m_j, k_j, c_j\}_{j=1}^8$.

Following some initial guesses, the system parameters Θ are iteratively updated through a constrained optimisation algorithm, based on prior knowledge of the 8DOF rig, such that,

$$\hat{\Theta} = \underset{\Theta}{\text{argmin}} \left\{ \left\| \tilde{H}_{18}(\omega; \Theta) - H_{18}(\omega) \right\|_2 \right\}, \quad (11)$$

The optimised parameters are presented in Table 1³.

3.2. Simulating strongly-homogeneous members

In the 8-DOF population, the parameters (or *attributes*) associated with each i th member S_i are considered to be,

$$\Theta_i = \left\{ m_j^{(i)}, k_j^{(i)}, c_j^{(i)} \right\}_{j=1}^8 \quad (12)$$

To simulate a *strongly* homogeneous group, according to Section 2.1, the densities $p(\Theta_i)$ that describe the underlying distribution of these parameters (across the population) should be unimodal with low dispersion; thus, Gaussian distributions are placed over each value in Table 1 to approximate the variance that might be observed over a strongly homogenous population, such that,

³ The spring constant k_1 is near zero due to base-excitation with absolute coordinates, the actual value of k_1 is $1 \times 10^{-6} \text{ kN/m}$.

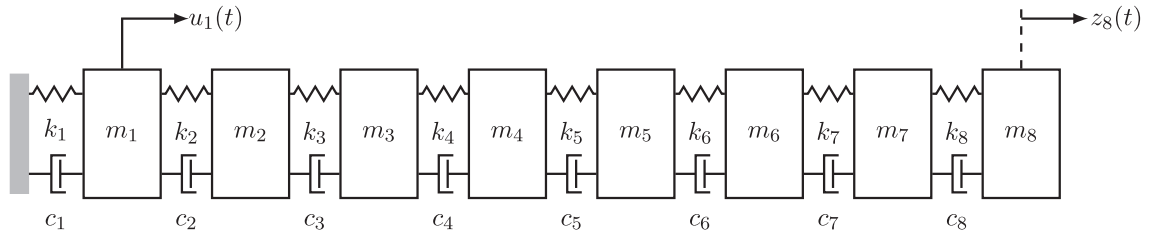


Fig. 2. The 8-DOF system.

Table 1
System parameters.

j	1	2	3	4	5	6	7	8
m_j (kg)	0.5318	0.4040	0.4101	0.4123	0.3960	0.3809	0.4086	0.3798
k_j (kN/m)	0.000	56.70	56.70	56.70	56.70	56.70	56.70	56.70
c_j (Ns/m)	8.7460	8.7916	8.8012	8.8512	8.7146	8.7378	8.5494	8.7521

$$\begin{aligned}
 m_j^{(i)} &\sim \mathcal{N}(m_j, 0.03 \times m_j) \\
 k_j^{(i)} &\sim \mathcal{N}(k_j, 0.01 \times k_j) \\
 c_j^{(i)} &\sim \mathcal{N}(c_j, 0.10 \times c_j)
 \end{aligned} \tag{13}$$

A strongly-homogeneous population is generated by sampling 19 parameter sets Θ_i from the distributions in (13), to define the 19 similar members.

3.2.1. The Frequency Response Function (FRF) for damage detection

In the context of vibration-based monitoring, it is generally expected that damage will manifest itself as alterations in the structural parameters – typically, a reduction in stiffness [11,19]. Changes in the structural stiffness will affect the dynamic characteristics, therefore, frequency domain features can be used to (indirectly) monitor physical changes that *could* relate to damage; it should be considered, however, that these features will also be sensitive to confounding influences [11].

Considering the above, and to match the experimental data, the empirical FRF H_{18} is used as a damage-sensitive feature, defined in (9). For the simulated members, the output time-series are modelled using the system parameters in Θ_i to represent the system equations in state-variable form [20,21]. As in the experiments [11], the input is a white-noise excitation over 8s, with a sample-rate of 400.25 Hz. A Hanning window is applied to the 8s input and output time series, and the FRF is calculated using (9). The resulting FRF is truncated, such that there are 1040 bins in the frequency domain, ranging from 0–130 Hz. Measurement noise is added to the outputs, to better represent the experimental data; the noise is assumed to be zero-mean normally-distributed, with a signal-to-noise ratio (in terms of variance) of 40 dB.

For simplicity, only the real part of the FRF is considered $\text{Re}(H_{18})$; this is convenient for computational reasons and model simplicity. Additionally, *for the purpose of novelty detection*, the real part should sufficiently characterise the system dynamics (for this example). In fact, for a linear system, the real part completely specifies the imaginary part (through the Hilbert transform) and vice versa [18].

In terms of the SHM strategy, each FRF is considered as an observation of the system, and these observations are used to inform damage detection. Following the experimental procedure in [11], the stiffness of k_5 is reduced to imitate damage. Reductions correspond to 7%, 14% and 24% for the simulated members, and a single reduction of 24% for the test rig.

3.2.2. Dataset summary

The dataset represents a population of *twenty* 8-DOF systems, $S_{1,19}$ are simulated, while S_{20} is the test-rig. Each system is observed over time, through FRFs estimated over 8s time-windows. For each FRF there are 1040 frequency bins. Measurements from the simulated members are defined such that:

- For the normal-condition data, there are 20 FRFs from each structure. These data are shown by black markers in Fig. 3.
- 20 additional normal-condition FRFs are simulated to test the form. These data are shown by magenta markers in Fig. 3.
- 20 FRFs are generated for each state of damage (7%, 14% and 24%), shown by red markers in Fig. 3.

The experimental data recorded from the test-rig include:

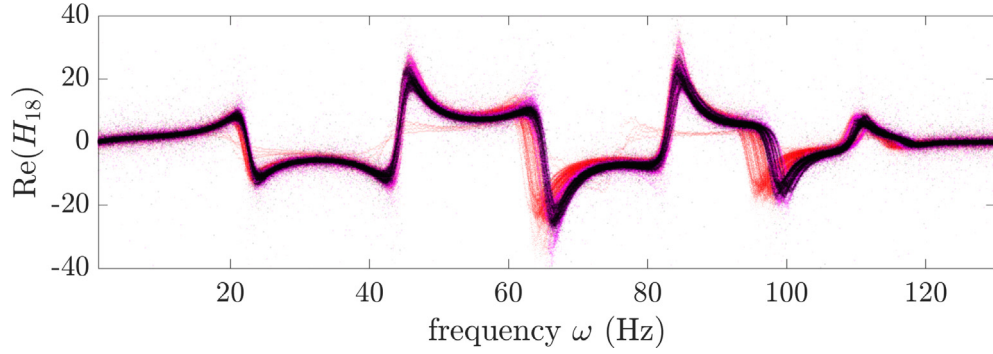


Fig. 3. FRFs data from a population of twenty 8DOF systems $\{S_1, \dots, S_{20}\}$, $S_{1,19}$ are simulated members and S_{20} the experimental rig. The normal-condition is shown by black markers (training-data) and magenta markers (test-data), while the damage data are shown by red markers (test-data).

- 8 FRFs – four corresponding to the normal condition, and four to damage (24% only). In Fig. 3, the normal and damaged test-data are also shown by magenta and red markers respectively.

While the normal-condition FRFs are similar in Fig. 3, population variation should still be observed within the associated FRFs.

3.3. Gaussian Process regression of the FRF as the population form

For the form in this case study, the latent function modelled by the GP regresses between *input frequency* x_i and *output response* y_i , such that,

$$\begin{aligned} \text{Re}[H_{18}(x_i)] &= y_i \\ &= f(x_i) + \epsilon_i \end{aligned} \quad (14)$$

$$f(x_i) \sim \mathcal{GP}(m(x_i), k(x_i, x_j)), \epsilon_i \sim \mathcal{N}(0, \sigma_n^2) \quad (15)$$

Given knowledge of the dynamic system, an appropriate prior mean-function $m(x_i)$ is the real part of the analytical FRF, which was defined under linear assumptions and proportional damping in (10). To reiterate,

$$\begin{aligned} m(x_i) &= \text{Real}\{\tilde{H}_{18}(x_i)\} \\ &= \text{Real}\left\{x_i^2 \sum_{j=1}^8 \frac{\Psi_{1k} \Psi_{8k}}{\tilde{m}_j x_i^2 - i \tilde{c}_j x_i - k_j}\right\} \end{aligned} \quad (16)$$

The corresponding function is parameterised by M, K and C through (10) – in turn, these variables are parameterised by Θ (Eq. (12)). Therefore, the explicit mean function is,

$$m(x_i; \Theta) = \text{Real}\{\tilde{H}_{18}(x_i, M, K, C)\} \quad (17)$$

The kernel function $k(x_i, x_j)$ is the squared-exponential, outlined in Section 2.2.1; thus, the complete set of hyperparameters for the GP are, $\theta = \{l, \sigma_o, \sigma_n, \Theta\}$; these are optimised while minimising the log-marginal-likelihood, according to (11) and (8).

