Knowledge sharing for improving damage identification across a population of heterogeneous laboratory-scale aircraft models

Knowledge sharing for improving damage identification across a population of heterogeneous aircraft models

G. Delo [1], **C. Surace** [2], **K. Worden** [3]
[1] Department of Mechanical and Aerospace Engineering, Politecnico di Torino,
Corso Duca degli Abruzzi, 24, Torino, 10129, Italy
e-mail: **giulia.delo@polito.it**

[2] Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino,
Corso Duca degli Abruzzi, 24, Torino, 10129, Italy

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

Abstract

Advances in machine learning and pattern recognition have led to the development of data-driven approaches for Structural Health Monitoring (SHM). However, their application in real case studies is often limited by the lack or incompleteness of experimental data. Thus, Population-Based Structural Health Monitoring (PBSHM) addresses these issues by promoting knowledge-sharing between similar structures. The PBSHM theory distinguishes between *homogeneous* and *heterogeneous* populations. Structures in a heterogeneous population include different sources of variability, which affect their dynamic response and could reduce the effectiveness of knowledge-sharing performance, leading to so-called *negative transfer*. This study investigates how attribute variations influence knowledge transfer in a population of heterogeneous laboratory-scale aircraft. The transfer-learning problem is solved via a domain-adaptation algorithm, i.e., the Joint Distribution Adaptation (JDA), considering damage-detection and localisation tasks.

1 Introduction

Vibration-based Structural Health Monitoring (SHM) has been introduced as a family of powerful strategies mainly distinguished into model-based and data-driven to analyse the dynamic behaviour of structures and eventually identify any damage occurring on them. This study expands on data-driven approaches, which directly try to perform damage identification from acquired measurements and extracted features, by adopting machine-learning and pattern-recognition algorithms [1]. Particular attention is given to transfer-learning algorithms, which can enrich damage-sensitive insights about a designated structure by exploiting the data acquired from a similar structure. These algorithms underlie the theory of Population-Based Structural Health Monitoring (PBSHM), as introduced in [2, 3, 4]. PBSHM promotes knowledge sharing to create a complete SHM framework at the population level, improving monitoring and maintenance in various engineering fields, such as mechanical and aerospace, where similar or identical components and structures are often produced. PBSHM can also be beneficial in civil engineering [5], facilitating the management of infrastructural networks consisting of non-identical bridges, which often share a certain degree of similarity, as implementing a comprehensive monitoring system for each structure is frequently unfeasible, leading to scarce or incomplete data.

In analysing potential applications, PBSHM distinguishes between *homogeneous* and *heterogeneous* populations. Homogeneous populations are comprised of nominally equivalent components that may solely differ in construction details [2, 6]. Conversely, heterogeneous populations may encompass structures that exhibit variations in topology, geometrical properties, materials, and dimensions [3, 4, 7, 8]. These sources of variability affect structures' dynamic response and could reduce the effectiveness of knowledge-sharing performance, leading to so-called *negative transfer* [9]. Thus, it is required to assess the similarity between the members of a population and define a threshold on this metric, as proposed in [10, 11], to understand to what extent the PBSHM theory can be exploited, the impact of different sources of heterogeneity on transfer-learning tasks, and investigate the value of positive transfer [12]. The current study analyses the attribute impact on PBSHM performance in a population of heterogeneous laboratory-scale aircraft. The population includes multiple models with similar topologies but different materials and dimensions. Each model has been tested via Experimental Modal Analysis (EMA) to identify the corresponding transfer functions and the modal parameters. All possible combinations of source and target structures are considered, thereby studying multiple combinations of attributes. The transfer-learning problem is addressed using a domain-adaptation algorithm known as Joint Distribution Adaptation (JDA). Subsequently, the two tasks of damage detection and localisation are performed, and the results are described via performance confusion matrices.

The layout of the paper is as follows. Section 2 describes the class of domain-adaptation algorithms and the JDA, Section 3 introduces the aircraft case study, Section 4 reports the knowledge-sharing results, and Section 5 provides some conclusions.

2 Domain Adaptation

Knowledge sharing can be achieved using different transfer-learning algorithms depending on the type and the quantity of data available. This investigation implements JDA, which belongs to the category of domain-adaptation algorithms [13]. Domain adaptation has been proposed for PBSHM because it proposes various algorithms to deal with the differences between a source and a target domain [9], where a domain $\mathcal{D}$ is defined as a feature space χ, and a marginal probability distribution $p(X)$, where $X = \{x_1, x_2, ...x_n\} \in \chi$ are the data samples [14]. In addition, a knowledge-transfer task $\mathcal{T}$ is the union of a label space $\mathcal{Y}$, where $Y = \{y_1, y_2, ...y_n\} \in \mathcal{Y}$, and a predictive function $p(y_i|x_i)$. In this framework, the algorithms learn a feature transformation to minimise the differences between the feature spaces and improve the target predictive function. Multiple approaches are defined according to the differences between the domains, considering the feature spaces or their distribution, and between the tasks. In domain adaptation, the differences regard the marginal probability distributions. However, some domain-adaptation algorithms, such as TCA [13] and JDA [15], additionally consider different conditional-probability distributions. This study adopts the first seven natural frequencies identified from classic EMA as input features, assuming that a complete labelled dataset is available only for the source domain. Instead, the target training dataset includes unlabelled data and some labelled data corresponding to undamaged-condition samples. Therefore, it is possible to pre-process the features via Statistic Alignment (SA) [16, 17], and subsequently apply JDA.

