
Book Section:
Gardner, P., Bull, L.A., Gosliga, J. et al. (2022) Population-based structural health monitoring. In: Cury, A., Ribeiro, D., Ubertini, F. and Todd, M.D., (eds.) Structural Health Monitoring Based on Data Science Techniques. Structural Integrity (21). Springer Cham, pp. 413-435. ISBN: 9783030817152.

https://doi.org/10.1007/978-3-030-81716-9_20

Population-Based Structural Health Monitoring

Paul Gardner[1], Lawrence A. Bull[1], Julian Gosliga[1], Nikolaos Dervilis[1], Elizabeth J. Cross[1], Evangelos Papatheou[2], and Keith Worden[1]

[1] Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 SJD, UK
[2] College of Engineering Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK
p.gardner@sheffield.ac.uk; l.a.bull@sheffield.ac.uk; j.gosliga@sheffield.ac.uk; n.dervilis@sheffield.ac.uk; e.j.cross@sheffield.ac.uk; e.papatheou@exeter.ac.uk; k.worden@sheffield.ac.uk

Abstract. One of the dominant challenges in data-based Structural Health Monitoring (SHM), is the scarcity of measured data corresponding to different damage states of the structures of interest. A new arsenal of advanced technologies is described here that can be used to solve this problem. This new generation of methods is able to transfer health inferences and information between structures in a population-based environment – Population-based SHM (PBSHM).
In the category of *homogenous* populations (sets of nominally-identical structures), the idea of a Form can be utilised, as it encodes information about the ideal or typical structure, together with information about variations across the population.
In the case of sets of different structures and thus *heterogeneous* populations, technologies of transfer learning are described as a powerful tool for sharing inferences, that is also applicable in the homogenous cases. In order to avoid negative transfer and assess the likelihood of a meaningful inference, an abstract representation framework for spaces of structures, will be analysed as it can capture similarities between structures via the framework of graph theory.
This chapter presents and discusses all of these very recent developments and provides illustrative examples.

Keywords: Population-based Structural Health Monitoring (PBSHM), Machine Learning, Graph Theory, Complex Networks, Transfer Learning, Forms.

1 Population-based Structural Health Monitoring

Despite significant successes in data-based approaches to structural health monitoring (SHM) [1], several limitations have prevented wide-scale adoption of these techniques in industry. One of these limitations, that prevents data-based approaches from progressing beyond novelty detection, is a scarcity of measured

data corresponding to the damage-states of interest for the structures in question. This sparsity of labelled health-state data means that supervised (and even semi-supervised) techniques are limited in their effectiveness, unable to classify observations on a structure that correspond to a health state not previously seen on the structure (unless inspections are performed for the particular observation in question). As a consequence, conventional data-based approaches that are developed for individual structures are often limited in industrial applications to performing novelty detection (in an absence of labelled health-state data), where these techniques typically are not only sensitive to damage, but detect novelty for a variety of reasons, such as due to confounding influences and other benign effects [2].

In the light of these challenges, population-based structural health monitoring (PBSHM) [3, 4, 5, 6, 7], provides a variety of tools that seek to expand the available data for performing SHM, by considering observations from a population of structures. By utilising data from multiple structures, observations of health-states from across the population can be shared in diagnosing different members of the population. A population-based viewpoint therefore overcomes problems associated with a scarcity of health-state (or usually, *damage-state*) labelled data, and enables diagnostic predictions from the start of an SHM campaign. This chapter introduces key concepts for PBSHM, such as population types and the tools most applicable for each type, with the focus of this chapter being on *learning* for PBSHM.

It is helpful at this stage to provide an illustration of a typical industrial setting for PBSHM. Imagine a scenario in which an asset manager of a wind farm is interested in performing SHM for each wind turbine in the farm, depicted in Figure 1. Each wind turbine in the farm is of the same model type and can be considered nominally identical – this type of population is termed a *homogeneous population* [3, 4, 5]. During the complete operational phase each structure may transition from its normal operating condition to a different health-state; however, it is unlikely that any one turbine will observe *all* health-states of interest to the asset manager (particularly as these structures are designed for low failure rates). The lack of observed labelled health-states for each wind turbine means that conventional data-based SHM is limited to novelty detection. In fact, even observing the complete normal condition for any one wind turbine may not be possible for a variety of reasons, such as local differences in weather, local interactions between structures, and different operational patterns. Despite the fact that some wind turbines may have limited or no labelled health-state data, the asset manager is still tasked with maintaining and monitoring the complete population. The asset manager therefore requires a population-based approach to SHM, where the information across the population is used to create a machine learner that will both generalise across the population and will allow label information to be transferred to any wind turbine in the farm, allowing robust health diagnostics for all wind turbines in the population. Figure 1 demonstrates this process, where data from across the farm are mapped into some space where a data-based model can be constructed utilising the population-level information.

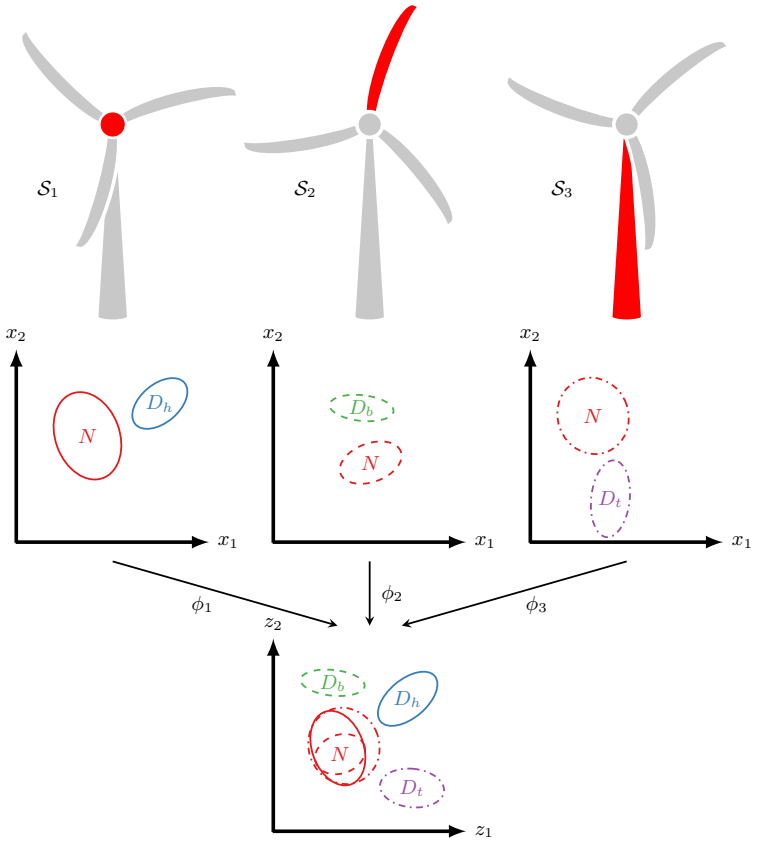


Fig. 1. A typical population-based structural health monitoring scenario. The scenario depicts a population of wind turbines ($\{\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3\}$) where different types of damage have been observed for each wind turbine over their operational phase (N - normal condition, D_h - damage to the rotor hub, D_b - damage to a blade, D_t - damage on the tower). Population-based SHM methods seek to define some model that captures information from across the population, generally via some mapping ϕ.