Importantly, only systems $\{S_1, \dots, S_{10}\}$ in the population contribute training data to learn the form. The remaining 10 systems $\{S_{11}, \dots, S_{20}\}$ (nine simulated systems and the experimental rig) are *held out* of the training process. Therefore, the held out systems test the generalisation of the form when applied to *new* unseen members (from the same population).

For computational reasons, the GP regression is trained using a random sub-sample of 2000 observations from the *simulated* normal-condition data, $\{S_1, \dots, S_{10}\}$. A more rigorous approach to deal with large data, such as sparse GPs [22], is being considered for future work. The resulting GP representation of the form, and the data used to train it, are illustrated in Fig. 4. Interestingly, the GP successfully models the residual between the analytical *linear* FRF (used as the mean-function of the prior (17)) and measurements from the population in practice. Corrections in the FRF through the GP are particularly notable for damping effects at higher modes, where nonlinearities of the rig in practice, appear to increase the discrepancy between the linear (proportional damping) approximation of the FRF and the measured data.

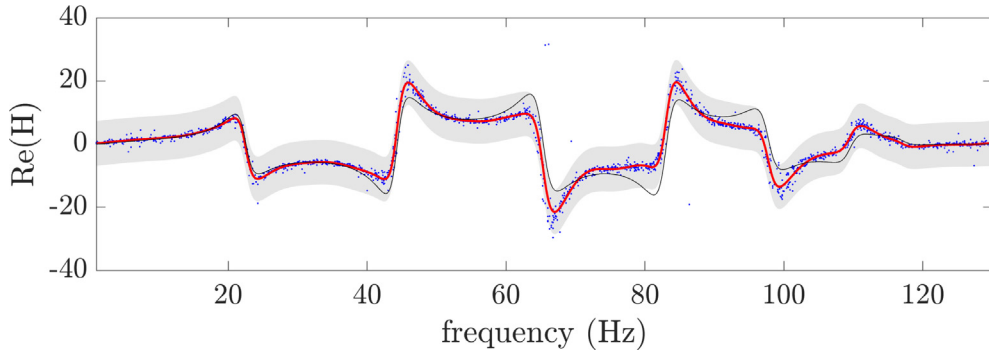


Fig. 4. Gaussian process regression of the FRF as the population-form. Blue markers indicate the training-set. The red line indicates the posterior predictive mean, $\mathbb{E}[y^*]$, and the shaded area indicates a $3\text{-}\sigma$ credible region, corresponding to the variance, $\mathbb{V}[y^*]$. The black line plots the mean function of the GP prior $m(x_i)$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3.1. Novelty detection via the form

The form can now be used to monitor future data and inform damage detection. In this example, test FRFs from *all* members in the population are compared to the form⁴. Any appropriate measure of *deviation* can be used here; for example, measures of discordancy, error, extreme function theory [23], or the predictive likelihood (as in the next case study). In this example, the multivariate *Mahalanobis Squared-Distance* (MSD) is used as a novelty metric, calculated for 1000-point random-samples from each test FRF. The MSD is useful in this application, as it scales well with the number of observations in the test sample, and it considers the mean and covariance from the full posterior-predictive distribution of the form model, $p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D})$ – where \mathbf{y}_* , \mathbf{x}_* represent the 1000 collected test observations, such that $[\mathbf{y}]_i = y_{*i}$ and $[\mathbf{x}]_i = x_{*i}$

$$MSD_i = (\boldsymbol{\mu}_* - \mathbf{y}_*)^\top \boldsymbol{\Sigma}_*^{-1} (\boldsymbol{\mu}_* - \mathbf{y}_*) \quad (18)$$

The mean and covariance are defined by the predictive equations from the GP form in (6), and the test-data are the experimental/simulated outputs.

In order to define a detection threshold, which flags observations as *inlying* or *outlying* (i.e. normal or novel), bootstrapping is used [16,19]. This defines the threshold by randomly sampling 1000 points from the normal-condition data used to train the form. The multivariate MSD is then calculated according to (18). These steps are repeated for a large number of trials, and the resulting MSDs are sorted in order of magnitude. The critical value is the threshold which contains 95.45% (two-sigma) of the MSD values beneath it.

3.4. Results

Results for novelty detection across the population via the form are shown in Figs. 5 and 6; these plots can be interpreted as control charts, each sub-figure representing an individual member. Fig. 5 presents the MSD values (corresponding to test FRFs) for members $\{S_1, \dots, S_{10}\}$; these members contributed (a separate set of) normal-condition data used to train the form, shown in Fig. 4. The MSD values for members $\{S_{11}, \dots, S_{20}\}$ are presented in Fig. 6, including the experimental rig, S_{20} ; importantly, these systems are in the held-out group, that did not contribute data to train the form.

For the normal-condition test-data, relating to members $\{S_1, \dots, S_{10}\}$, the discordancy measure generally falls below the detection threshold, for all members in the population. This is expected, as variations in these data (compared to the form) relate to small perturbations in the system parameters Θ_i (Eq. (12)), due to resampling members from $p(\Theta_i)$ (Eq. (13)), as well as measurement noise.

There are some false positives present, corresponding to the normal condition FRFs; for example, S^9 , S^{11} , S^{13} and S^{18} . As well as noise effects, these false-positives most likely correspond to ‘extreme’ parameter sets being drawn from the underlying distribution $p(\Theta_i)$ of the homogeneous population. Notably, for the experimental member S^{20} , one normal condition FRF is flagged as an outlier, and the rest remain close to the threshold. This is unsurprising, however, considering that the experimental member did not contribute training-data to learn the form; additionally, errors in the estimated parameters from Table 1 will add to the discordancy – as these were used to generate simulated members.

For the damaged-condition FRFs, the number of true positives increases as the severity of damage increases. Generally speaking, this form fails to highlight 7% damage, with increased sensitivity to 14% damage, and successfully flagging 24% damage observations as outlying. False negatives for 7% occur because, at low levels of damage, the variation across the population, defined by $p(\Theta)$, is similar to (or more severe than) the variations due to damage. As a result, with the current form,

⁴ To reiterate, only $\{S_1, \dots, S_{10}\}$ contributed to the training-set.

