White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/175070/

Version: Accepted Version

Proceedings Paper:

Gardner, P., Bull, L.A., Gosliga, J. et al. (2020) Towards population-based structural health monitoring, part IV : heterogeneous populations, transfer and mapping. In: Mao, Z., (ed.) Model Validation and Uncertainty Quantification, Volume 3 : Proceedings of the 38th IMAC, A Conference and Exposition on Structural Dynamics 2020. 38th IMAC : A Conference and Exposition on Structural Dynamics, 10-13 Feb 2020, Houston, TX USA. Conference Proceedings of the Society for Experimental Mechanics Series. Springer International Publishing, pp. 187-199. ISBN: 9783030487782. ISSN: 2191-5644. EISSN: 2191-5652.

https://doi.org/10.1007/978-3-030-47638-0_20

© 2020 The Society for Experimental Mechanics, Inc. This is an author-produced version of a paper subsequently published in Proceedings of the 38th IMAC. Uploaded in accordance with the publisher's self-archiving policy. For the version of record please see: https://doi.org/10.1007/978-3-030-47638-0_20

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Towards population-based structural health monitoring, Part IV: Heterogeneous populations, transfer and mapping

P. Gardner, L. A. Bull, J. Gosliga, N. Dervilis, & K. Worden

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

Abstract

Population-based structural health monitoring (PBSHM) involves utilising knowledge from one set of structures in a population and applying it to a different set, such that predictions about the health states of each member in the population can be performed and improved. Central ideas behind PBSHM are those of *knowledge transfer* and *mapping*. In the context of PBSHM, knowledge transfer involves using information from a structure, defined as a source domain, where labels are known for a given feature, and mapping these onto the unlabelled feature space of a different, target domain structure. If the mapping is successful, a machine learning classifier trained on the transformed source domain data will generalise to the unlabelled target domain data; i.e. a classifier built on one structure will generalise to another, making Structural Heath Monitoring (SHM) cost-effective and applicable to a wide range of challenging industrial scenarios. This process of mapping features and labels across source and target domains is defined as *domain adaptation*, a subcategory of *transfer learning*. However, a key assumption in conventional domain adaptation methods is that there is consistency between the feature and label spaces. This means that the features measured from one structure must be the same dimension as the other (i.e. the same number of spectral lines of a transmissibility), and that labels associated with damage locations, classification and assessment, exist on both structures. These consistency constraints can be restrictive, limiting to which types of population domain adaptation can be applied. This paper, therefore, provides a mathematical underpinning for when domain adaptation is possible in a structural dynamics context, with reference to topology of a graphical representation of structures. By defining when conventional domain adaptation is applicable in a structural dynamics setting, approaches are discussed that could overcome these consistency restrictions. This approach provides a general means for performing transfer learning within a PBSHM context for structural dynamics-based features.

Keywords: Population-based structural health monitoring; transfer learning; domain adaptation

1 Introduction

The process of utilising information across a population of structures in order to perform and improve inferences that generalise for the complete population is known as *Population-based* Structural Health Monitoring (PBSHM) [1]. This approach to Structural Health Monitoring (SHM) clearly provides significant benefits, as any knowledge, whether about the behaviour of features, or any damage-labelled data, that may have been obtained from one member of the population, aids predictions across the whole population. For example, in the case of an aeroplane fleet, an assortment of damage-labelled data may be available for different members of the fleet, all under differing operational conditions. The concept of PBSHM is to incorporate the feature and label data from each aeroplane to generate a machine learning-based approach that generalises across the complete fleet for all damage scenarios, especially when many members of the fleet have no labelled data associated with them. This work is Part IV of a series of papers on PBSHM [2, 3, 4, 5, 6], where for an overview of PBSHM and further motivation the reader is referred to earlier parts [2, 3, 4].

Central concepts for performing PBSHM are those of *knowledge transfer* and *mapping*. These two processes are crucial for realising PBSHM for various reasons. Firstly, conventional approaches to data-driven SHM, whether supervised, unsupervised or semi-supervised machine learning methods, assume that the training and test data are drawn from the same distribution. This assumption breaks down in PBSHM, as each member of the population will have differences in their data distributions, whether due to environmental variations, manufacturing and assembly differences, operational conditions etc. As a result, conventional machine learning approaches fail to generalise; for example, a classifier trained on data from one aeroplane in the fleet will generally fail to classify for a different aeroplane from the same population. Mapping data from one member of the population to another is therefore required such that a general classifier can be generated. Another issue is that damage-labelled data, for all possible damage states of interest, are often not obtainable for each individual in the population. In fact, it is generally infeasible to obtain a complete label set for all possible damage states, even for one structure. However, it may be possible to obtain a relatively complete damage-label set by combining labels from all structures in a population, or from a historic database from other SHM campaigns. Leveraging this label knowledge from across a population and mapping it onto a consistent space means that knowledge transfer is possible, aiding the generation of a general machine learning method for the complete population. By utilising these two processes, SHM methods can be generated that are cost-effective and applicable across a wider variety of challenging industrial applications.

Transfer learning is one approach that aims to improve the performance of a learner by transferring knowledge between different domains. Within this branch of machine learning, various methods exist with differing assumptions about consistency between domains and what knowledge is being transferred. This paper outlines the mapping and knowledge transfer problems within PBSHM, with a particular focus on modal-based features. Specifically, the work presented here focuses on recent successes in the application of *domain adaption*, a sub-category of transfer learning, in PBSHM. Domain adaption assumes labelled data are available in a *source* domain and that this can be used to aid classification in an unlabelled (or partially-labelled) *target* domain, by mapping the two domains onto a common subspace. However, conventional approaches to domain adaptation assume consistency between the feature and label space for each domain. This means that the features from one structure must be the same dimension as the other structure (i.e. the same number of natural frequencies), and that any damage labels for detection, location, classification or assessment can exist for both structures. These constraints can be limiting in the context of PBSHM. With reference to the topology of a graphical representation of a structure (discussed in Parts II and III of the paper series [3, 4]), this paper presents a mathematical underpinning for when domain adaption is possible within the context of PBSHM.

The outline of this paper is as follows. Section 2 outlines the PBSHM contexts where transfer learning is applicable, providing key definitions and an overview of the mapping and knowledge transfer problems. Following the mapping problem descriptions, Section 3 demonstrates the contexts in which domain adaptation is applicable, referencing graphical representations of structures. Finally, conclusions are presented highlighting where technological advances are required in order to produce general learning tools for PBSHM.

