

A novel interpretable domain adaptive framework for robust damage detection in composite structures under environmental variability

Nima Rezazadeh^{1,2} , Alessandro De Luca² , Giuseppe Lamanna² ,
Fawaz Annaz¹  and Mario de Oliveira¹ 

Structural Health Monitoring

1–37

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14759217261433879

journals.sagepub.com/home/shm



Abstract

Structural health monitoring of mechanical assets can be hindered by environmental variability that causes distribution shifts between training and deployment. Many domain adaptation (DA) methods mitigate these shifts but behave as black boxes with limited insight into how representations change. This work introduces a novel interpretable framework, scattering-based prototype-aligned DA, that combines physics-guided feature extraction, synthetic data generation and prototype-based alignment for robust damage detection under temperature variation. A convolutional conditional variational autoencoder, trained on healthy data across temperatures with multi-domain reconstruction losses, generates temperature-conditioned synthetic damaged guided-wave signals from limited baseline damage measurements and healthy responses, creating a controlled testbed when damaged data at other temperatures are unavailable. Prototype-based domain adversarial training with gradient reversal and entropy-gated pseudo-labelling aligns source and target feature manifolds while preserving damage-sensitive patterns. Interpretability modules based on prototype trajectories, instance to prototype similarities and low-dimensional visualisations reveal how decision boundaries and latent representations evolve. Experiments on composite structures across temperatures show that the framework improves robustness over baselines and maintains high diagnostic accuracy while providing actionable insight into the adaptation process and enabling informed diagnostic assessment by domain experts in safety-critical contexts.

Keywords

Structural health monitoring, composite structures, physics-guided approach, prototype-based learning, environmental variability

Highlights

- Interpretable domain adaptation improves damage detection under temperature change.
- Physics-guided features reduce environmental influence while keeping damage patterns.
- Synthetic temperature-conditioned signals enable training when damage data are scarce.
- Prototype-based alignment explains how decisions shift across operating conditions.
- Validated on composite plates and wind turbine blades across multiple temperatures.

Introduction

Structural health monitoring (SHM) of mechanical assets faces fundamental challenges when deploying damage detection models across data scarcity and

¹School of Architecture, Built Environment, Computing and Engineering, Birmingham City University, Birmingham, UK

²Department of Engineering, University of Campania “L. Vanvitelli,” Aversa, Italy

Corresponding author:

Nima Rezazadeh, School of Architecture, Built Environment, Computing and Engineering, Birmingham City University, Birmingham, B4 7XG, UK.
Email: Nima.Rezazadeh@mail.bcu.ac.uk

varying operational conditions.¹ Temperature fluctuations, humidity variations, boundary condition changes and manufacturing tolerances introduce distributional shifts that severely degrade the performance of models trained in laboratory settings when applied to field conditions.² This domain shift problem is particularly acute in safety-critical applications such as wind turbine blades (WTBs), aerospace composites and civil infrastructure, where acquiring labelled damage data for each individual asset is economically prohibitive, while the consequences of missed detections can be catastrophic.³ The core challenge lies not merely in achieving robust performance but in providing transparent evidence that damage-sensitive features are preserved during the transition from controlled environments to operational deployment.⁴

Environmental and operational variabilities (EOVs) have motivated extensive research into robust feature extraction methods. Classical time and frequency domain descriptors, while computationally efficient, exhibit brittleness under temperature and operational variations.⁵ Wavelet transform and its variants such as wavelet time scattering (WTS) have emerged as a promising physics-guided approach, providing deformation-stable representations through cascaded wavelet transforms and modulus operations without requiring learned parameters.⁶ Ojha et al.⁷ demonstrated improved robustness compared with conventional scalogram features in composite impact localisation, while Rezazadeh et al.⁸ showed effective extraction of damage-sensitive features under operational variations in rotor systems using WTS combined with long short-term memory networks. Ma et al.⁹ coupled scattering transform features with least squares proximal twin support vector machines, demonstrating robust fault diagnosis under noise in rotating machinery. However, these applications treat WTS merely as preprocessing without exploiting its invariance properties to facilitate domain adaptation (DA). Critically, existing work does not leverage the well-characterised stability properties of WTS to provide robust features that support knowledge transfer between operational conditions.

The challenge of transferring knowledge between different operational regimes has driven the development of DA techniques for SHM. Bull et al.^{10,11} established population-based SHM concepts for knowledge sharing across similar assets through statistical mixture models. Wang et al.¹² employed adversarial DA between finite element simulation and experimental data for fatigue crack detection, while da Silva et al.¹³ used transfer component analysis to stabilise impedance-based diagnostics under temperature changes. Yang et al.¹⁴ introduced multi-source dynamic adaptive generalisation for composite crack

detection without requiring target domain labels. These methods achieve domain alignment through statistical criteria but most provide limited transparency regarding which signal characteristics are being aligned or whether physically meaningful patterns are preserved during adaptation. More recent efforts have begun to incorporate physical knowledge into the adaptation process, for instance by embedding physics-informed constraints within generative adversarial domain alignment for SHM,¹⁵ while contrastive learning strategies have been integrated with adversarial DA to improve class-level feature discriminability during cross-domain fault diagnosis.¹⁶ Notwithstanding these advances, existing approaches rarely offer systematic tools for tracking how class-level representations reorganise throughout the adaptation process or for verifying that damage-sensitive structure is maintained after domain alignment.

Environmental compensation strategies have evolved from baseline subtraction towards reference-free approaches that exploit physical invariants.¹⁷ Salmanpour et al.^{18,19} developed comprehensive temperature correction procedures including minimum residual alignment, single baseline correction and instantaneous baseline mapping for guided wave monitoring, demonstrating effective compensation across thermal ranges. Amer and Kopsaftopoulos²⁰ embedded structural damage indices within Gaussian process regression to achieve baseline-free inference with probabilistic outputs. Yue et al.²¹ proposed relative referencing that compares paths to each other rather than historical records, effectively reducing environmental drift at the feature level. While these compensation techniques successfully mitigate environmental effects, they operate on raw signals or hand-crafted features without leveraging the structured mathematical invariances that methods like WTS inherently provide, and function independently of DA mechanisms.

The deployment of machine learning models in safety-critical infrastructure has intensified the need for interpretable diagnostic systems.²² Post hoc attribution techniques have been widely applied, with Kim and Kim²³ adapting Gradient-weighted Class Activation Mapping for vibration data, Yan et al.²⁴ employing SHapley Additive exPlanations to identify physics-guided features and Hanchate et al.²⁵ applying Local Interpretable Model-agnostic Explanations to highlight influential time-frequency regions. Ante-hoc interpretable architectures have received more limited attention, with Chen and Dong²⁶ proposing Sparse Temporal Logic Networks and Li et al.²⁷ introducing Variational Attention-based Transformers with sparse Dirichlet priors. However, these interpretability approaches focus exclusively on explaining predictions within single operational conditions, failing to address how

representations transform during DA. Rezazadeh et al.²⁸ presented one of the few attempts to visualise internal changes during adaptation through activation pattern heat maps, although this post hoc visualisation does not guide the adaptation process itself. Prototype-based learning offers an interpretable alternative by representing each class through exemplar patterns corresponding to specific damage modes,²⁹ however its integration with DA for SHM remains largely unexplored.

The convergence of these research streams reveals critical gaps preventing confident deployment of domain-adaptive SHM systems. While WTS provides theoretically grounded invariances for environmental robustness, current applications do not exploit these properties to support DA, missing the opportunity to build adaptation upon physics-guided representations. DA techniques achieve statistical alignment without maintaining interpretable correspondence between source and target representations, leaving practitioners unable to verify that damage classes preserve their physical meaning. Environmental compensation and DA are treated separately rather than as jointly optimised processes. Most critically, no framework provides interpretable evidence of how damage-sensitive features transform during adaptation, preventing validation that transferred models retain physical validity rather than learning spurious correlations. These limitations raise fundamental research questions concerning how mathematical invariances of physics-guided feature extraction can be leveraged to facilitate interpretable DA in SHM. Furthermore, it remains unclear how damage-sensitive features preserve their physical meaning when transferred across operational conditions, and what mechanisms enable practitioners to verify that adaptation maintains structural rather than purely environmental patterns.

This work addresses these challenges through a novel unsupervised domain adaptation (UDA) framework (scattering-based prototype-aligned DA (SPADA)) integrating physics-guided feature extraction, prototype-based adaptation and dedicated diagnostic modules. The framework employs WTS as a principled front end, leveraging its mathematical stability to deformations and translation invariance to suppress certain EOVs while preserving damage-related modulations. The resulting scattering coefficients feed into a prototype-based UDA mechanism maintaining explicit correspondence between source and target exemplars. Separate interpretability modules, operating as post-training inspection tools without altering the optimisation dynamics, quantify prototype trajectories, instance-prototype similarity patterns, attention weight dynamics and low-dimensional separability, enabling practitioners to examine how the latent decision

structure evolves during adaptation and to identify potential failure modes before deployment decisions are made.

The primary contributions are fourfold.

- Demonstration of the mathematical invariances of WTS to provide stable, physics-guided features facilitating prototype-based UDA, ensuring feature representations are less sensitive to temperature while capturing damage-related modulations. This constitutes one of the first integrations of WTS-based feature extraction with interpretable transfer learning in SHM.
- An interpretable prototype tracking mechanism that reveals how damage classes reorganise between temperature conditions while maintaining correspondence to physical damage modes, providing practitioners with evidence that adaptation preserves structural rather than purely environmental patterns.
- An interpretability diagnostic method based on prototype trajectories, instance-prototype similarity matrices, attention weight evolution and low-dimensional visualisations that enable domain experts to assess whether changes remain consistent with known material behaviour under temperature variation.
- The framework achieved classification accuracy comparable to black-box approaches while equipping practitioners with diagnostic tools to assess whether damage-sensitive patterns are preserved during adaptation, offering an interpretable alternative to opaque statistical alignment procedures.

The remainder of this paper is structured as follows. The second section presents the methodology, detailing the fundamentals of WTS and the proposed interpretable UDA framework. The third section describes the case studies, including the composite plate dataset and the WTB benchmark. The fourth section provides the results and discussion, examining the effects of EOVs on signal characteristics, feature extraction results, data augmentation outcomes, damage detection performance with and without DA and the internal activities of the proposed framework. The fifth section summarises the conclusions and outlines future research directions.