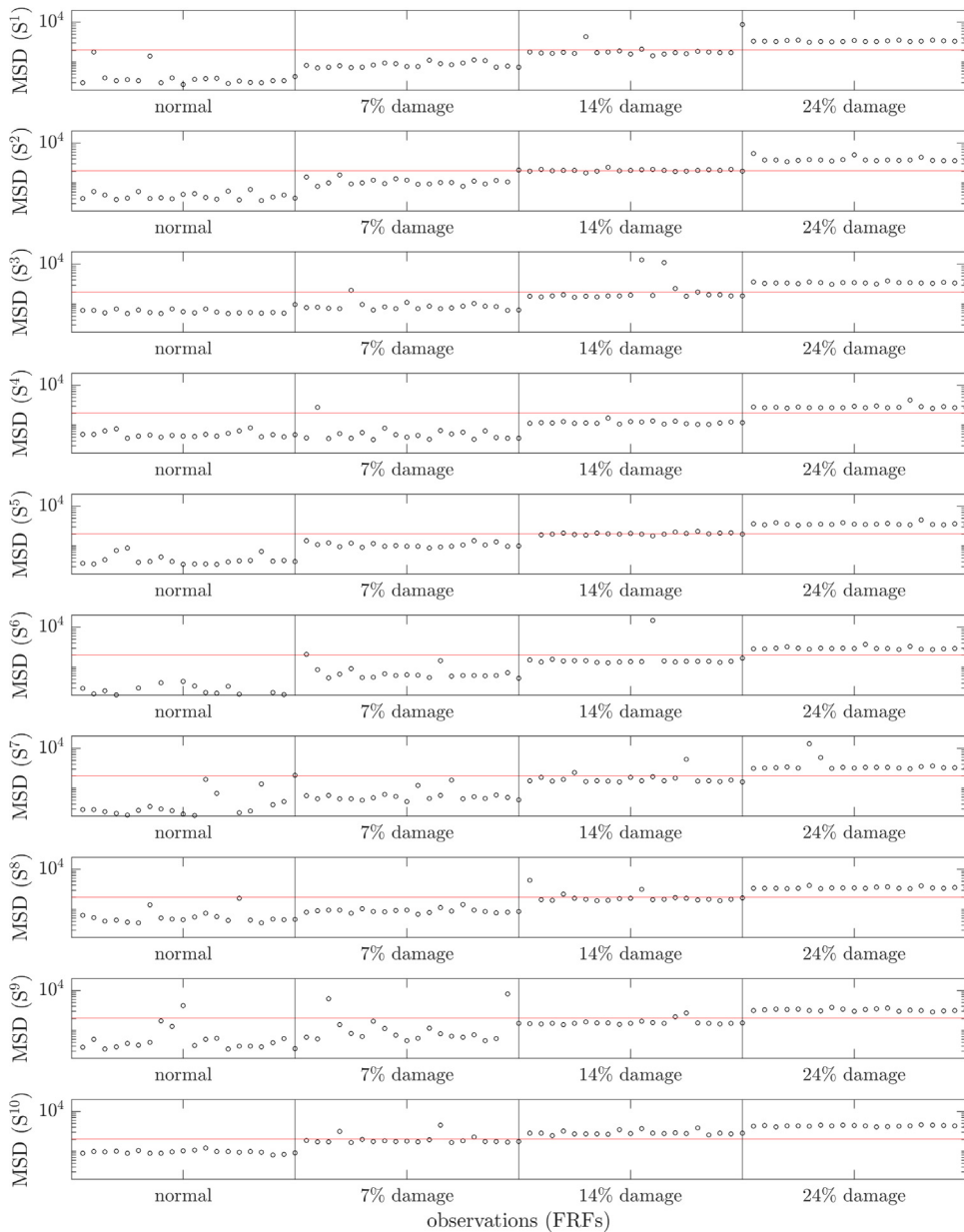


Fig. 5. MSD novelty index for the test-data FRFs, comparing members $\{S^1, \dots, S^{11}\}$ to the form. These members contributed a separate set of training data, used to learn the form. Damage/normal conditions are separated by vertical lines. The red line indicates the detection threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

population variance *masks* the variations due to low-level damage. To expose low-level damage, another definition of the form is required; this can be done by defining an alternative feature-space, an alternative model, or both.

3.5. Discussion

This case study has demonstrated that the form can be used as a general representation of a strongly-homogeneous population. Given training data from a *subset* of members, the form can be used to model missing information from the hold-out group, to aid diagnostic decisions.

The success of this initial approach, however, depends greatly on $p(\Theta_i)$, which, in turn, depends on the type of population. If the underlying density $p(\Theta_i)$ across members is expected to be dispersed and/or multi-modal (unlike the Gaussian distributions in this example) it is likely that the population variance will *mask* changes in the feature space that are due to dam-

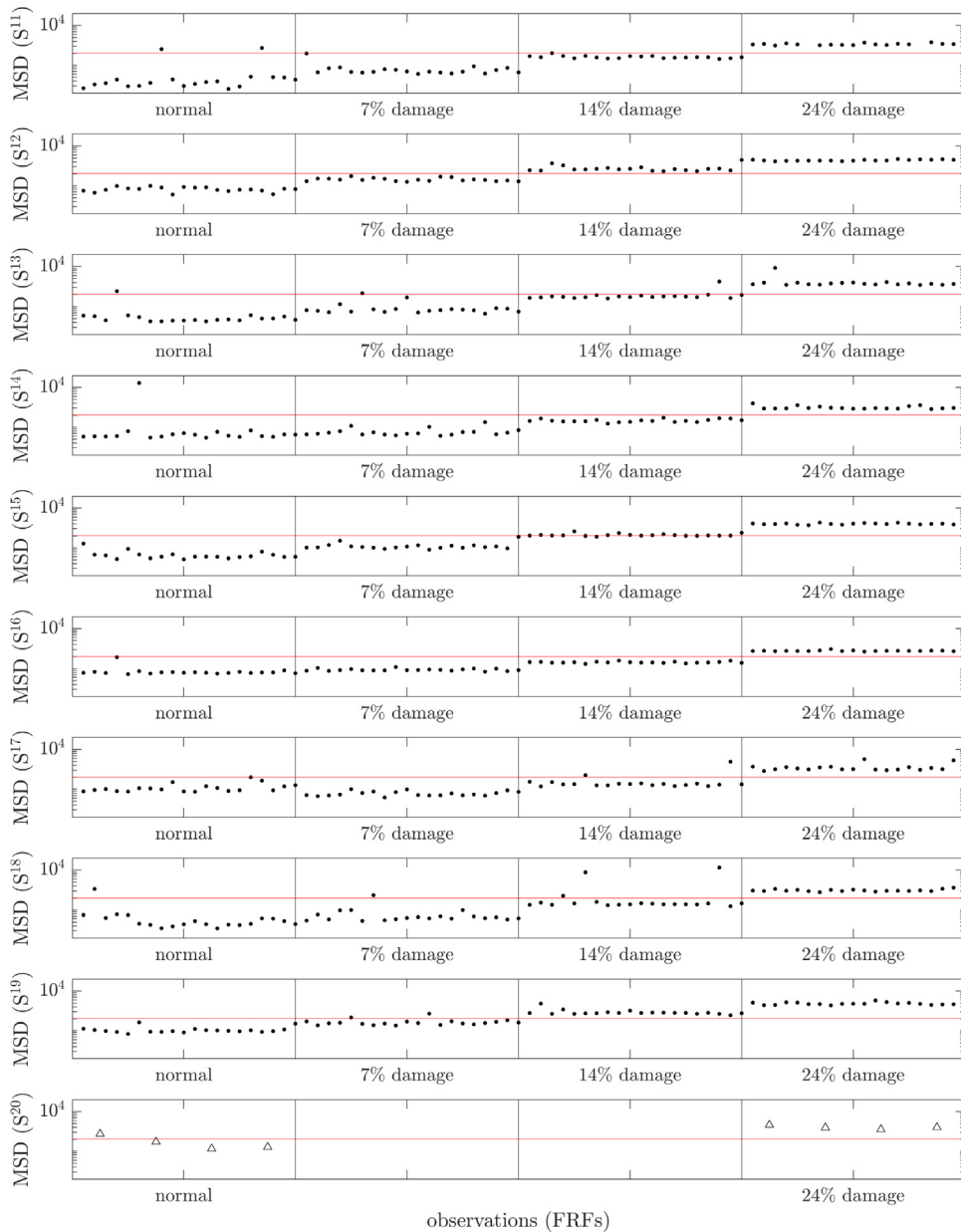


Fig. 6. MSD novelty index for the test-data FRFs, comparing members $\{S^{11}, \dots, S^{20}\}$ to the form. These members did not contribute training data to learn the form. Damage/normal conditions are separated by vertical lines. The \bullet markers are used for simulated members $\{S^{11}, \dots, S^{19}\}$, while \triangle markers are used for the test-rig, member S^{20} . The red line indicates the detection threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

age, leading to false negatives. On the other hand, some less frequent population behaviour could fail to be captured in the training-set (and the form) leading to false positives. As a result, in scenarios where $p(\Theta_i)$ is complex – which, unsurprisingly, proves to be common in practical examples – conventional SHM can no longer be applied to population data. In this case, more complex models of the form should be investigated – an alternative technique is proposed in the next case study.

4. Case study II: Beyond strongly-homogeneous populations – updating the form

In practice, the *strongly-homogeneous* case breaks down for several reasons; for example, variations in the operational ‘mass’ would be expected for offshore structures, such as oil rigs, due to changes in the *variable load*; this could include additional loading from the workforce, extracted materials, or helicopter landings. As a result, complex and multi-modal distri-

butions would be associated with the parameters that define population variation, $p(\Theta)$. Alternatively, when monitoring composite components (such as wind-turbine blades), while the mass should remain relatively consistent between structures, manufacturing tolerances are likely to lead to complicated distributions over the stiffness and damping parameters – before any potentially inconsistent boundary conditions are considered.

