Cluster-based Joint Distribution Adaptation Method for Debonding Quantification in Composite Structures

Xuan Zhou*, Daniele Oboe†, Dario Poloni‡, Claudio Sbarufatti§
Polytechnic University of Milan, Milan, Italy, 20156

Leiting Dong¶
Beihang University, Beijing, China, 100191

Marco Giglio‖
Polytechnic University of Milan, Milan, Italy, 20156

Adhesive bonding is widely adopted in aeronautic structures to join composite materials or to repair damaged substrates. However, one of the most common failure modes for this type of joint is debonding under fatigue loading. In the past years, it has been proven that deboning quantification is feasible, given abundant experimental data is available. In this context, using domain adaptation to assist diagnostic tasks based on labeled data from similar structures or simulations would be thoroughly beneficial. Though, most domain adaptation methods are designed for classifications and cannot efficiently address regressions. A fuzzy-set-based Joint Distribution Adaptation for Regression method has been developed by the authors, tackling regression problems but being limited to single-outputs. The novelty presented in this paper exploits clustering techniques to approach multi-output problems, adopting a modified multi-kernel maximum mean discrepancy to improve the domain discrepancy metric. The proposed method is applied to Cracked Lap Shear specimens to assist debonding quantification. Several domain adaptations are investigated: from simulations to experiments, and from one specimen to another, proving the accuracy of damage quantification can be improved significantly in realistic environments. It is envisioned that the proposed approach could be integrated into fleet-level Digital Twins for nominally-identical but heterogeneous systems.

Nomenclature

A = adaptation matrix

*Ph.D. student, Department of Mechanical Engineering, also Ph.D. student at the School of Aeronautic Science and Engineering, Beihang University, Beijing, 100191, China
†Ph.D. student, Department of Mechanical Engineering
‡Ph.D. student, Department of Mechanical Engineering
§Associate Professor, Department of Mechanical Engineering; Corresponding author: claudio.sbarufatti@polimi.it
¶Professor & Vice Dean, School of Aeronautic Science and Engineering; Corresponding author: ltdong@buaa.edu.cn.
‖Professor, Department of Mechanical Engineering.

C = number of clusters

$\mathcal{D}$ = domain

f = adapted regressor

$\mathbf{H}$ = center matrix

k = number of subspace bases

$\mathbf{K}$ = X kernel matrix

m = number of shared features in two domains

$\mathbf{M}_{(c)}$ = MMD matrices, $c \in 0, \cdots, C$

$\widetilde{\mathbf{M}}_{(c)}$ = modified MMD matrices, $c \in 0, \cdots, C$

n = number of samples

n_{run} = number of random sampling operations for MK-MMD calculation

$\mathbf{P}$ = marginal probability distribution

$\mathbf{Q}$ = conditional probability distribution

tr = trace

$\mathbf{x}$ = features of samples

$\mathbf{X}$ = input (or feature) matrix

$\mathbf{y}$ = labels of samples

$\hat{\mathbf{y}}$ = predicted labels of samples

$\mathbf{Y}$ = label matrix

λ = regularization parameter

$\boldsymbol{\mu}$ = membership degree matrix

Subscripts

c = index of the class

exp = experiment

s = source domain

sim = simulation

t = target domain

tl = labeled data in the target domain

tu = unlabeled part in the target domain

I. Introduction

The growing adoption of composite materials in aeronautical structures has opened new horizons and research fields, both for civil and military applications. Adhesive bonding is a state-of-the-art technique utilized to join composite materials, preventing corrosion and avoiding stress concentrations around holes associated with bolted/riveted joints [1]; It is also suitable to join patch repairs on damaged substrates [2]. However, the main issue of adhesive bonding is debonding under fatigue loading: the adhesion between the adhered and the substrate fades and the joint is compromised, especially for aging structures. Embedding sensors in structures, Structural Health Monitoring (SHM) techniques detect, identify and estimate the Remaining Useful Life of damaged structures, avoiding catastrophic failure [3–10]. The ultimate goal is to build an Airframe Digital Twin [11] to boost the damage prognostic capabilities and improve decision-making in a Condition-Based Maintenance framework.

However, two main challenges affect the application of SHM to debonding quantification. First, the evolution of debonding experienced by each component differs due to variabilities in manufacturing and usage, making prognostic tasks challenging even when data from similar components is available. Secondly, due to epistemic uncertainties and modelling assumptions, the accuracy of numerical models is limited, hindering the performance of models trained on simulations in the deployment stage [12].

Domain adaptation methods could be utilized to address the concerns mentioned above, as previously carried out in other application scenarios [13–16]. However, current applications are still mainly limited to single-output regression problems. Domain adaptation assumes that labeled data available in a source domain could help improve the task (classification or regression) in an unlabeled or partially labeled target domain by mapping the two domains onto a common latent space on which the data distributions coincide [17]. Classical methods include Transfer Component Analysis (TCA) [18], which evaluates the marginal distribution difference in two domains, and Joint Distribution Adaptation (JDA) [19] which considers the marginal and conditional distribution discrepancy simultaneously. In the context of SHM, many applications [20–25] focus on the damage detection, which is a classification problem. A few works concentrate on visual-based damage quantification [26], which are not the topic in this study. TCA can be used for regression, yet it might work well only for adaptation of very similar datasets, since it only adapts the marginal distribution. JDA also adapts the data conditional distribution, but it is limited to classification problems since class labels are required. In a previous study, taking inspiration from [27], we proposed a Fuzzy-set-based Joint Distribution Adaptation for Regression (FJDAR) [28], in which the regression problem is converted to a classification problem using fuzzy sets so that the marginal and conditional distribution could be adapted simultaneously. However, applying FJDAR to multi-output regression problems is challenging because of the difficulties in constructing reasonable multi-dimensional fuzzy sets.

In this study, a Cluster-based Joint Distribution Adaptation for Regression (CJDAR) is proposed: clustering methods are used to replace the fuzzy set in FJDAR, making it feasible to obtain class labels for multi-output problems. Moreover,

a modified Multi-kernel Maximum Mean Discrepancy (MK-MMD) [29] is adopted to improve the domain discrepancy metric. The proposed approach has been tested on data acquired from two experiments and numerical simulations of debonding of large Cracked Lap Shear (CLS) specimens. Four adaptation scenarios are considered: two inter-specimen adaptations and two adaptations from the simulated data to the specimens. The results demonstrate that the proposed method can effectively improve the accuracy of debonding quantification. The remainder of this paper is arranged as follows. Sec. II describes the proposed domain adaptation method in detail and incorporates it into an online damage quantification and prognostic framework with domain adaptation. In Sec. III we present the application scenario, the experimental setup, and the numerical model. The results are compared and discussed in Sec. IV. Sec. V completes this study with some concluding remarks.

II. Cluster-based Joint Distribution Adaptation for Regression

In this section, a new Cluster-based Joint Distribution Adaptation for Regression (CJDAR) is proposed to tackle domain adaptation for multi-output regression problems. The previous developed FJDAR method is presented in Sec. II.A. Sec. II.B describes the Multi-kernel Maximum Mean Discrepancy, while in Sec. II.C, a discussion on clustering methods for Conditional Distribution Adaptation is given. The complete CJDAR method is illustrated in Sec. II.D. Finally, the proposed CJDAR is integrated into a framework to conduct the online damage quantification.

A. Fuzzy-set-based Joint Distribution Adaptation for Regression

FJDAR [28] follows the structure of the primal JDA that adapts both the marginal and conditional distribution between a source domain $\mathcal{D}^s$ and a target domain $\mathcal{D}^t$. In the adaptation of the conditional distribution, a fuzzy set is adopted to convert the continuous labels to fuzzy class labels, by which the conditional distribution discrepancy of the source and target domains can be measured.

1. Domain discrepancy calculations

In FJDAR, the Maximum Mean Discrepancy (MMD) is adopted to measure the difference between two distributions. The MMD $d(p, q)$ between distributions p and q is defined as the reproducing kernel Hilbert space (RKHS) distance between the mean embeddings of p and q. The squared formulation of MMD is defined as

$$d^2(p, q) \triangleq \left\| \mathbf{E}_p \left[\phi \left(\mathbf{x}^s \right) \right] - \mathbf{E}_q \left[\phi \left(\mathbf{x}^t \right) \right] \right\|^2_{\mathcal{H}} \tag{1}$$

where $\mathbf{x}^s$ and $\mathbf{x}^t$ are the source and target domain labels, $\mathbf{H}$ is the endowed RKHS with a characteristic kernel ϕ, so that $k(\mathbf{x}^s, \mathbf{x}^t) =< \phi(\mathbf{x}^s), \phi(\mathbf{x}^t) >$, $\mathbf{E}_{\mathbf{x}\sim p} f(\mathbf{x}) = \langle f(\mathbf{x}), \mu(p) \rangle_{\mathcal{H}}$ is the mean embedding of distribution p in $\mathcal{H}$, $d^2_k(p, q) = 0$ if and only if $p = q$. [30]

2. Marginal distribution

The marginal distribution adaptation aims at reducing the difference between the marginal distributions of the source and target domain, $P^s(\mathbf{x}^s)$ and $P^t(\mathbf{x}^t)$ [18]. After performing adaptation, MMD is calculated as:

$$\left\| \frac{1}{n_s} \sum_{\mathbf{x}_i \in \mathcal{D}^s} \mathbf{A}^{\mathrm{T}} \phi(\mathbf{x}_i) - \frac{1}{n_t} \sum_{\mathbf{x}_j \in \mathcal{D}^t} \mathbf{A}^{\mathrm{T}} \phi(\mathbf{x}_j) \right\|^2 = \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}} \mathbf{K} \mathbf{M}_0 \mathbf{K}^{\mathrm{T}} \mathbf{A} \right) \tag{2}$$

where $\mathbf{A}$ is a desired adaptation matrix that is the core of the domain adaptation. ϕ is the kernel function (here the Radial Basis Function (RBF) is adopted) to map the original data to a higher dimensional space, and $\mathbf{K}$ is the kernel matrix, which is computed by:

$$\mathbf{K}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2} \right) = \exp\left(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2 \right). \tag{3}$$