Methodology and methods

This section presents two complementary frameworks for temperature-dependent SHM under EOVs. First, a convolutional conditional variational autoencoder

generates temperature-conditioned synthetic damaged signals from limited baseline damage and multi-temperature healthy data, addressing data scarcity. Second, the SPADA framework performs fault diagnosis under fully UDA, where target domain labels are unavailable. The integration of WTS, prototype-based attention and interpretability mechanisms enables robust cross-domain generalisation while maintaining diagnostic transparency.

Convolutional conditional variational autoencoder

A generative framework was developed to synthesise temperature-conditioned synthetic damaged guided-wave signals. This architecture integrates convolutional feature extraction with conditional variational inference to model the joint distribution of signals and temperature conditions. Latent-space offsets and amplitude masks produce synthetic signal variations reflecting temperature and structural damage influences inferred from the baseline condition, extending limited experimental datasets across operational environments.

Let $\mathcal{D}_s = \{(x^{(n)}, y^{(n)})\}_{n=1}^{N_s}$ and $\mathcal{D}_t = \{\tilde{x}^{(m)}\}_{m=1}^{N_t}$ represent labelled source and unlabelled target datasets, with $x^{(n)} \in \mathbb{R}^{N_{\text{time}} \times C}$ where C is the sensor count. The baseline condition at T_{Baseline} serves as reference. The encoder maps input signal $x^{(n)}$ and temperature T to latent representation z :

$$q_\phi(z|x^{(n)}, T) = \mathcal{N}(\mu_\phi(x^{(n)}, T), \sigma_\phi^2(x^{(n)}, T)I) \quad (1)$$

where μ_ϕ and σ_ϕ are convolutional layer parameters representing mean and per-dimension variance; $q_\phi(\cdot)$ is the posterior distribution; $\mathcal{N}(\cdot)$ is Gaussian distribution; I is the identity matrix ensuring diagonal covariance. The decoder reconstructs signals from latent variables and temperature:

$$\hat{x}^{(n)} = f_\theta(z, T) \quad (2)$$

For damage level l , latent offset Δz_l is computed relative to healthy baseline at T_{Baseline} :

$$\Delta z_l = \frac{1}{N} \sum_{i=1}^N z_{l,i}(T_{\text{Baseline}}) - \frac{1}{N} \sum_{i=1}^N z_{\text{healthy},i}(T_{\text{Baseline}}) \quad (3)$$

This offset characterises latent-space displacement from structural damage, serving as a transferable feature across temperatures. Per-channel and per-time amplitude masks are constructed from the analytic envelopes of the damaged and healthy signals at T_{Baseline} . For each damage level l and sensor channel, the Hilbert transform yields the instantaneous envelope of each

observation, and the element-wise ratio of the damaged envelope to the median healthy envelope is computed. The median of these ratios across observations is smoothed with a Gaussian kernel to obtain the deterministic mask $\mathcal{M}_l \in \mathbb{R}^{C \times N_{\text{time}}}$.

A stochastic component is derived from the interquartile range (IQR) of the same ratios: the per-element dispersion is set as a fraction of the estimated standard deviation (approximated as IQR divided by 1.35), capped at a fixed proportion of the deterministic mask value to prevent extreme perturbations. During generation (Equation (5)), the decoded signal is first modulated by the deterministic mask via element-wise multiplication, and an additive Gaussian noise term ϵ with per-element standard deviation set by this dispersion is then added to the masked decoded signal. This stochastic augmentation procedure was chosen over alternatives such as fixed additive noise or dropout-based augmentation because it preserves the spatial and temporal structure of the damage-induced amplitude modulation while introducing physically motivated variability that scales with the local signal-to-damage ratio. Implementation parameters are reported in ‘‘Conv-VAE’’ section.

The training objective is a weighted composite of five terms:

$$\begin{aligned} \mathcal{L}_{\text{CVAE}} = & \alpha_{\text{MSE}} \mathcal{L}_{\text{MSE}} + \alpha_{\text{spec}} \mathcal{L}_{\text{spec}} + \alpha_{\text{wav}} \mathcal{L}_{\text{wav}} \\ & + \alpha_{\text{env}} \mathcal{L}_{\text{env}} + \beta(e) \mathcal{L}_{\text{KL}} \end{aligned} \quad (4)$$

where \mathcal{L}_{MSE} is the mean squared error between reconstructed and input signals for temporal fidelity; $\mathcal{L}_{\text{spec}}$ is the log-magnitude spectral loss computed from the real fast Fourier transform for frequency-domain preservation; \mathcal{L}_{wav} is the mean cosine distance between multi-level wavelet decomposition coefficients of the input and reconstruction for multi-resolution structural similarity; and \mathcal{L}_{env} is the L_1 deviation of the peak analytic envelope ratio from unity, penalising amplitude distortion in the primary wave packet. The Kullback–Leibler divergence \mathcal{L}_{KL} regularises the approximate posterior towards the standard normal prior. The coefficient $\beta(e)$ follows a linear warm-up schedule to prevent posterior collapse during early training. The fixed weights $\alpha_{(\cdot)}$ are determined by preliminary grid search on reconstruction quality at the baseline temperature

Final synthesised damaged signals at temperature T are:

$$\bar{x}_{\text{damaged}}(T) = f_\theta(z_{\text{healthy}}(T) + \Delta z_l, T) \odot \mathcal{M}_l + \epsilon \quad (5)$$

where \mathcal{M}_l is the learned amplitude mask for damage level l , \odot denotes elementwise multiplication and ϵ is Gaussian noise for stochasticity.

This method generates temperature-conditioned synthetic damaged signals using baseline damaged data at T_{Baseline} together with healthy responses across temperatures. Through latent-space transformations, amplitude masks and temperature-conditioned decoding, the convolutional conditional variational autoencoder produces synthetic responses that closely match experimental damaged signals at the baseline temperature and are consistent with the observed temperature dependence of healthy waveforms, enriching datasets for studying DA under environmental variability.

A limitation of this synthesis strategy is that the latent offset Δz_l and the amplitude mask \mathcal{M}_l are both derived exclusively from damaged and healthy signal pairs at the baseline temperature T_{Baseline} . Consequently, the generated signals at other temperatures inherit the damage morphology observed at the baseline condition. In practice, damage-induced changes in guided-wave signals are not strictly temperature-invariant: attenuation coefficients, mode conversion ratios and scattering patterns at a delamination boundary depend on the local stiffness and damping of the surrounding laminate, both of which vary with temperature. The synthesis procedure captures the temperature dependence of the healthy waveform through the conditioned decoder $f_{\theta}(z_{\text{healthy}}(T) + \Delta z_l, T)$ but assumes that the incremental effect of damage on the latent representation remains constant across temperatures. This assumption is most likely to hold when the damage perturbation is small relative to the baseline signal energy and when temperature-induced property changes are approximately uniform across the sensing path. For larger damage severities or strongly heterogeneous thermal fields, the fixed-offset assumption may introduce systematic bias in the synthesised signals at temperatures far from T_{Baseline} . The experimental validation at T_{Baseline} (“Conv-CVAE” section) and the CWI consistency checks across temperatures (“Conv-CVAE” section) provide partial support for the plausibility of the generated signals, but cannot fully substitute for experimental damaged measurements at off-baseline temperatures. Accordingly, conclusions drawn from CONCEPT transfers at non-baseline temperatures should be interpreted within this constructed evaluation context.

SPADA framework

SPADA addresses UDA where target labels are unavailable by integrating WTS for stable, translation-invariant feature extraction with prototype-based attention mechanisms that maintain class-level structure through source prototypes (using true labels) and target prototypes (employing entropy-gated pseudo-labels). Training combines four objectives: source

classification via weighted cross-entropy with label smoothing, target pseudo-labelling on confident samples, adversarial domain discrimination through gradient reversal and prototype compactness encouraging source clustering. The pseudo-labelling mechanism with confidence gating and dynamic prototype updating distinguishes SPADA from purely adversarial approaches by actively generating and refining target labels while managing noise. Interpretability modules track prototype trajectories, similarity structure, attention dynamics and latent separation throughout training to diagnose adaptation quality and identify failure modes.

Wavelet time scattering. WTS processes signals through hierarchical operations: wavelet convolution for localised frequency content, modulus operation for stability enhancement, and averaging via scaling functions for translation invariance. For channel $x_i(t)$, the zero-order coefficient retains coarse, low-frequency components³⁰:

$$S_{(\text{scat}, 0)}(x_i) = x_i * \phi_J \quad (6)$$

where $*$ denotes convolution and ϕ_J represents low-pass filter at scale J . First-order coefficients arise from band-pass filtering, modulus extraction, and low-pass filtering:

$$S_{(\text{scat}, 1)}(x_i, \omega_1) = |x_i * \psi_{\omega_1}| * \phi_J \quad (7)$$

where ψ_{ω_1} is band-pass wavelet at frequency ω_1 and $|\cdot|$ is complex modulus. Higher-order coefficients iterate wavelet-modulus-averaging along path $(\omega_1, \dots, \omega_\eta)$:

$$S_{(\text{scat}, \eta)}(x_i, \omega_1, \dots, \omega_\eta) = |\dots| |x_i * \psi_{\omega_1}| * \psi_{\omega_2} \dots * \psi_{\omega_\eta} | * \phi_J \quad (8)$$

where $\eta \geq 0$ is scattering order. Increasing η captures progressively complex temporal dependencies. Concatenating coefficients up to maximum order η yields:

$$s_i^{\text{scat}} = [S_{(\text{scat}, 0)}(x_i), S_{(\text{scat}, 1)}(x_i, \omega_1), \dots, S_{(\text{scat}, \eta)}(x_i, \omega_1, \dots, \omega_\eta)] \in \mathbb{R}^{D_{\text{scat}}} \quad (9)$$

The scattering transform possesses formally established stability properties that underpin its suitability for cross-domain SHM. Mallat³¹ proved that the scattering coefficients satisfy a Lipschitz continuity bound with respect to diffeomorphic deformations: for a signal $x \in L^2(\mathbb{R}^d)$ and its deformed counterpart $x_\tau(t) = x(t - \tau(t))$, the distance between their scattering representations is bounded by $C(\sup_t \|\nabla \tau(t)\| +$

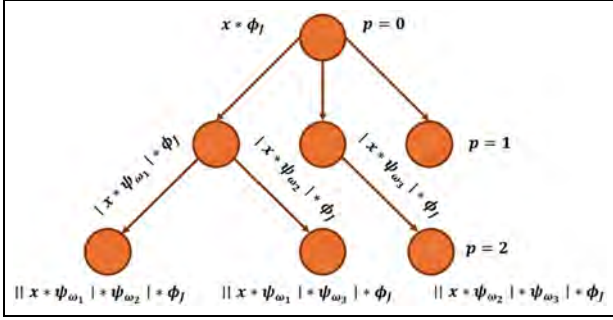


Figure 1. Schematic representation of the WTS process up to the second order for a single channel. WTS: wavelet time scattering.