As well as variations that can be described in terms of the structural parameters, the form should be capable of modelling *operational* variations, that do not relate to damage. If these effects are ignored, benign events will be flagged as outlying, leading to false positives. For example, the data used to learn the form might change significantly upon operator involvement/control, or during maintenance procedures. While these changes are important, they do not indicate damage (or important novelty for SHM); therefore, they should be captured by the form appropriately.

4.1. Wind turbine population data

Using measured data from an operational wind farm, a practical example of the form and the *variation* it should represent is introduced. The data were recorded from wind turbines owned by Vattenfall, using a Supervisor Control and Sensory Data Acquisition (SCADA) system [24,23]. For confidentiality reasons, information regarding the specific type, location, and number of turbines cannot be disclosed. The data were recorded from a *homogeneous population* of systems over a period of 125 weeks [25,23]. The mean value of the *power* produced and the measured *wind speed* are available over ten minute intervals. Through population-based SHM, the goal is to determine whether individual wind turbines within the farm are operating in a *permitted* normal state, or not.

In order to monitor the performance of turbines given the available data, as in previous work [25,23], the *power curve* method is used. Specifically, wind turbines are designed by manufacturers to have a specific relationship between the power produced and wind speed; therefore, deviation in the measured power curves can be used to monitor the performance of a turbine [25,23]. To visualise the method, power curves that would be considered normal (i.e. ‘good’) as well as ‘bad’ are presented in Fig. 7 – provided by Vattenfall [23]. Importantly, for the SCADA data presented here, information regarding the specific status of each wind turbine is unavailable; therefore, when labels are required, the power curves must be classified (somewhat) subjectively, according to engineering judgement, and a criteria provided by Vattenfall experts [23] – examples will be provided.

Considering the framework from Case Study I, the power curve presents an ideal candidate (functional) feature to learn the form, in order to monitor members within a population⁵. Various methods have been used to model the power curve in the literature [26,27]; in line with previous studies [25,23], as well of the first case study of this work, a method for GP regression will be used here.

4.1.1. Population variation in the power curve data

As the dataset contains 125 power curves from each turbine in the population (one curve per week), it would be unreasonable for an engineer to examine the entire dataset, to label measurements that correspond to normal operation. (In practice, the monitoring period will be longer than the data presented here, and there may be significantly more members in the population.) Therefore, given expert knowledge, it is desirable to categorise power curves from a *subset* of systems, for a subset of weeks, and use these data to learn a general model of the population as a whole (i.e. a form) – following the framework proposed in the strongly-homogeneous case.

An example of population data that should represent the *ideal* power curves is presented in Fig. 8a, corresponding to measurements from *three* members over three weeks. Given the feature space in Fig. 8a, a single GP regression could be learnt (as demonstrated in Case Study I) in the hope that it would represent the general population behaviour. Unfortunately, however, there are important variations across the population that are not represented in these (ideal) data.

For example, Fig. 8b plots a larger set of population data, over another three weeks and one additional turbine. *Curtailements* can now be observed in the shared feature-space; these effects correspond to the power-output of wind turbines being limited to 50% (or otherwise controlled) by the operator. Interventions like this can occur frequently, and for various reasons – including maintenance or power requirements – they do not (normally) represent damage or reduced performance. Therefore, while the data in Fig. 8b do not represent the ideal case, curtailments that are *known* to correspond to the normal operation should be included in the form – to prevent similar (future) activity from being flagged as outlying.

4.2. OMGP regression of the power curve as the population form

To model the data in Fig. 8b and approximate a population form, an overlapping mixture of GP regressions (OMGP) is proposed [10,28]. Unlike the single GP applied in Case Study I, the OMGP can automatically model multiple latent functions, to represent more complex behavior in the population in the feature space. A brief overview of the model is provided here; further details can be found in [10,28].

⁵ Conveniently, this functional feature should be *more general* across the population than the FRF, as it is less sensitive to certain parameters (such as those associated with the boundary conditions) which introduce (less important) variation between members for the form.

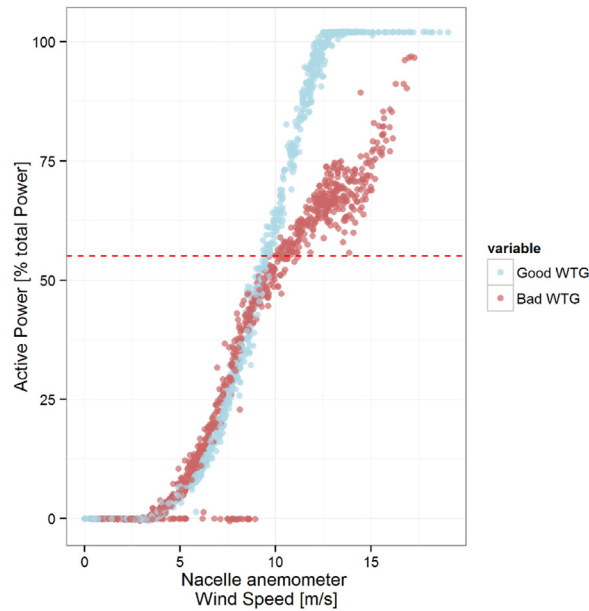


Fig. 7. An example of a 'good' and 'bad' power curve. WTG stands for wind turbine generator. Figure provided by Vattenfall [23].

The overlapping mixture of Gaussian processes (OMGP) assumes that there are K latent functions that describe the feature space,

$$y_i^{(k)} = \left\{ f^{(k)}(x_i) + \epsilon_i^{(k)} \right\}_{k=1}^K \quad (19)$$

i.e. each output observation y_i is found by evaluating one of these functions, with additive Gaussian noise (as with a single GP, where $K = 1$ (Eq. (1))). The function that generated each observation is unknown, therefore, a binary indicator matrix \mathbf{Z} is introduced (as a latent variable), which defines the specific function associated with each observation: if the entry $[\mathbf{Z}]_{ik}$ is non-zero, this implies that the i th data point was generated by the latent function k . Each observation can belong to one function only, thus, there is only one non-zero term per row in \mathbf{Z} [10].

Continuing with the Bayesian framework, priors are placed over the latent variables and functions,

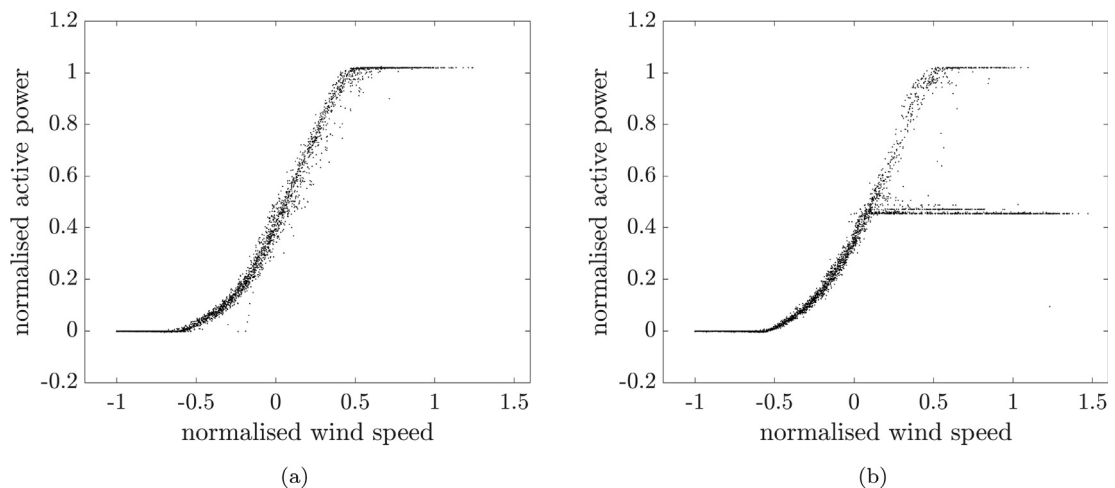


Fig. 8. Population power curve data. (a) Ideal data from four turbines over three weeks. (b) The ideal data, as well as data measured over three additional weeks; 50% power curtailments can be observed in the feature-space.