$\mathbf{M}_0$ is the MMD matrix and is computed as follows:

$$(M_0)_{ij} = \begin{cases} \frac{1}{n_s n_s}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}^s \\ \frac{1}{n_t n_t}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}^t \\ \frac{-1}{n_s n_t}, & \text{otherwise} \end{cases} \tag{4}$$

3. Conditional distribution

In FJDAR, the continuous labels for the regression task are transformed into fuzzy class labels using the fuzzy set method. The MMD of the conditional distribution is then computed as follows:

$$\left\| \sum_{\mathbf{x}_i \in \mathcal{D}_s} \bar{\mu}_{ic}^s \mathbf{A}^{\mathrm{T}} \phi(\mathbf{x}_i) - \sum_{\mathbf{x}_j \in \mathcal{D}_t} \bar{\mu}_{ic}^t \mathbf{A}^{\mathrm{T}} \phi(\mathbf{x}_j) \right\|^2 = \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}} \mathbf{K} \widetilde{\mathbf{M}}_{(c)} \mathbf{K}^{\mathrm{T}} \mathbf{A} \right) \tag{5}$$

where $\widetilde{\mathbf{M}}_{(c)}$ is a modified version of $\mathbf{M}_{(c)}$, which is computed as:

$$\left(\widetilde{\mathbf{M}}_{(c)} \right)_{ij} = \begin{cases} \bar{\mu}_{ic}^s \bar{\mu}_{jc}^s, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}^s \\ \bar{\mu}_{ic}^t \bar{\mu}_{jc}^t, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}^t \\ -\bar{\mu}_{ic}^s \bar{\mu}_{jc}^t, & \begin{cases} \mathbf{x}_i \in \mathcal{D}^s, \mathbf{x}_j \in \mathcal{D}^t \\ \mathbf{x}_j \in \mathcal{D}^s, \mathbf{x}_i \in \mathcal{D}^t \end{cases} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$\bar{\mu}_{ic}$ is the class-normalized membership:

$$\bar{\mu}_{ic} = \frac{\mu_{ic}}{\sum_{i=1}^{n} \mu_{ic}}, \quad i = 1, \ldots, n;\ c = 1, \ldots, C \tag{7}$$

where μ_{ic} is the membership of the i_{th} sample for the class c obtained by the fuzzy set method.

4. Finding the adaptation matrix by solving a constrained optimization problem

After computing $\widetilde{\mathbf{M}}_{(c)}$, the following constrained optimization is adopted to find the optimal adaptation matrix $\mathbf{A}$. The interested reader can refer to [19] for more detail.

The constrained optimization problem is expressed as follows:

$$\min_{\mathbf{A}^{\mathrm{T}}\mathbf{K}\mathbf{H}\mathbf{K}^{\mathrm{T}}\mathbf{A}=\mathbf{I}} \sum_{c=0}^{C} \operatorname{tr}\left(\mathbf{A}^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{M}}_{(c)}\mathbf{K}^{\mathrm{T}}\mathbf{A}\right) + \lambda\|\mathbf{A}\|_F^2 \tag{8}$$

where $\mathbf{A}^{\mathrm{T}}\mathbf{K}\mathbf{H}\mathbf{K}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$ constrains the variance of the data to be equal before and after the transformation, $\mathbf{H}$ is the central matrix, and $\mathbf{I}$ is the identity matrix. $\widetilde{\mathbf{M}}_{(0)}$ is the MMD matrix of the marginal distribution, which is the same as that in primal JDA [19], thus $\widetilde{\mathbf{M}}_{(0)} = \mathbf{M}_{(0)}$, λ is a regularization parameter set to guarantee the optimization problem to be well-posed.

The constrained optimization in Eq. 9 is solved by constructing its Lagrange function:

$$L = \operatorname{tr}\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{K}\sum_{c=0}^{c}\widetilde{\mathbf{M}}_{(c)}\mathbf{K}^{\mathrm{T}} + \lambda\mathbf{I}\right)\mathbf{A}\right) + \operatorname{tr}\left(\left(\mathbf{I} - \mathbf{A}^{\mathrm{T}}\mathbf{K}\mathbf{H}\mathbf{K}^{\mathrm{T}}\mathbf{A}\right)\mathbf{\Phi}\right) \tag{9}$$

where $\mathbf{\Phi} = \operatorname{diag}(\phi_1, \ldots, \phi_k) \in \mathbb{R}^{k\times k}$ is a matrix containing the Lagrange multipliers.

Setting $\frac{\partial L}{\partial \mathbf{A}} = 0$, we can obtain the generalized eigenproblem:

$$\left(\mathbf{K}\sum_{c=0}^{c}\widetilde{\mathbf{M}}_{(c)}\mathbf{K}^{\mathrm{T}} + \lambda\mathbf{I}\right)\mathbf{A} = \mathbf{K}\mathbf{H}\mathbf{K}^{\mathrm{T}}\mathbf{A}\mathbf{\Phi} \tag{10}$$

Finally, the optimal adaptation matrix $\mathbf{A}$ is computed by solving Eq. 10 for the k smallest eigenvectors.

B. Measure domain differences by a modified MK-MMD

The performance of MMD is affected significantly by the selection of the kernel. In [28], we proposed an optimization procedure for the selection of the optimal length scale γ for the RBF (Radial Basis Function) kernel in each step. However, leveraging on multiple kernels should enhance the domain discrepancy metric, as proposed by Gretton et al. [29] with the Multi-kernel Maximum Mean Discrepancy (MK-MMD), which is widely used in deep domain adaptation methods [30, 31].

In MK-MMD, the characteristic kernel associated with the feature map ϕ is defined as the convex combination of m positive-definite kernels $\{k_u\}$:

$$\mathcal{K} \triangleq \left\{ k = \sum_{u=1}^{m} \beta_u k_u : \sum_{u=1}^{m} \beta_u = 1, \beta_u \geqslant 0, \forall u \right\} \tag{11}$$

where the constraints on coefficients $\{\beta_u\}$ are imposed to guarantee that the derived multi-kernel k is characteristic [32].

As studied theoretically in Gretton et al.[33], the optimal kernel parameter β for the MK-MMD is learned by jointly maximizing the test power and minimizing the Type II error, leading to the optimization problem:

$$\max_{k \in \mathcal{K}} d_k^2 \left(\mathcal{D}_s^l, \mathcal{D}_t^l \right) \sigma_k^{-2} \tag{12}$$

where $\sigma_k^2 = \mathbf{E}_{\mathbf{z}} g_k^2(\mathbf{z}) - [\mathbf{E}_{\mathbf{z}} g_k(\mathbf{z})]^2$ is the estimation variance and $\mathbf{d} = (d_1, d_2, \cdots, d_m)^{rmT}$, where each d_u is the MMD via the kernel k_u. The covariance $\mathbf{Q} = \operatorname{cov}(g_k) \in \mathbb{R}^{m \times m}$ can be computed in $O(m^2 n)$ cost, i.e. $\mathbf{Q}_{uu'} = \frac{4}{n_s} \sum_{i=1}^{n_s/4} g_{k_u}^{\Delta}(\overline{\mathbf{z}}_i) g_{k_{u'}}^{\Delta}(\overline{\mathbf{z}}_i)$, where $\mathbf{z}_i \triangleq \left(\mathbf{x}_{2i-1}^s, \mathbf{x}_{2i}^s, \mathbf{x}_{2i-1}^t, \mathbf{x}_{2i}^t \right)$, $g_k(\mathbf{z}_i) \triangleq k\left(\mathbf{x}_{2i-1}^s, \mathbf{x}_{2i}^s \right) + k\left(\mathbf{x}_{2i-1}^t, \mathbf{x}_{2i}^t \right) - k\left(\mathbf{x}_{2i-1}^s, \mathbf{x}_{2i}^t \right) - k\left(\mathbf{x}_{2i}^s, \mathbf{x}_{2i-1}^t \right)$, $\overline{\mathbf{z}}_i \triangleq (\mathbf{z}_{2i-1}, \mathbf{z}_{2i})$, $g_{k_u}^{\Delta}(\overline{\mathbf{z}}_i) \triangleq g_{k_u}(\mathbf{z}_{2i-1}) - g_{k_u}(\mathbf{z}_{2i})$. Hence Eq. 12 reduces to a quadratic programming (QP) minimization problem:

$$\min_{\mathbf{d}^{\mathrm{T}}\boldsymbol{\beta}=1, \boldsymbol{\beta} \geqslant \mathbf{0}} \boldsymbol{\beta}^{\mathrm{T}} (\boldsymbol{Q} + \varepsilon \mathbf{I}) \boldsymbol{\beta} \tag{13}$$

where $\varepsilon = 10^{-3}$ is a small regularizer to make the problem well-posed. By solving Eq. 13, we obtain a multi-kernel $k = \sum_{u=1}^{m} \beta_u k_u$.