2.1 Statistic Alignment

The natural frequencies in the training sets are standardised using Normal Condition Alignment (NCA) [17]. The NCA is a Statistic Alignment (SA) algorithm which transforms the target data to align its undamaged-condition samples to the source undamaged-condition samples. Given the means and the standard deviations from the undamaged condition samples in the source domain ($\mu_{s,n}, \sigma_{s,n}$) and in the target domain ($\mu_{t,n}, \sigma_{t,n}$), the transformation is given by,

$$z_t^{(j)} = \frac{x_t^{(j)} - \mu_{t,n}}{\sigma_{t,n}} \cdot \sigma_{s,n} + \mu_{s,n} \tag{1}$$

2.2 Joint Domain Adaptation

JDA maps the source and target features onto a shared latent space with reduced dimensionality via a non-linear transformation, where the difference between the marginal probability distributions and the difference between the conditional probability distributions are minimised [15]. Afterwards, a classifier on the target domain can be trained using the transformed features of the source domain. This approach adopts the Maximum Mean Discrepancy (MMD) [18] with an empirical kernel embedding to measure the distance between the domains,

$$\mathrm{Dist}(p(\psi(X_S)), p(\psi(X_T))) + \mathrm{Dist}(p(Y_S|\psi(X_S)), p(Y_T|\psi(X_T))) \approx tr(W^T K M_C K W)) \qquad (2)$$

where $k(x_i, x_j) = \psi(x_i)^T \psi(x_j)$ is the Radial Basis Function (RBF) kernel [19], and K is the kernel matrix. $W \in \mathbb{R}^{\mathrm{N}\times\mathrm{k}}$ is the transformation function, which additionally reduces feature dimensionality to k, and M_C is the class-conditional MMD matrix, where $c = \{0, 1, ..., C\}$ are the possible classes. M_C is expressed as it follows [15],

$$M_C(i,j) = \begin{cases} \frac{1}{\mathrm{N}_S^{(c)}\mathrm{N}_S^{(c)}}, & \text{if } x_i, x_j \in \mathcal{D}_S^{(c)} \\ \frac{1}{\mathrm{N}_T^{(c)}\mathrm{N}_T^{(c)}}, & \text{if } x_i, x_j \in \mathcal{D}_T^{(c)} \\ \frac{-1}{\mathrm{N}_S^{(c)}\mathrm{N}_T^{(c)}}, & \text{if } x_i \in \mathcal{D}_S^{(c)},\ x_j \in \mathcal{D}_T^{(c)} \\ \frac{-1}{\mathrm{N}_T^{(c)}\mathrm{N}_S^{(c)}}, & \text{if } x_i \in \mathcal{D}_T^{(c)},\ x_j \in \mathcal{D}_S^{(c)} \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

where $\mathcal{D}_S^{(c)}$ is the set of source samples in class c and $\mathrm{N}_S^{(c)}$ is the number of source samples in class c. Similarly, $\mathcal{D}_T^{(c)}$ is the set of target samples in class c and $\mathrm{N}_T^{(c)}$ is the number of target samples in class c. This study assumes that no labelled samples are available in the target domain for any damage-condition class. Thus, the algorithm estimates some pseudo-labels using a KNN classifier trained on the source data, and the optimisation problem is applied iteratively until convergence, refining the labels and providing W. Afterwards, a KNN classifier is trained on the transformed features. For further details on the JDA algorithm, the reader may refer to [9, 15]. The first task is damage detection ($Y = \{0, 1\}$), while the second is damage localisation ($Y = \{0, 1, 2, 3\}$ to distinguish damage on the wings, fuselage and tail respectively.

3 Experimental aircraft case study

Knowledge sharing is applied to a population of eight laboratory-scale aircraft. These models are designed based on the benchmark study conducted by the Structures and Materials Action Group (SM-AG19) of the Group for Aeronautical Research Technology in EURope (GARTEUR) [20]. Each structure consists of beam and plate elements connected via bolted connections to model the fuselage, the wings, the vertical and the horizontal tail, as shown in Figure 1.

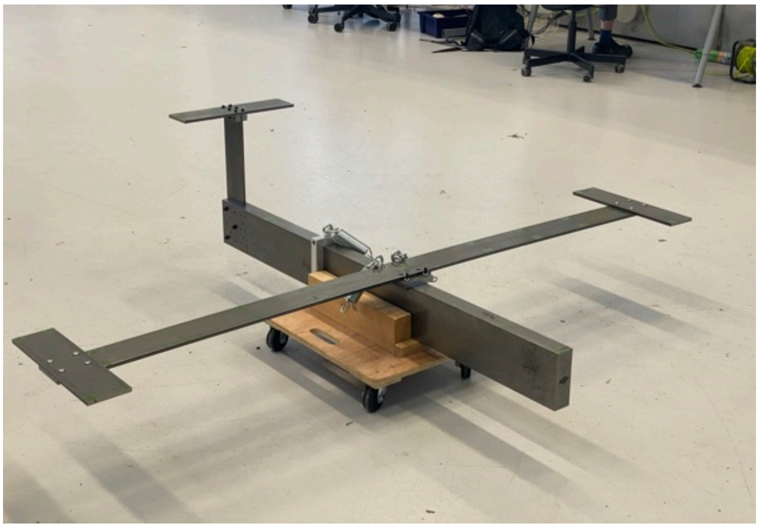

Figure 1: Large steel aircraft with end winglets (#2).