$$p(\mathbf{Z}) = \prod_{i=1, k=1}^{N, K} [\mathbf{\Pi}]_{ik}^{Z_{ik}} \quad (20)$$

$$f^{(k)}(x_i) \sim \mathcal{GP}(m^{(k)}(x_i), k^{(k)}(x_i, x_i)), \epsilon_i^{(k)} \sim \mathcal{N}(\mathbf{0}, \sigma_n^{(k)2}) \quad (21)$$

i.e. a multinomial distribution over the indicator matrix, such that $\mathbf{\Pi}$ is a histogram over the k th component for the i th observation, and $\sum_{k=1}^K [\mathbf{\Pi}]_{ik} = 1$ [15]; Gaussian distributions over the noise terms; and independent GP priors over each latent function [14]. Given knowledge of the expected functions (power curves), an appropriate prior mean $m^{(k)}(x_i)$ for the GPs is a scaled sigmoid, such that,

$$m^{(k)}(x_i; \boldsymbol{\alpha}_k) = \alpha_{k1} \frac{\exp(\alpha_{k2} x_i + \alpha_{k3})}{1 + \exp(\alpha_{k2} x_i + \alpha_{k3})} \quad (22)$$

$$\boldsymbol{\alpha}_k \triangleq \{\alpha_{k1}, \alpha_{k2}, \alpha_{k3}\}$$

As in the first case study, $m^{(k)}(x_i)$ serves as an approximation of the expected functions only; the discrepancy between the measured data and the prior should be automatically modelled by each GP – this is visualised in the results. Again, the squared-exponential function is used as the kernel $k(x_i, x_i)$, defined in Section 2.2.1. This completes the collected hyperparameters for each latent function, $\boldsymbol{\theta} = \{l_k, \sigma_0^{(k)}, \sigma_n^{(k)}, \boldsymbol{\alpha}_k\}_{k=1}^K$.

As the computation of the posterior distribution $p(\mathbf{Z}, \{f^{(k)}\} | \mathcal{D})$ is intractable, methods for approximate inference must be applied – specifically, a variational inference and Expectation Maximisation (EM) scheme proposed in [29]. This strategy iteratively computes the *approximate* posterior, and optimises the hyperparameters of the model, while an (improved) lower bound on the marginal likelihood is maximised [10]; the learning scheme is outlined below.

4.2.1. E-step – variational approximation

During the E-step, the hyperparameters $\boldsymbol{\theta}$ are fixed; therefore, it is possible to compute an approximation of the true posterior $p(\mathbf{Z}, \{f^{(k)}\} | \mathcal{D})$ through variational inference. The variational approximation is denoted $q(\mathbf{Z}, \{f^{(k)}\})$, and the search is restricted to distributions that factorise, such that $q(\mathbf{Z}, \{f^{(k)}\}) = q(\mathbf{Z})q(\{f^{(k)}\})$ [10].

Firstly, assuming that $q(\{f^{(k)}\})$ is known (and, therefore, the marginals for each component, $q(f^{(k)}) = \mathcal{N}(f^{(k)}(\mathbf{x}) | \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)})$) it is possible to analytically maximise a lower bound on the (log) marginal likelihood $p(\mathbf{y} | \mathbf{x}, \boldsymbol{\theta})$, with respect to $q(\mathbf{Z})$ [10],

$$q(\mathbf{Z}) = \prod_{i=1, k=1}^{N, K} [\hat{\mathbf{\Pi}}]_{ik}^{Z_{ik}}, \text{ s.t. } [\hat{\mathbf{\Pi}}]_{ik} \propto [\mathbf{\Pi}]_{ik} \exp(a_{ik}) \quad (23)$$

$$a_{ik} \triangleq -\frac{1}{2\sigma_n^{(k)2}} \left(([\mathbf{y}]_i - [\boldsymbol{\mu}^{(k)}]_i)^2 + [\boldsymbol{\Sigma}^{(k)}]_{ii} \right) - \frac{1}{2} \log(2\pi\sigma_n^{(k)2})$$

i.e. \mathbf{x}, \mathbf{y} represent the collected training observations, $\{x_i, y_i\}_{i=1}^N$.

Similarly, assuming $q(\mathbf{Z})$ is known, it is possible to analytically define a variational approximation that maximises the same lower bound, with respect to the latent functions [10],

$$q(f^{(k)}(\mathbf{x})) = \mathcal{N}(f^{(k)}(\mathbf{x}) | \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}) \quad (24)$$

$$\boldsymbol{\Sigma}^{(k)} \triangleq \left(K_{\mathbf{xx}}^{-1(k)} + \mathbf{B}^{(k)} \right)^{-1}$$

$$\boldsymbol{\mu}^{(k)} \triangleq m^{(k)}(\mathbf{x}) + \boldsymbol{\Sigma}^{(k)} \mathbf{B}^{(k)} (\mathbf{y} - m^{(k)}(\mathbf{x}))$$

where $\mathbf{B}^{(k)}$ is a diagonal matrix (off-diagonals are zero) with elements,

$$\left\{ \frac{[\hat{\mathbf{\Pi}}]_{1k}}{\sigma_n^{(k)2}}, \dots, \frac{[\hat{\mathbf{\Pi}}]_{Nk}}{\sigma_n^{(k)2}} \right\} \quad (25)$$

To learn an optimal approximation of the posterior, $q(\{f^{(k)}\})$ and $q(\mathbf{Z})$ can be initialised from their priors (21), and q is then iteratively refined by alternating updates (23) and (24). Both updates are optimal with respect to the distribution that they compute, therefore, they are guaranteed to increase the lower bound on the (log) marginal likelihood [10]. The following (implementation-friendly) expression can be used to monitor the lower bound⁶,

⁶ Note, this is a improved lower bound, proposed in [10], which remains stable during M-steps of the learning scheme.

$$\begin{aligned}
\mathcal{L}_{VB} &= \sum_{k=1}^K \left(-\frac{1}{2} \left\| \mathbf{R}^{(k)\top} \setminus \left(\mathbf{B}^{(k)\frac{1}{2}} (\mathbf{y}^{(k)} - m^{(k)}(\mathbf{x})) \right) \right\|^2 - \sum_{i=1}^N \log \left(\left[\mathbf{R}^{(k)} \right]_{ii} \right) \right) \dots \\
&- \text{KL}(q(\mathbf{Z}) \| p(\mathbf{Z})) - \frac{1}{2} \sum_{i=1, k=1}^{N, K} [\hat{\Pi}]_{ik} \log \left(2\pi \sigma_n^{(k)^2} \right) \\
\mathbf{R}^{(k)} &\triangleq \text{chol} \left(\mathbf{I} + \mathbf{B}^{(k)\frac{1}{2}} K_{\mathbf{xx}}^{(k)} \mathbf{B}^{(k)\frac{1}{2}} \right) \\
\text{KL}(q(\mathbf{Z}) \| p(\mathbf{Z})) &\triangleq \sum_{i=1}^N \left(\sum_{k=1}^K [\hat{\Pi}]_{ik} \log \left(\frac{[\Pi]_{ik}}{[\hat{\Pi}]_{ik}} \right) \right)
\end{aligned} \tag{26}$$

The term $\text{KL}(q(\mathbf{Z}) \| p(\mathbf{Z}))$ is the KL-divergence between the approximate posterior $p(\mathbf{Z})$, and the corresponding prior. The term 'chol' indicates a Cholesky decomposition.

4.2.2. M-step – hyperparameter optimisation

In the E-step, the hyperparameters θ were fixed; therefore, an optimisation procedure with respect to θ can improve the likelihood of the model given the data (via. empirical Bayes, as with the conventional GP (11)). Following the learning scheme suggested in [10], the hyperparameters are estimated through an optimisation given the lower bound defined in (26),

$$\hat{\theta} = \underset{\theta}{\text{argmax}} \{ \mathcal{L}_{VB} \}, \tag{27}$$

Eq. (27) defines the M-step, which follows E-steps. E and M steps are alternated until convergence (in the final \mathcal{L}_{VB}).

4.2.3. Predictive equations

Having learnt the model, the OMGP can be used to predict latent variables/functions. The predictions can be used to inform diagnostics when utilising the form in population-based SHM.