However, the above presented procedures require the target and source domain to have the same number of samples, i.e. $n_s = n_t$, which cannot be satisfied for online damage quantification and prognosis. Although $\mathbf{z}_i$ is selected as $\left(\mathbf{x}_{2i-1}^s, \mathbf{x}_{2i}^s, \mathbf{x}_{2i-1}^t, \mathbf{x}_{2i}^t \right)$ in the above procedure, as indicated in [29], it can be selected randomly. Hence, a modified MK-MMD is proposed in this study, which consists of an averaging on several runs. In each run i, a random sampling operation is adopted to extract samples from the domain containing the larger number of samples, so that the two domains have the same number of samples: a β_i is then computed by the standard MK-MMD. Finally, the β is obtained by averaging:

$$\boldsymbol{\beta} = \frac{1}{n_{\mathrm{run}}} \sum_{i=1}^{n_{\mathrm{run}}} \boldsymbol{\beta}_i \tag{14}$$

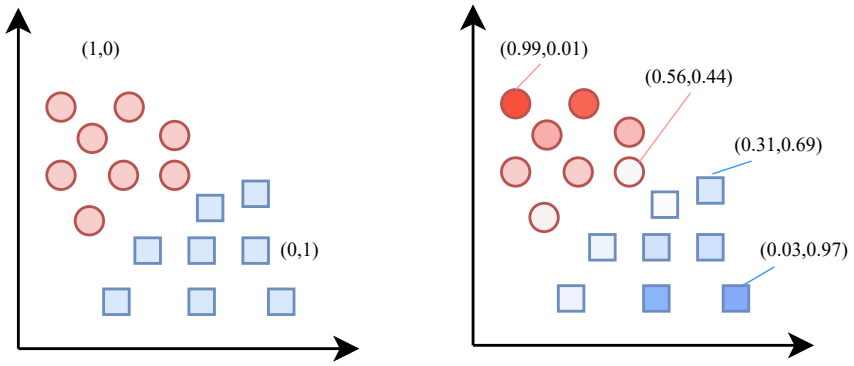
where n_{run} is the number of sampling operations.

C. Conditional Distribution Adaptation using clustering methods

In this study, the fuzzy set method in FJDAR is replaced with clustering methods, which can account for the internal relationship of samples in the source and target domains.

Cluster analysis aims to group a set of objects in such a way that objects in the same group (called a cluster) are

more similar (according to some features and metric) to each other than to those in other clusters [34]. This method has been widely applied in the aerospace field [35–37]. It can be categorized into two groups: hard clustering and soft clustering., As shown in Fig. 1, the main difference between hard and soft clustering is that soft clustering assign probabilities of belonging to a group, while in hard clustering a sample is clustered into one class exclusively.



(a) hard clustering (b) soft clustering

Fig. 1 Hard clustering and soft clustering.

In this study, four clustering methods are considered for comparison, including two hard clustering methods (k-means [38] and spectral cluster [39]) and two soft clustering methods (Fuzzy C-means (FCM) [40] and Gaussian Mixture Model (GMM) [41]). For the sake of brevity these clustering methods are not discussed in detail, and interested readers can refer to [34, 42] for more information. The output of these clustering methods are either hard or soft labels. For soft labels, all the labels participate in the computation of $\widetilde{\mathbf{M}}_{(c)}$ which is the same as that in Eq. 6, while for hard labels, Eq. 6 is simplified to traditional class-wise MMD as in [19].

D. Cluster-based Joint Domain Adaptation for Regression (CJDAR)

The CJDAR algorithm takes as input the features $\mathbf{x}^s$ and labels y^s in the source domain, the available features $\mathbf{x}^{tl}$ and labels y^{tl} in the target domain, and current unlabeled features $\mathbf{x}^{tu}$. The algorithm is described in Algorithm 1.

In the first step, the matrix $\mathbf{X}$ contains all features, $\mathbf{X}^l$ contains all labeled features, $\mathbf{Y}^l$ contains all true labels are constructed by stacking the inputs. $\mathbf{M}_0$ is calculated by Eq. 4 since the marginal distribution adaptation only relies on features that have been provided. Then the pseudo labels $\mathbf{y}^{tu}$ are obtained by training a regressor based on available labeled samples. Even though $\mathbf{y}^{tu}$ are pseudo labels, it allows us to construct the matrix $\mathbf{Y}$. The $\mathbf{X}$ and $\mathbf{Y}$ matrices are then inputted to the clustering methods to obtain the class labels, which are used to compute the MMD matrix $\widetilde{\mathbf{M}}_{(c)}$. In the sixth step, the MK-MMD is computed and the weight vector β is obtained. By solving the generalized eigen decomposition problem, the adaptation matrix $\mathbf{A}$ is obtained. Finally, the adapted regressor f and predicted label $\mathbf{y}^{tu}$ are outputted as the result.

Algorithm 1 Cluster-based joint distribution adaptation for regression

Input: $\mathbf{x}^s$, $\mathbf{x}^{tl}$, $\mathbf{x}^{tu}$, $\mathbf{y}^s$, $\mathbf{y}^{tl}$, m RBF kernels with various γ
Output: Adaptation matrix $\mathbf{A}$, predicted label $\hat{\mathbf{y}}^{tu}$, Adapted regressor f

1: Construct matrix $\mathbf{X} = [\mathbf{x}^s, \mathbf{x}^{tl}, \mathbf{x}^{tu}]$, $\mathbf{X}^l = [\mathbf{x}^s, \mathbf{x}^{tl}]$, $\mathbf{Y}^l = [\mathbf{y}^s, \mathbf{y}^{tl}]$
2: Calculate the kernel matrix $\mathbf{K}$ by the RBF kernel function.
3: Construct MMD matrix $\mathbf{M}_0$ for each RBF kernel by Eq. 4;
4: Train a regressor f on $(\mathbf{X}^l, \mathbf{Y}^l)$, predict the pseudo label $\hat{\mathbf{y}}^{tu}$, and construct $\mathbf{Y} = [\mathbf{y}^s, \mathbf{y}^{tl}, \hat{\mathbf{y}}^{tu}]$;
5: Obtain the class labels by the cluster method in Sec. II.C.
6: Construct MMD matrices $\{\widetilde{\mathbf{M}}_{(c)}\}_{c=1}^{C}$) for each RBF kernel by Eq. 6.
7: Calculate $\boldsymbol{\beta}$ by Eq. 14 and the MK-MMD by Eq. 11.
8: Solve the generalized eigen decomposition problem by Eq. 10 and select the k smallest eigenvectors to construct the adaptation matrix $\mathbf{A}$
9: Train a regressor on $(\mathbf{A}^{\mathrm{T}}\mathbf{K}^l, \mathbf{Y}^l)$, and obtain the predicted label $\hat{\mathbf{y}}^{tu}$;
10: **Return** an adapted regressor f, predicted label $\hat{\mathbf{y}}^{tu}$.

E. The framework of online damage quantification with proposed algorithm

The proposed CJDAR method is integrated to an online damage quantification framework with domain adaptation, as in [28]. The framework, referred as Online Cluster-based Domain Adaptation (OCJDAR), is shown in Fig. 2.

Prior to the online structural damage quantification stage, the source domain data $\mathcal{D}^s$ is extracted from $\mathcal{D}^h$, a database containing all available historical data constructed by collecting damage quantification data obtained from similar structures, damages or simulations. To start the online stage, sequential monitoring data is collected from the deployed sensors, features are extracted producing the unlabeled features in the target domain $\mathbf{x}^{tu}$. It is worth mentioning that a few labels in the target domain are needed to start the process. The proposed domain adaptation methods are then utilized to map the features from the labeled source domain data and the labeled and unlabeled data from the target domain into a common latent space. A regressor is then trained upon the adapted labeled data. Following the construction of the regressor, the damage size $\mathbf{y}^{tu}$ is obtained, which can be utilized to support the decision-making in a realistic environment. Once the true damage size $\mathbf{y}^{tu}$ is measured, the prediction accuracy can be evaluated, and the $\mathbf{y}^{tu}$, paired with the $\mathbf{x}^{tu}$, would constitute new labeled samples that could be stored to enrich the database.

III. Application to the delamination of composite Cracked Lap Shear specimens

In this study, the proposed method is applied to large CLS specimens. The data acquired from two experiments and numerical simulations are considered as source and target domains, and domain adaptations are conducted. Sec. III.A describes the CLS specimens and the numerical models, while Sec. III.B presents the experimental and numerical datasets. Sec. III.C describes the practical implementation and the hyperparameters setting. Sec. III.D recaps the baseline methods, while Sec. III.E shows the performance metric.