2 Population-based structural health monitoring and transfer learning

Population-Based Structural Health Monitoring (PBSHM) involves mapping data and labels from different structures within a population, such that a general learner can be inferred across the complete population. As a consequence, health monitoring can be performed online for any member of the population. This section seeks to define the mapping and knowledge transfer scenarios within PBSHM and demonstrate the applicable forms of learning for each scenario, with an emphasis on transfer learning.

2.1 Population types

Key definitions are required in order to outline the mapping problems within PBSHM. A *population*, in the context of PBSHM, is a group of structures (the smallest being a group of two structures) that provide information required for performing health monitoring. This general definition of a population can be further divided into two categories: *homogenous* and *heterogeneous* populations [3]; these groupings relate to the level of dissimilarity within a population, where both population types benefit from knowledge transfer. Colloquially, a homogeneous population is one in which each structure in the population can be deemed nominally identical for a given context [2, 3]. A heterogeneous population is thus a group of non-identical, and therefore different structures. Conceptually, differences will occur for a multitude of reasons, and structures will be deemed different due to various properties, based on specific contexts. This leads to a spectrum of different types of heterogeneous populations, which as similarities increase, will approach a homogenous population. One method of quantifying these differences, utilising Irreducible Element (IE) models and Attributed Graphs (AGs), has been discussed in Part II of this series of papers [3]. This approach highlights three main sources of difference within a structure, relating to the topology of the graph and types of attribute stored within the nodes and edges:

> *Geometry*: relates to the shape, size and scaling of a structure in a population, e.g. a group of aluminium rectangular plates where each plate has a different length, width and thickness.
>
> *Material*: relates to different material classes, specific materials, and their properties for structures in a population, e.g. a pair of the same size rectangular plates where one is made from aluminium and another from steel.
>
> *Topology*: relates to different topologies for graphical representations of structures in a populations, e.g. a pair of aluminium beams where one is a cantilever and the other is éncastre, where there is a difference between the two lumped-mass representations due to the difference in boundary conditions.

These classes of differences are summarised in Figure 1, where each of these categories may overlap forming a different category of heterogeneous population. For example, consider a heterogeneous population of two beam-like structures, these may have geometric differences, e.g. structure one is a beam with a tapered rectangular geometry and structure two is a beam of the same length with a uniform I-beam cross-section. In addition to these geometric differences, aspects of the materials may be different, e.g. Structure One is constructed from a unidirectional carbon fibre layup and Structure Two is formed from the same carbon fibre but in a plain weave layup. This population is therefore a heterogeneous population due to geometric and material differences, which are properties stored in the attributed graph, as defined in Part II [3], meaning the population would be located in the Venn diagram between material and geometry.

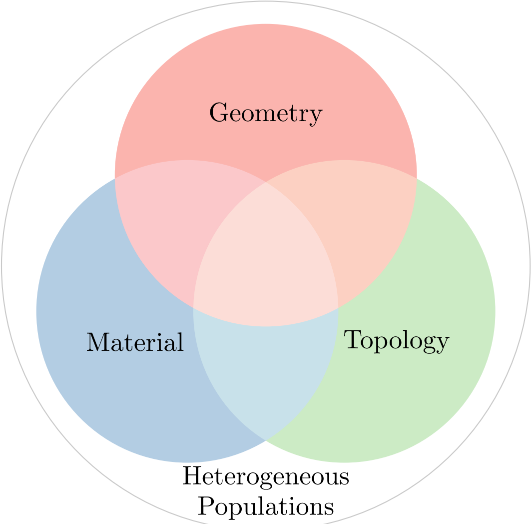

Figure 1: Categories of heterogeneous population within population-based SHM.

More formal definitions may also be attached to these types of population, as discussed in paper II [3]. A homogeneous population can be defined as a group of structures that are *topologically homogeneous* and where the geometric $\boldsymbol{\theta}_g$ and material $\boldsymbol{\theta}_m$ properties (collectively $\boldsymbol{\theta} = \{\boldsymbol{\theta}_g, \boldsymbol{\theta}_m\}$) in the nodes and edges of the attributed graph are defined as being random draws from a concentrated base distribution i.e. $p(\boldsymbol{\theta})$. It is noted that topologically homogeneous, as defined in paper II [3], relates to a group of structures that can be considered pairwise topologically equivalent, with respect their graphical representation, e.g. two structures that could be modelled as five degree-of-freedom lumped mass models where the first mass is fixed to ground. The probability mass in the base distribution $p(\boldsymbol{\theta})$ therefore defines the small differences between members in the population. A *strongly-homogeneous population* would have a unimodal concentrated base distribution for the geometric and material properties, where the strictest, perfect, form of homogeneous population is one in which the base distribution is a Dirac delta function, i.e. each member of the population is exactly the same, a scenario that never occurs in reality, but is the assumption in applying conventional machine learning methods trained on one structure, to another. Using these population definitions and categories of difference helps determine the difficulty of the mapping problem for PBSHM.

It is also noted that differences observed in data may also occur outside the structural properties of individuals in a population, not captured in the attributes and topology discussed above. These differences will relate to how data acquisition and any further processing to obtain features were performed. For example, differences in sensor placement will lead to differences in data distributions, even if those sensors are placed in the 'same' location. Properties of the data acquisition may also affect the obtained feature for each member of the population, e.g. the transmissibilities from two members of a homogeneous population may be different due to differences in the raw data sample rate. These sources of difference related to data acquisition and processing will need to be considered in defining the mapping problems in PBSHM; however, this paper simplifies the scenarios such that only structural differences in a population are considered.

2.2 Transfer learning

Transfer learning offers several techniques for dealing with scenarios where domains, tasks and distributions used within training and testing a learner are different [7]. Distinct from multi-task learning, where the objective is to learn multiple tasks across different domains [8], transfer learning leverages knowledge from *source* tasks in improving specific *target* tasks, i.e. the focus is on *target* tasks rather than all tasks (both source and target) equally [7]. This type of learning is applicable to PBSHM, even in the homogeneous population scenario, as variations in each structure within the population will lead to differences in the data distribution, meaning that a learner trained on one structure will not apply to another structure in the population. Before formally defining transfer learning, several other objects are required.