Table 1: Summary of the GARTEUR population's properties.

	Material	Scale	Topology
G1	copper alloy	s0	end winglets
G2	aluminium	s1	end winglets
G3	steel	s1	end winglets
G4	steel	s1	end winglets
G5	aluminium	s0	middle engine
G6	steel + composite	s0	no winglets
G7	steel	s0	no winglets
G8	steel	s1	no winglets

Variations in geometries, materials, and dimensions have been included to form a comprehensive and varied experimental dataset, albeit with limited topological differences. Besides undamaged conditions, multiple damage conditions have been simulated by applying masses concentrated on the wings, tail and fuselage, creating three damage classes. The resulting population is summarised in Table 1. The tests have been carried out in the "LAQ-AERMEC Aeromechanical Structural Systems" laboratory of the Department of Mechanical and Aerospace Engineering, Politecnico di Torino, and the "Laboratory for Verification and Validation " (LVV), of the University of Sheffield. For a more comprehensive description of one of the structures and the performed experimental activities, the reader is directed to [21].

4 Results

The domain-adaptation input features are sampled from an uncertain representation to build a transfer-learning dataset, by considering the mean experimental natural frequencies and polluting them with additive Gaussian noise ($\sigma = 0.1$). The training and test datasets have been formulated as follows.

- The source training set includes 50 labelled samples from the undamaged condition and 50 labelled samples from each damaged condition.
- The target training set includes 25 labelled samples from the undamaged condition and 25 unlabelled samples from each damaged condition.
- The target test set includes 25 unlabelled samples from the undamaged condition and 25 unlabelled samples from each damaged condition.

This approach is implemented on the 56 possible combinations of source and target structures, and the damage-identification performance is described via F1-score confusion matrices. In addition, the domain-adaptation algorithm is iterated ten times to compute the averaged performance and reduce the influence of the sampling process on the findings. The JDA latent space is $k = 2$; the regularisation parameter, $\lambda = 0.1$ and the maximum number of iterations is 5.

4.1 Damage detection

Damage detection is performed by reducing the labels to $Y = \{0, 1\}$, where 0 indicates an undamaged condition, and 1 indicates the presence of damage. The results of the binary KNN classification on the aircraft population are shown in Figure 2 in terms of F1-score.

Figure 2 depicts various combinations with good damage-detection performance, optimal along the diagonal but consistently exceeding 0.90 despite differences between source and target structures. These results demonstrate the strength of domain adaptation to detect damage in heterogeneous populations, even in the case of multiple differences between source and target. For a more detailed breakdown of the knowledge-sharing result, it is possible to analyse the different steps of domain harmonisation in one of the 56 structure

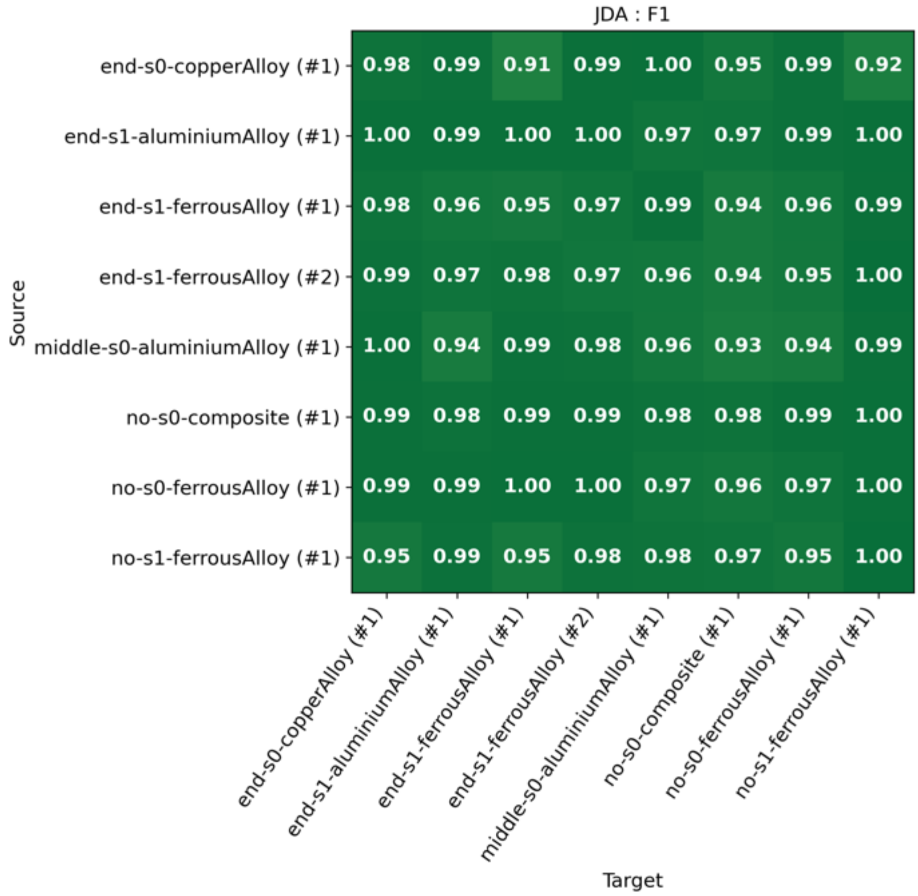

Figure 2: Damage-detection performance across the population: F1-score confusion matrix.