$2^{-J} \sup_t \|\tau(t)\| \|x\|$, where C is a wavelet-dependent constant and \sup_t denotes the supremum (maximum) over all time instants. The first term controls sensitivity to local warplings while the second term captures residual translation sensitivity that diminishes with increasing scale J . These properties do not eliminate temperature-induced changes in wave propagation or vibration response but structure them as smooth, bounded displacements in feature space, ensuring that temperature-driven distributional shifts remain amenable to downstream domain alignment rather than producing erratic or discontinuous feature variations. Multi-channel features concatenate all C channels:

$$z = [s_1^{\text{scat}}; s_2^{\text{scat}}; \dots; s_C^{\text{scat}}] \in \mathbb{R}^D \quad (10)$$

where $D = C \cdot D_{\text{scat}}$. Figure 1 illustrates WTS to second order for single channel.

Features are invariant to small temporal shifts, noise-robust and maintain discriminative properties essential for cross-domain fault diagnosis.

Domain adaptation. The DA block comprises four loss functions: source classification, target pseudo-label, domain adversarial and prototype alignment. A domain-adversarial neural network with shared feature extraction projects WTS features into shared latent space, calculating source classification and adversarial losses.

Prototype construction and attention-based updating. Class prototypes in latent space h achieve class-level alignment. Latent space derives from projecting scattering features z through shared feature extractor g_θ :

$$h = g_\theta(z) \quad (11)$$

Source prototypes use true labels:

$$\mu_k^{(s)} = \frac{1}{|\mathcal{D}_s^k|} \sum_{h \in \mathcal{D}_s^k} h \quad (12)$$

where \mathcal{D}_s^k contains source features for class k . Target prototypes employ pseudo-labelled samples, updated iteratively:

$$\mu_k^{(t)} = \frac{1}{|\mathcal{D}_t^k|} \sum_{h \in \mathcal{D}_t^k} h \quad (13)$$

Target pseudo-labels derive from temperature-scaled softmax:

$$p^t(h) = \text{softmax} \left(\frac{\text{logits}(h)}{\tau_{\text{temp}}} \right) \quad (14)$$

Temperature parameter τ_{temp} controls distribution sharpness. Normalised entropy $H(p^t(h))$ measures confidence; binary gating admits only samples with entropy below threshold τ for prototype updating and target loss computation. Prototype-based attention weights modulate sample influence. Source weights employ cosine similarity between features and corresponding prototypes:

$$w^{(s)} = (1 - \beta_w) + \beta_w \cos(h, \mu_k^{(s)}) \quad (15)$$

where β_w controls similarity contribution. Target weights incorporate entropy-based confidence:

$$w^{(t)} = [(1 - \beta_w) + \beta_w \cos(h, \mu_k^{(t)})] (1 - H(p^t(h))) \quad (16)$$

This mechanism ensures representative, confident samples contribute strongly. Prototypes are initialised at zero vectors in latent space. Since all features are standardised to zero mean and unit variance prior to training, zero initialisation places prototypes at the centroid of the standardised feature distribution rather than at an arbitrary location, providing a neutral starting point. Upon encountering the first mini-batch containing samples of class k , the corresponding prototype is overwritten with the batch mean for that class, after which exponential moving average (EMA) updates apply for all subsequent batches. This first-batch override ensures that prototypes acquire meaningful class-specific locations before EMA smoothing begins. During early epochs, the entropy gate admits only a small fraction of target samples, as classifier predictions are initially uncertain and most samples exceed the entropy threshold τ . Consequently, target prototypes

update slowly and from a conservative subset of high-confidence instances, reducing the risk of pseudo-label collapse to a single dominant class. As training progresses and the classifier sharpens, coverage increases and target prototypes receive updates from a progressively broader and more representative set of samples. The combination of EMA momentum, entropy gating, and prototype-based attention weighting (which downweights low-confidence samples even when admitted) provides three complementary safeguards against degenerate prototype behaviour.

Training objectives. The learning process employs four complementary losses.

- (a) Source classification loss (\mathcal{L}_{cls})

Weighted cross-entropy applies prototype-based attention weights to labelled source samples:

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N_s} \sum_{n=1}^{N_s} w_n^{(s)} \sum_{k=1}^K \mathbb{1}[y^{(n)} = k] \log(p(y_{(n_s)} = k | h_{(n_s)})) \quad (17)$$

where $p^k(h^{(n)})$ is predicted probability for class k and $\mathbb{1}[\cdot]$ is indicator function.

- (b) Target pseudo-label loss (\mathcal{L}_{tgt})

This loss guides learning on unlabelled target data through confident pseudo-labels. Only samples passing confidence gate ($H(p^t(h)) \leq \tau$) contribute:

$$\mathcal{L}_{\text{tgt}} = -\frac{1}{\sum_{m=1}^{N_t} \delta^{(m)}} \sum_{m=1}^{N_t} \delta^{(m)} w_{(m)}^{(t)} \log(p(y_{(m_t)} = \hat{y}_{(m_t)} | h_{(m_t)})) \quad (18)$$

where $\hat{y}^{(m)} = \arg \max_k p(y_{(m_t)} = k | h_{(m_t)})$.

- (c) Domain adversarial loss (\mathcal{L}_{adv})

Domain discriminator distinguishes source and target while feature extractor prevents discrimination:

$$\mathcal{L}_{\text{adv}} = -\frac{1}{N_s + N_t} \left[\sum_{n=1}^{N_s} \log(p(D_{(n_s)} = 0 | h_{(n_s)})) + \sum_{m=1}^{N_t} \log(p(D_{(m_t)} = 1 | h_{(m_t)})) \right] \quad (19)$$

- (d) Prototype alignment loss ($\mathcal{L}_{\text{proto}}$)

This enforces intra-class compactness within source domain:

$$\mathcal{L}_{\text{proto}} = \frac{1}{N_s} \sum_{n=1}^{N_s} \left(1 - \cos(h_{(n_s)}, \mu_{y_{(n_s)}}^{(s)}) \right) \quad (20)$$

promoting class cohesion and stabilising prototype-based weighting.

Overall objective function. Complete training objective combines four losses with weighting coefficients:

$$\mathcal{L} = \lambda_{\text{cls}} \mathcal{L}_{\text{cls}} + \lambda_{\text{tgt}} \mathcal{L}_{\text{tgt}} + \lambda_{\text{adv}} \mathcal{L}_{\text{adv}} + \lambda_{\text{proto}} \mathcal{L}_{\text{proto}} \quad (21)$$

in which $\lambda_{(\cdot)}$ is the weight value related to the corresponding loss function. This multi-objective optimisation ensures: (i) discriminative performance on labelled source data, (ii) effective adaptation to unlabelled target domain, (iii) domain invariance preservation and (iv) class-level structural consistency.

Summary of the training process. Extracted WTS features are standardised before mini-batch training alternates between source and target batches. Prototypes update online using true labels (source) and entropy-gated pseudo-labels (target). Four losses combine with gradient reversal for adversarial learning.

Model selection employs unsupervised validation on two disjoint, class-balanced target subsets A and B, drawn from the held-out validation partition. For each subset, four quantities are computed from the unlabelled predictions: (i) prediction diversity, defined as the entropy of the mean predicted class distribution $p = N^{-1} \sum_j p_j$, (ii) mean prediction entropy across samples, (iii) prototype compactness, defined as the mean cosine similarity between confident samples and their assigned blended prototype $P_{\text{blend}} = \gamma P^{(s)} + (1 - \gamma) P^{(t)}$, where γ is the prototype blend coefficient and prototypes are L_2 -normalised before blending and (iv) coverage, defined as the fraction of samples whose maximum predicted probability exceeds the confidence threshold θ_{conf} . These components are combined into a scalar score per subset:

$$S = (\mathcal{H}(p) - \mathcal{H}) + \omega_{\text{comp}} \cdot \text{compact} + \omega_{\text{cov}} \cdot \text{cover} \quad (22)$$

where $\mathcal{H}(p)$ is the diversity term, \mathcal{H} is the mean sample entropy, and the fixed combination weights are reported in the fourth section. The epoch-level selection score is the sum $S_A + S_B$. Across training, the checkpoint with the highest cumulative score is retained. Across configurations, selection retains the configuration with the highest mean score over 15 seeds; ties are

broken by lower standard deviation, then by earlier sample index.

Figure 2 presents the SPADA schematic.

Interpretability of SPADA internal activity. SPADA incorporates four interpretability views monitoring internal activity to verify source-target latent alignment while preserving class discriminability. These views address WTS features, domain-adversarial latent space, prototype-based attention and unsupervised selection. Crucially, these views serve purely diagnostic purposes without altering training dynamics.