Definition 1. *A* ***domain*** $\mathcal{D} = \{\mathcal{X}, p(X)\}$ *is an object that consists of a feature space* $\mathcal{X}$ *and a marginal probability distribution* $p(X)$ *over the feature data* $X = \{\boldsymbol{x}_i\}_{i=1}^{N} \in \mathcal{X}$*, a finite sample from* $\mathcal{X}$*.*

Definition 2. *A* ***task*** $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$ *is an object that consists of a label space* $\mathcal{Y}$ *and a predictive function* $f(\cdot)$ *(*$p(y \,|\, x)$ *in a probabilistic setting) which can be inferred from training data* $\{\boldsymbol{x}_i, y_i\}_{i=1}^{N}$ *where* $\boldsymbol{x}_i \in \mathcal{X}$ *and* $y_i \in \mathcal{Y}$*.*

In the case of a single source and task domain, there will be a source domain data set $D_s = \{\boldsymbol{x}_{s,i}, y_{s,i}\}_{i=1}^{N_s}$ where $\boldsymbol{x}_{s,i} \in \mathcal{X}_s$ and $y_{s,i} \in \mathcal{Y}_s$, and a target domain data set $D_t = \{\boldsymbol{x}_{t,i}, y_{t,i}\}_{i=1}^{N_t}$ where $\boldsymbol{x}_{t,i} \in \mathcal{X}_t$ and $y_{t,i} \in \mathcal{Y}_t$, where it is generally assumed that $0 \leq N_t \ll N_s$.

Definition 3. ***Transfer learning*** *states that given a source domain* $\mathcal{D}_s$ *and task* $\mathcal{T}_s$*, and a target domain* $\mathcal{D}_t$ *and task* $\mathcal{T}_t$*, it is the process of improving the target predictive function* $f_t(\cdot)$ *in* $\mathcal{T}_t$ *by using knowledge from* $\mathcal{D}_s$ *and* $\mathcal{T}_s$*, whilst assuming* $\mathcal{D}_s \neq \mathcal{D}_t$ *and/or* $\mathcal{T}_s \neq \mathcal{T}_t$*.*

Transfer learning methods then differ in their assumptions about whether $\mathcal{X}$, $p(X)$, $\mathcal{Y}$ or $p(y \,|\, X)$ are consistent across the source and target. Another distinction is also made within transfer learning about the dimensions of the feature spaces.

Definition 4. ***Homogeneous transfer learning*** *assumes that the feature and label spaces represent the same attributes,* $\mathcal{X}_s = \mathcal{X}_t$ *and* $\mathcal{Y}_s = \mathcal{Y}_t$*, and therefore that the dimensions of the feature space are equal,* $d_s = d_t$*.*

Definition 5. ***Heterogeneous transfer learning*** *assumes that the features spaces are non-equivalent,* $\mathcal{X}_s \neq \mathcal{X}_t$*, and often that the source and target domains share no common features, meaning* $d_s \neq d_t$*. In addition, heterogeneous transfer learning can also assume that* $\mathcal{Y}_s \neq \mathcal{Y}_t$*.*

Clearly heterogeneous transfer learning is the more challenging, but more general, category of approach, with the mapping having to account for dimensionality transformations, a more complex mapping of the joint distributions, as well as the potential to handle inconsistent label spaces.

In order to clarify terminology, it is noted that the categories of homogeneous and heterogeneous *transfer learning* do not directly relate to homogeneous and heterogeneous *populations*, i.e. *homogeneous transfer learning* is *not* the only way of performing knowledge transfer for a *homogeneous population* and *heterogeneous transfer learning* is *not* the only way of performing knowledge transfer for a *heterogeneous population.* This is clarified further in Section 2.3, in

	Homogeneous	Heterogeneous		
		Geometry	Material	Topology
Feature Consistency	Yes	Yes	Yes	No
Detection Label Consistency	Yes	Yes	Yes	Yes
Location Label Consistency	Yes	Yes	Yes	No
Classification Label Consistency	Yes	Yes	Yes/No	Yes/No
Assessment Label Consistency	Yes	Yes	Yes/No	Yes/No

Table 1: Feature and label consistencies for types of population when considering modal features and location resolution at the graphical representation level.

which different transfer learning approaches are linked to a variety of population types, and in Section 3 with two case studies in which homogeneous transfer learning is applied to both a homogeneous and heterogeneous population.

2.3 Mapping scenarios

Before defining mapping scenarios within a *population-based* view of SHM, it is first important to define the categories of problem established in conventional SHM. There are several types of challenge within SHM, categorised by Rytter into a hierarchy [9], where generally difficulty increases with each level in the hierarchy.

- *Detection*: identify the presence of damage.
- *Location*: identify the position of damage.
- *Classification*: identify the type of damage.
- *Assessment*: identify the extent of damage.
- *Prognosis*: identify the safety and residual life of the structure.

This hierarchy provides a framework for defining the mapping problems within a population-based SHM context, as each level describes the required label agreement. By grounding the mapping scenarios within PBSHM to the hierarchy of SHM problems, it shows that the degree of similarity required is problem dependent. Table 1 states the level of feature and label[1] consistency for the categories of population outlined in Section 2.1 from a physical viewpoint. The table considers modal features and that the location label resolution is at the level of the graphical representation of the structure [3]. Each column in the heterogeneous category assumes that all other columns are equivalent, and that more complex heterogeneous populations are formed by combining columns, as indicated by Figure 1.

Two types of heterogeneous population in Table 1, material and topology, are stated as having scenarios where label consistency may or may not occur. Both of these categories have a hierarchy of additional attributes (discussed in Part II [3]) that will define whether these labels are consistent. For example, if populations are heterogeneous in terms of their material class the classification and assessment labels will be inconsistent, e.g. a metallic structure will not experience delamination, whereas there are delamination labels associated with a composite. Furthermore, even when the material class and specific materials are the same, the grade or batch properties will cause variation that could cause assessment label inconsistency. In a similar manner, classification and assessment label consistency can change within the topology category. If the topologies between a population are different, but all the joints within the graph are equivalent, then the classification and assessment labels will be consistent. In contrast, if the joints change between structures, e.g. structure one is assembled via bolted joints and the other structure welded, then the classification and assessment label spaces will be inconsistent.

In scenarios where the label spaces are not consistent, some form of label space matching is required. This is particularly important in an unlabelled (or partially-labelled) target space where information on cluster relationships between the source and target domain are not defined in the training data and the label space mismatch means the problem is ill-defined. For example, in a SHM location problem, where some locations do not actually exist on both structures.

---

[1] Prognosis labels are not considered here, as generally a physics-based model is required in order to perform prognosis [10].