combinations. The following results adopt the end-s1-ferrousAlloy (#1) GARTEUR as the source domain and the end-s1-aluminiumAlloy(#1) GARTEUR as the target domain. The standardised input features are shown in Figure 3, and the effect of NCA is shown in Figure 4. The figure only displays the first three natural frequencies for better visualisation, although the original input includes the first seven. It is evident that NCA correctly aligns undamaged-conditions samples. However, the clusters for damaged conditions are not well aligned and separated from the undamaged conditions, indicating the necessity for further domain adaptation, whose results are shown in Figure 5. The combination of NCA and JDA leads to 93% damage-detection accuracy and an F1-score of 96%.

4.2 Damage localisation

Damage detection is performed as a multinomial classification problem reducing the labels to $Y = \{0, 1, 2, 3\}$, where 0 indicates an undamaged condition; 1, 2 and 3 specify the presence of damage in the different main components, respectively, in the wings, the tail and the fuselage. The results of the KNN classification on the aircraft population are shown in Figure 6, in terms of F1-score.

Upon comparison of Figure 2 and Figure 6, it becomes evident that the localisation task is more affected by the attribute variation in the aircraft population. Optimal classification is achieved along the diagonal and between the end-s1-ferrousAlloy (#1) and end-s1-ferrousAlloy (#2), which share the same properties. However, there is a more pronounced decrease in the performance because of the differences in topology, materials and dimensions. When the end-s0-copperAlloy aircraft (#1) is involved in knowledge transfer and the second structure depicts a different topology, the poorest performance is obtained. In addition, it is worthwhile to examine the behaviour of the middle-s0-aluminumAlloy (#1) aircraft, which comprises

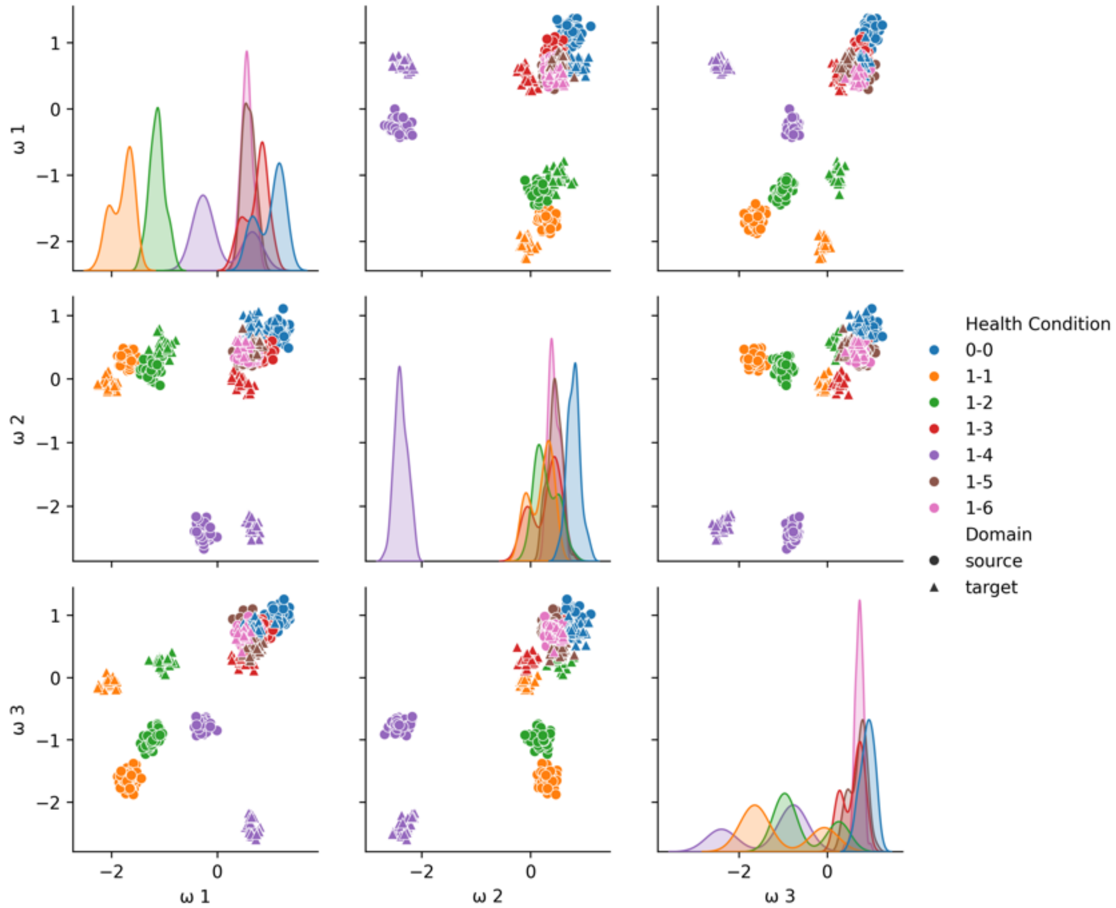

Figure 3: First three natural frequencies from the end-s1-ferrousAlloy (#1) GARTEUR (source), and the end-s1-aluminiumAlloy(#1) GARTEUR (target) in the different experimental health conditions, before NCA.

some elements to simulate the weight of the engines, leading to a different topology. This structure can be combined with similar-dimension steel and composite aircraft to achieve good results, especially when it is paired with two larger steel models with end-winglets.