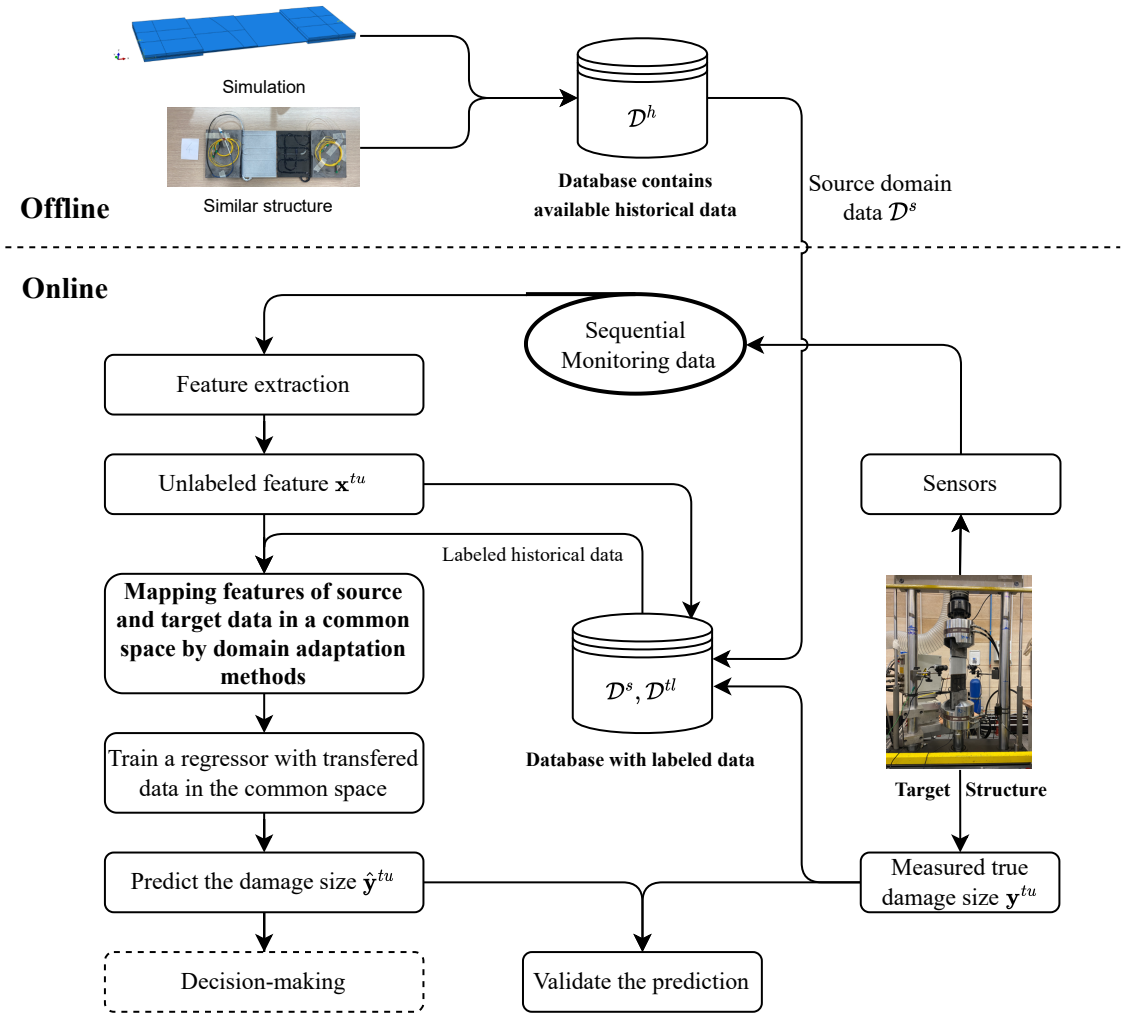
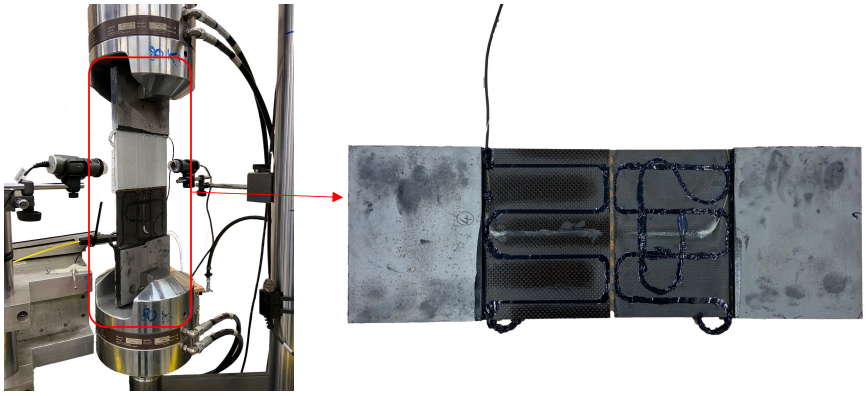


Fig. 2 The framework of online damage quantification with domain adaptation.

A. Cracked Lap Shear Specimens with Fiber Bragg Grating Sensors and Corresponding Numerical Models

Two debonding experiments of the large CLS specimens are conducted in a laboratory environment, as shown in Fig. 3. The layout and overall dimensions of the specimens are shown in Fig. 4. The adherent is realized with a laminate composed of 8 plies of woven fabric, with the stacking sequence $[0/+45/90/-45]_s$ and a total thickness after curing of 2 mm. The two adherents are joined together by a double layer of a film adhesive, and a Kapton film with a zig-zag profile is inserted in between them to guarantee the debonding initiation. The load is applied by means of four bonded steel tabs, which have been designed to guarantee a uniform stress distribution along the specimen width, then, the testing machine applies a cyclic load of 45 KN with a stress ratio $R = 0.1$.



(a) (b)

Fig. 3 Experimental setup and specimens. (a) experiment setup, (b) one of the specimens.

The large CLS is equipped with 12 fiber Bragg grating (FBG) sensors as shown in Fig. 5. Strains are statically acquired at defined intervals (generally every 20,000 cycles) at the load of 10 kN to avoid peak splitting issues, then, each acquisition is averaged to reduce the noise contribution and the temperature effect is compensated with proper calibration. In this study, the debonding lengths are measured during the experiments at the two specimen sides by means of two microscopes, as shown in Fig. 3. The two debonding lengths are represented as a_1 and a_2 respectively in Fig. 4.

The finite element model is built in the ABAQUS software and it is shown in Fig. 6. Several parts are modeled independently and then joined together in the assembly with the tie constraints between the surfaces. Both the laminate and the adhesive are modeled by continuum shell elements of type SC8R to describe the material behavior near the debonding front. The debonding is modeled by not constraining together the two adhesive surfaces on the disbanded area and hard-contact interaction properties are set to avoid any possible penetration between parts. The steel tabs are modeled with solid elements (C3D8R), with a mesh size of 2 mm on the XY plane and seven elements to describe the 5 mm thickness. In order to generate a database with multiple debonding scenarios, the debonding front shape is systematically modified and a simulation is run for each configuration.

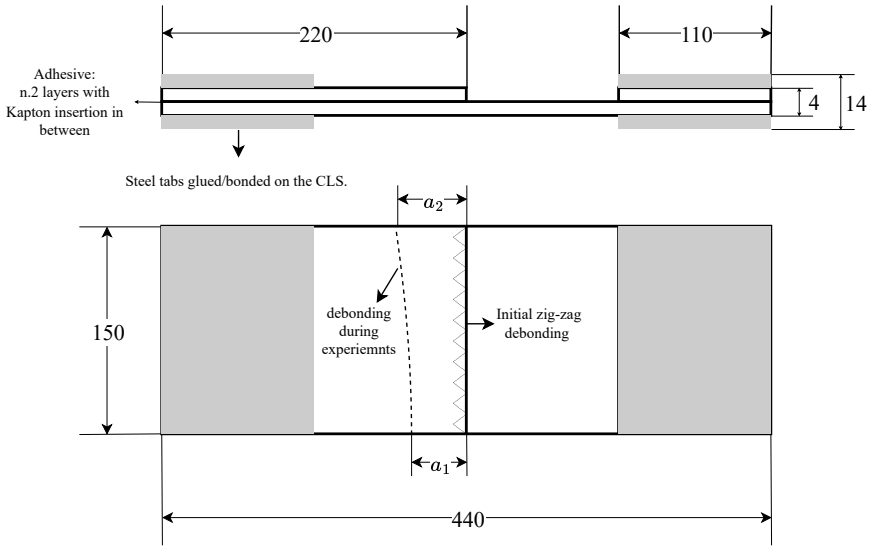


Fig. 4 Big CLS layout with steel tabs, dimensions in millimeters.

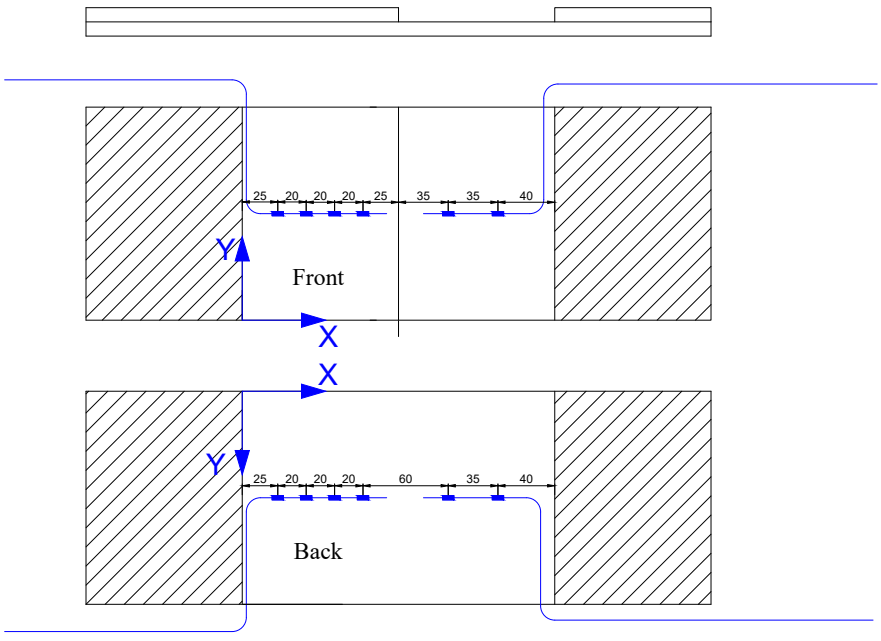


Fig. 5 The layout of the FBG sensor network.

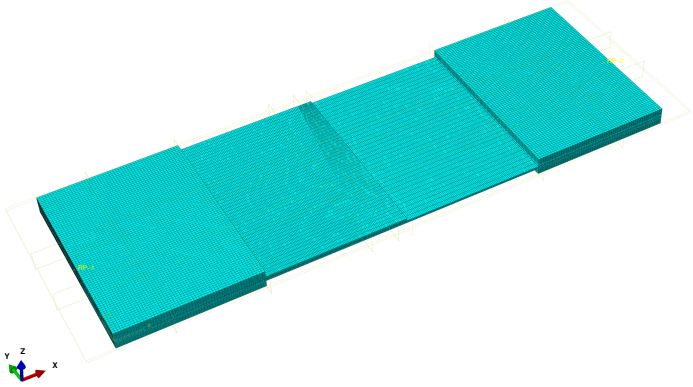


Fig. 6 The numerical model of the large CLS specimen.

B. Experimental and simulation datasets for domains adaptation

In this study, the strain values of 12 FBG sensors have been selected as the damage-dependent features. The comparison between strains measured in the two experiments is shown in Fig. 7, from which the effect of a different debonding front on the strain distribution can be appreciated.