Even when the damage label space between members of the population is inconsistent globally, a subset of these labels may be consistent, as shown in Figure 2. This demonstrates the power of graph and attribute matching between members of a population (see Part III [4]), as by locating the common attributes/subsystems S_c between members of a population helps identify the level of information that can be exchanged through a mapping ϕ (where this mapping may be through a latent space [7], as in Joint Domain Adaptation (JDA) [11]). Within an IE and AG representation of structures, these common subsystems S_c become common subgraphs in the AGs, which can be determined by graph matching, as demonstrated in Part III [4]. There will be SHM scenarios in which only part of a population needs to be similar for mapping to be possible. For example, if the SHM problem is performing damage location on a five and three degree-of-freedom structure, then locally features and location labels can be considered consistent. This would occur if, for example, the first three natural frequencies are considered, and that damage labels are only used for the first three degrees-of-freedom for each structure. It is also noted that global feature consistency can be obtained, even for topologically heterogeneous populations, by selecting an appropriate feature. For example, if in the previous example the features were transmissibilities, with the same number of frequencies bins over the same frequency range, then the feature spaces are consistent. However, this leads to questions over whether the key damage sensitive information is contained within the feature for both structures, and may lead to a phenomenon known as *negative transfer*, discussed in Section 2.4.

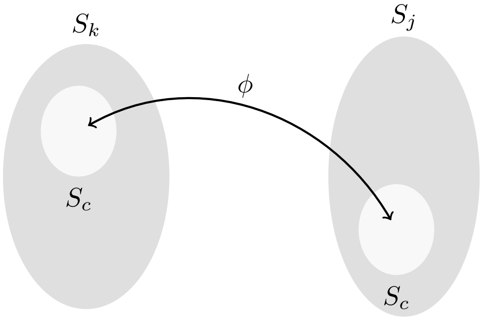

Figure 2: The potential mapping ϕ between common attributes/subsystems S_c for a population of two structures S_k and S_j.

From Table 1 it is clear that the start of any population-based SHM campaign is considering what SHM scenario is required and then obtaining the level of consistency and similarities between members of the population. Once this has been obtained, there are then four mapping problems that occur: consistency in feature and label spaces, consistency in the feature space and inconsistency in the label space, inconsistency in the feature space and consistency in the label space, and inconsistency in feature and label spaces. Table 2 defines the transfer learning approach appropriate for each assumption. Label-space matching refers to the process of identifying and pairing equivalent label classes in the transfer learning mapping, where some classes will be left unmatched between the source and target domains. This process must be performed correctly to prevent negative transfer of class labels, and is challenging, as often the task domain is unlabelled (or minimally labelled with undamaged labels only) requiring some form of semi-supervised learning [12]. It is important to state that this risk of negative transfer increases when label inconsistency is assumed, for both homogenous and heterogeneous transfer learning, and so as far as possible, a PBSHM scenario should be posed as one that is consistent in the label space.

	Label Consistency	**Label Inconsistency**
Feature Consistency	Homogeneous Transfer Learning	Homogeneous Transfer Learning with Label Space Matching
Feature Inconsistency	Heterogeneous Transfer Learning	Heterogeneous Transfer Learning with Label Space Matching

Table 2: The type of transfer learning approach for each feature and label consistency scenario.

At this point it is useful to compare transfer learning to other forms of learning that are also appropriate for a population-based approach to SHM. As mentioned in the first paper of this series [2], a *form* is one approach to PBSHM for a homogeneous population, and will assume consistent feature and label spaces. This technique seeks to create a general representation of a population within the data-domain. A form can be inferred from a group of structures and used to transfer this knowledge to remaining members of a population. When used in this manner, it can be seen as a type of transfer learning. However, a form can also be inferred across a complete population and in this scenario aims to infer an improved general learner for the complete population, in which it is a type of multi-task learning. The other difference with a form is that the representation is usually in the data-domain, and therefore the variance of general learner will often be inflated, unlike transfer learning or multi-task learning that often project to a latent space, with the aim of removing uncertainty not related to the latent process. As stated previously, multi-task learning differs from transfer learning in that the aim is to infer a general learning over all domains rather than transfer from one group to another. Multi-task learning can be both consistent and inconsistent in the feature space but will be consistent in the label space. In a PBSHM context, multi-task learning may be useful in removing the effects of confounding influences [13] and improving a learner for a structure with minimal labels within a population.

2.4 Negative transfer

A major concern when performing transfer learning, is if information is incorrectly mapped from one domain to another and reduces the performance of a learner when compared to learning from the target domain alone [7]. This phenomenon is known as *negative transfer*, and often occurs when source domains are *very* dissimilar to the target domain [14]. The fundamental idea behind transfer learning is that the source domains contain useful and relevant information about a target domain. This can be hard to determine before transfer, and becomes especially problematic when label inconsistency is assumed, the target domain is unlabelled, or that an example for each label class is not available. A binary classification scenario where negative transfer occurs is when source domain classes are mapped on to the opposite classes on the target domain, due to the target domain being unlabelled, and the source classes being dissimilar to the target domain.

Negative transfer highlights the question of when to transfer, and motivates the reasoning behind developing a measure of similarity between structures [4], as this provides an informative method for determining if a dataset will harm the learning process. Within the machine learning literature, several approaches have been proposed that seek to address negative transfer [15, 16, 17]. Graph-based methods that define the relationships between source domains by stating them graphically using a transferability metric have been developed [15]. This provides similar motivation for the Irreducible Element and Attributed Graph approach [3, 4]; however, in an SHM context these graphical representations can be formed from a physics-based viewpoint, aiding the strength of knowledge about similarity. Other approaches in avoiding negative transfer have sought to weight each source domain based on its relevance to the target domain; this is known as instance weighting [18]. Local cluster-based weighting has also been proposed, meaning that for each class on a source domain, an individual weighting is provided, stating that informativeness may not be globally shared in a particular source domain [16]. As a result, it is important to consider and account for the possibility of negative transfer when identifying what structures, and their corresponding datasets, to use in transfer learning, as well as developing algorithms to reduce or remove the possibility of negative transfer within PBSHM.