The above example describes a scenario involving a *homogeneous population* where the structures are nominally identical. A second category of population also exists, termed a *heterogeneous population*, where every structure in the population is different for various reasons, broadly categorised as geometric, material and topological differences [3, 4, 5]. Staying with the wind-farm illustration, imagine the asset manager is tasked with overseeing multiple wind farms situated around the world, with each farm containing wind turbines of different model types. A population-based approach can be extended to consider a larger population, covering all the wind turbines in the portfolio. However, more care must be taken in scenarios where the differences between structures in a population are large. For example, the asset manager must consider the physics of the

structures that are grouped into a population; are there common failure types that each turbine will experience such that labelled data can be shared, are there sub-systems or components that are common across the population? To answer these questions it will be necessary to demonstrate an abstract representation framework for spaces of structures, as this will allow an engineer to quantify and capture similarities between structures via the framework of graph theory.

A key component of population-based SHM concerns how learning algorithms are constructed in a population setting. Clearly, data from one structure cannot naively be used to classify data from another structure, without some mapping to harmonise datasets, as the generative distributions from each member of the population will be different. In the context of population types, this chapter outlines two key technology types for performing population-based SHM, the concept of a *Form* and *transfer learning*. Each are suited to different problem types in PBSHM, which are discussed within the chapter. Briefly, a *Form* seeks to capture the essence of a population, typically by defining a data-based model that captures the expected normal condition as well as the variability across the population. By contrast, transfer learning seeks to leverage information from source structures where label information is known and *transfer* this knowledge via some mapping to partially or unlabelled target structures. These types of learning are outlined in the context of each population type along with illustrative examples.

2 Homogeneous Populations

2.1 The concept of a Form

As discussed, homogeneous populations are groups of structures that can be considered *nominally identical* [3]. Some examples include same-model vehicle fleets, or turbines within a wind farm. In this special case, a general and *shared* representation – referred to as the *Form* – may be be used to monitor the collected group of systems.

The concept of the *Form* for PBSHM is motivated by the work of Plato. Initially in Meno [8] and later The Republic [9], Plato considers Forms to be the essence of things, existing as abstract entities: eternal, immutable, and representative of the highest level of reality. Ordinary objects derive their nature and properties by 'participating' in the Forms. For example, all cats in the world are recognisable as such because they participate in the Form of *cat*.

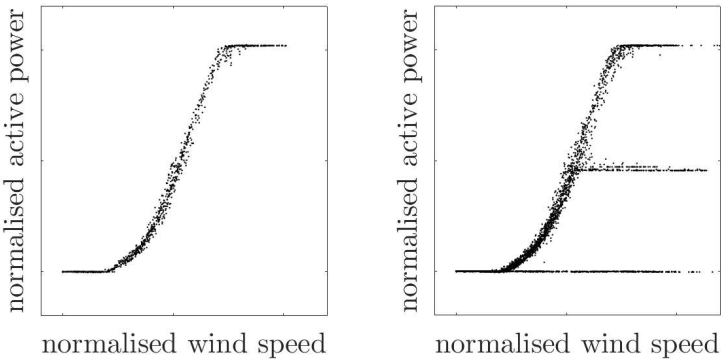
To apply Plato's Form to a group of structures for monitoring purposes, an extension is needed: not only to capture the essence of things, but the extent of variations in their participants [3]. For example, consider a specific model of vehicle; one could argue that the essential nature of the vehicle is captured in the complete design specification; in reality, variations will occur over the production-run: manufacturing tolerances, changes in operating environments, as well as other inconsistencies.

Herein, the term *form* will be used in a mathematical sense to denote a model, in some feature space, of an object (the uncapitalised *form* distinguishes

the model from the conceptual Form). The model attempts to capture the two ingredients of the extended Form: the *essential nature* of the object and the *variations encountered* when the object is embodied in the real world. The *object of interest* need not be the structure itself, but rather a feature or measurement vector – which represents the structure for SHM purposes. The feature, therefore, is part of the description of the form.

A motivating example: wind turbine power curves For a specific model of turbine, the *power curve* captures the relationship between wind speed and power output; the associated function can be used as a indicator of performance [10, 11]. For a wind farm consisting of identical turbines, this trend should be relatively consistent across the group. Variations in the power curve can exist for an array of reasons; the results of operator control, or shadowing and wake effects from other turbines [12], for example.

Intuitively, the power-curve defines a convenient object (i.e. feature) to consider as the form for a wind-farm. This relationship captures the *essential nature* of power production, while also the *variations* across the group. To demonstrate, operational power-curve data (SCADA [11]) from a wind farm are presented in Figure 2a. A regression of these data should generalise to future measurements, given optimal power-generation for turbines within the farm.



(a) Optimal power curves: three turbines over three weeks.

(b) Curtailed power curves: seven turbines over nine weeks.

Fig. 2. Wind farm power curves (normalised).

In practice, however, only a subset of measurements or turbines would be approximated by a form modelled on the data in Figure 2a. This is because, in actuality, *power-curtailments* will appear as additional functional components; that is, further variations in the form. An example of operational data including curtailments is shown in Figure 2b; here, three trends can be observed in the

power data: (i) the ideal power curve (ii) ≈50%-limited output, and (iii) zero-limited output. Curtailments usually correspond to the output being controlled (or limited) by the operator for various reasons; e.g. responding to requirements of the electrical grid [13] or the mitigation of loading/wake effects [12]. As a result, the form object is multi-valued, differing significantly from the ideal curve. However, as these data capture important *variations* that are expected in practice, they should be useful to model a more complete form for the wind farm.

The power curve form as a mixture of Gaussian Processes As the functional feature (Figure 2b) is multi-valued, conventional regression would prove inappropriate for this expression of the form. The work in [14] proposes that an overlapping mixture of probabilistic regression models (Gaussian Processes) is used to approximate the power-curve relationship.

Specifically, the overlapping mixture of regression models (introduced by Lázaro-Gredilla et al. [15]) assumes that there are K latent functions to approximate the form,

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

i.e. the power y_i at each input x_i (wind speed) is found by evaluating one of K latent functions $f^{(k)}(x_i)$ with additive noise ϵ_i. From the power-curve data, it should be clear that an appropriate number of components is $K = 3$: (i) ideal, (ii) 50% limited, and (iii) zero power.