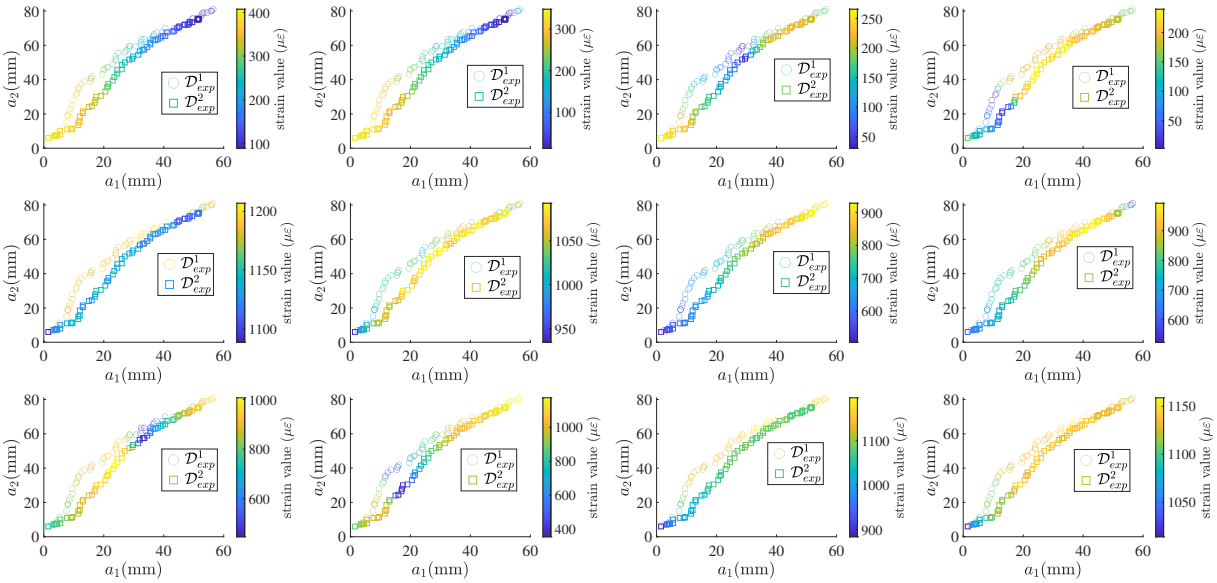


Fig. 7 Comparison of the strains of each FBG sensor between the first and the second CLS specimens. a_1 and a_2 are the debonding length at the lower and upper sides, respectively.

Three datasets are constructed in this study, which forms three domains. Two experimental datasets are extracted from the experiment of the two specimens. The first dataset $\mathcal{D}^1_{\text{exp}}$ consists of 71 samples, while the second one $\mathcal{D}^2_{\text{exp}}$ consists of 68 samples. The simulation dataset $\mathcal{D}_{\text{sim}}$ is generated from a numerical simulation with different debonding

fronts, which consists of 18 samples with a parallel debonding front ($a_1 = a_2$) and 174 samples with the debonding front being modeled as a straight line between a_1 and a_2, which is a simplification of the realistic debonding.

C. Implementation and parameters setting

The validation of the proposed method in an online damage quantification framework is carried out by shifting the source domain $\mathcal{D}_s$ and the target domain $\mathcal{D}^t$, performing adaptation in between the experimental datasets $\mathcal{D}^1_{\text{exp}}$ and $\mathcal{D}^2_{\text{exp}}$ and from the simulated dataset $\mathcal{D}_{\text{sim}}$ to the experimental ones. The practical implementation of domain adaptation is performed as follows: at each step there are n_{tl} labeled samples in the target domain, denoted as $\mathcal{D}_{tl}$. n_{tu} newly unlabeled samples, denoted as $\mathcal{D}^{tu}$, are processed conducting domain adaption, and their labels are predicted. Finally, the labels of the n_{tu} samples are revealed and added to the $\mathcal{D}_{tl}$ domain. n_{tu} is also referred to as the prediction step-size ΔN, and $n_{tl} = n_{tu}$ at the beginning of the task. In other words, $n_{tl} = \Delta N, 2\Delta N, 3\Delta N, \ldots, n_t$ respectively during the online damage quantification. In Sec. IV the method is compared with respect to three different prediction step sizes. It is worth noting that the debonding lengths in $\mathcal{D}_{tl}$ are smaller than those in $\mathcal{D}^{tu}$ because of the nature of damage growth.

For each analyzed case, the values of parameters of OCJDAR used for this application are shown in Tab. 1. λ is set to 10 so that more weight is given to the variance of the data with respect to the MMD difference. k is the number of eigenvalues in Eq. 10, and it is set to 10, which is a little smaller than the number of original features and may remove some noise. The number of random sampling operations in the MK-MMD calculation n_run is set to 100. Four clustering methods and a different number of clusters C are compared in Sec. IV.C, while in Sec. IV.A and IV.B, spectral cluster is chosen as the clustering method and the number of clusters is set to $C = 8$.

Gaussian Process Regression (GPR) [43], which is widely used [44, 45] and works well on small datasets, is chosen as the regressor for all tasks. The Radial Basis Function kernel is selected for the regressor, whose hyperparameters are estimated by Maximum Likelihood Estimate.

Table 1 Parameters of the OFJDAR algorithm for the large CLS application.

Parameter	Value
m	20
λ	10
k	10
n_{run}	100
Cluster algorithm	Spectral Cluster
C	8

D. Baseline methods

Four baseline methods are presented in this subsection to compare the performance of the proposed method.

Only source domain (OSD): Only the labeled strain data in the source domain $\mathcal{D}_s$ are chosen to train the regressor.

Only target domain (OTD): Only the available labeled strain data $\mathcal{D}_{tl}$ (with smaller damage sizes) in the target domain are chosen to train the regressor.

Combine two domains (CTD): the labeled strain data in the source domain $\mathcal{D}_s$ and target domain $\mathcal{D}_{tl}$ of the two previous cases are directly combined to train the regressor.

Online Transfer Component Analysis for Regression (OTCAR): the adaptation is conducted by the TCA method.

E. Performance evaluation

The accuracy of the debonding quantification is evaluated by the Root Mean Square Error (RMSE) metric:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathbf{y}_i - \hat{\mathbf{y}}_i)^2} \tag{15}$$

where $\mathbf{y}_i$ is the i_{th} true label, $\hat{\mathbf{y}}_i$ is the i_{th} prediction. n is the number of tested samples.

During the online debonding prediction, the predicted debonding lengths are stored at each step. When all the predictions are made, the stored predictions are compared with the true lengths by calculating the RMSE with Eq. 15.

IV. Results and Discussion

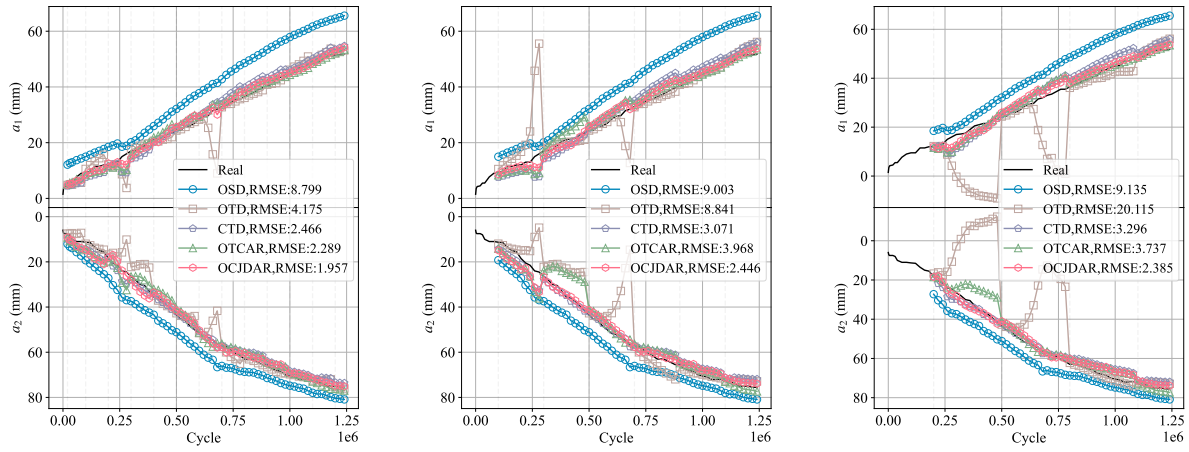
This section reported the performances of domain adaptation with two types of adaptation scenarios. In Sec. IV.A, domain adaptation is conducted between two experiments with nominally identical specimens. Meanwhile, domain adaptations from simulations to experiments are also carried out in Sec. IV.B, which includes more domain differences. The sensitivity to the prediction step sizes is addressed in both scenarios. Sec. IV.C investigates the effects of the four clustering methods and varying numbers of clusters.

A. Domain adaptation across two experiments

Since two specimens are available, two scenarios of domain adaptation are carried out: from $\mathcal{D}^1_{\text{exp}}$ to $\mathcal{D}^2_{\text{exp}}$ and vice versa.

The adaptation performances from $\mathcal{D}^1_{\text{exp}}$ to $\mathcal{D}^2_{\text{exp}}$ are reported in Fig. 8, It is obvious that the Only Source Domain (OSD), which directly uses the model trained on $\mathcal{D}^1_{\text{exp}}$ to quantify the debonding state of $\mathcal{D}^2_{\text{exp}}$, results in large biases regardless of the step size. The Only Target Domain (OTD), which only uses available labeled data in $\mathcal{D}^2_{\text{exp}}$ to quantify the next debonding state, has a large error in some steps, since OTD is essentially an extrapolation: in fact, the error of OTD grows significantly increasing the prediction step size. CTD methods directly combine the available labeled data in two domains without any adaption and obtain better results than OSD and OTD. However, as reported in [28], in some cases, simple fusion might lead to worse prediction: in this case, the performances of CTD are inferior to the proposed method. The performance of the baseline domain adaptation method OTCAR is not stable and it results in

high discontinuities. It performs better than CTD when $\Delta N = 5$, but worse when $\Delta N = 10$ and $\Delta N = 15$. Our proposed OCJDAR method outperforms baseline methods for all step-sizes, and the results are relatively stable.