Analogously to the damage-detection task, it is possible to analyse a more detailed breakdown of the knowledge-sharing result on the same combination of GARTEUR structures (i.e., end-s1-ferrousAlloy (#1) GARTEUR and end-s1-aluminiumAlloy(#1)), which share the same topology, the same dimensions but are made of different materials. After NCA, Figure 7 reports the bi-dimensional latent space produced by JDA. The samples showing damage on the wings ($Y = 1$) and damage on the tail ($Y = 2$) are well aligned, resulting in almost perfect performance. However, some target samples indicating damage on the fuselage ($Y = 3$) are incorrectly classified as undamaged. This is illustrated in Figure 8, which displays the KNN confusion matrix normalised with respect to the true labels. In this case, the combination of NCA and JDA leads to 91% damage-localisation accuracy and an F1-score of 90%. Therefore, it can be concluded that the material difference has a relatively low influence on damage detection but a higher one in the distinction of the three damage classes, where performance is lower, albeit satisfactory. Indeed, if the scenarios where the differences are multiple are analysed, such as no-s1-ferrousAlloy (#1) GARTEUR and middle-s0-aluminiumAlloy (#1) GARTEUR, the results are poorer. These structures display different topologies, materials and dimensions. Consequently, although damage-detection performs well, the algorithms fail in harmonising the damage classes, producing 66% damage-localisation accuracy, and 63% F1-score.

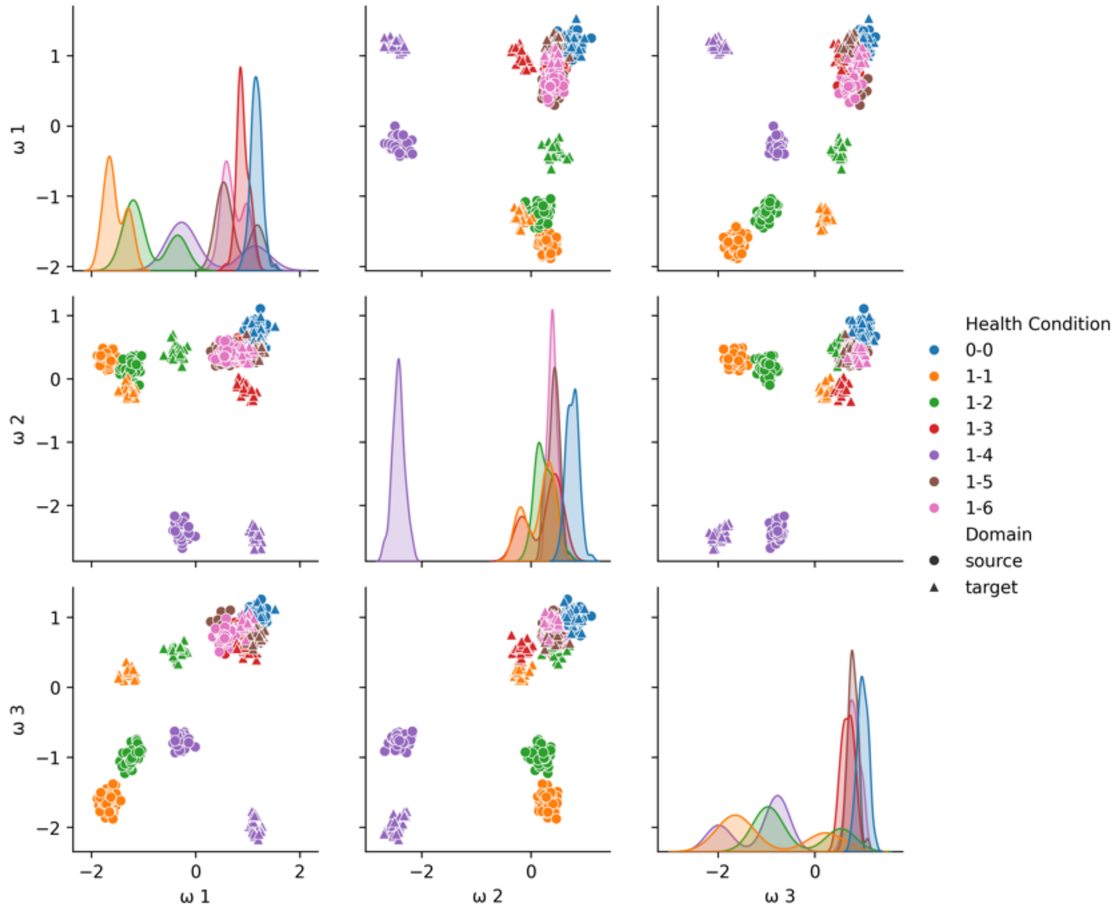

Figure 4: First three natural frequencies from the end-s1-ferrousAlloy (#1) GARTEUR (source), and the end-s1-aluminiumAlloy(#1) GARTEUR (target) in the different experimental health conditions, after NCA.