Labels to assign each observation $\{x_i, y_i\}$ to function k are unknown, so a latent variable is introduced to the model, $\boldsymbol{Z}$; this is a binary indicator matrix, such that $\boldsymbol{Z}[i,k] \neq 0$ indicates that observation i was generated by function k. There is only one non-zero entry per row in $\boldsymbol{Z}$ (each observation is found by evaluating one function only). Therefore, for N data, the likelihood of the model is [15],

$$p\left(\boldsymbol{y} \mid \left\{\boldsymbol{f}^{(k)}\right\}_{k=1}^{K}, \boldsymbol{Z}, \boldsymbol{x}\right) = \prod_{i,k=1}^{N,K} p\left(y_i \mid f^{(k)}(x_i)\right)^{\boldsymbol{Z}[i,k]} \tag{2}$$

A Gaussian Process (GP) is associated with each of the (three) latent functions $f^{(k)}$. Briefly, each GP can be described by its mean and kernel function [16], which can be specified for power-curve modelling; for details, refer to [14]. Unlike a conventional GP, the computation of the posterior distribution $p\left(\mathbf{Z}, \{f^{(k)}\} \mid \mathcal{D}\right)$ is intractable; thus, methods for approximate inference are implemented to infer the latent variables and functions – a variational inference and Expectation Maximisation (EM) approach was proposed in [17]. Additionally, input-dependant noise is approximated according to the scheme [18]. The resulting form is presented in Figure 3.

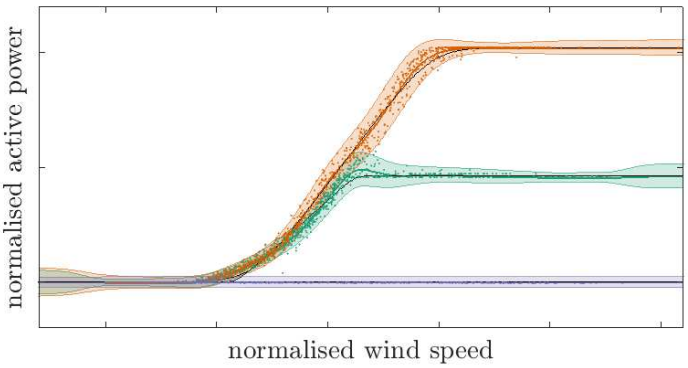
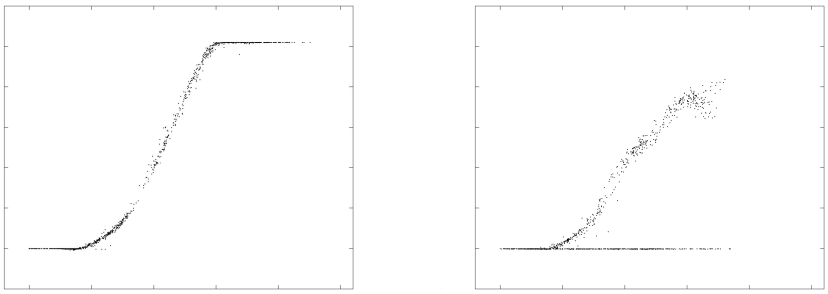


Fig. 3. The wind farm form as an overlapping mixture of Gaussian processes.

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

In the experiments presented in [3], a similar power-curve model was used to inform outlier analysis, by measuring the deviation of future data from the form (via the combined predictive-likelihood of the mixture model). Examples of (weekly) data (from across the wind farm) that appeared as inlying or outlying with respect to the form are shown in Figure 4. In other words, data that appear *likely* or *unlikely* when compared to the currently-modelled population form respectively. The examples are sampled at random from the most extreme inlying and outlying weeks of data in the test set [3].



(a) Inlying (high likelihood). (b) Outlying (low likelihood).

Fig. 4. Weekly datasets, compared to the form in [3].

Specifically, Figure 4b resembles a typical *sub-optimal* power curve [10]. On the other hand, the inlying example in Figure 4a resembles one of the permitted normal conditions associated with the form – in this case, the ideal curve.

Form difficulties: increased population variance As variation across individuals increases, variation in the population-data is also likely to increase; thus it becomes progressively difficult to approximate the *form*. In particular, when the underlying distributions of data vary dramatically between individuals, more involved techniques are required to infer a *shared* model. An idea presented in [19] suggests that dissimilar population data might be projected into a shared and more consistent space, where the form can then be inferred. These concepts align closely with those of *transfer learning*.

2.2 Transfer learning for homogeneous populations

Transfer learning – a branch of machine learning – provides an alternative viewpoint to the concept of a form. Rather than seeking to capture the essence and variation of a population, transfer learning seeks to leverage label information from source structures in aiding classification of unlabelled (or partially-labelled) target structures. Transfer learning is applicable across a variety of PBSHM scenarios, including homogeneous populations, where the data vary too greatly to be modelled by the form framework in the previous section, preventing a learner trained on one (or more) structure(s) from generalising to another.

Two objects are required to formally define transfer learning,

- A **domain** $\mathcal{D} = \{\mathcal{X}, p(X)\}$, consists of a feature space $\mathcal{X}$ and a marginal probability distribution $p(X)$ over the feature data $X = \{\mathbf{x}_i\}_{i=1}^{N} \in \mathcal{X}$, a finite sample from $\mathcal{X}$.
- A **task** $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$, consists of a label space $\mathcal{Y}$ and a predictive function $f(\cdot)$ which can be inferred from training data $\{\mathbf{x}_i, y_i\}_{i=1}^{N}$ where $\mathbf{x}_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$.

Using these objects, transfer learning between a single source domain and single target domain is defined as [20],

Definition 1. ***Transfer learning*** *is the process of improving the target prediction function $f(\cdot)$ in the target task $\mathcal{T}_t$ using knowledge from a source domain $\mathcal{D}_s$ and a source task $\mathcal{T}_s$ (and a target domain $\mathcal{D}_t$), whilst assuming $\mathcal{D}_s \neq \mathcal{D}_t$ and/or $\mathcal{T}_s \neq \mathcal{T}_t$.*

Within the field of transfer learning, *domain adaptation* arguably offers one of the most useful tools for PBSHM and is defined as,

Definition 2. ***Domain adaptation*** *is the process of improving the target prediction function $f(\cdot)$ in the target task $\mathcal{T}_t$ using knowledge from a source domain $\mathcal{D}_s$ and a source task $\mathcal{T}_s$ (and a target domain $\mathcal{D}_t$), whilst assuming $\mathcal{X}_s = \mathcal{X}_t$ and $\mathcal{Y}_s = \mathcal{Y}_t$, but that $p(X_s) \neq p(X_t)$ and typically that $p(Y_s \mid X_s) \neq p(Y_t \mid X_t)$.*

Domain adaptation is appropriate for scenarios where a classifier will not generalise across domains because of differences in the underlying data distributions, such as the example outlined in Figure 1. For this reason, the transfer learning methods demonstrated in this chapter are therefore all forms of domain adaptation.