3 Domain adaptation

Domain adaptation is one form of transfer learning that seeks to transfer feature spaces between source and target domains, assuming that their marginal distributions over the finite sample set are not equal $p(X_s) \neq p(X_t)$. These techniques are primarily designed for homogeneous transfer learning, where the feature space and label space are consistent [7, 19, 11, 20]; however, heterogeneous transfer learning forms of domain adaptation do exist [21, 22, 23]. This section outlines domain adaptation and its assumptions, before demonstrating its applicability to homogenous populations, and a heterogeneous population; that contain both geometric and material differences. These case studies motivate what aspects of PBSHM are currently achievable, highlighting the required areas of further research in making PBSHM applicable across the complete range of problems outlined in Section 2.3.

Domain adaptation is formally defined (for a single source and target domain) as:

Definition 6. ***Domain adaptation*** *applies when a given inference is required for a target domain* $\mathcal{D}_t$ *and task* $\mathcal{T}_t$*, and is the process of improving the target predictive function* $f_t(\cdot)$ *in* $\mathcal{T}_t$ *given a source domain* $\mathcal{D}_s$ *and task* $\mathcal{T}_s$*, 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 one can further assume* $p(Y_s \mid X_s) \neq p(Y_t \mid X_t)$*.*

Various algorithms have been developed for this scenario [19, 11, 20], and several have been applied in a PBSHM context [24]. One approach is implemented in this section, Joint Domain Adaption (JDA) [11], in order to visually motivate the PBSHM problems in which a domain adaptation approach is applicable, although it is noted that other techniques may offer different levels of classification performance.

Briefly, JDA is a technique that assumes the joint distributions between the source and target are different $p(Y_s, X_s) \neq p(Y_t, X_t)$, and finds the optimal latent mapping in which $p(\phi(X_s)) \approx p(\phi(X_t))$ and $p(\phi(X_s) \,|\, Y_s = c) \approx p(\phi(X_t) \,|\, Y_t = c)$ for each class $c \in \{1, ..., C\}$ in $\mathcal{Y}$ (the class conditionals are matched as the conditionals are often challenging and computationally expensive to compute). This match is performed by leveraging the empirical form of the Maximum Mean Discrepancy (MMD) distance [25]; the difference between two mean embeddings in a Reproducing Kernel Hilbert Space (RKHS), induced by a kernel K [26] (where a variety of kernels can be implemented). The summation of the marginal and class conditional distributions can be formed as,

$$\mathrm{Dist}(p(\phi(X_s))\,,\, p(\phi(X_t))) + \mathrm{Dist}(p(\phi(X_s) \,|\, Y_s)\,,\, p(\phi(X_t) \,|\, Y_t)) \approx \mathrm{tr}\big(W^{\mathsf{T}} K M_c K W\big) \tag{1}$$

where W are weights associated with the empirical MMD distance and M_c is the MMD matrix, defining the mean embedding; this summation can be minimised to find the optimal latent mapping. It is noted that JDA assumes an unlabelled target domain and utilises a simplistic form of semi-supervised learning, using a classifier trained on the source domain projection to predict pseudo-labels $\hat{Y}_t$ for the projected target domain. The MMD matrix is defined as,

$$M_c(i,j) = \begin{cases} \frac{1}{N_s^{(c)} N_s^{(c)}}, & x_i, x_j \in \mathcal{D}_s^{(c)} \\ \frac{1}{N_t^{(c)} N_t^{(c)}}, & x_i, x_j \in \mathcal{D}_t^{(c)} \\ \frac{-1}{N_s^{(c)} N_t^{(c)}}, & \begin{cases} x_i \in \mathcal{D}_s^{(c)} x_j \in \mathcal{D}_t^{(c)} \\ x_i \in \mathcal{D}_s^{(c)} x_j \in \mathcal{D}_t^{(c)} \end{cases} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $\mathcal{D}_s^{(c)} = \{\boldsymbol{x}_i : \boldsymbol{x}_i \in \mathcal{D}_s \wedge y(\boldsymbol{x}_i) = c\}$ are the instances that belong in class c given the true source label $y(\boldsymbol{x}_i)$ of $\boldsymbol{x}_i$ and $N_s^{(c)} = |\mathcal{D}_s^{(c)}|$; and $\mathcal{D}_t^{(c)} = \{\boldsymbol{x}_i : \boldsymbol{x}_i \in \mathcal{D}_t \wedge \hat{y}(\boldsymbol{x}_i) = c\}$ are the instances that belong in class c given the pseudo-target label $\hat{y}(\boldsymbol{x}_i)$ of $\boldsymbol{x}_i$ and $N_t^{(c)} = |\mathcal{D}_t^{(c)}|$ (where $\wedge$ is the logical AND symbol). The optimisation problem is then formed (subject to the regularisation constraint, where μ governs the level of regularisation, and kernel Principle Component Analysis (PCA)) and becomes an eigenvalue problem where the optimal weights W are obtained from the eigenvectors corresponding to the k smallest eigenvalues from,

$$\left(K \sum_{c=0}^{C} M_c K + \mu \mathbb{I} \right) W = K H K W \phi. \tag{3}$$

Due to the pseudo-labelling of the target features, [11] recommends running several iterations of the optimisation to find the optimal W. A k-dimensional transformed feature space can be calculated by $Z = KW \in \mathbb{R}^{(N_s + N_t) \times k}$, and a classifier trained on the transformed source data can be applied to the transformed target data.

3.1 Homogeneous populations

A homogenous population is one in which the label spaces will be consistent, as stated in table 1, and depending on the feature utilised, may also have consistent feature spaces. This makes homogenous populations an ideal scenario for applying JDA, and demonstrating the benefits of transfer learning. This section presents a homogenous population of two, five storey shear structures depicted in figure 3. For both structures the stiffness elements $\{k_i\}_{i=1}^{5}$ are constructed from the summation of the tip stiffness of four rectangular cantilever beams i.e. $4k_b$, presented in 3b, where damage is applied to one beam of the four beams at a given degree-of-freedom, in the form of an open crack using the reduction in stiffness model from [27]. The geometric and material properties are nominally the same and are shown in Table 3.

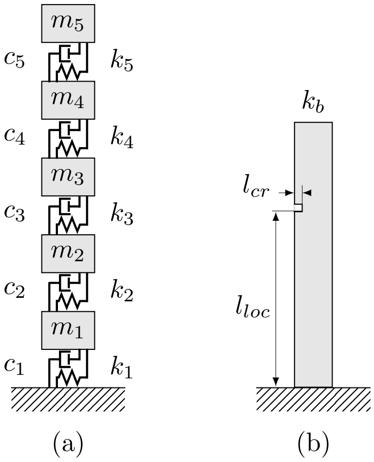

Figure 3: Schematic of the five degree-of-freedom shear structure: panel (a) is of the full system and panel (b) depicts the cantilever beam component where $\{k_i\}_{i=1}^{5} = 4k_b$ i.e. the stiffness coefficients in (a) are generated from four times the tip bending stiffness in (b).