Let $h = g_\theta(z) \in \mathbb{R}^{D_{\text{lat}}}$ denote latent features from scattering coefficients z and K be class count. Source class subset $D_s^k = \{i \in D_s | y_{(i_s)} = k\}$ remains fixed, where $y_{(i_s)}$ is true label. Source prototype at epoch e :

$$\mu_k^S(e) = \frac{1}{|D_s^k|} \sum_{i \in D_s^k} h_i(e) \quad (23)$$

Source labels ensure stable computation despite evolving representations $h_i(e)$. Target class subset constructs dynamically via pseudo-labels and confidence filtering:

$$\{j \in D_t | \hat{y}_{j_t}(e) = k, H(p^T(h_{(j_t)}(e))) \leq \tau\} \quad (24)$$

where $\hat{y}_{j_t}(e)$ is pseudo-label at epoch e , $H(\cdot)$ is Shannon entropy of temperature-scaled softmax $p^T(h)$ and τ is confidence threshold. Only instances with pseudo-label k and entropy below τ contribute. Target prototype:

$$\mu_k^T(e) = \frac{1}{|D_t^{(k,\tau)}(e)|} \sum_{j \in D_t^{(k,\tau)}(e)} h_j(e) \quad (25)$$

ensuring confident predictions avoid ambiguous sample contamination. Coverage quantifies confident target sample proportion:

$$C(e) = \frac{1}{N_t} \sum_{j \in D_t} \mathbb{1}\{H(p^T(h_{(j_t)}(e))) \leq \tau\} \quad (26)$$

where $\mathbb{1}\{\cdot\}$ is indicator function and N_t is total target samples. Cosine similarity between feature h and prototype μ measures angular alignment:

$$\text{sim}(h, \mu) = \frac{h^\top \mu}{\|h\|_2 \|\mu\|_2} \quad (27)$$

This normalised similarity ranges from -1 to $+1$, with values near $+1$ indicating strong alignment. Instance-

level attention leverages similarity to modulate sample contribution. Source attention at epoch e :

$$a_i^S(e) \propto \exp\{\alpha \cdot \text{sim}(h_i(e), \mu_{y_{(i_s)}}^S(e))\} \quad (28)$$

where α controls sharpness and $y_{(i_s)}$ is true label. Target attention incorporates confidence:

$$a_j^T(e) \propto (1 - H(p^T(h_{(j_t)}(e)))) \cdot \exp\{\alpha \cdot \text{sim}(h_j(e), \mu_{\hat{y}_{(j_t)}(e)}}^T(e))\} \quad (29)$$

Term $(1 - H(p^T(h_{(j_t)}(e))))$ downweights high-entropy instances; exponential term upweights prototype-aligned features. Attention normalises within minibatches before loss application.

Tracking prototype trajectory evolution. Characterising class structure evolution and target migration towards source centres while preserving inter-class distinctness requires per-class drift and alignment metrics. Source prototype drift quantifies Euclidean displacement between consecutive epochs:

$$\Delta_k^s = \|\mu_k^s(e) - \mu_k^s(e-1)\|_2 \quad (30)$$

Target prototype drift measures movement:

$$\Delta_k^t = \|\mu_k^t(e) - \mu_k^t(e-1)\|_2 \quad (31)$$

Alignment gap quantifies distance between source and target prototypes:

$$A_k(e) = \|\mu_k^s(e) - \mu_k^t(e)\|_2 \quad (32)$$

Metrics compute independently per class k throughout training. Diminishing $\Delta_k^t(e)$ with reducing $A_k(e)$ indicates target prototype stabilisation and source convergence, signalling successful alignment. Persistent large $\Delta_k^t(e)$ or increasing $A_k(e)$ suggests unreliable pseudo-labels, excessive domain-specific information removal or insufficient adaptation. Trajectory crossings where different-class prototypes converge indicate collapsing boundaries requiring intervention. Early-stage movement is expected as gradient reversal strengthens, and confidence gate admits more samples. Desired outcome: convergence with stability, wherein prototypes settle into distinct locations across classes, preserving discriminative margins while achieving corresponding class cross-domain alignment.

The trajectory metrics defined above can be interpreted in relation to known physical effects of temperature on structural dynamic response. As demonstrated in the EOVS analysis (“Effects of EOVS” section),

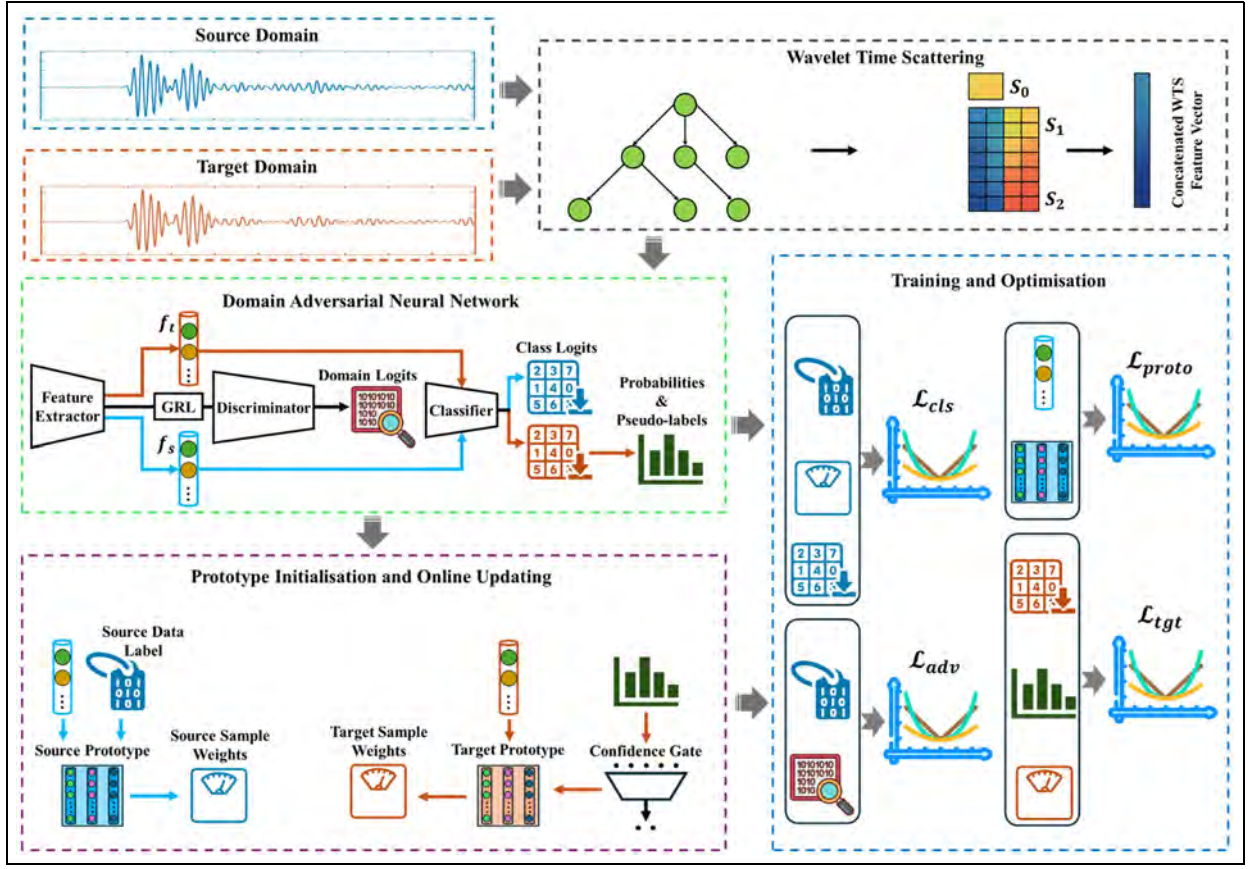


Figure 2. A schematic of SPADA framework. SPADA: scattering-based prototype-aligned domain adaptation.

temperature variation induces systematic changes in wave propagation speed and modal frequencies consistent with thermal softening of the host material. When source and target domains correspond to different temperatures, a reducing alignment gap $A_k(e)$ for a given damage class k indicates that the adaptation mechanism is compensating for these temperature-induced feature displacements while retaining the class identity associated with that damage mode. Conversely, persistent or increasing $A_k(e)$ for specific classes may signal that the corresponding damage signatures interact with temperature effects in ways that simple feature-space alignment cannot resolve, for instance when temperature-dependent attenuation disproportionately affects certain damage severities. Domain experts can therefore use the trajectory diagnostics in conjunction with the physical characterisation of the EOV to assess whether adaptation behaviour is consistent with expected material response rather than reflecting artefacts of the alignment procedure.

Computing instance-prototype cosine similarity. Examining within-class compactness and between-class confusion

requires similarity matrix quantifying instance-prototype relationships at selected epochs. For N instances and K classes, similarity matrix $M \in \mathbb{R}^{N \times K}$ has entries:

$$M_{i,k} = \text{sim}(h_i(e), \mu_k(e)) \quad (33)$$

where $h_i(e)$ is latent feature at epoch e , $\mu_k(e)$ denotes either $\mu_k^s(e)$ or $\mu_k^t(e)$ depending on domain, and $\text{sim}(u, v) = u^\top v / (\|u\|_2 \|v\|_2)$. Row i represents instance i similarity across prototypes; column k reflects all instance relationships to prototype k .

Computation: (1) select analysis epoch e (typically best epoch from unsupervised validation), (2) extract latent features $h_i(e)$ for domain instances, (3) retrieve prototypes $\mu_k(e)$ for all classes and (4) compute pairwise cosine similarities populating M . Heatmap visualisation sorts rows by domain then class.

Sharp diagonal block structure in source matrix indicates high similarity to true-class prototypes and low similarity elsewhere, confirming compact, well-separated clusters. Effective adaptation produces similar target patterns: high on-diagonal values for assigned pseudo-label classes and low off-diagonal

values. Off-diagonal bands identify confusable class pairs, flagging attention or confidence imbalances. Progressive diagonal enhancement across epochs signals sharpening structure and improving discriminability. Persistent off-diagonal mass on particular classes indicates prototype attractor behaviour, potentially absorbing neighbouring class instances, meriting examination for collapse or overlap.

Monitoring prototype attention weight dynamics. Assessing appropriate target sample weighting and controlled target data reliance growth requires epoch-wise attention weight and coverage summary statistics. Robust statistics avoid outlier sensitivity. For attention weights $\{a_j(e)\}$ at epoch e , median and IQR:

$$\tilde{a}(e) = \text{median}_j\{a_j(e)\}, \text{IQR}(a)(e) = \text{IQR}_j\{a_j(e)\} \quad (34)$$

where IQR computes as 75th minus 25th percentile difference. Small IQR indicates even attention distribution; large IQR signifies concentration on sample subsets. Median $\tilde{a}(e)$ provides robust central tendency.