To contextualise transfer learning for homogeneous populations, a case study is provided. The case study considers a special case of the homogeneous population type, when the source and target structure are exactly equivalent, i.e. they are the same structure. Even in this context, PBSHM provides a useful framework for overcoming challenges with data-based SHM; in this particular instance, the problem of structural repairs and how they change the underlying data distribution of a system. Structural repairs introduce modifications that change (even if locally) the mass, stiffness and damping of the structure, causing shifts in the underlying generating distributions and manifest as drift in the output feature space. This dataset shift means that a data-based model trained on pre-repair labelled data will not generalise to the post-repair structure because the distributions of the dataset in training and testing are not the same. The implication of this dataset shift is that a new labelling campaign would be required every time structural repairs are made. Instead, a PBSHM viewpoint can be taken by treating the two datasets as coming from a population of two homogeneous structures. The following case study considers two datasets from the same Gnat trainer aircraft, before and after the inspection panels have been removed and reattached (simulating a repair scenario).

The Gnat aircraft dataset was collected as part of an experimental campaign in which a network of uni-axial accelerometers were used to obtain transmissibility features (under white-noise excitation) from the starboard wing of a Gnat trainer aircraft *in situ* [21]. During the experiments, psuedo-damage was introduced into the structure by removing individual inspection panels – the locations of which are represented in Figure 5. The sensor network was designed such that each transmissibility path targeted a specific inspection panel (i.e. the transmissibility targeting panel 1 (P_1), denoted T_1, is computed from the reference accelerometer AR and response $A1$), with each transmissibility covering a frequency range of 1024-2048Hz containing 1024 spectral lines, with the magnitudes being utilised as the feature data. In the following analysis the feature data are seven stacked transmissibilities (i.e. $\mathbb{R}^{1024\times 7}$) covering panels $\{P1, P2, P4, P5, P7, P8, P9\}$ where the label space is $\mathcal{Y} \in \{1, 2, 4, 5, 7, 8, 9\}$ (with only the panels with a large surface area being considered in this analysis). For more details about the experiments, the reader is referred to [21].

The dataset simulates a repair scenario, because the experimental sequence (removing all panels one-by-one and replacing them on the structure with the same applied torque on the fasteners) was repeated twice. In fact, the maintenance process on the Gnat aircraft wing typically involves removing and reattaching inspection panels. In order to visualise the problem caused by the repair action, Figures 6 and 7 present two depictions of the dataset. Figure 6 demonstrates the changes in the first transmissibility path before (show in red) and

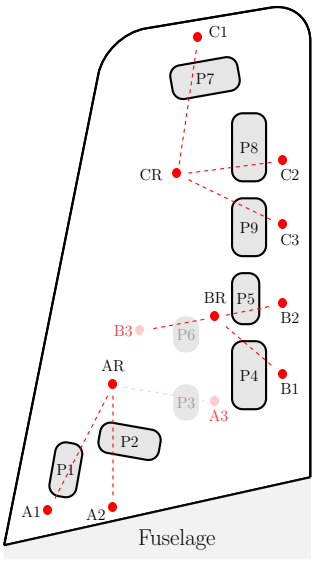


Fig. 5. A representative schematic of the Gnat aircraft starboard wing (not to scale), depicting inspection panel, accelerometer and transmissibility path locations. Recreated from [21].

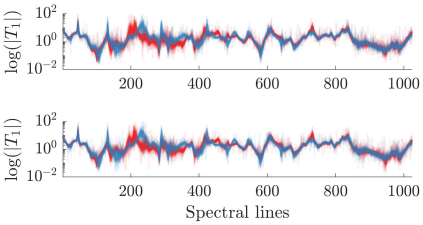


Fig. 6. A comparison of T_1 for the pre- (red) and post-repair (blue) scenarios; top and bottom sub-panels depict the normal condition and the removal of Panel One.

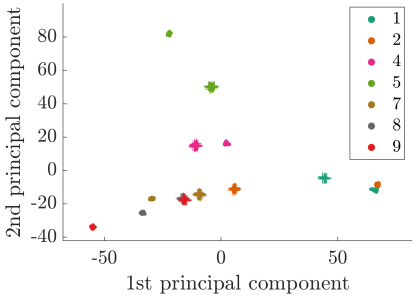


Fig. 7. Visualisation of the pre- and post-repair datasets; the first two principal components of the pre-repair (•) and post-repair (+) datasets.

after (depicted in blue) repairs had taken place for two health states, the undamaged normal condition (top sub-panel) and when Panel One has been removed (bottom sub-panel). It is clear from this figure that there are larger changes due to the repair action than due to damage (for example the large differences around 200 to 400 spectral lines between pre- and post-repair transmissibilities), and hence a classifier will not generalise from the pre- to post-repair scenarios. Figure 7 shows the first two principal components of the complete feature space, demonstrating that the data distributions have changed significantly between the pre- and post-repair structural states, and hence $p(Y_s, X_s) \neq p(Y_t, X_t)$ (where the pre-repair data has been denoted as the source domain, subscript $_s$, and the post-repair data is denoted as the target domain, subscript $_t$), so domain adaptation is applicable.

In order to harmonise the pre- and post-repair datasets, such that label information from the pre-repair dataset can be used to diagnose the unlabelled post-repair data, a domain adaptation algorithm has been applied. The algorithm, *metric-informed joint domain adaptation* [22], seeks to find a mapping $Z = KW \in \mathbb{R}^{(N_s+N_t \times k)}$, by learning some weight matrix $W \in \mathbb{R}^{(N_s+N_t \times k)}$ that

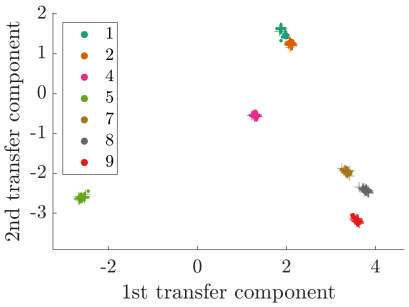


Fig. 8. Transfer components of the pre-repair (•) and post-repair (+) from the metric-informed joint domain adaptation approach.