(a) $\mathcal{D}^1_{\text{exp}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 5$ (b) $\mathcal{D}^1_{\text{exp}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 10$ (c) $\mathcal{D}^1_{\text{exp}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 15$

Fig. 8 Domain adaptation from $\mathcal{D}^1_{\text{exp}}$ to$\mathcal{D}^2_{\text{exp}}$, which is the experimental dataset from the first and second large CLS specimens respectively. The vertical dashed lines indicate the start of each step.

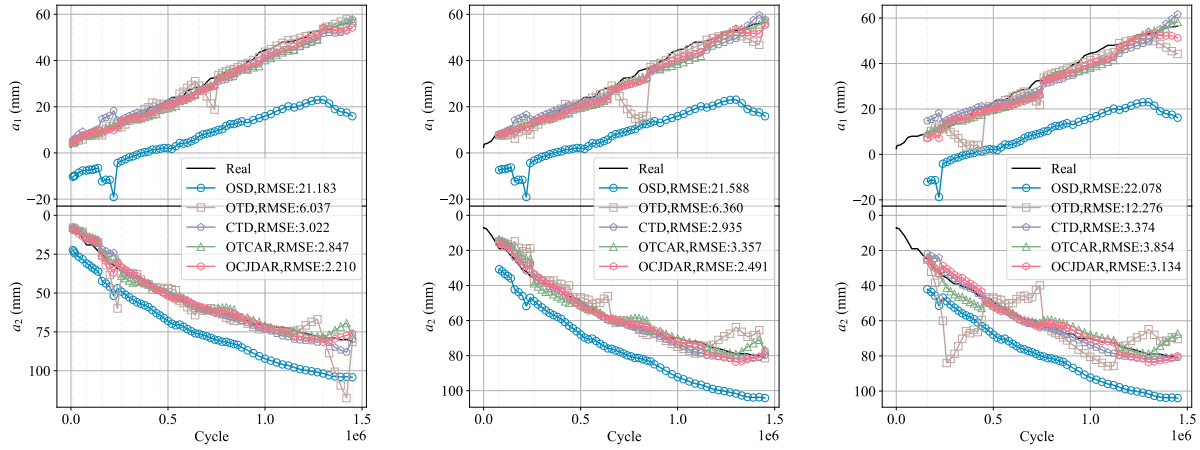
In the adaptation from $\mathcal{D}^2_{\text{exp}}$ to $\mathcal{D}^1_{\text{exp}}$, which is shown in Fig. 9, the performance of all these methods is similar. The bias in the OSD prediction changed from positive to negative. OCJDAR retains the best performance. Towards the end of the prediction, all the methods show a large deviation from the true target, which is linked to unexpected FBG measurements.

The results of the two adaptations cross experiments show that our proposed method can improve the debonding prediction by utilizing the damage data from similar structures, which has great potential for application in structures of the same type in a fleet, since due to the manufacturing variabilities and different usage and environmental conditions the damage propagation may vary. With the proposed method, it is feasible to use data obtained from structures with more severe damage to assist in damage quantification of structures with less severe damage.

B. Domain adaptation from simulation to experiments

Two domain adaptations are also conducted from simulations to experiments, which is also a common scenario in damage quantification field.

The results of the first adaptation from $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$ are shown in Fig. 10. It is obvious that the performance of all the methods is worse than those in the adaptation crossing experiments, which is caused by the much larger domain difference between the simulation and the experimental dataset. The OSD is not plotted since the RMSE (blue text in Fig. 10) is very large. The discrepancy between the numerical simulation and the experiment can be reasonably related to a



(a) $\mathcal{D}^2_{\text{exp}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 5$ (b) $\mathcal{D}^2_{\text{exp}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 10$ (c) $\mathcal{D}^2_{\text{exp}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 15$

Fig. 9 Domain adaptation from $\mathcal{D}^2_{\text{exp}}$ to $\mathcal{D}^1_{\text{exp}}$, which is the experimental dataset from the second and first large CLS specimens respectively. The vertical dashed lines indicate the start of each step.

strain bias between the FE model results and the experimental test, which is mainly due to modeling assumptions, such as the linear debonding front. However, the result of CTD is still better than OTD, and OCJDAR still demonstrates the best accuracy. Hence, the simulation dataset still contains some useful information that is helpful for domain adaptation.

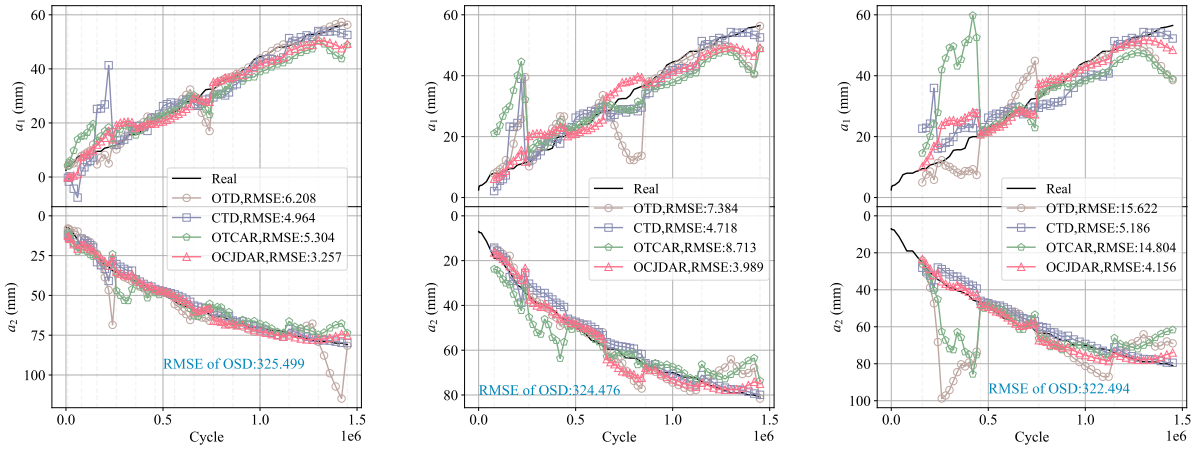
The overall pattern of the adaptation from $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^2_{\text{exp}}$ is similar to the one shown above, as shown in Fig. 11. However, the RMSE of CTD and OCJDAR is smaller than that in $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$. This may be because $\mathcal{D}_{\text{sim}}$ differs less from $\mathcal{D}^2_{\text{exp}}$ than $\mathcal{D}^1_{\text{exp}}$, thus achieving better domain adaptation.

The results of the two adaptations from simulations to experiments show that our proposed method can improve the debonding prediction utilizing the damage data from simulations, which has great potential for application in digital twins. The digital twin model, usually built from simulated data, will inevitably have some differences from physical entities. The use of observed data from physical entities with domain adaptation methods can reduce the discrepancy between the real world and the simulation space, improving the accuracy of Digital Twin predictions.

C. Effects of clustering methods and the number of clusters

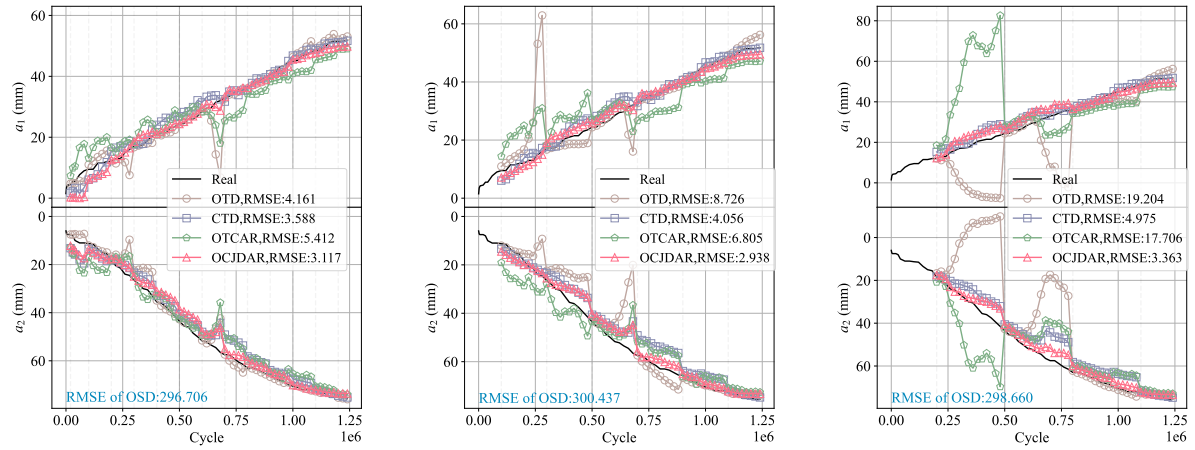
The accuracy of clustering methods affects the measurement of the conditional distribution difference and therefore impacts the adaptation quality. The comparison of a different number of clusters with four clustering methods is shown in Fig. 12. The average value of each clustering method at a different number of clusters C is marked with dotted lines.

Regarding the clustering method, we can see that in most cases spectral clustering achieves the lowest average RMSE, especially for the adaptation cross experiments, and the RMSE decreases increasing C. The performance of Fuzzy C-mean (FCM) is stable with respect to the change of C in all the four domain adaptations. However, its RMSE



(a) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 5$ (b) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 10$ (c) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$, $\Delta N = 15$

Fig. 10 Domain adaptation from $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^1_{\text{exp}}$. $\mathcal{D}_{\text{sim}}$ is the simulation dataset, and $\mathcal{D}^1_{\text{exp}}$ is the experimental dataset from the first large CLS specimen. The RMSE of OSD is too large, so the curve is removed in the three figures, and its value is marked in blue. The vertical dashed lines indicate the start of each step.