The two structures are a homogeneous population as they are topologically[2] and structurally[3] equivalent in their lumped-mass representation and the material and geometry parameters can be defined by a concentrated unimodal distribution. The SHM problem considered here is a location problem, where a crack that is 10% of the beam width is applied at a distance 15% of the beam length from the base of the beam. This damage scenario is simulated for each of the degrees-of-freedom meaning that there are six labelled scenarios: undamaged labelled '1', damage at k_1 labelled '2', damage at k_2 labelled '3', damage at k_3 labelled '4', damage at k_4 labelled '5', and damage at k_5 labelled '6'. The features considered in this case study, depicted in Figure 4, are the five damped natural frequencies meaning $X_s = \mathbb{R}^{N_s \times 5}$ and $X_t = \mathbb{R}^{N_t \times 5}$, where $N_s = 1800$ and $N_t = 1200$ and each class has an equal weighting of data points. In order to test the inferred mapping, a separate test data set from the target data was obtained where $X_{test} = \mathbb{R}^{N_{test} \times 5}$ and $N_{test} = 1500$, again where each class has an equal number of data points.

Table 3: Properties of the source and target structures for the homogenous population case study.

Property	**Unit**	**Source**	**Target**
Beam geometry, $\{l_b, w_b, t_b\}$	mm	{5000.0, 350.0, 350.0}	{4999.0, 351.2, 349.7}
Mass geometry, $\{l_m, w_m, t_m\}$	mm	{12000.0, 12000.0, 350.0}	{12001.0, 11998.8, 351.6}
Elastic modulus, E	GPa	$\mathcal{N}\left(210.00, 1 \times 10^{-9}\right)$	$\mathcal{N}\left(210.89, 1 \times 10^{-9}\right)$
Density, ρ	kg/m^3	$\mathcal{N}(8000.0, 50)$	$\mathcal{N}(8019.4, 10)$
Damping coefficient, c	Ns/m	$\mathcal{G}(8.0000, 0.8)$	$\mathcal{G}(7.9981, 0.8)$

JDA was implemented with ten iterations, a linear kernel, and a k-Nearest Neighbour (k-NN) classifier (five neighbours), such that the emphasis is on the domain adaptation mapping and not the classifier. Cross-validation (5-folds) was implemented in order to identify the regularisation parameter, $\mu = 1 \times 10^{-3}$, and the number of transfer components, $k = 4$. The inferred mapping is presented in Figure 5, where it can be seen that the source and target datasets have been successfully mapped onto each other, where the histograms of the marginal distributions are shown on the diagonal. The classification results are given in Table 4, where it can be seen that compared to a k-NN without JDA, domain adaptation provides a substantial improvement in classification performance.

---

[2] *Topologically-equivalent* structures are those with the same graphical representation. In terms of IE and AGs, two graphs are topologically-equivalent if the underlying unattributed graphs are identical.

[3] *Structurally-equivalent* structures are those with the same graphical representation with the same locations of ground nodes. In terms of IE and AGs, two graphs are structurally equivalent if they are topologically equivalent with ground nodes in the AGs occuring in the same place.

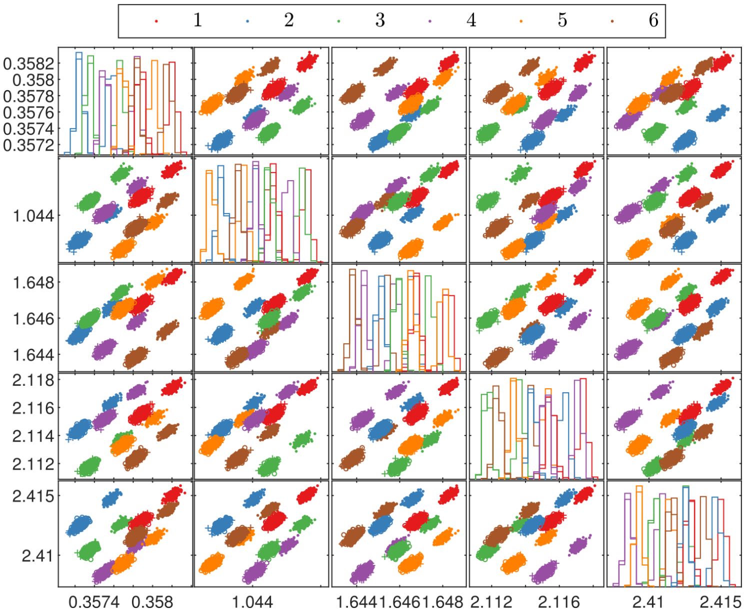

Figure 4: Source and target domain features for the homogeneous population case study. The source domain data are denoted by ($\cdot$), the target-domain training and testing data are denoted by ($+$) and ($\circ$) respectively.

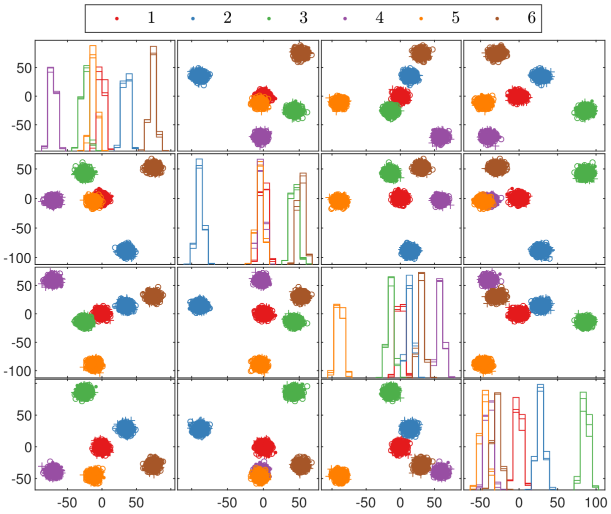

Figure 5: Transfer components for the source and target domains in the homogeneous population case study. The source domain transfer components are denoted by ($\cdot$), the target-domain training and testing transfer components are denoted by ($+$) and ($\circ$) respectively.

Table 4: Classification results for the homogenous population case study trained on the labelled source domain and applied to the unlabelled target domain.