Fig. 9. Comparison of classification accuracy given feature spaces with no transfer learning (Data and PCA) and the feature space after transfer learning (M-JDA).

projects a kernel matrix, formed from the joint dataset $K \in \mathbb{R}^{(N_s+N_t \times N_s+N_t)}$, onto a k-dimensional space. The weight matrix is inferred by minimising the distances between the joint distributions from the source and target data, formed as an optimisation problem that minimises the maximum mean discrepancy distance between the marginal and class conditional distributions (for more information about the algorithm, the reader is referred to [22]). The projected space, defined by a set of *transfer components*, where $k = 2$, is shown in Figure 8, where the algorithm has matched the joint distributions in the projected space. In this space, label data from the pre-repair state can be used to classify the post-repair data, transferring the label information. A k-nearest neighbour classifier was trained on the pre-repair data from three different feature spaces, the original transmissibilities (Data), the principal components of the transmissibilities (PCA) and the transfer components (M-JDA). The classification results in Figure 9 show that domain adaptation has allowed for label information to be successfully transferred to the target domain, signified by classification accuracies in the target domain of 100%, and that, because of differences in between the pre- and post-repair datasets, a classifier trained on either the pre-repair transmissibilities or principal components, does not generalise to the post-repair data.

3 Heterogeneous Populations

3.1 Transfer learning for heterogeneous populations

In the field of population-based SHM, heterogeneous populations provide a more complex set of challenges for transfer learning. The reason for this increased difficulty is that transfer learning assumes that there is some shared commonality between the source and target domains. Heterogeneous population push these assumptions towards their limits, and as structures in a population become more

dissimilar, the risk of negative transfer increases. The term *negative transfer* describes the scenario where transfer learning incorrectly maps information from one domain onto another, reducing the performance of the learner (discussed in more detail in Section 3.2). It is therefore important in heterogeneous populations to understand and quantify the level of similarity between structures, such that transfer learning is only attempted in contexts where transfer will be successful and beneficial. Later on in this chapter, an approach for assessing the similarities between structures (before attempting transfer learning) is introduced. Briefly, the approach converts structures into an abstract representation, called Irreducible Element models, where a graph theory framework can be used to quantity similarities. The remainder of this section looks at illustrating transfer learning in the context of heterogeneous populations via two example populations of n-storey buildings.

The first example considers a population of two n-storey structures, a three-storey structure where the feature data are labelled, denoted the *source* structure, and a four-storey target structure, where the feature data is unlabelled. The two structures form a heterogeneous population as they have different topologies – with their nominal geometries and material properties being the same. Each structure is modelled as a lumped-mass model, shown in Figure 10, where the spring stiffness between each floor is modelled as four springs in parallel. The SHM problem is that of locating damage, in the form of open cracks at one of the beams at a particular floor using lateral bending natural frequencies of the whole structure as features.

In this example, the PBSHM problem is that of transferring localisation labels from the three-storey structure to the four-storey structure. An interesting challenge arises in the context of a heterogeneous population when performing a localisation task between structures with different topology, namely that the labels spaces between the two structure are not exactly equivalent, termed label inconsistency. This phenomenon means that care must be taken when transferring information between members of a heterogeneous population. In this example, both structures have an undamaged condition ($Y = 1$), and can be damaged at floors one to three ($Y = \{2, 3, 4\}$ respectively). As a result, the complete label set from the three-storey structure can be transferred to the four-storey structure, where the algorithm should not try to pair data points relating to damage at the fourth floor ($Y = 5$) of the target structure with data from the source structure. This type of PBSHM problem is termed an $L + 1$-problem, as there is one more class label in the target domain than in the source domain (i.e. $Y_s \in \{1, 2, 3, 4\}$ and $Y_t \in \{1, 2, 3, 4, 5\}$. Furthermore, most domain adaptation techniques require that both the source and target feature spaces to be of the same dimension, such that negative transfer is minimised. In this example, the feature spaces are the first three bending natural frequencies such that both feature spaces are $\mathbb{R}^3$, with Figure 11 presenting a visualisation of the first two natural frequencies for each structure.

The approach taken to solve this particular $L + 1$-problem is outlined next. Firstly, an unsupervised clustering method (namely a Gaussian mixture model)

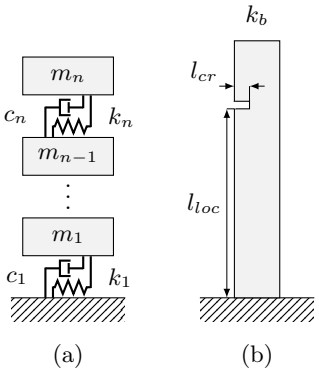


Fig. 10. Schematic of the n-storey building structures lumped-mass models, panel (a), and beam component, panel (b).

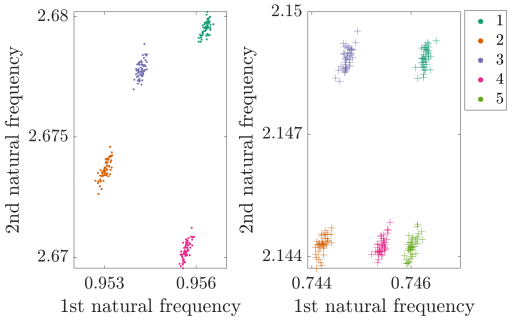


Fig. 11. A visualisation of the first two natural frequencies (in Hz) from the three- (•) and four (+) storey structures.

is used to identify and group the unlabelled target domain feature data. Once unlabelled target clusters are identified, each target cluster is removed iteratively from the domain adaptation training dataset. A mapping, in the form of $Z = KW \in \mathbb{R}^{(N_s+N_t \times k)}$ is subsequently identified from the (complete) source dataset to the particular target dataset (where one cluster has been removed); where the domain adaptation algorithm uses the maximum mean discrepancy as a cost function and $k = 2$. The algorithm then selects the mapping that produces the smallest distance between the source and target training datasets. This methodology is based on a naive form of manifold assumption, i.e. it is expected that the manifold of the source and target clusters is the 'same'. Figure 12 presents the 'optimal' mapping, where the correct target clusters were used in training, and the distance – in the form of a maximum mean discrepancy distance (MMD) – between the mappings with different target clusters removed. It can be seen that the 'optimal' mapping is selected by this approach, as the smallest MMD distance is produced when the target cluster corresponding to $Y = 5$ (i.e. damage at the fourth storey) is removed in training. A classifier trained on the source domain data in the transfer component space in Figure 12 can be shown to classify the target structure with 100% accuracy using a semi-supervised Gaussian mixture model. This case study demonstrates the challenges heterogeneous populations cause for transfer learning techniques, with care being needed in order to minimise negative transfer.