The whole process can be summarised as: (1) collect target attention weights $a'_j(e)$ for all $j \in D_t$ each epoch; (2) compute median $\tilde{a}(e) = \text{median}(\{a'_j(e)\})$; (3) compute 25th percentile $Q_1(e)$ and 75th percentile $Q_3(e)$, then $\text{IQR}(a)(e) = Q_3(e) - Q_1(e)$; (4) simultaneously compute coverage $C(e)$ as fraction with entropy below τ . Statistics track separately for domains, plotting across epochs.

Healthy adaptation shows rising or stable $\tilde{a}(e)$ concurrent with increasing $C(e)$, suggesting meaningful attention as more samples admit via confidence gate. Declining $\tilde{a}(e)$ while $C(e)$ rises indicates newly admitted samples receive low attention, possibly from unreliable pseudo-labels or noisy features. Large IQR with concentrated attention suggests narrow instance-driven prototype updates, potentially causing instability or bias. Comparing $\tilde{a}(e)$ with $C(e)$, co-movement (both rising) implies entropy-admitted samples are judged informative by attention; divergence ($C(e)$ increasing but $\tilde{a}(e)$ flat or declining) flags mismatch, such as admitting many low-quality or ambiguous samples contributing poorly to adaptation.

Visualising decision boundaries in feature space. SPADA provides qualitative class separation and source-target overlap visualisation through dimensionality reduction. Projection $\Pi : \mathbb{R}^{D_{\text{lat}}} \rightarrow \mathbb{R}^2$ fits on reference set comprising source and target features at chosen epoch (typically best), using t-distributed stochastic neighbour embedding (t-SNE) or principal component analysis. Procedure: (1) collect latent features $h_i(e)$ for all

domains at epoch e ; (2) concatenate into single matrix; (3) fit projection Π ; (4) transform instances and prototypes to two dimensional (2D): $u_i = \Pi(h_i(e))$ and $v_k = \Pi(\mu_k(e))$.

Resulting scatter displays source and target instances with distinct markers (circles for source, squares for target), colour-coded by class or pseudo-label. Source prototypes $\Pi\mu_k^s(e)$ and target prototypes $\Pi\mu_k^t(e)$ overlay with distinct markers (X for source, + for target), colour-coded by class. Ideal structure exhibits class-wise overlap between domains, that is, same-class different-domain instances intermingle with clear inter-class separation, confirming domain-invariant features while preserving discriminability.

Excessive invariance manifests as reduced inter-class separation where different-class clusters merge, indicating over-suppression of task-relevant information. Insufficient adaptation produces disjoint source-target clusters within classes, signalling inadequate domain shift mitigation. Alignment between visualisation and quantitative trends in $A_k(e)$ and $\Delta_k^t(e)$ reinforces interpretation: converging prototypes (small $A_k(e)$) should correspond to overlapping source-target clusters; stable prototypes (small $\Delta_k^t(e)$) should appear as fixed cluster centres.

Algorithm 1 presents the algorithm for SPADA.

Case studies

In the present study, two publicly available datasets were employed to assess the effectiveness of the proposed damage detection frameworks. The selection of these datasets was informed by the common experimental setup, in which a temperature chamber was used to regulate the ambient temperature during data collection. In addition, the datasets were derived from two different types of signals, that is, guided wave and vibration signals, thereby enabling a comprehensive evaluation across diverse sensing approaches. The datasets are described in detail in the following sections.

Small-scale WTB under varying climate conditions (WTB-VibClimate)

The first dataset contains experimental signals of a small-scale WTB for the blade of a Windspot 3.5 kW WT model manufactured by Sonkyo Energy disclosed by Qu et al.³² This blade is made of a three-layered sandwich composite configuration; it has a length of 1.75 m and a mass of 5.0 kg.

Experiments were conducted at 12 temperature conditions from -15 to 40°C in five-degree increments (Wn15, Wn10, Wn5, Wp0, Wp5, Wp10, Wp15, Wp20, Wp25, Wp30, Wp35, Wp40) at 60% humidity. Two

Algorithm 1. SPADA framework for SHM.

Input: Source domain signals (labelled), target domain adaptation signals (unlabelled), target domain test signals (labelled, held out)

Output: Domain-adaptive model, performance metrics, interpretability visualisations

1. Data processing

- 1.1 Load source and target domain time-series signals.
- 1.2 Extract wavelet scattering transform features using Kymatio.
- 1.3 Apply log-stabilisation and concatenate multi-channel features.

2. Data preparation

- 2.1 Split source features into train/validation/test sets (stratified).
- 2.2 Partition target adaptation features into DA set and validation set (unlabelled).
- 2.3 Further split target validation into subsets A and B for unsupervised scoring.
- 2.4 Standardise all features using combined source-target statistics.

3. Model architecture

- 3.1 Feature extractor with batch normalisation and dropout.
- 3.2 Label classifier for class prediction.
- 3.3 Domain discriminator with gradient reversal layer.
- 3.4 Prototype module with exponential moving average updates.

4. Training

- 4.1 Sample mini-batches from source (labelled) and target DA (unlabelled).
- 4.2 Generate target pseudo-labels via temperature-scaled softmax.
- 4.3 Update prototypes using source labels and confident target pseudo-labels.
- 4.4 Compute instance weights based on feature-prototype similarity and entropy.
- 4.5 Optimise losses: source classification, target pseudo-labelling, domain adversarial, prototype alignment.
- 4.6 Model selection using target validation subsets A/B with diversity, confidence and prototype metrics.

5. Hyperparameter optimisation

- 5.1 Define search space over adaptation, prototype and scattering parameters.
- 5.2 Select configuration with highest mean unsupervised score across seeds.
- 5.3 Retrain best configuration for final evaluation.

6. Evaluation

- 6.1 Test on held-out target test set, compute accuracy and confusion matrix.

7. Interpretability analysis

- 7.1 Visualise prototype trajectory evolution via t-SNE.
- 7.2 Generate instance-prototype similarity heatmaps.
- 7.3 Plot attention weight dynamics and pseudo-label coverage.
- 7.4 Illustrate decision boundaries in feature space.

excitation modes were applied: white noise (0–400 Hz) and sine sweep (1–300 Hz). Both signals were applied for approximately 120 s with a constant sampling frequency of 1666 Hz at a fixed point on the blade surface. Two sensor types recorded signals: accelerometers and strain gauges, with different configurations. This study used accelerometer data assuming white noise excitation. Although eight sensors recorded data, only three accelerometers (channels 1, 4 and 8) were retained to reduce computational complexity and assess the SHM framework under sensor-limited conditions.

The selection of three from eight available accelerometers was motivated by two considerations. First, it tests the framework under realistic deployment constraints, where cost, cabling and maintenance limit the number of sensors that can be sustained over the operational life of a blade. Second, the three retained channels (1, 4 and 8) span distinct positions along the blade, providing spatial diversity in the captured dynamic response without redundancy from closely spaced sensors. A controlled comparison of the three-sensor and eight-sensor configurations falls outside the scope of

the present study, and the reported results should therefore be interpreted as representative of sensor-limited conditions rather than as the upper bound of performance achievable with the full sensor array.

Within the SPADA framework's WTB-VibClimate case study, these are referenced as channels 1, 2 and 3. Figure 3 shows these accelerometer positions, excitation points and locations of unbalancing mass and cracks on the WTB.

Thirteen health scenarios were considered for this WTB: one intact state, nine crack cases (one to three cracks with varying lengths), and three icing scenarios (one to three unbalanced masses of 44 g each). Table 1 summarises these health scenarios with fault quantity and severity; the introduced index is used for classification.

Carbon–epoxy composite plate

The second dataset, carbon–epoxy composite plate (CONCEPT)³³ contains Lamb wave measurements from a unidirectional carbon-epoxy laminate plate in

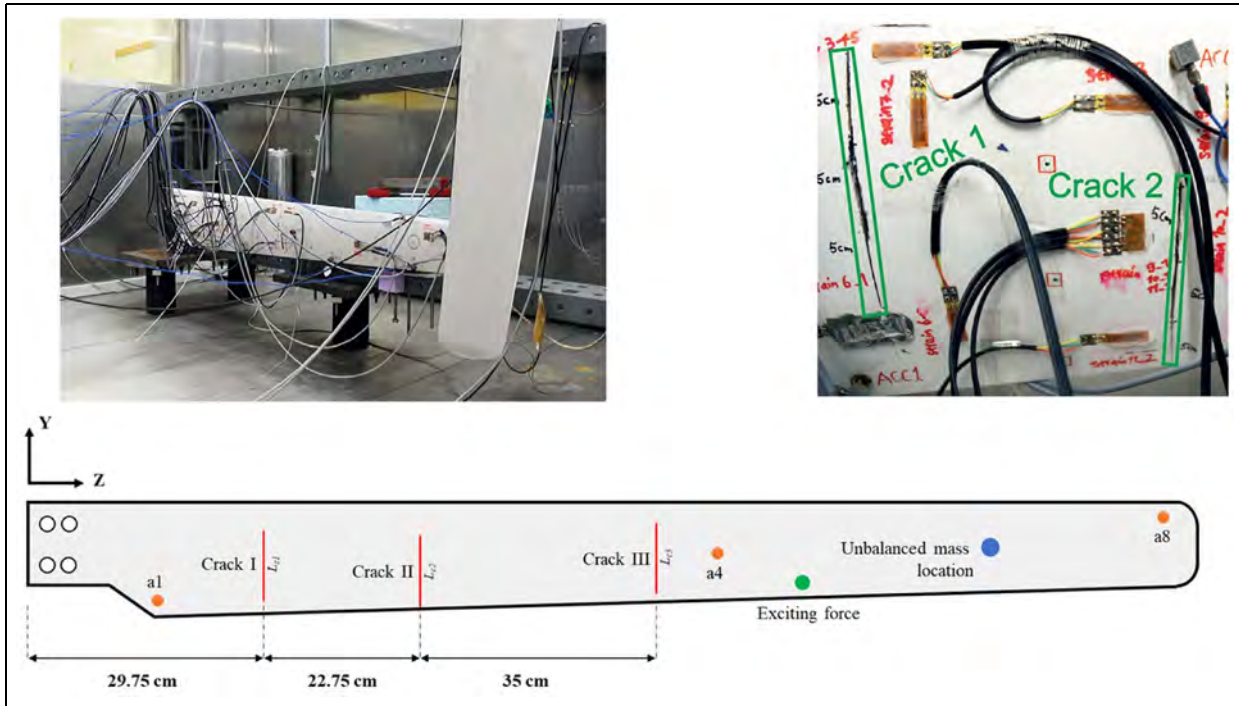


Figure 3. Test rig and sensor configuration in WTB-VibClimate. WTB-VibClimate: small-scale wind turbine blade under varying climate conditions.