(a) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 5$ (b) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 10$ (c) $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^2_{\text{exp}}$, $\Delta N = 15$

Fig. 11 Domain adaptation from $\mathcal{D}_{\text{sim}}$ to $\mathcal{D}^2_{\text{exp}}$. $\mathcal{D}_{\text{sim}}$ is the simulation dataset, and $\mathcal{D}^2_{\text{exp}}$ is the experimental dataset from the second large CLS specimen. The RMSE of OSD is too large, so the curve is removed in the three figures, and its value is marked in blue. The vertical dashed lines indicate the start of each step.

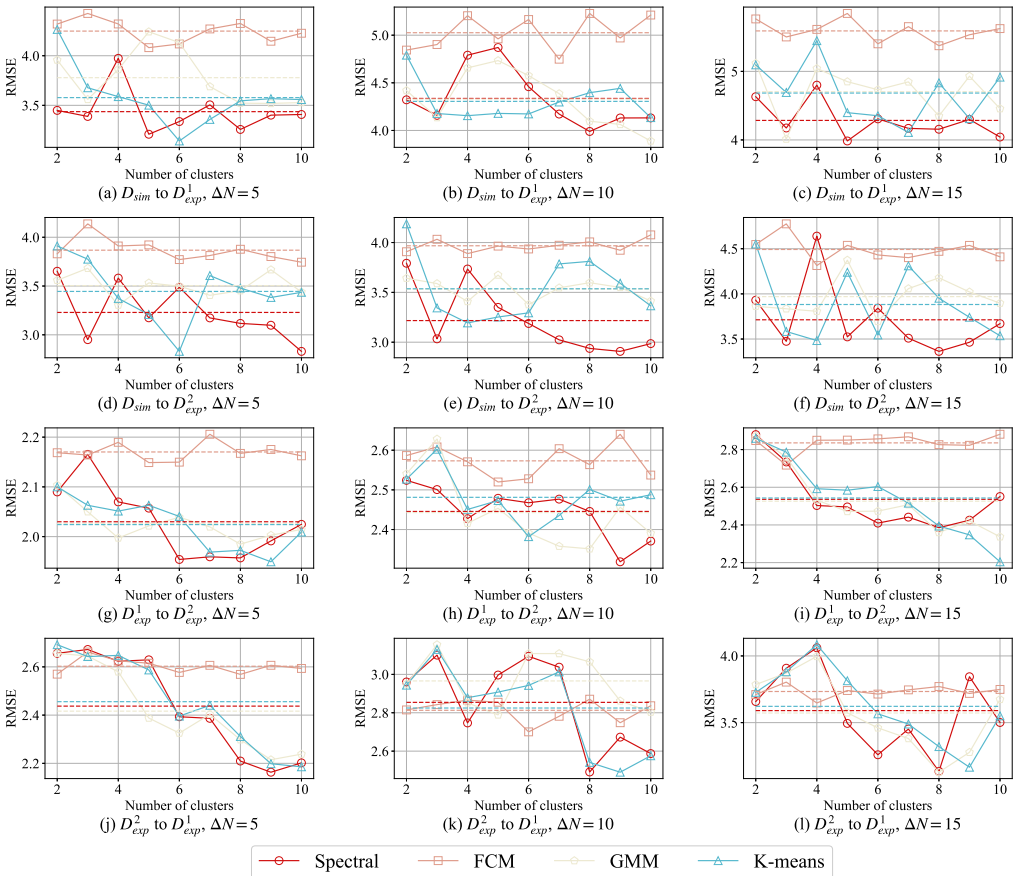


Fig. 12 Comparison of different clustering methods and a number of clusters for the four adaptation scenarios. The dot lines indicate the average value of a different number of clusters for each clustering method.

is the highest. The performance of K-means and Gaussian Mixture Models lie in-between FCM and Spectral clustering. Overall, soft clustering methods (GMM and FCM) do not show superiority over hard clustering methods (spectral cluster and k-means), indicating that the accuracy of clustering is more important than the ambiguity of clustering labels for the MK-MMD calculation. The results show that the selection of clustering method is critical for the proposed method, which can also be confirmed by [39]. The spectral cluster is deemed suitable for this study, thus it has been selected for Sec. IV.A and Sec. IV.B. In terms of the number of clusters C, the RMSEs vary with the change of the C, indicating that an improper C may decrease the quality of domain adaptations. It can be seen that for this study, a large C for the spectral cluster can get the best performance in most cases, and $C = 8$ is selected. However, the selection of the optimal C is not a trivial task in general and future research could be devoted to it [42].

V. Conclusion

In this study, a novel domain adaptation method for multi-output regression, OCJDAR, is proposed. The key novelty of this method relies on obtaining the class labels by clustering methods and to measure the discrepancy of the conditional distribution by class-wise MK-MMD. The proposed method is integrated into a framework for online structural damage quantification with domain adaptation.

Three datasets containing debonding samples of Cracked Lap Shear specimens from experiments and simulations are exploited to evaluate the proposed method's performance, which is evaluated performing two types of domain adaptation: simulation to experiment and across experiments. The results demonstrate that direct use of data from simulations or other structures for damage quantification of the current structure is highly inaccurate, which advocates the need for domain adaptive methods. The results are also benchmarked against various state-of-the-art methods, accounting for distinct prediction step-sizes, demonstrating that the proposed domain adaptation methods and framework can significantly improve the online damage quantification accuracy and perform robustly in the real environment. A sensitivity analysis has been performed on different clustering methods and the number of clusters, so that the spectral cluster is suggested to be used in CJDAR at least for this application scenario. However, choosing the appropriate number of clusters for a general application is still an open problem that needs further exploration.

Although the debonding application in this study is a two-output problem, the proposed method is general for multi-output problems. Future work could be devoted to improving the robustness of domain adaptation with respect to hyperparameter variations and exploring its wide range of applications. The authors are working on combining the proposed framework with a physics-based Reduced-Order Model Digital-Twin, so that domain adaptation may improve the accuracy of Model Updating strategies in a nominally-identical but heterogeneous population of structures.

Funding Sources

This work has been developed based on the results from PATCHBOND II project (Certification of adhesive bonded repairs for Primary Aerospace composite structures), a Cat.-B project (No B.PRJ.RT.670) coordinated by the European Defense Agency (EDA) and financed by the following nations: The Netherlands, Norway, Italy, Germany, Finland, Belgium, and Czech Republic. The composite specimens were produced by NLR - Netherlands Aerospace Center.

References

[1] Pantelakis, S., and Tserpes, K. I., "Adhesive Bonding of Composite Aircraft Structures: Challenges and Recent Developments," *Science China Physics, Mechanics and Astronomy*, Vol. 57, No. 1, 2014, pp. 2–11. https://doi.org/10.1007/s11433-013-5274-3.

[2] Davis, M., and Bond, D., "Principles and Practices of Adhesive Bonded Structural Joints and Repairs," *International Journal of Adhesion and Adhesives*, Vol. 19, No. 2, 1999, pp. 91–105. https://doi.org/10.1016/S0143-7496(98)00026-8.

[3] Bland, S. M., and Kapania, R. K., "Damage Identification of Plate Structures Using a Hybrid Genetic-Sensitivity Approach," *AIAA Journal*, Vol. 43, No. 2, 2005, pp. 439–442. https://doi.org/10.2514/1.3857.

[4] Soejima, H., Nakamura, N., Ogisu, T., Wakai, H., Okabe, Y., Takeda, N., and Koshioka, Y., "Development of a Structural Health Monitoring System to Evaluate De-Bonding in Composite Adhesive Structures," *ICAF 2009, Bridging the Gap between Theory and Operational Practice*, edited by M. J. Bos, Springer Netherlands, Dordrecht, Netherlands, 2009, pp. 1187–1197. https://doi.org/10.1007/978-90-481-2746-7_66.

[5] Reddy Lyathakula, K., and Yuan, F.-G., "Fatigue Damage Diagnostics–Prognostics Framework for Remaining Life Estimation in Adhesive Joints," *AIAA Journal*, Vol. 60, No. 8, 2022, pp. 4874–4892. https://doi.org/10.2514/1.J060979.

[6] Namura, N., Muto, K., Ueki, Y., Ueta, R., and Takeda, N., "Fatigue Damage Monitoring of Wind Turbine Blades by Using the Kalman Filter," *AIAA Journal*, Vol. 59, No. 12, 2021, pp. 5130–5140. https://doi.org/10.2514/1.J060254.

[7] Cadini, F., Sbarufatti, C., Corbetta, M., Cancelliere, F., and Giglio, M., "Particle Filtering-Based Adaptive Training of Neural Networks for Real-Time Structural Damage Diagnosis and Prognosis," *Structural Control and Health Monitoring*, Vol. 26, No. 12, 2019, p. e2451. https://doi.org/10.1002/stc.2451.

[8] Li, T., Sbarufatti, C., Cadini, F., Chen, J., and Yuan, S., "Particle Filter-Based Hybrid Damage Prognosis Considering Measurement Bias," *Structural Control and Health Monitoring*, Vol. 29, No. 4, 2022, p. e2914. https://doi.org/10.1002/stc.2914.

[9] Colombo, L., Oboe, D., Sbarufatti, C., Cadini, F., Russo, S., and Giglio, M., "Shape Sensing and Damage Identification with iFEM on a Composite Structure Subjected to Impact Damage and Non-Trivial Boundary Conditions," *Mechanical Systems and Signal Processing*, Vol. 148, 2021, p. 107163. https://doi.org/10.1016/j.ymssp.2020.107163.

[10] Sbarufatti, C., Manes, A., and Giglio, M., "Application of Sensor Technologies for Local and Distributed Structural Health Monitoring," *Structural Control and Health Monitoring*, Vol. 21, No. 7, 2014, pp. 1057–1083. https://doi.org/10.1002/stc.1632.

[11] Tuegel, E. J., Ingraffea, A. R., Eason, T. G., and Spottswood, S. M., "Reengineering Aircraft Structural Life Prediction Using a Digital Twin," *International Journal of Aerospace Engineering*, Vol. 2011, 2011, p. 154798. https://doi.org/10.1155/2011/154798.