Method		***k*-NN**	**JDA**
Mapping Training	Accuracy	83.8%	100.0%
Mapping Testing	Accuracy	83.9%	100.0%

3.2 Heterogeneous populations

The heterogeneous population case study considers a population with geometry and material differences. The two structures in the population are structurally equivalent and are both five-storey shear structures (Figure 3). The

Table 5: Properties of the source and target structures for the heterogeneous population case study.

Property	**Unit**	**Source**	**Target**
Beam geometry, $\{l_b,\, w_b,\, t_b\}$	mm	$\{5000,\, 350,\, 350\}$	$\{4300,\, 500,\, 500\}$
Mass geometry, $\{l_m,\, w_m,\, t_m\}$	mm	$\{12000,\, 12000,\, 350\}$	$\{10000,\, 10000,\, 500\}$
Elastic modulus, E	GPa	$\mathcal{N}\left(210, 1\times10^{-9}\right)$	$\mathcal{N}\left(71, 75\times10^{-7}\right)$
Density, ρ	kg/m^3	$\mathcal{N}\left(8000, 10\right)$	$\mathcal{N}\left(2700, 2.5\right)$
Damping coefficient, c	Ns/m	$\mathcal{G}\left(50, 0.1\right)$	$\mathcal{G}\left(8, 0.8\right)$

material differences exist, as the first structure is steel and the other aluminium; the geometric dissimilarities occur due to differences in the dimensions of the structure, displayed in Table 5. The SHM problem is the same as in the homogeneous case study i.e. a six-class location problem where damage is introduced as open cracks located at each degree-of-freedom. The features are damped natural frequencies such that $X_s = \mathbb{R}^{N_s\times5}$ and $X_t = \mathbb{R}^{N_t\times5}$, $X_{test} = \mathbb{R}^{N_{test}\times5}$; $N_s = 1800$, $N_t = 1200$ and $N_{test} = 1500$ all with equal weighting of data points in each class. The source and target domain features are presented in Figure 6, the magnitudes of which are *very* different highlighting the need for transfer learning.

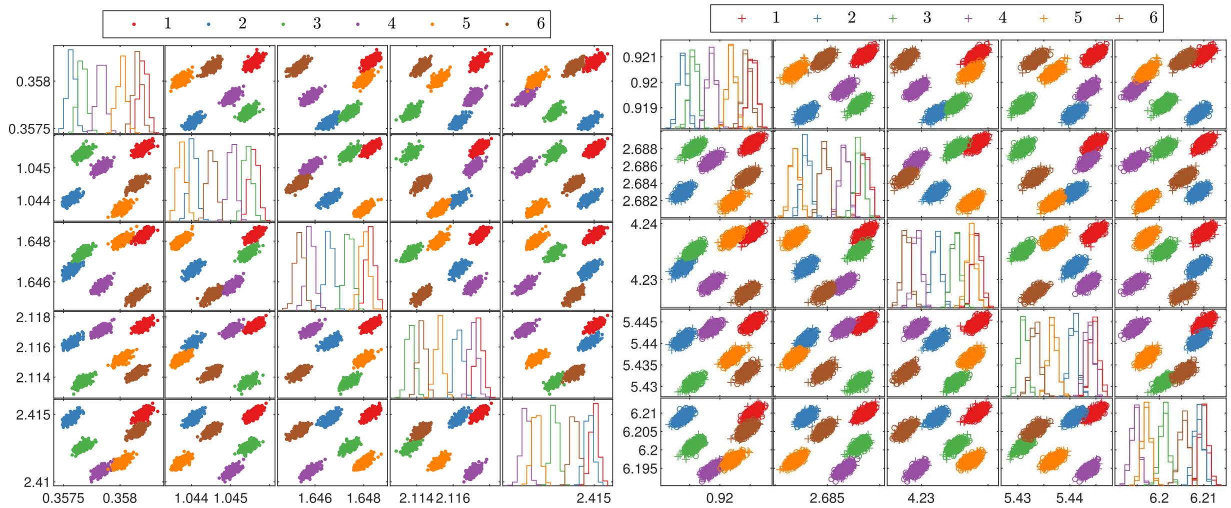

Figure 6: Source (left) and target (right) domain features for the heterogeneous population case study. The source domain data are denoted by $(\cdot)$, the target-domain training and testing data are denoted by $(+)$ and $(\circ)$ respectively.

In a similar manner to the homogeneous case study, JDA was implemented with ten iterations, a linear kernel and a k-NN classifier (five neighbours). Cross-validation (5-folds) identified the regularisation parameter, $\mu = 1\times10^{-3}$, and number of transfer components, $k = 4$, used to generate the transfer components in Figure 7. Once more, domain adaptation in the form of JDA, shows vast improvements in classification performance when compared to training a classifier on the source domain and applying it to the target, which has a classification performance equivalent to random guessing, as demonstrated in Table 6. This case study demonstrates that for a selection of heterogeneous problems, homogeneous transfer learning, in the form of domain adaptation, is still applicable.

Table 6: Classification results for the heterogeneous population case study trained on the labelled source domain and applied to the unlabelled target domain.

Method		**k-NN**	**JDA**
Mapping Training	Accuracy	16.7%	100.0%
Mapping Testing	Accuracy	16.7%	100.0%

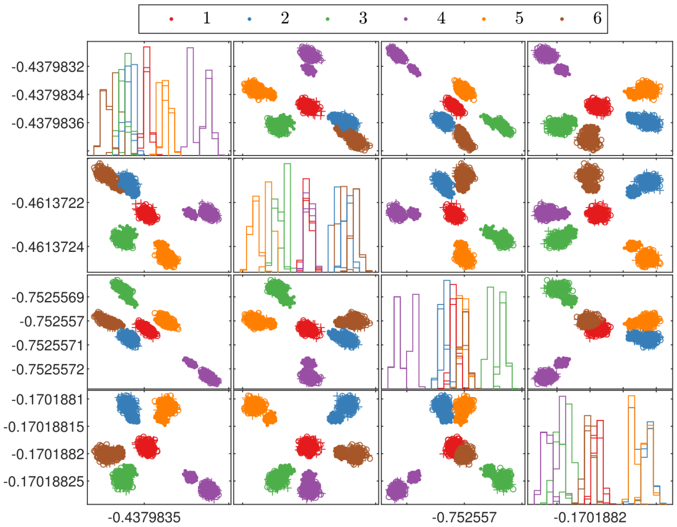

Figure 7: Transfer components for the source and target domains in the heterogeneous population case study. The source domain transfer components are denoted by $(\cdot)$, the target-domain training and testing transfer components are denoted by $(+)$ and $(\circ)$ respectively.