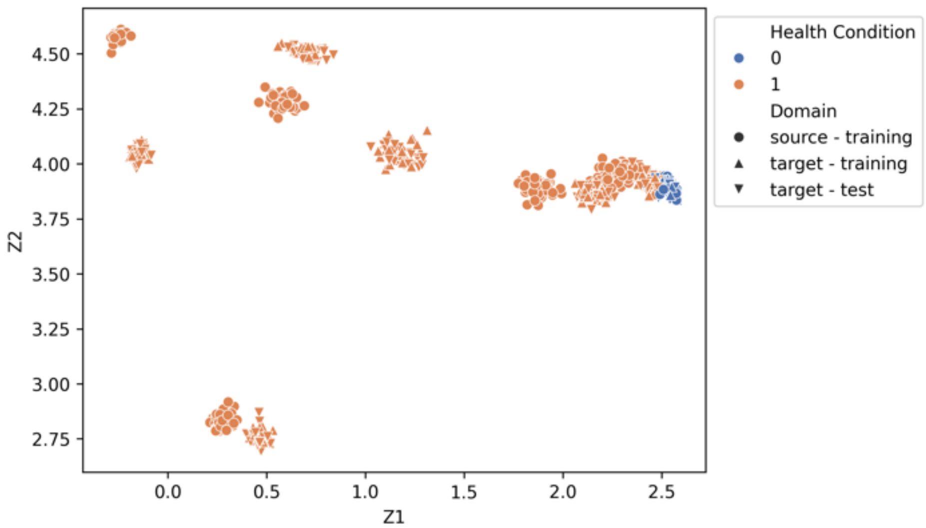

Figure 5: Transformed features in the latent space $Z = \in \mathbb{R}^{\mathrm{N} \times \mathrm{k}}$, where N is the number of samples and $\mathrm{k} = 2$. These features are computed for the damage-detection task considering JDA on the end-s1-ferrousAlloy (#1) GARTEUR (source domain) and the end-s1-aluminiumAlloy(#1) GARTEUR (target domain).

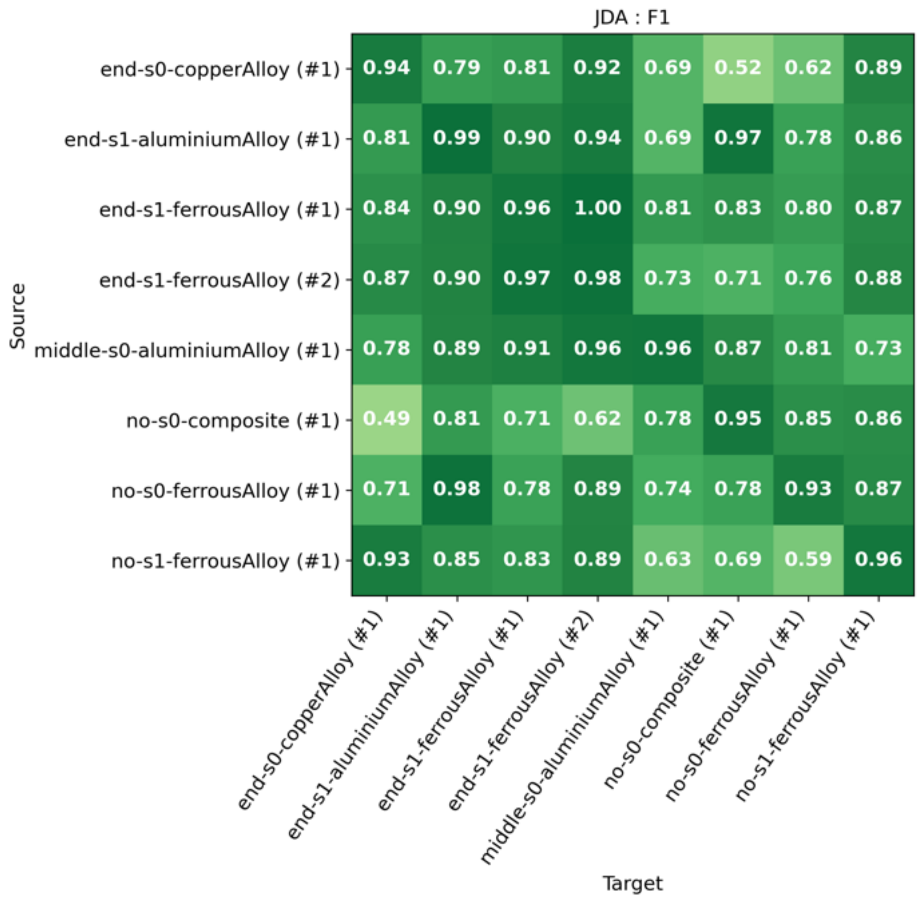

Figure 6: Damage-localisation performance across the population: F1-score confusion matrix.