[12] Sbarufatti, C., Manes, A., and Giglio, M., "Performance Optimization of a Diagnostic System Based upon a Simulated Strain Field for Fatigue Damage Characterization," *Mechanical Systems and Signal Processing*, Vol. 40, No. 2, 2013, pp. 667–690. https://doi.org/10.1016/j.ymssp.2013.06.003.

[13] Soother, D. K., Ujjan, S. M., Dev, K., Khowaja, S. A., Bhatti, N. A., and Hussain, T., "Towards Soft Real-Time Fault Diagnosis for Edge Devices in Industrial IoT Using Deep Domain Adaptation Training Strategy," *Journal of Parallel and Distributed Computing*, Vol. 160, 2022, pp. 90–99. https://doi.org/10.1016/j.jpdc.2021.10.005.

[14] Liu, B., Zhang, Y., Lv, J., Majeed, A., Chen, C.-H., and Zhang, D., "A Cost-Effective Manufacturing Process Recognition Approach Based on Deep Transfer Learning for CPS Enabled Shop-Floor," *Robotics and Computer-Integrated Manufacturing*, Vol. 70, 2021, p. 102128. https://doi.org/10.1016/j.rcim.2021.102128.

[15] Truong, J., Chernova, S., and Batra, D., "Bi-Directional Domain Adaptation for Sim2Real Transfer of Embodied Navigation Agents," *IEEE Robotics and Automation Letters*, Vol. 6, No. 2, 2021, pp. 2634–2641. https://doi.org/10.1109/LRA.2021.3062303.

[16] Guan, H., and Liu, M., "Domain Adaptation for Medical Image Analysis: A Survey," *IEEE Transactions on Biomedical Engineering*, Vol. 69, No. 3, 2022, pp. 1173–1185. https://doi.org/10.1109/TBME.2021.3117407.

[17] Pan, S. J., and Yang, Q., "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 22, No. 10, 2010, pp. 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

[18] Pan, S. J., Tsang, I. W., Kwok, J. T., and Yang, Q., "Domain Adaptation via Transfer Component Analysis," *IEEE Transactions on Neural Networks*, Vol. 22, No. 2, 2011, pp. 199–210. https://doi.org/10.1109/TNN.2010.2091281.

[19] Long, M., Wang, J., Ding, G., Sun, J., and Yu, P. S., "Transfer Feature Learning with Joint Distribution Adaptation," *Proceedings of the 2013 IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers, New York City, United States, 2013, pp. 2200–2207. https://doi.org/10.1109/ICCV.2013.274.

[20] Chakraborty, D., Kovvali, N., Chakraborty, B., Papandreou-Suppappola, A., and Chattopadhyay, A., "Structural Damage Detection with Insufficient Data Using Transfer Learning Techniques," *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2011*, Vol. 7981, International Society for Optics and Photonics, Bellingham WA, 2011, p. 798147. https://doi.org/10.1117/12.882025.

[21] Zhang, Z., Sun, C., and Guo, B., "Transfer-Learning Guided Bayesian Model Updating for Damage Identification Considering Modeling Uncertainty," *Mechanical Systems and Signal Processing*, Vol. 166, 2022, p. 108426. https://doi.org/10.1016/j.ymssp.2021.108426.

[22] Buisman, S. v. B., Michau, G., Rene, A., and Fink, O., "Domain Adaptations for Guided Wave SHM of Composites: Towards Fleet Monitoring," *Proceedings of the 6th European Conference of the Prognostics and Health Management Society 2021*, Vol. 6, Prognostics and Health Management Society, State College, PA, 2021, pp. 430–438. https://doi.org/10.3929/ethz-b-000506104.

[23] Bull, L. A., Gardner, P. A., Gosliga, J., Rogers, T. J., Dervilis, N., Cross, E. J., Papatheou, E., Maguire, A. E., Campos, C., and Worden, K., "Foundations of Population-Based SHM, Part I: Homogeneous Populations and Forms," *Mechanical Systems and Signal Processing*, Vol. 148, 2021, p. 107141. https://doi.org/10.1016/j.ymssp.2020.107141.

[24] Gosliga, J., Gardner, P. A., Bull, L. A., Dervilis, N., and Worden, K., "Foundations of Population-based SHM, Part II: Heterogeneous Populations – Graphs, Networks, and Communities," *Mechanical Systems and Signal Processing*, Vol. 148, 2021, p. 107144. https://doi.org/10.1016/j.ymssp.2020.107144.

[25] Gardner, P., Bull, L. A., Gosliga, J., Dervilis, N., and Worden, K., "Foundations of Population-Based SHM, Part III: Heterogeneous Populations – Mapping and Transfer," *Mechanical Systems and Signal Processing*, Vol. 149, 2021, p. 107142. https://doi.org/10.1016/j.ymssp.2020.107142.

[26] Kim, I.-H., Jeon, H., Baek, S.-C., Hong, W.-H., and Jung, H.-J., "Application of Crack Identification Techniques for an Aging Concrete Bridge Inspection Using an Unmanned Aerial Vehicle," *Sensors*, Vol. 18, No. 6, 2018, p. 1881. https://doi.org/10.3390/s18061881.

[27] Wu, D., Lawhern, V. J., Gordon, S., Lance, B. J., and Lin, C.-T., "Driver Drowsiness Estimation From EEG Signals Using Online Weighted Adaptation Regularization for Regression (OwARR)," *IEEE Transactions on Fuzzy Systems*, Vol. 25, No. 6, 2017, pp. 1522–1535. https://doi.org/10.1109/TFUZZ.2016.2633379.

[28] Zhou, X., Sbarufatti, C., Giglio, M., and Dong, L., "A Fuzzy-set-based Joint Distribution Adaptation Method for Regression and Its Application to Online Damage Quantification for Structural Digital Twin," *arXiv:2211.02656 [cs]*, 2022. https://doi.org/10.48550/arXiv.2211.02656.

[29] Gretton, A., Sejdinovic, D., Strathmann, H., Balakrishnan, S., Pontil, M., Fukumizu, K., and Sriperumbudur, B. K., "Optimal Kernel Choice for Large-Scale Two-Sample Tests," *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Curran Associates, Inc., Red Hook, NY, 2012, pp. 1205–1213.

[30] Long, M., Cao, Y., Wang, J., and Jordan, M., "Learning Transferable Features with Deep Adaptation Networks," *Proceedings of the 32nd International Conference on International Conference on Machine Learning*, Vol. 37, Proceedings of Machine Learning Research, Cambridge MA, 2015, pp. 97–105.

[31] Liu, W., and Qin, R., "A MultiKernel Domain Adaptation Method for Unsupervised Transfer Learning on Cross-Source and Cross-Region Remote Sensing Data Classification," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58, No. 6, 2020, pp. 4279–4289. https://doi.org/10.1109/TGRS.2019.2962039.

[32] Simon-Gabriel, C.-J., and Schölkopf, B., "Kernel Distribution Embeddings: Universal Kernels, Characteristic Kernels and Kernel Metrics on Distributions," *The Journal of Machine Learning Research*, Vol. 19, No. 1, 2018, pp. 1708–1736.

[33] Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A., "A Kernel Two-Sample Test," *The Journal of Machine Learning Research*, Vol. 13, No. 1, 2012, pp. 723–773.

[34] Romesburg, H. C., *Cluster Analysis for Researchers*, reprint of 1984 edition ed., Lulu Press, Morrisville, NC, 2004.

[35] Ban, N., and Yamazaki, W., "Black-Box Function Aerodynamic Topology Optimization Algorithm via Machine Learning Technologies," *AIAA Journal*, Vol. 59, No. 12, 2021, pp. 5174–5185. https://doi.org/10.2514/1.J059605.

[36] Barone, M., Ray, J., and Domino, S., "Feature Selection, Clustering, and Prototype Placement for Turbulence Datasets," *AIAA Journal*, Vol. 60, No. 3, 2022, pp. 1332–1346. https://doi.org/10.2514/1.J060919.

[37] Gao, Z., Puranik, T. G., and Mavris, D. N., "Probabilistic REpresentatives Mining (PREM): A Clustering Method for Distributional Data Reduction," *AIAA Journal*, Vol. 60, No. 4, 2022, pp. 2580–2596. https://doi.org/10.2514/1.J061079.

[38] MacQueen, J., "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California Press, Oakland, CA, USA, 1967, pp. 281–297.

[39] von Luxburg, U., "A Tutorial on Spectral Clustering," *Statistics and Computing*, Vol. 17, No. 4, 2007, pp. 395–416. https://doi.org/10.1007/s11222-007-9033-z.

[40] Dunn, J. C., "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," *Journal of Cybernetics*, Vol. 3, No. 3, 1973, pp. 32–57. https://doi.org/10.1080/01969727308546046.

[41] Reynolds, D., "Gaussian Mixture Models," *Encyclopedia of Biometrics*, edited by S. Z. Li and A. Jain, Springer US, Boston, MA, 2009, pp. 659–663. https://doi.org/10.1007/978-0-387-73003-5_196.

[42] Jain, A. K., "Data Clustering: 50 Years beyond K-means," *Pattern Recognition Letters*, Vol. 31, No. 8, 2010, pp. 651–666. https://doi.org/10.1016/j.patrec.2009.09.011.

[43] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, Mass, 2006.

[44] Zhou, X., He, S., Dong, L., and Atluri, S. N., "Real-Time Prediction of Probabilistic Crack Growth with a Helicopter Component Digital Twin," *AIAA Journal*, Vol. 60, No. 4, 2022, pp. 2555–2567. https://doi.org/10.2514/1.J060890.

[45] Lin, Q., Hu, J., Zhang, L., Jin, P., Cheng, Y., and Zhou, Q., "Gradient-Enhanced Multi-Output Gaussian Process Model for Simulation-Based Engineering Design," *AIAA Journal*, Vol. 60, No. 1, 2022, pp. 76–91. https://doi.org/10.2514/1.J060728.