The second example involves a heterogeneous population comprised of a numerical physics-based model and an experimental structure, as displayed in Figure 13. The aim in this example is to transfer damage-extent label information from an unvalidated (and in this case a deliberately poor-performing) numerical model to unlabelled data from an experimental structure. This case demonstrates the flexibility of a PBSHM approach in which a variety of sources of label

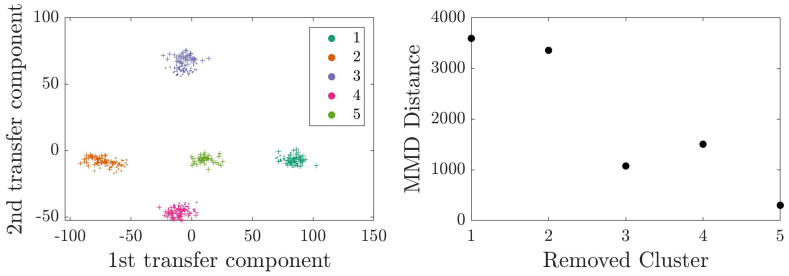


Fig. 12. Transfer learning results for the three- to four-storey example. Left panel displays the transfer components for the three- (•) and four-storey (+) structures, with the right panel demonstrating the maximum mean discrepancy (MMD) distances in the transformed space when one cluster has been removed from the target domain.

information can be utilised, and shows a significant advantage of the PBSHM viewpoint, namely that damage-state data can be generated in a cost-effective manner from physics-based models, even when computer model validation is challenging. The numerical model, constructed using the approach in Figure 10 as a three degree-of-freedom lumped-mass model, was formed using the measured dimensions of the experimental structure and with typical material properties that matched those from the structure.

Fig. 13. Experimental three-storey building structure [5].

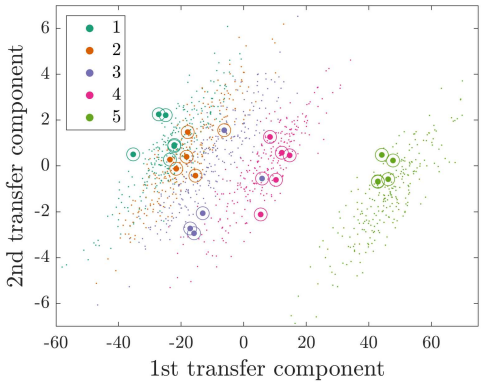


Fig. 14. The transfer components from the numerical (·) and experimental (• predicted label and ◯ true label) datasets.

The SHM problem was to classify the extent of damage, in the form of open cracks from 0mm to 20mm, where label $Y = 1$ denotes the undamaged condition, $Y = 2$ refers to a 5mm crack, $Y = 3$ for a 10mm crack etc. Again, it is reiterated

that the numerical model has been oversimplified such that the example demonstrates the effectiveness of domain adaptation in utilising physics-based models in labelling real world structures. The simplified physics-based model therefore reflects that physics-based models are challenging to validate in SHM contexts and may not fully agree with observational data, because of model form-errors. The feature data in this example were the first three lateral bending natural frequencies of the system.

Domain adaptation was applied such that a mapping in the form of $Z = KW \in \mathbb{R}^{(N_s+N_t \times k)}$ could be identified (where $k = 2$), using an MMD-based cost function [5]. The inferred mapping is visualised in Figure 14 where it can be seen that the numerical and experimental data have been aligned. A k-nearest neighbour classifier was trained using the numerical model data, both before and after the transfer mapping. The classifier trained on the untransformed data produced a testing accuracy in the target domain of 48%; this is compared to a testing accuracy 88% from the classifier trained on the transfer components. This result demonstrates the effectiveness and applicability of utilising both physics-based models and observational datasets within a PBSHM framework.

Both of these case studies highlight the potential of PBSHM beyond populations of nominally-similar structures. Of course, there are a number of research questions that are posed by considering the full extent of heterogeneous populations. It is therefore important to explore similarities between structures and datasets, and to form groupings of PBSHM problems that address each scenario, while monitoring the potential for negative transfer.

3.2 The problem of negative transfer

Negative transfer has been mentioned several times in this chapter because of the risk it poses to making inferences of health states in a PBSHM approach. As previously stated, negative transfer occurs when class data from a source domain has been incorrectly and confidently paired with class data from a target domain, i.e. Class One in the source domain is mapped onto Class Two in the target domain. Generally, negative transfer is any scenario where transfer learning reduces the performance on a classifier when compared to not performing any transfer. Negative transfer is a particular concern for a PBSHM viewpoint, as typically there are no labelled data points in the target domain that can be used to validate the inferred transfer mapping. As such, it is important to quantify how related the source and target structures are and to only perform transfer when positive transfer is likely. This is a significant and open research question, with the following section outlining one approach to identified the similarities between members of a population.

To illustrate the effect of negative transfer, the three- to four-storey example is reintroduced. The problem in this example was to transfer localisation labels from the three-storey source structure to the unlabelled four-storey target structure. The problem is an $L + 1$-problem, meaning that there is one more class label in the target domain than in the source domain, and that all other class labels exist in both domains. Figure 15 illustrates the identified mapping when

data corresponding to classes $Y_s \in \{1, 2, 3, 4\}$ and $Y_t \in \{1, 2, 4, 5\}$ were used to train the domain adaptation mapping – this gave the second smallest distance in Figure 15 and therefore was not selected as the optimal mapping, but will be useful for this discussion. Of particular note, negative transfer has occurred between the source data for class $Y = 3$ and the target data for class $Y = 4$. This is interesting, as source data for class $Y = 4$ were used in training, but the domain adaptation algorithm, given no label information about the target, has inferred a mapping that incorrectly pairs these classes. Negative transfer has also occurred between class $Y = 4$ in the source domain and $Y = 5$ in the target, which arises as the cost function in the domain adaptation approach only considers a mapping with the smallest distance between the two domains.

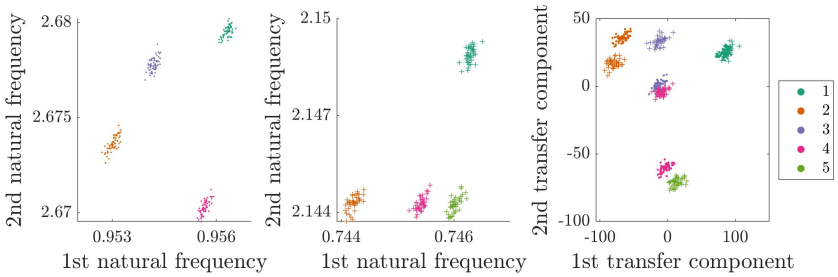


Fig. 15. An example of negative transfer between the three- and four-storey structures. The left panel shows the first two natural frequencies of the source domain data (•) used in training, the middle panel shows the first two natural frequencies of the training target data (+), and the right panel shows the inferred transfer components after domain adaptation, where negative transfer has occurred.

In this example, it should be clear that the risk of negative transfer was high, as the data from the source and target domains did not correspond to the same classes in training. However, it is worth noting that negative transfer can occur in scenarios where both the source and target training data do refer to the same classes, and where the two datasets are more dissimilar than can be accounted for given the assumptions and adaptation method (i.e. the type of mapping) used in the transfer learning algorithm. For PBSHM to be applied with confidence using transfer learning approaches, it will be important to estimate the probability of negative transfer from an algorithm, such that the risk of negative transfer is always a minimum, and if possible, zero.