Table I. WTB health conditions.

Health scenario	Number of observations	Label	Commentary
Intact	20	0	Healthy
Unbalanced	5	1	1 added mass
	5	2	2 added masses
	5	3	3 added masses
Cracked	5	4	1 crack, $L_{c1} = 5$ cm
	5	5	2 cracks, $L_{c1} = 5$ cm, $L_{c2} = 5$ cm
	5	6	3 cracks, $L_{c1} = 5$ cm, $L_{c2} = 5$ cm, $L_{c3} = 5$ cm
	5	7	3 cracks, $L_{c1} = 10$ cm, $L_{c2} = 5$ cm, $L_{c3} = 5$ cm
	5	8	3 cracks, $L_{c1} = 10$ cm, $L_{c2} = 10$ cm, $L_{c3} = 5$ cm
	5	9	3 cracks, $L_{c1} = 10$ cm, $L_{c2} = 10$ cm, $L_{c3} = 10$ cm
	5	10	3 cracks, $L_{c1} = 15$ cm, $L_{c2} = 10$ cm, $L_{c3} = 10$ cm
	5	11	3 cracks, $L_{c1} = 15$ cm, $L_{c2} = 15$ cm, $L_{c3} = 10$ cm
	5	12	3 cracks, $L_{c1} = 15$ cm, $L_{c2} = 15$ cm, $L_{c3} = 15$ cm

WTB: wind turbine blade.

healthy and damaged states. The experiments investigate how temperature fluctuations and damage progression affect the laminate's structural behaviour.

Four lead zirconate titanate (PZT) transducers from Accellent Technologies were bonded to the plate. PZT1 acted as the actuator and PZT2, PZT3 and PZT4 as sensors. The plate was tested under free-free boundary conditions to minimise constraints on wave propagation and capture a representative dynamic response. The test rig and instrumentation are shown in Figure 4.

The experiments were conducted in a Thermotron thermal chamber to provide precise temperature control using an integrated cascade refrigeration system, so that 0°C was reached by mechanical cooling rather than ambient freezing. A sinusoidal tone burst served as the excitation signal, and the responses were sampled using dedicated data acquisition systems operated via LabVIEW. For the intact plate, 100 measurements were collected at each of seven temperature levels from 0 to 60°C , labelled Cp0, Cp10, Cp20, Cp30, Cp40,

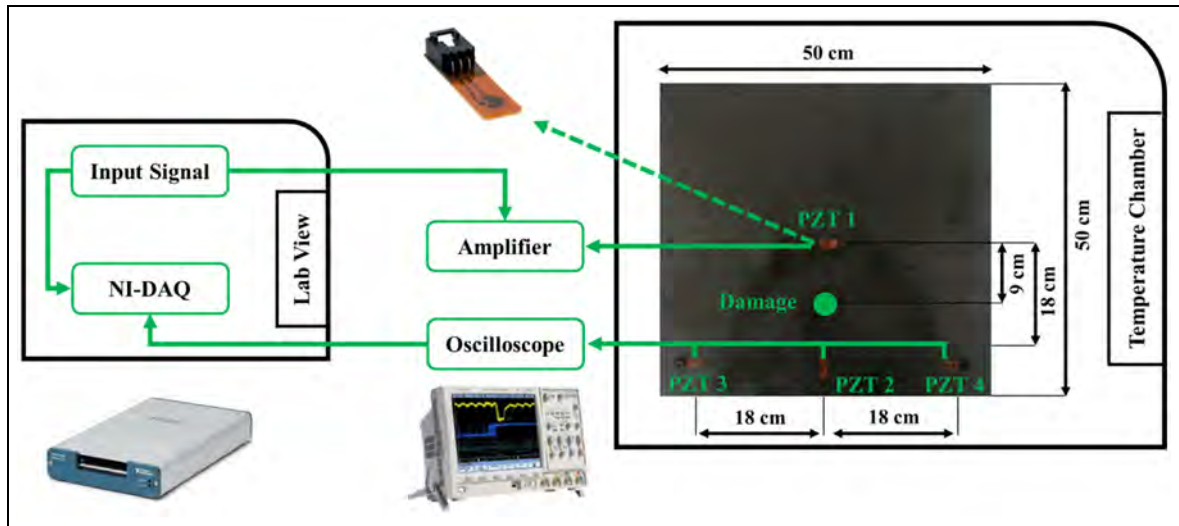


Figure 4. The test rig setup in the CONCEPT test. CONCEPT: carbon–epoxy composite plate.

Table 2. Damage scenarios and severities for simulated defects in the CONCEPT experiment.

Damage scenario	Severity (%)	Label	Description	Temperature (°)				
Healthy	0	C0	No damage	0				
				10				
				20				
				30				
				40				
				50				
Damaged D1	0.196	C1	Industrial putty	30				
				Damaged D2	0.282	C2	Increased coverage of putty	
				Damaged D3	0.384	C3	Further increase in coverage	
				Damaged D4	0.502	C4	Progressive increase	
				Damaged D5	0.785	C5	Larger area covered	
				Damaged D6	1.13	C6	Substantial coverage	
				Damaged D7	1.53	C7	Continued increase	
				Damaged D8	1.95	C8	Different progression pattern	
				Damaged D9	2.01	C9	Extensive coverage	
				Damaged D10	2.27	C10	High severity	
				Damaged D11	2.54	C11	Maximum simulated severity	

CONCEPT: carbon–epoxy composite plate.

Cp50 and Cp60. For the damaged plate, 100 measurements were acquired only at 30°C, which served as the baseline, with no damaged data at other temperatures.

Damage scenarios were simulated by applying industrial adhesive putty to the plate surface to create delamination-like defects. The damage severity was progressively increased in a localised region between PZT1 and PZT2 to study the resulting changes in wave attenuation and propagation. Table 2 summarises the

health scenarios, their severities and brief descriptions, and reports the labels assigned to the different temperatures.

Results and discussion

This section presents the effects of EOVs on guided wave and vibration signals. Data augmentation results are discussed through two standard metrics. Results

with and without the SPADA domain-adaptation stage are compared. The internal activities of SPADA throughout DA are examined. The applied feature extraction methods and the effects of domain shift on extracted features are presented and compared with convolutional neural networks (CNNs).

Effects of EOVs

Because EOVs mainly alter stiffness, damping and wave-propagation speed in the monitored structure, its influence is often not apparent in raw time-domain signals. More diagnostic representations are therefore required, such as frequency-response functions (FRFs),³⁴ mode shapes and coda wave interferometry (CWI).³⁵ In FRF analysis, the structural response to a known input is expressed in the frequency domain as $H(\omega) = Y(\omega)/F(\omega)$, from which shifts in resonance frequencies, changes in modal damping and variations in modal residues can be estimated to quantify EOVs. CWI estimates small fractional changes in wave speed, dv/v , by correlating the late, multiply scattered coda of repeat waveforms; positive dv/v indicates faster propagation and negative dv/v slower propagation.

Method suitability depends on actuation and sensing capabilities, wavefield properties, linearity, access constraints, computational cost and baseline availability. For composite plates with PZT-guided waves, FRF approaches fail because bonded PZTs provide distributed frequency-dependent tractions, dispersive multimodal fields mix modes in single-input-output measurements, closely spaced lightly damped modes require dense sampling and temperature drifts violate time-invariance. Wavefield-based methods like CWI are more effective. In this section, mode-shape and FRF analyses were applied to the low-frequency vibration responses in WTB-VibClimate, while CWI was used for the guided-wave measurements in CONCEPT.

WTB-VibClimate. The purpose of this analysis is to quantify how temperature affects mode shapes and FRFs of the WTB-VibClimate system. FRFs were estimated in MATLAB® using the H1 estimator from healthy measurements at all temperatures labelled Wn15 to Wp40. For each temperature, 20 runs were processed with channel 1 as the response and the force channel as the input. Signals were trimmed by discarding the first 10,000 and last 20,000 samples, then detrended and band pass filtered between 0.5 and 380 Hz using a zero-phase finite impulse response filter of order 800. Spectral estimates for FRF computation used Welch's method with a sampling rate of 1666 Hz, a Hann window of 4 s, 50% overlap and a fast Fourier transform length of 8192.

FRFs from repeated runs were then pooled using weights proportional to the squared coherence, retaining only frequency bins with squared coherence at or above 0.8 for averaging, while plots were masked below 0.7. Modal candidates at the 10°C baseline were identified by peak picking with at most four modes, minimum peak prominence 3% of the baseline maximum, minimum inter-peak spacing 1 Hz and matching tolerance 1 Hz for ridge initialisation. Modal ridges were tracked across temperatures within non-overlapping frequency bands around the baseline peaks, and half power bandwidth calculations yielded damping estimates for each mode. When multiple response channels were available, mode shapes were obtained from the complex FRFs at each modal peak and normalised across channels at the baseline, with an optional output only frequency domain decomposition check based on the first singular value of the response spectral density matrix. The corresponding mode shape plots and FRF maps are presented in Figure 5(a) and (b), respectively.

Figure 5(a) shows four modal frequencies that are highest at -15°C and decrease approximately linearly as temperature rises to $+40^{\circ}\text{C}$. The largest shift occurs near 300 Hz, with a smaller but clear shift around 215–220 Hz and more modest changes near 140 and 120 Hz, strongest for the two highest frequency modes. The dashed horizontal markers indicate the 10°C baseline, with the curves above the baseline at sub-zero temperatures and below it at warmer conditions. Figure 5(b) confirms these trends: bright FRF ridges remain in the same modal bands but migrate to lower frequency with increasing temperature, and their slight thickening and reduced intensity at higher temperatures suggest modest peak broadening consistent with increased damping. No mode crossing is observed, so mode ordering is preserved. Together, the figures show a systematic temperature dependence of the dynamic response, consistent with thermal softening, which motivates temperature compensation when using these data for classification.