4 Discussion and conclusions

Knowledge transfer and mapping are important processes in developing a population-based approach to SHM. The idea of mapping data from different structures within a population such that knowledge about health states are transferred can be achieved via transfer learning. It is important when applying transfer learning to determine what similarities exist between structures within a population, and what information should be transferred, such that negative transfer is avoided. With this aim in mind, it is helpful to categorise structures based on their differences. Here two main types of population within PBSHM have been considered: *homogenous populations* and *heterogeneous populations*. Further to these two main categories, three more subdivisions have been discussed for heterogeneous populations: *geometry*, *material* and *topology*. These sources of dissimilarity form part of an Irreducible Element and Attributed Graph representation of structures, [3, 4].

By starting from what type of SHM problem is required, it is possible to categorise the level of consistency that exists for each type of population. This helps define what assumptions are required for a given transfer learning method, and whether homogenous or heterogeneous transfer learning should be implemented. Furthermore, label inconsistency between members of the population will increase the likelihood of negative transfer. This issue shows the power of an Irreducible Element and Attributed Graph representation of structures, as this approach means that similarities between members of the population can be extracted, reducing the transfer learning problem to one in which labels are considered consistent.

Transfer learning, in the form of domain adaptation has been demonstrated to be applicable for problems when both feature and label spaces are consistent. By defining the mapping problem within an SHM and population type context, several PBSHM scenarios can be considered to fulfil these requirements. This paper has demonstrated that domain adaptation is applicable for both a homogeneous population, and heterogeneous population context; where the geometric differences are changes in dimensions and two different metallic materials are used. This work has highlighted that heterogeneous populations, where topology differences occur, represent a significant challenge. When label spaces overlap, and can be considered consistent, heterogeneous domain adaptation may provide a solution to SHM problems for this type of population, and this is emphasised as an area of further research.

Acknowledgements

The authors would like to acknowledge the support of the UK Engineering and Physical Sciences Research Council via grants EP/R006768/1 and EP/R003645/1.

References

[1] E. Papatheou, N. Dervilis, A.E. Maguire, I. Antoniadou, and K. Worden. Population-based SHM: a case study on an offshore wind farm. In *Proceeding of the 10th International Workshop on Structural Health Monitoring*, 2015.

[2] L.A. Bull, P.A. Gardner, J. Gosliga, N. Dervilis, E. Papatheou, A.E. Maguire, C. Campos, T.J. Rogers, E.J. Cross, and K. Worden. Towards a 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.

[3] J. Gosliga, P.A. Gardner, L.A. Bull, N. Dervilis, and K. Worden. Towards a 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*, 2020.

[4] J. Gosliga, P.A. Gardner, L.A. Bull, N. Dervilis, and K. Worden. Towards a population-based structural health monitoring. Part III: Graphs, networks and communities. In *Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX*, 2020.

[5] K. Worden. Towards a population-based structural health monitoring. Part VI: Structures as geometry. In *Proceedings of IMAC XXXVIII – the 38th International Modal Analysis Conference, Houston, TX*, 2020.

[6] W. Lin, K. Worden, A.E. Maguire, and E.J. Cross. Towards a 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.

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

[8] Y. Zhang and Q. Yang. An overview of multi-task learning. *National Science Review*, 5:30–43, 2018.

[9] A. Rytter. *Vibrational Based Inspection of Civil Engineering Structures*. PhD thesis, Aalborg University, Denmark, 1993.

[10] C.R. Farrar and N.A.J. Lieven. Damage prognosis: The future of structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365:623–632, 2007.

[11] M. Long, J. Wang, G. Ding, J. Sun, and P.S. Yu. Transfer feature learning with joint distribution adaptation. In *2013 IEEE International Conference on Computer Vision*, pages 2200–2207, 2013.

[12] O. Chapelle, B. Schölkopf, and A. Zien. *Semi-Supervised Learning*. MIT press, 2006.

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

[14] M.T. Rosenstein, Z. Marx, L.P. Kaelbling, and T.G. Dietterich. To transfer or not to transfer. In *NIPS 2005 Workshop on Transfer Learning*, 2005.

[15] E. Eaton, M. DesJardins, and T. Lane. Modeling transfer relationships between learning tasks for improved inductive transfer. In *Lecture Notes in Computer Science*, 2008.

[16] L. Ge, J. Gao, H. Ngo, K. Li, and A. Zhang. On handling negative transfer and imbalanced distributions in multiple source transfer learning. In *Proceedings of the 2013 SIAM International Conference on Data Mining, SDM 2013*, 2013.

[17] Z. Wang, Z. Dai, B. Póczos, and J. Carbonell. Characterizing and avoiding negative transfer. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11293–11302, 2019.

[18] J. Quiñonero-Candela, M. Sugiyama, A. Schwaighofer, and N.D. Lawrence. *Dataset Shift in Machine Learning*. Mit press edition, 2009.

[19] S.J. Pan, I.W. Tsang, J.T. Kwok, and Q. Yang. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 22:199–210, 2011.

[20] M. Long, J. Wang, G. Ding, S.J. Pan, and P.S. Yu. Adaptation regularization: a general framework for transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 26:1076–1089, 2014.

[21] L. Duan, D. Xu, and I.W. Tsang. Learning with augmented features for heterogeneous domain adaptation. In *Proceedings of the 29th International Conference on Machine Learning, ICML 2012*, 2012.

[22] W. Li, L. Duan, D. Xu, and I.W. Tsang. Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36:1134–1148, 2014.

[23] O. Day and T.M. Khoshgoftaar. A survey on heterogeneous transfer learning. *Journal of Big Data*, 4:29, 2017.

[24] P.A. Gardner and K. Worden. On the application of domain adaptation for aiding supervised SHM methods. In *Proceedings of the 12th International Workshop on Structural Health Monitoring*, pages 3347–3357, 2019.

[25] B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299–1319, 1998.

[26] A. Gretton, K.M. Borgwardt, M.J. Rasch, B. Schöolkopf, and A. Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13:723–773, 2012.

[27] S. Christides and A.D.S. Barr. One-dimensional theory of cracked Bernoulli-Euler beams. *International Journal of Mechanical Sciences*, 1984.