3.3 Abstract representation framework for spaces of structures

The abstract representation mentioned previously in this chapter is a method for describing engineering structures systematically in a way that lends itself to comparison. This abstract representation focusses on three areas believed to be important for avoiding negative transfer in an SHM context: geometry,

material properties, and topology. This abstract representation is known as an *Irreducible Element* (IE) model. The Irreducible Elements which give the modelling approach its name describe the constituent parts of a given structure. In a simplified example describing a wind turbine, these elements could be considered to be the blades, the hub, the nacelle, the tower and the foundations (Figure 16). These elements possess attributes which describe their geometry and material properties.

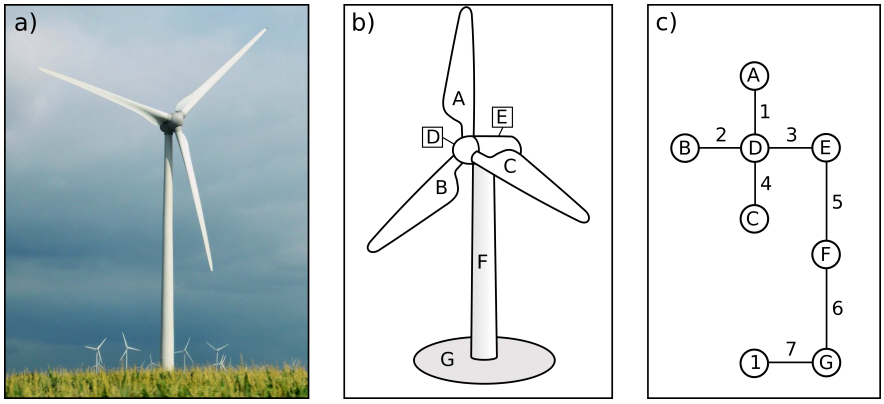


Fig. 16. A simple illustrative example of how a wind turbine (a) may be conceptualised as a series of elements (b). These elements then form the nodes in a graph representation of the structure (c), with the physical connections between elements represented by the edges in said graph.

These elements are connected by joints, labelled with numerals in Figure 16. Joints describe the physical connections between elements. Describing the physical connections between elements requires a description of which elements are connected by a particular joint, as well as the nature of a particular connection; for example, whether a connection is welded or bolted. Combining the joints with the elements within a structure allows one to determine the topology of the structure.

Boundary conditions, describing how a structure interacts with its environment, are also included in the IE model. The boundary conditions constitute an element-joint pair (element 1 and joint 7 in Figure 16), where a special element describes the nature of the boundary, for example, the ground, and the joint describes the nature of the connection between the structure and the boundary. Together, the element-joint pair fully describes the effect of the boundary condition on the behaviour of the structure.

Isomorphic to the information within the IE model is the attributed graph (AG) representation of the physical structure; that is, the representations in panel b) and c) within Figure 16, contain the same physical structural information. One could consider that the AG provides a more structured form of the

data describing an IE model, and as such facilitates the storage and creation of IE models within a database. The formal data structure of an AG also facilitates the use of graph-matching algorithms to find similar physical structures within said database.

In PBSHM, IE models can be used to compare two structures to determine the overall level of similarity. If two models are sufficiently similar, then transfer between the two structures should be possible. What 'sufficiently similar' means however, is a complicated question; one way of determining this is to examine the largest substructure common to two structures. If the largest common substructure is in fact as large as the two structures in question, then the two structures can be said to be homogeneous. In which case, as discussed, not only should transfer learning be possible, but in some cases may not even be required. Of course, for the majority of comparisons between structures, this situation will not be the case, with the largest substructure only representing only a small part of the overall structure. In this case, transfer learning may be possible within the substructure. For example, in Figures 17 and 18, which show the AGs for two different beam and slab bridges, the intermediate pier is topologically identical in both. If upon further examination, the attributes show that this pier is identical in terms of geometrical and material properties, then it would seem that some form of information transfer should be possible between the intermediate piers within these two bridges, making allowances for the influence from the rest of the structure. At the very least, the label spaces would be consistent between the two.

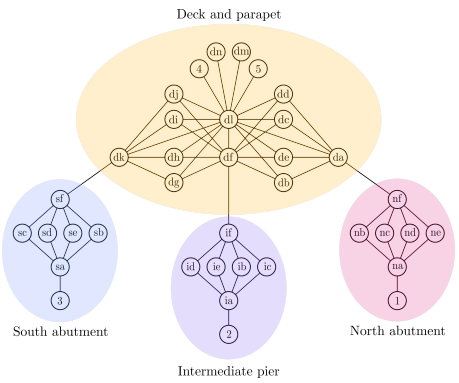


Fig. 17. A graph representation of a beam and slab bridge, located near Castledawson in Northern Ireland. This bridge features a deck supported by four longitudinal beams, as well as columns at the North and South abutments.

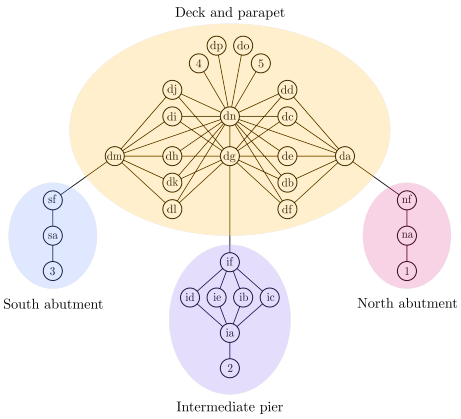


Fig. 18. A graph representation of a slightly different beam and slab bridge, located near Randalstown in Northern Ireland. This bridge features a deck supported by five longitudinal beams, and the cap beams sit directly on top of the foundations at either abutment.

A more complex problem would be transferring between the deck and parapet sections for the two bridges, since in Figure 18 there is an additional support beam. Here the label spaces are not consistent if one were to attempt a damage localisation problem. However, in theory, both deck sections should still exhibit some similar behaviour, since both are plate type elements supported with longitudinal beams for support and intermediate piers and so some physics should be common to the two decks. Therefore, there would be an expectation that some information could be transferred from a classifier trained on one bridge to another.

The largest common substructure found for the two bridges in fact involves not only the intermediate pier, but also the four longitudinal beams within the deck structure, the parapets, the deck itself and associated boundary conditions, as well as the cap beams linking the deck to either the supports or foundations. This situation creates a problem where there is still a label mismatch between the substructure and the original substructure, shown in Figure 18. How to cope with these issues and how to perform meaningful similarity comparisons within a transfer learning context remain open research questions.