5 Conclusions

This paper investigated the influence of multiple sources of heterogeneity on the knowledge-transfer performance within the framework of PBSHM. A heterogeneous population of laboratory-scale GARTEUR aircraft has been analysed. These structures have been designed, built and tested to provide a comprehensive dataset for validating PBSHM in the aerospace field. The GARTEUR population provided 56 knowledge-transfer instances, solved by employing SA and JDA as domain-adaptation algorithms. Afterwards, the harmonised features have been used to perform the tasks of damage detection and localisation, and the results have been compared by means of performance indicators.

The structures all have a similar topology, with some variations in the wings. As a result, the damage-detection performance is generally satisfactory, particularly because of SA introduction as the initial step of domain adaptation. Conversely, damage localisation is more affected by attribute and topological variations, which leads to higher variability of results, with performance decreasing when several variations, e.g. of material and topology, are coupled.

These findings confirm the necessity to consider attribute variations when analysing heterogeneous populations in PBSHM, which becomes essential when performing SHM with a higher level of identification. Hence, their impact must be properly addressed in similarity assessment by developing distance metrics capable of estimating both topological and attribute distances. Therefore, further developments will focus on knowledge-sharing results integration with distance metrics to determine which structures are appropriate for knowledge transfer and reduce the risk of negative transfer.

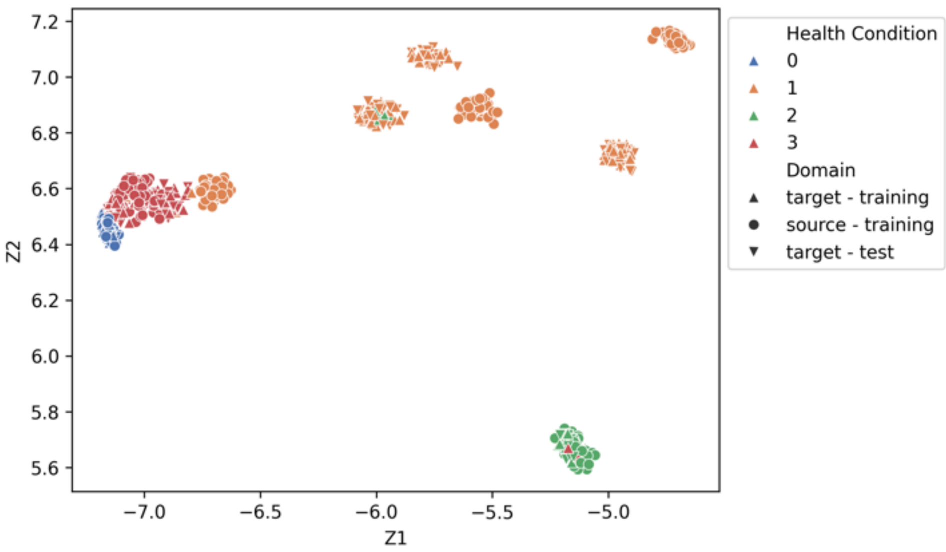

Figure 7: Transformed features in the latent space $Z =\in \mathbb{R}^{\mathrm{N}\times\mathrm{k}}$, where N is the number of samples and $\mathrm{k} = 2$. These features are computed for the damage-localisation task considering JDA on the end-s1-ferrousAlloy (#1) GARTEUR (source domain) and the end-s1-aluminiumAlloy (#1) GARTEUR (target domain).

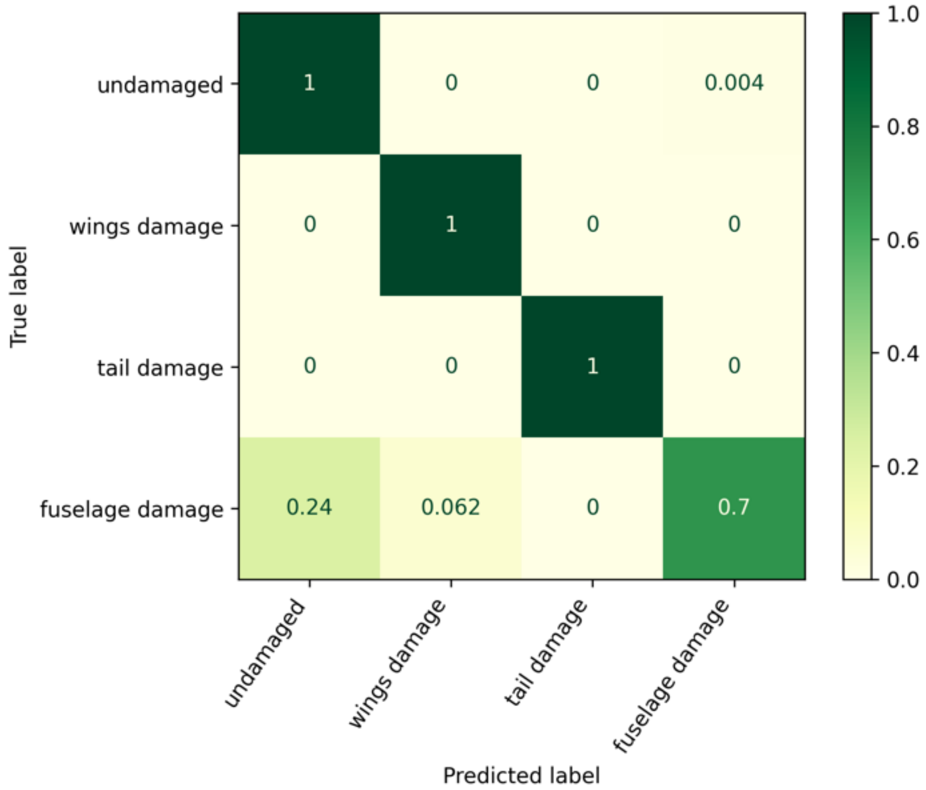

Figure 8: KNN confusion matrix for the damage-localisation task considering JDA on the end-s1-ferrousAlloy (#1) GARTEUR (source domain) and the end-s1-aluminiumAlloy (#1) GARTEUR (target domain).