The posterior predictive likelihood given some unseen inputs $\{\mathbf{x}_*, \mathbf{y}_*\} = \{\mathbf{x}_{*i}, \mathbf{y}_{*i}\}_{i=1}^M$ is,

$$\begin{aligned}
p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D}) &= \sum_{k=1}^K \mathcal{N}(\mathbf{y}_* | \boldsymbol{\mu}_*^{(k)}, \boldsymbol{\Sigma}_*^{(k)}) [\Pi]_{*k} \\
\boldsymbol{\mu}_*^{(k)} &\triangleq m^{(k)}(\mathbf{x}_*) + K_{\mathbf{x}, \mathbf{x}}^{(k)} \left(K_{\mathbf{xx}}^{(k)} + \mathbf{B}^{(k)-1} \right)^{-1} (\mathbf{y} - m^{(k)}(\mathbf{x})) \\
\boldsymbol{\Sigma}_*^{(k)} &\triangleq K_{\mathbf{x}, \mathbf{x}_*}^{(k)} - K_{\mathbf{x}, \mathbf{x}}^{(k)} \left(K_{\mathbf{xx}}^{(k)} + \mathbf{B}^{(k)-1} \right)^{-1} K_{\mathbf{xx}_*}^{(k)} + \sigma_n^{(k)^2} \mathbf{I}_M
\end{aligned} \tag{28}$$

i.e. a mixture of Gaussians given the approximated posterior. The prior mixing proportion of new observations $[\Pi]_{*k}$ is a fixed hyperparameter, which weights class components equally. Interestingly, the predictive equations for the OMGP are similar to the conventional GP (6), however, the noise component ($\mathbf{B}^{(k)-1}$) is scaled, according to $[\hat{\Pi}]_{ik}^{-1}$ [10]; this essentially weights the contribution of each observation in \mathcal{D} to its posterior predictive component in the mixture.

Another useful prediction categorises observations according to the most likely component k . For the training data in \mathcal{D} , this is simply the *maximum a posteriori* (MAP) estimate, given the approximated posterior (23),

$$\hat{k}_i = \underset{k}{\text{argmax}} \{ [\hat{\Pi}]_{ik} \} \tag{29}$$

For a set of test-data (i.e. weekly power curves $\{\mathbf{x}_*, \mathbf{y}_*\}$), the posterior predictive class component is,

$$p(k_* | \mathbf{x}_*, \mathbf{y}_*, \mathcal{D}) = \frac{p(\mathbf{y}_* | \mathbf{x}_*, k_*, \mathcal{D}) p(k_*)}{p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D})} \tag{30}$$

where the denominator was defined in (28), and the numerator is,

$$\begin{aligned}
p(\mathbf{y}_* | \mathbf{x}_*, k_*, \mathcal{D}) p(k_*) &\triangleq \mathcal{N}(\mathbf{y}_* | \boldsymbol{\mu}_*^{(k_*)}, \boldsymbol{\Sigma}_*^{(k_*)}) [\Pi]_{*k_*} \\
\hat{k}_* &= \underset{k_*}{\text{argmax}} \{ p(k_* | \mathbf{x}_*, \mathbf{y}_*, \mathcal{D}) \}
\end{aligned} \tag{31}$$

To summarise, the OMGP model automatically finds K latent functions, given a set of *unlabelled* input and output data; this is achieved using a variational approximation, to construct a (corrected) lower bound on the marginal likelihood (26), and then iteratively maximising this bound via. Eqs. (23), (24), and (27).

4.2.4. The power curve form

When modelling the form given the data in Fig. 8b, the number of latent functions is $K = 2$, as prior knowledge informs us that there are two key characteristics in the feature space: data that represent the ideal power curve, and data that (likely) represent an operator limiting the turbine to 50% power, leading to curtailments.

The resulting OMGP representation of the form is shown in Fig. 9a; the model has automatically found two distinct latent functions, capturing the 'ideal' power curve from the population data ($k = 1$), as well as the curtailment behaviour ($k = 2$) – both of which have been assumed to represent (acceptable) normal conditions. To reiterate, only a *subset* of turbines from the wind farm contributed data to learn the form shown in Fig. 9a, which is used as a general representation of all systems in the population. As in Case Study I, Fig. 9a demonstrates that the GP successfully models the residual between the prior belief, encoded in the GP via the mean function (22), and the measured data.

For comparison, a conventional (single) GP was learnt for the same data, shown in Fig. 9b; clearly, such a form is a poor representation of the population feature-space, and would mask variations due to damage, leading to false negatives during the monitoring regime. Additionally, data corresponding to power curtailments (which are assumed to represent an acceptable normal condition) are likely to be indicated as outliers.

4.3. Results & discussion

To demonstrate the OMGP as a diagnostic (or performance-monitoring) tool for population-based SHM, the form is compared to test-data from *all* turbines within the population. As such, the form is being treated as a general model, and used to make predictions across the wind farm.

The form is now a mixture of GPs, therefore, novelty is assessed via the posterior predictive likelihood $p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D})$ (28), rather than the MSD of Eq. (18). This is a useful indication of novelty, as the class component k has been marginalised out from the predictive equations. In other words, it considers the complete mixture model, including multiple components. Measuring deviation from the complete mixture model through the MSD is less interpretable, as it considers the (normalised) distance to one component only. For the 125 (weekly) power curves from *all* turbines (other than those used for training), the likelihoods were calculated and ranked; then, the *most likely* (inlying) and *most unlikely* (outlying) weeks were extracted from all turbines across the farm.

To visualise diagnostics, Fig. 10 plots four example power curves, sampled from the inlying and outlying groups over the population. An example of data that appear significantly different to the form are shown in Fig. 10b and d; clearly these data are outlying, and do not correspond the ideal case. Specifically, Fig. 10d most likely corresponds to a turbine that is regularly *inactive*, while Fig. 10b resembles a *bad* power curve, according to the examples provided by Vattenfall (Fig. 7). On the other hand, inlying examples are shown in Fig. 10a and c; these data present a high likelihood (28) given the population form. As expected, the power curves resemble the permitted normal-conditions, shown in Fig. 9a.

Importantly, the form can also be used to classify the inlying data, according to (31). For example, Fig. 10a plots inlying data that were classified as *ideal* ($\hat{k}_* = 1$), and Fig. 10c plots an inlying example of 50% limited power ($\hat{k}_* = 2$).

4.3.1. Control chart example

To demonstrate the OMGP form for diagnostics via. control charts (as in Case Study I), weekly power curves were manually labelled for *one* turbine within the population. Data were labelled as *good* or *bad* by visual inspection and employing the criteria provided by Vattenfall experts [23]. As in previous work [23], an additional *ambiguous/unidentified* group is included, corresponding to power curves for which the *good* or *bad* classification – based simply on visual inspection – is debatable [25].

The turbines presented in these tests did not contribute training data to learn the form, so predictions should test the generalisation of the model across the population. The *negative* log predictive likelihood – $-\log p(\mathbf{y}_* | \mathbf{x}_*, \mathcal{D})$ (28) is used here as a measure of novelty, and the corresponding control chart is presented in Fig. 11. The detection threshold is defined according to the bootstrap-sampling scheme proposed in Case Study I (calculating the predictive likelihood, rather than the MSD) and represents the 95.45% (two-sigma) threshold.

In Fig. 11, the negative likelihood is a good indication of novelty, when comparing population data to the OMGP form. All observations that were classified as *good* (following engineering judgement) are inlying – this is true for both ideal ($\hat{k}_* = 1$) and 50% limited ($\hat{k}_* = 2$) examples. Additionally, the majority *bad* power curves are flagged as outlying.

For the *ambiguous* category, outlier analysis is less conclusive. To investigate this further, an example of an *inlying* ambiguous case is presented in Fig. 12a, while an *outlying* example is presented in Fig. 12b. Both power curves are considered ambiguous (via. engineering judgement) due to zero-power curtailment behavior, which can be observed in the feature-space. For the example corresponding to an inlying likelihood (Fig. 12a), only a few zero-curtailment observations are present; therefore, the feature space is visually similar to the ideal case, somewhat justifying the *inlying* classification by the form. For the outlying example (Fig. 12b), the zero-power curtailment is more significant, which, unsurprisingly, leads to higher measures of novelty.