A definition of similarity for PBSHM needs to be designed with transfer learning in mind and will likely vary depending on the SHM problem in question. A set of criteria informed by cases where negative transfer is likely to occur would be a useful thing to have. Currently it is hoped that by ensuring material, geometric and topological similarity, the criteria for transfer learning yielding improved classifier performance are met.

4 Conclusions and Future Directions

Clearly, the subject of PBSHM is in its infancy. While there have been scattered papers on SHM in systems-of-systems or on fleet-based maintenance etc. the formal framework presented in this chapter has only the theoretical foundations established [3, 4, 5, 6, 7]. This situation means that the scope for 'future work' is very open indeed; however, to bring a little focus to the discussion, it will be suggested here that there are three main areas in which research is needed.

The general framework is based on knowledge transfer in populations, where the structures of the population are represented in an abstract representation space; this suggest the three areas for development. In the first case, the abstract representation of the structures requires research. There will need to be rules for building IE models with consistency and for developing the theory of the representation space in which the structures are embedded. Perhaps most important is that the representation space should have a metric (or metrics), so that structures can be compared; transfer of knowledge will be contingent on structures being 'close together' in some sense, so that negative transfer is avoided.

The second main area for research is on *transfer*. As discussed in the main body of this chapter, transfer learning has only recently been considered for PBSHM; the problems which arise are difficult. In the worst case scenario, different

structures will have different label spaces and current technology can not be applied with confidence. New transfer learning algorithms are needed, perhaps informed by physics, as a great deal is known about the dynamics of structures.

Finally, in order to transfer knowledge, one must have knowledge. Data for diagnostics can come from sensors on the structures or from model predictions; these data should be optimised for diagnostics across populations and this produces demands and constraints that have not been seen before in conventional SHM. In particular, the role of physics-based models is very interesting in PBSHM. Because the main algorithms will come from transfer learning, it will not be necessary that models conform perfectly to their physical counterparts (although digital twin technology may offer prospects in terms of high-fidelity representation); it will only be required that transfer is possible. In fact, the way that models of structures will be embedded in populations means that they are not distinct from real structures. PBSHM therefore offers the prospect of a final convergence of model-based and data-based SHM.

Acknowledgements

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) via Grant references EP/R003645/1, EP/R004900/1 and EP/R006768/1. They would also like to thank their colleagues David Hester and Andrew Bunce of Queen's University Belfast, for their help and advice in building IE models of bridges and for access to the details of specific bridges.

Bibliography

[1] C.R. Farrar and K. Worden. *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley and Sons, Ltd, Chichester, UK, 2012.

[2] H. Sohn. Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society, Series A*, 365:539–560, 2007.

[3] L.A. Bull, P. Gardner, J. Gosliga, N. Dervilis, E. Papatheou, A.E. Maguire, C. Campos, T.J. Rogers, E.J. Cross, and K. Worden. Foundations of population-based structural health monitoring, Part I: Homogeneous populations and forms. *Mechanical Systems and Signal Processing*, 148:107141, 2021.

[4] J. Gosliga, P. Gardner, L.A. Bull, N. Dervilis, and K. Worden. Foundations of population-based structural health monitoring, Part II: Heterogeneous populations and structures as graphs, networks, and communities. *Mechanical Systems and Signal Processing*, 148:107144, 2021.

[5] P. Gardner, L.A. Bull, J. Gosliga, N. Dervilis, and K. Worden. Foundations of population-based structural health monitoring, Part III: Heterogeneous populations, transfer and mapping. *Mechanical Systems and Signal Processing*, 149:107142, 2021.

[6] G. Tsialiamanis, C. Mylonas, E. Chatzi, N. Dervilis, D.J. Wagg, and K. Worden. Foundations of population-based structural health monitoring, Part IV: The geometry of spaces of structures and their feature spaces. *Mechanical Systems and Signal Processing*, 157:107142, 2021.

[7] K. Worden, L.A. Bull, P. Gardner, J. Gosliga, T.J. Rogers, E.J. Cross, E. Papatheou, W. Lin, and N. Dervilis. A brief introduction to recent developments in population-based structural health monitoring. *Frontiers in Built Environment*, 6:146, 2020.

[8] Plato. *Protagoras and Meno*. Penguin Classics, 2005.

[9] Plato. *The Republic*. Penguin Classics, 2007.

[10] E. Papatheou, N. Dervilis, A.E. Maguire, I. Antoniadou, and K. Worden. A performance monitoring approach for the novel Lillgrund offshore wind farm. *IEEE Transactions on Industrial Electronics*, 62:6636–6644, 2015.

[11] W. Yang, R. Court, and J. Jiang. Wind turbine condition monitoring by the approach of SCADA data analysis. *Renewable Energy*, 53:365–376, 2013.

[12] M.P.C. Bontekoning, S. Sanchez Perez-Moreno, B.C. Ummels, and M.B. Zaaijer. Analysis of the reduced wake effect for available wind power calculation during curtailment. In *Journal of Physics: Conference Series*, volume 854, page 012004, 2017.

[13] M. Waite and V. Modi. Modeling wind power curtailment with increased capacity in a regional electricity grid supplying a dense urban demand. *Applied Energy*, 183:299–317, 2016.

[14] L.A. Bull, P. Gardner, T.J. Rogers, N. Dervilis, E.J. Cross, E. Papatheou, A.E. Maguire, C. Campos, and K. Worden. Statistical modelling of curtail-

ments in wind turbine power curves. *Submitted to: Mechanical Systems and Signal Processing*, 2021.

[15] M. Lázaro-Gredilla, S. Van Vaerenbergh, and N.D. Lawrence. Overlapping mixtures of Gaussian processes for the data association problem. *Pattern Recognition*, 45:1386–1395, 2012.

[16] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[17] D.M. Blei, A. Kucukelbir, and J.D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112:859–877, 2017.

[18] K. Kersting, C. Plagemann, P. Pfaff, and W. Burgard. Most likely heteroscedastic Gaussian process regression. In *Proceedings of the 24th International Conference on Machine learning*, pages 393–400, 2007.

[19] L.A. Bull, P. Gardner, N. Dervilis, E. Papatheou, M. Haywood-Alexander, R.S. Mills, and K. Worden. On the transfer of damage detectors between structures: An experimental case study. *To appear: Journal of Sound and Vibration*, 2021.

[20] S.J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22:1345–1359, 2010.

[21] G. Manson, K. Worden, and D.J. Allman. Experimental validation of a structural health monitoring methodology: Part III. Damage location on an aircraft wing. *Journal of Sound and Vibration*, 259:365–385, 2003.

[22] P. Gardner, L.A. Bull, N. Dervilis, and K. Worden. Overcoming the problem of repair in structural health monitoring: metric-informed transfer learning. *Submitted to: Journal of Sound and Vibration*, 2021.