Acknowledgements

The authors of this paper gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) via grant reference EP/W005816/1. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

[1] K. Worden and G. Manson, "The application of machine learning to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 515–537, 2007.

[2] L. A. Bull, P. A. Gardner, J. Gosliga, T. J. Rogers, N. Dervilis, E. J. Cross, E. Papatheou, A. E. Maguire, C. Campos, and K. Worden, "Foundations of population-based SHM, Part I: homogeneous populations and forms," *Mechanical Systems and Signal Processing*, vol. 148, p. 107141, 2021.

[3] J. Gosliga, P. A. Gardner, L. A. Bull, N. Dervilis, and K. Worden, "Foundations of population-based SHM, Part II: Heterogeneous populations–graphs, networks, and communities," *Mechanical Systems and Signal Processing*, vol. 148, p. 107144, 2021.

[4] P. A. Gardner, L. A. Bull, J. Gosliga, N. Dervilis, and K. Worden, "Foundations of population-based SHM, Part III: heterogeneous populations–mapping and transfer," *Mechanical Systems and Signal Processing*, vol. 149, p. 107142, 2021.

[5] J. Gosliga, D. Hester, K. Worden, and A. Bunce, "On population-based structural health monitoring for bridges," *Mechanical Systems and Signal Processing*, vol. 173, p. 108919, 2022.

[6] G. Tsialiamanis, C. Mylonas, E. Chatzi, N. Dervilis, D. J. Wagg, and K. Worden, "Foundations of population-based SHM, Part IV: The geometry of spaces of structures and their feature spaces," *Mechanical Systems and Signal Processing*, vol. 157, p. 107692, 2021.

[7] G. Delo, C. Surace, K. Worden, and D. S. Brennan, "On the influence of structural attributes for assessing similarity in population-based structural health monitoring," in *Proceedings of the 14th International Workshop on Structural Health Monitoring 2023: Designing SHM for Sustainability, Maintainability, and Reliability*. DEStech Publications, Inc., 2023.

[8] G. Delo, D. S. Brennan, C. Surace, and K. Worden, "On the influence of structural attributes for transferring knowledge in population-based structural health monitoring," in *Proceedings of the 42nd International Conference on Modal Analysis (IMAC-XLII)*, 2024.

[9] P. Gardner, X. Liu, and K. Worden, "On the application of domain adaptation in structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 138, p. 106550, 2020.

[10] K. Worden, D. Hester, A. Bunce, and J. Gosliga, "When is a Bridge not an Aeroplane?" in *10th International Conference on Structural Health Monitoring of Intelligent Infrastructure, SHMII 10*, Porto, Portugal, 2021.

[11] G. Delo, A. Bunce, E. Cross, J. Gosliga, D. Hester, C. Surace, K. Worden, and D. Brennan, "When is a bridge not an aeroplane? Part II: A population of real structures," in *European Workshop on Structural Health Monitoring: EWSHM 2022-Volume 2*. Springer, 2022, pp. 965–974.

[12] A. J. Hughes, J. Poole, N. Dervilis, P. Gardner, and K. Worden, "Quantifying the value of information transfer in population-based shm," in *Proceedings of the 42nd International Conference on Modal Analysis (IMAC-XLII)*, 2024.

[13] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2010.

[14] Q. Yang, Y. Zhang, W. Dai, and S. J. Pan, *Transfer learning*. Cambridge University Press, 2020.

[15] 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*, 2013, pp. 2200–2207.

[16] P. Gardner, L. Bull, J. Gosliga, J. Poole, N. Dervilis, and K. Worden, "A population-based shm methodology for heterogeneous structures: Transferring damage localisation knowledge between different aircraft wings," *Mechanical Systems and Signal Processing*, vol. 172, p. 108918, 2022.

[17] J. Poole, P. Gardner, N. Dervilis, L. Bull, and K. Worden, "On statistic alignment for domain adaptation in structural health monitoring," *Structural Health Monitoring*, vol. 22, no. 3, pp. 1581–1600, 2023.

[18] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel method for the two-sample-problem," *Advances in Neural Information Processing Systems*, vol. 19, 2006.

[19] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 723–773, 2012.

[20] E. Balmes and J. R. Wright, "Garteur group on ground vibration testing: results from the test of a single structure by 12 laboratories in europe," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 80401. American Society of Mechanical Engineers, 1997, p. V01AT03A004.

[21] G. Delo, R. Roy, K. Worden, and C. Surace, "Using the inverse finite-element method to harmonise classical modal analysis with fibre-optic strain data for robust population-based structural health monitoring," *Strain*, vol. e12481, 2024.