Considering Fig. 12, the outlier analysis in Fig. 11 is intuitive. If desired, the zero-power curtailment could be included in as an additional latent function in the OMGP (i.e. $K = 3$). This is an opportunity to *update* the form with another permitted

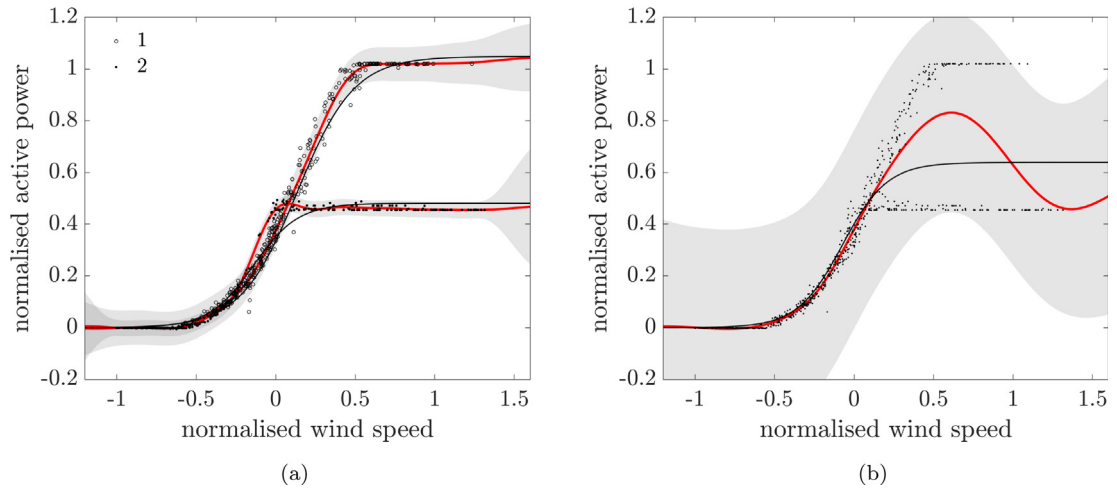


Fig. 9. Gaussian process regression of the power curve as the population form (a) Overlapping mixture of GPs (b) conventional GP regression. The red line indicates the posterior predictive mean, $E^{(k)}[y^*]$, and the shaded area indicates a $3\text{-}\sigma$ credible region, corresponding to the variance, $V^{(k)}[y^*]$. The black line plots the mean function of the GP prior $m^{(k)}(x_i)$.

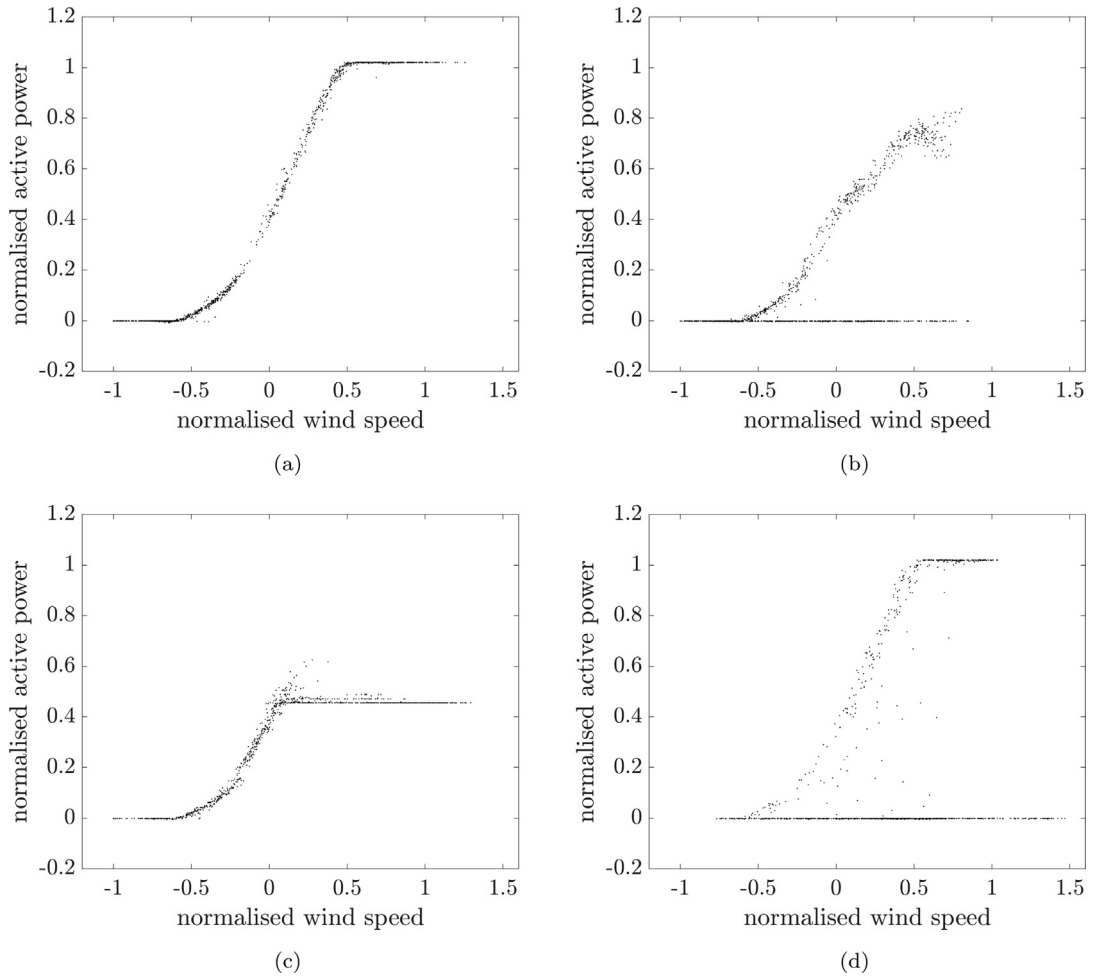


Fig. 10. Examples of power curves (weekly) that were (a) similar to the OMGP modelled form (b) significantly different to the OMGP form.

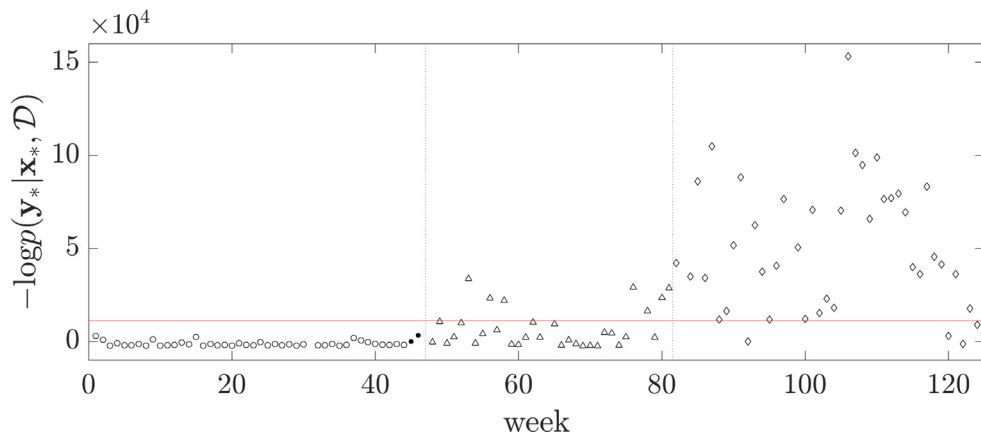


Fig. 11. Negative log likelihood of the OMGP form as a measure of novelty. \circ markers represent *good* power curves (i.e. ideal, $\hat{k}_i = 1$) and \bullet markers *good* power curves, 50% power curtailments ($\hat{k}_i = 2$); \triangle markers represent the *ambiguous/unidentified* class, and \blacklozenge markers represent *bad* power curves.

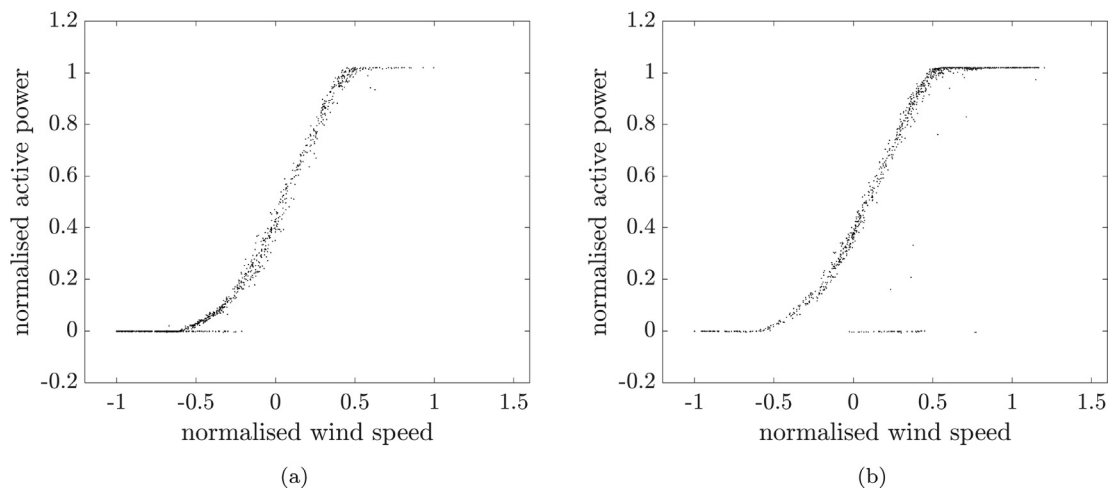


Fig. 12. Power curves classified as *ambiguous* according to engineering judgment. (a) an *inlying* example, according to the likelihood given the form (b) an *outlying* example, according to the likelihood given the form.

normal condition – to prevent zero-power behavior across the wind farm from being flagged as outlying. Clearly, for this to be appropriate, the zero-power curtailment must correspond to an accepted normal condition.