Concept. For the CONCEPT guided-wave data, temperature-induced EOVs were quantified using CWI. A baseline at 30°C was formed by median-stacking the kept runs, and for each temperature level (0 – 60°C) every observation that passed force-channel quality screening was pre-processed (detrending and zero-phase band-pass filtering around the burst) and compared against the baseline through a stretch-correlation search over $\epsilon \in [\pm 0.5\%]$. Estimation was carried out using a moving-window approach (40 μs windows, 10 μs hop) with a window-quality gate (median $\rho \geq 0.90$). For each temperature, the resulting correlation-versus-stretch curves were averaged across

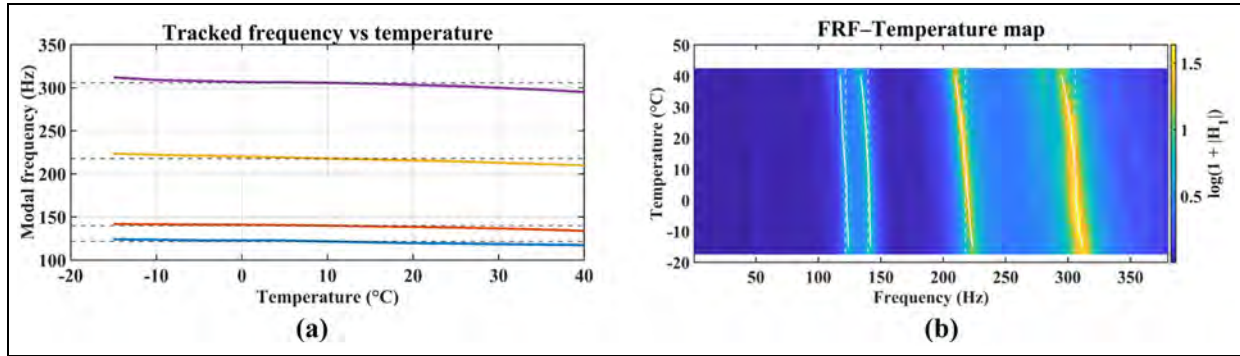


Figure 5. (a) Mode shapes identified at the baseline temperature of 10°C, based on normalised FRFs and (b) FRF–temperature maps for the WTB-VibClimate dataset showing the variation of response magnitude across temperature levels. FRF: frequency–response function; WTB-VibClimate: small-scale wind turbine blade under varying climate conditions.

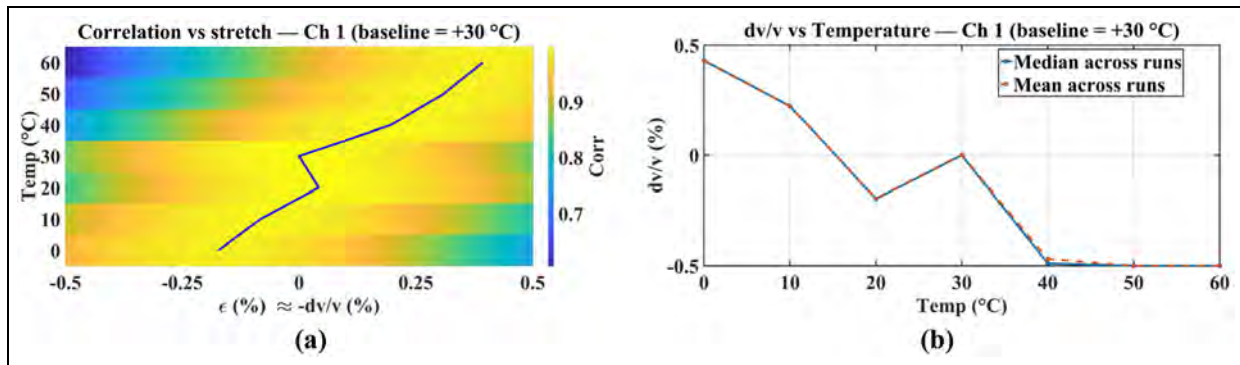


Figure 6. (a) CWI correlation–stretch map and (b) fractional wave-speed change versus temperature for CONCEPT. CWI: coda wave interferometry; CONCEPT: carbon–epoxy composite plate.

runs, and the corresponding dv/v estimates were aggregated using the median, with the mean also reported for completeness. The analysis was performed only on the healthy plate and channel 1, as this sensor was later shown to have the highest importance (owing to its augmented amplitude). At each temperature, all 100 observations were processed, and Figure 6(a) and (b) report, respectively, the aggregated correlation–stretch map and the fractional wave-speed change versus temperature obtained from these observations.

From Figure 6(a), similarly, it can be understood that the ridge of maximum correlation, $\epsilon(T)$, moved from negative stretch at low temperatures to positive stretch at higher temperatures, consistent with the dv/v trend (since $dv/v = -\epsilon$). Correlation along the ridge remained high (0.9 or above), suggesting that temperature primarily induced a quasi-uniform phase dilation rather than substantial waveform distortion. This behaviour indicated that the dominant EOV mechanism in the guided-wave setting was a small fractional change in wave speed.

Analysing Figure 6(b), one can observe that a clear, near-monotonic decrease in dv/v with temperature was observed, crossing approximately 0 at the 30°C baseline. Approximate values were +0.43% (0°C), +0.22% (10°C), -0.20% (20°C), close to 0% (30°C), about -0.47% (40°C) and roughly -0.50% (50–60°C). The median and mean curves were essentially coincident across all temperatures, indicating low between-run dispersion and a stable, repeatable temperature effect. The sign pattern implied faster propagation at sub-baseline temperatures and slower propagation at elevated temperatures, with a mild saturation of the negative dv/v beyond 40°C.

Data augmentation

The two different data augmentation techniques, that is, signal windowing and Conv-CVAE implemented on WTB-VibClimate and CONCEPT case studies, respectively, are elaborated in the following.

Signal windowing. A windowing strategy expanded the WTB-VibClimate dataset, feasible given uniform sampling. Each observation was divided into five equal 39,200-point segments. To avoid edge effects, the first 4000 data points and the 4000 last data points were excluded, resulting in 196,000 clean points divided into five non-overlapping windows. This produced 25 observations per damaged condition. The same procedure was applied to the healthy condition (using the first five original observations). With 13 health scenarios, this generated a $325 \times 3 \times 39,200$ dataset per temperature level. To prevent data leakage, 15 target-domain observations (from three original recordings) were allocated for DA, while 10 observations (from two separate recordings) remained entirely unseen for testing.

A methodological caveat applies to the windowing procedure. Because all windows from a given original recording share the same excitation event, boundary conditions and sensor coupling state, they are not statistically independent realisations. The reported uncertainty, therefore, reflects variability due to random data splits and model initialisation rather than variability across independent experimental repetitions. This distinction does not invalidate the reported accuracies, which remain valid point estimates of classification performance on the held-out windows, but it means that the associated confidence intervals may underestimate the true variability that would be observed across fully independent measurement campaigns. Future work should incorporate recording-level resampling or leave-recording-out cross-validation to provide uncertainty estimates that account for this dependence structure.

Conv-CVAE. Conv-CVAE as a temperature-conditioned generative augmentation method was implemented in PyTorch to augment CONCEPT. The model was trained for 150 epochs with batch size 128, latent dimension 64, learning rate 3×10^{-4} and AdamW optimiser, using 100 healthy observations at each of the seven temperatures (Cp0–Cp60, 10°C intervals), normalised using global statistics.

The composite loss weights were set to $\alpha_{\text{MSE}} = 1.0$, $\alpha_{\text{spec}} = 0.5$, $\alpha_{\text{wav}} = 0.2$ and $\alpha_{\text{env}} = 0.05$ based on a preliminary grid search over reconstruction quality at Cp30. The KL coefficient followed $\beta(e) = \min(1, e/20)$, warming up linearly over the first 20 epochs. Wavelet decomposition used a four-level Daubechies-4 basis, and the envelope term was computed over the first 600 samples of the Hilbert analytic envelope, corresponding to the primary wave packet arrival window.

The Gaussian smoothing kernel for the amplitude masks used $\sigma = 6$ samples. The stochastic dispersion was set to 15% of the estimated standard deviation per

element, capped at 30% of the deterministic mask value.

The encoder comprised three one-dimensional convolutional layers (filters 7, 5, 5; stride 2) with batch normalisation and GELU activations; the decoder used mirrored transposed convolutions. Temperature was encoded as a normalised scalar concatenated at fully connected layers. For each of 12 damage levels, a latent offset vector was computed at reference temperature Cp30. Per-sensor and per-time amplitude masks derived from Hilbert envelopes used Gaussian-smoothed median ratios ($\sigma = 6$), with dispersion set to 15% of the estimated per-element standard deviation, where $\hat{\sigma}$ is robustly approximated as $\text{IQR}/1.35$.

Loss combined mean squared error (MSE), log-magnitude fast Fourier transform (FFT) spectral terms, wavelet similarity (four-level Daubechies-4), envelope consistency and Kullback–Leibler (KL) divergence

Damaged signals at temperature T were generated by encoding healthy seeds at T , adding damage offsets, decoding with T conditioning, then modulating by learned masks with Gaussian-sampled stochasticity (Equation (5)). Each damage condition and temperature resulted in 30 synthetic observations alongside 30 randomly selected healthy observations ($360 \times 3 \times 1000$ per temperature). At Cp30, 30 original observations were selected randomly; synthetics were discarded.

To examine the extent to which the synthesised signals replicate the behaviour of the experimental data, two quantitative metrics were employed: dynamic time warping (DTW)³⁶ and cross correlation (CC).³⁷ A lower DTW (ideally 0) value and $|\text{CC}|$ close to 1 indicates strong similarity between the synthesised and original data. The data were generated for the baseline temperature of 30°C, which includes observations from all health scenarios. For each condition, 50 observations were used to produce the synthesised data, while the remaining 50 observations per class were retained for comparison with the corresponding synthesised signals. This approach ensured a balanced and consistent evaluation process. The results are shown in Figure 7.

Observing Figure 7, Conv-CVAE demonstrates strong agreement at Cp30. Mean DTW values are extremely low, that is, PZT2 (0.000362), PZT3 (0.000103) and PZT4 (0.000094) while CC values remain high: PZT2 (0.990627), PZT3 (0.983301) and PZT4 (0.984319). Conv-CVAE's low DTW value and high (near to 1) CC magnitude reflect its capacity to capture non-linear temporal variations through temperature-conditioned latent representations and learned amplitude and phase masks.