5. Concluding remarks

Population-based SHM considers that valuable information might be transferred, in some sense, between groups of *similar* structures. In view of this, the concept of the *form* has been introduced, used to represent a population of nominally identical systems. In two case studies, a statistically-modelled form was used to achieve novelty detection across a simulated population, and measured data recorded from an operational wind farm. In these examples, Gaussian process models were used to learn functional features as the form; however, the choice of model (and feature space) is flexible and application dependent. Importantly, the form was trained using normal-condition data recorded from *subset* of the total population only – this information is used in an attempt to learn a *shared* model, to represent the *general* population behaviour. Finally, novelty detection was achieved through comparisons between population (test) data and the form – via. the Mahalanobis squared-distance and likelihood measures.

5.1. Alternative technologies

The forms in this work (modelled by a standard Gaussian process, as well as a mixture of Gaussian process regression models) are just two methods to represent shared information between structures; alternative technologies should be inves-

tigated within the same general framework. More specifically, by considering population data in an alternative feature space (or latent space), various tools for pattern recognition become naturally applicable.

As discussed, if the population measurements are pointwise multivariate, rather than functional, methods for density estimation (such as Gaussian Mixture models, or Dirichlet Process mixture models [29]) might prove appropriate to represent the population form. Alternatively, by considering data from different structures in a shared latent space, methods for *transfer learning* [30] and *multi-task learning* [31] become relevant. These ideas of *transfer* are explored through domain adaptation [30] in Part III of this series [2]; however, the development of a *latent space* form representation is perhaps more applicable to multi-task learning frameworks [31] – where supervised training data are shared across several structures (i.e. domains) to improve the predictive model(s) (i.e. the form, or *tasks*). It is important to consider, however, that the model must still approximate the variance associated with the population Form.

Development of the theory of PBSHM will now proceed to the case of *heterogeneous populations* in Parts II and III of this sequence [1,2].

CRedit authorship contribution statement

L.A. Bull: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **P.A. Gardner:** Methodology, Writing - review & editing. **J. Gosliga:** Methodology, Writing - review & editing. **T.J. Rogers:** Methodology, Software. **N. Dervilis:** Methodology, Data curation, Writing - review & editing, Supervision. **E.J. Cross:** Methodology, Writing - review & editing. **E. Papatheou:** Data curation, Methodology. **A.E. Maguire:** Data curation. **C. Campos:** Data curation. **K. Worden:** Conceptualization, Data curation, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) through grant references EP/R003645/1, EP/R004900/1, EP/R006768/1, and EP/S001565/1.

References

- [1] J. Gosliga, P. Gardner, L. Bull, N. Dervilis, K. Worden, Foundations of population-based structural health monitoring, part II: Heterogeneous populations – graphs, networks and communities. Submitted to, Mech. Syst. Signal Process. 148 (2021) 107144.
- [2] P. Gardner, L. Bull, J. Gosliga, N. Dervilis, K. Worden, Foundations of population-based structural health monitoring, part III: Heterogeneous populations – mapping and transfer. Submitted to, Mech. Syst. Signal Process. 149 (2021) 107142.
- [3] L. Bull, P. Gardner, J. Gosliga, T. Rogers, M. Haywood-Alexander, N. Dervilis, E. Cross, K. Worden. Towards population-based structural health monitoring, Part I: Homogeneous populations and forms, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020..
- [4] J. Gosliga, P. Gardner, L. Bull, N. Dervilis, K. Worden. Towards population-based structural health monitoring, Part II: Heterogeneous populations and structures as graphs, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020b..
- [5] L. Bull, T. Rogers, N. Dervilis, E. Cross, K. Worden. A Gaussian process form for population-based structural health monitoring, in: proceedings of the 13th International Conference on Damage Assessment of Structures (DAMAS 2019). Springer, 2019..
- [6] J. Gosliga, P. Gardner, L. Bull, N. Dervilis, K. Worden. Towards population-based structural health monitoring, Part III: Graphs, networks and communities, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020c..
- [7] P. Gardner, K. Worden. Towards population-based structural health monitoring, Part IV: Heterogeneous populations, matching and transfer, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020..
- [8] K. Worden. Towards population-based structural health monitoring, Part VI: Structures as geometry, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020..
- [9] W. Lin, K. Worden, E. Cross. Towards population-based structural health monitoring, Part VII: EoV fields: environmental mapping, in: Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX, 2020..
- [10] M. Lázaro-Gredilla, S.V. Vaerenbergh, N. Lawrence, Overlapping mixtures of gaussian processes for the data association problem, Pattern Recogn. 45 (2012) 1386–1395.
- [11] C. Farrar, K. Worden, Structural Health Monitoring: a Machine Learning Perspective, John Wiley and Sons, 2012.
- [12] Plato. Protagoras and Meno. Penguin Classics, 2005..
- [13] Plato. The Republic. Penguin Classics, 2007..
- [14] C. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, The MIT Press, 2005.
- [15] A. Gelman, H. Stern, J. Carlin, D. Dunson, A. Vehtari, D. Rubin, Bayesian Data Analysis, Chapman and Hall/CRC, 2013.
- [16] K. Murphy, Machine Learning: A Probabilistic Perspective, MIT press, 2012.
- [17] D. MacKay. Bayesian model comparison and backprop nets, in: Advances in Neural Information Processing Systems, pages 839–846, 1992..
- [18] K. Worden, G. Tomlinson, Nonlinearity in Structural Dynamics: Detection, Identification and Modelling, CRC Press, 2000.
- [19] K. Worden, G. Manson, N. Fieller, Damage detection using outlier analysis, J. Sound Vib. 229 (2000) 647–667.
- [20] L. Ljung, System Identification: Theory for the User, Prentice-Hall PTR (1987).
- [21] R. Clark, Control System Dynamics, Cambridge University Press, 1996.
- [22] J. Quiñero-Candela, C. Rasmussen, A unifying view of sparse approximate gaussian process regression, J. Mach. Learn. Res. 6 (2005) 1939–1959.
- [23] E. Papatheou, N. Dervilis, A. Maguire, C. Campos, I. Antoniadou, K. Worden, Performance monitoring of a wind turbine using extreme function theory, Renewable Energy 113 (2017) 1490–1502.
- [24] W. Yang, R. Court, J. Jiang, Wind turbine condition monitoring by the approach of SCADA data analysis, Renewable Energy 53 (2013) 365–376.

- [25] E. Papatheou, N. Dervilis, A. Maguire, I. Antoniadou, K. Worden, A performance monitoring approach for the novel Lillgrund offshore wind farm, *IEEE Trans. Industr. Electron.* 62 (2015) 6636–6644.
- [26] T. Ouyang, A. Kusiak, Y. He, Modeling wind-turbine power curve: a data partitioning and mining approach, *Renewable Energy* 102 (2017) 1–8.
- [27] V. Thapar, G. Agnihotri, V. Sethi, Critical analysis of methods for mathematical modelling of wind turbines, *Renewable Energy* 36 (2011) 3166–3177.
- [28] C. Tay, C. Laugier, Modelling smooth paths using Gaussian processes, in: *Proceedings of the International Conference on Field and Service Robotics*, Chamonix, France, 2007..
- [29] D. Blei, A. Kucukelbir, J. McAuliffe, Variational inference: a review for statisticians, *J. Am. Stat. Assoc.* 112 (2017) 859–877.
- [30] P. Gardner, X. Liu, K. Worden, On the application of domain adaptation in structural health monitoring, *Mech. Syst. Signal Process.* 138 (2020) 106550.
- [31] Y. Zhang, Q. Yang, An overview of multi-task learning, *National Sci. Rev.* 5 (1) (2018) 30–43.