For the Conv-CVAE-generated damaged scenario C11, CWI was deployed to assess whether the temperature-induced EOV behaviour in the synthetic

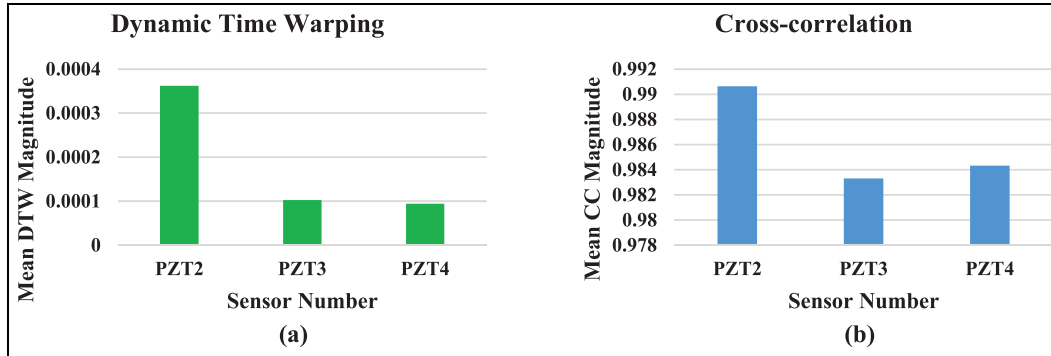


Figure 7. Comparison of synthesised data (through Conv-CVAE) and real data across sensors (PZT2, PZT3, PZT4) for Cp30 using (a) mean DTW values and (b) mean CC values. PZT: lead zirconate titanate; DTW: dynamic time warping; CC: cross correlation.

responses remains consistent with that observed on the experimental healthy plate. Temperature-induced EOV was quantified using the same CWI configuration as in the healthy CONCEPT analysis, except that the force-channel-based quality screening could not be applied because excitation traces are not available for the synthesised data. For each temperature level (0–60°C), all 30 synthesised observations for each PZT were pre-processed and compared with the 30°C damaged baseline; run-wise dv/v estimates were then aggregated across the 30 observations using the median, with the mean also reported for consistency with the healthy case, and the corresponding correlation-versus-stretch curves were averaged in the same manner. The resulting dv/v -temperature curves for the three PZTs are presented in Figure 8(b). The associated correlation-versus-stretch maps, referenced to the 30°C damaged baseline, are presented in Figure 8(a).

The dv/v -temperature curves in Figure 8(b) and the ε^* ridges in the correlation-versus-stretch maps of Figure 8(a) are observed to follow the same qualitative behaviour as in the experimental healthy plate: dv/v is positive at low temperatures, crosses zero near the 30°C baseline, and becomes negative at higher temperatures, with similar magnitude across all three PZTs. The smooth shift of the correlation maximum from negative to positive stretch with increasing temperature, together with consistently high correlation values, indicates that the synthetic damaged responses undergo a physically plausible temperature-induced time-stretch rather than arbitrary distortions. On this basis, the Conv-CVAE is regarded as capturing the underlying temperature dependence of guided-wave propagation in the composite plate, providing a physics-guided semi-synthetic extension of the experimental dataset.

To provide indirect validation of Conv-CVAE synthesis fidelity beyond the DTW and CC metrics reported at Cp30, bidirectional cross-classification was conducted using a Random Forest classifier with 200

trees. Training on real Cp30 data with 1200 observations and testing on synthetic data with 360 observations achieved 81.67% accuracy, whereas the reverse direction achieved 75.50%. These moderate accuracies indicate a distributional gap at the raw-waveform level, where the classifier must resolve fine inter-class boundaries that the Conv-CVAE does not replicate perfectly. However, per-class cross-correlation remained above 0.98 for all 12 damage classes, with a mean CC of 0.991, confirming that the dominant waveform morphology and damage-severity ordering are preserved. The contrast between high CC and moderate cross-classification accuracy is consistent with the SPADA pipeline design, since the Conv-CVAE is not intended to generate waveform-identical copies, but rather to generate signals whose WTS features preserve class-discriminative structure across temperatures.

Damage detection

To assess SPADA's effectiveness in detecting damage under EOVs, two case studies (WTB-VibClimate and CONCEPT) were analysed separately. This section presents results from intermediate and final evaluation stages, including damage detection without DA and with complete SPADA. Ablation and comparative studies were conducted for both cases. Internal SPADA activities were visualised to demonstrate domain adjustment for EOV mitigation, and computational efficiency was evaluated for real-time deployment capability.

Unless otherwise stated, all reported target-domain accuracies correspond to the mean across 15 independent seeds for the selected configuration, rather than the single highest accuracy across seeds.

WTB-VibClimate. To evaluate SPADA's performance on small-scale WTB damage detection, the augmented WTB-VibClimate dataset was employed following the

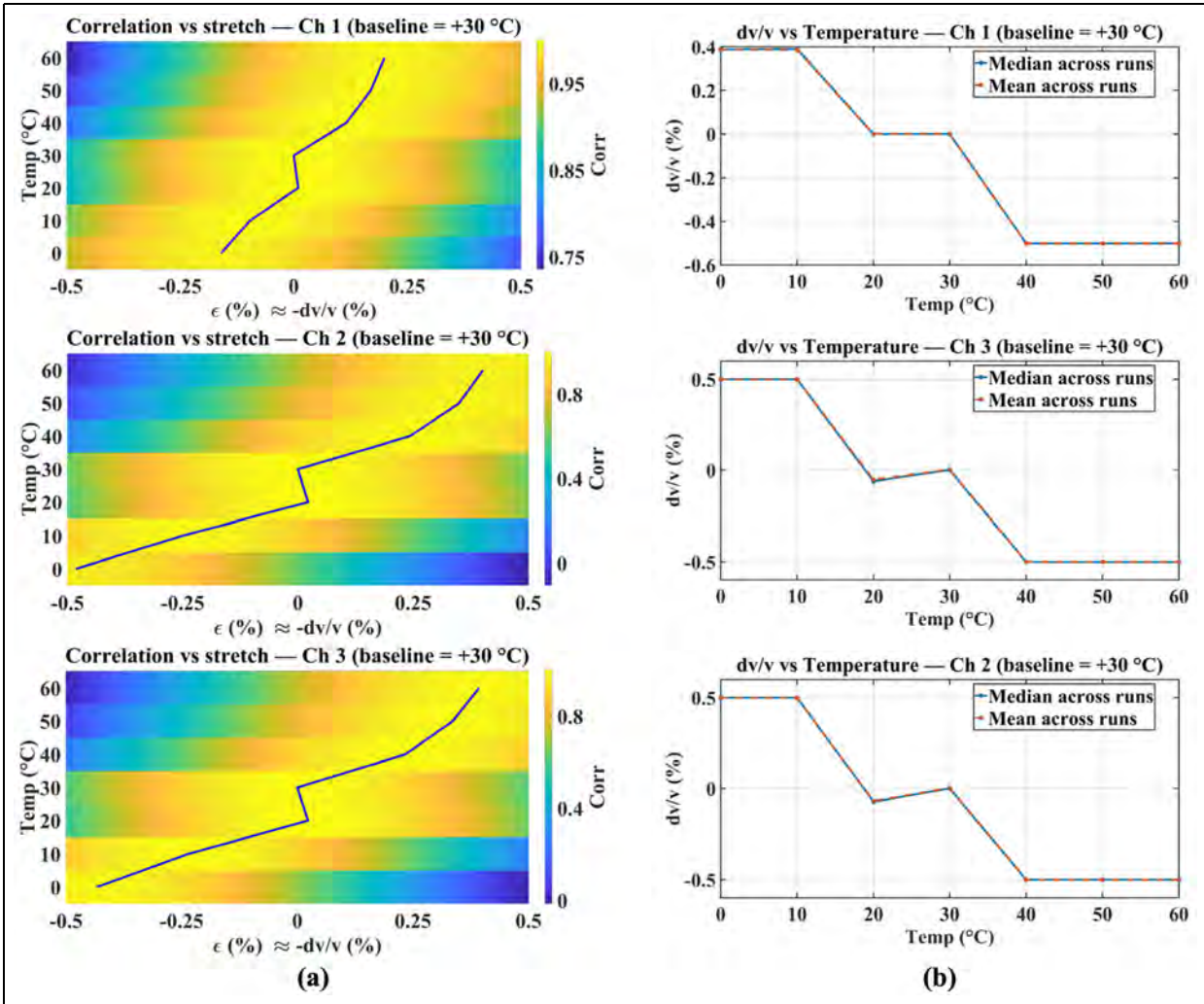


Figure 8. CWI results for the Conv-CVAE-synthesised damaged case C11 in CONCEPT: (a) correlation–stretch map; (b) fractional wave-speed change dv/v versus temperature. CWI: coda wave interferometry; CONCEPT: carbon–epoxy composite plate.

windowing procedure. To prevent test-set leakage during DA, 10 samples (from two original observations) were reserved for testing, while 15 samples (from three originals) were allocated for training and validation: 11 for DA and 4 for unsupervised validation. This stratified allocation ensured class balance and complete separation between adaptation and testing data.

Feature extraction. SPADA utilised a high-level WTS to extract discriminative features. Each channel was standardised to remove mean offsets and normalise variance. The scattering transform employed Morlet wavelets, suitable for vibration analysis due to their balanced frequency localisation and time resolution, capturing transient and oscillatory behaviours.³⁸ The transform used maximum scale 2^{10} for long-range

temporal dependencies and two wavelets per octave for balanced frequency resolution. Coefficients up to second scattering order captured primary spectral content and nonlinear frequency-band interactions.

Scattering coefficients underwent logarithmic compression to reduce large fluctuations and enhance stability. Temporal averaging produced fixed-length, time-shift-invariant descriptors sensitive to oscillatory structure. Features from three sensor channels were concatenated to form unified representations. WTS resulted in compact, translation-invariant, noise-robust features preserving fine-scale transients and broader structural variations, providing reliable input for subsequent DA and classification. To evaluate feature extraction efficacy, scattering transforms were applied to three-channel vibration signals at Wp20 (representative operational condition). Features were extracted

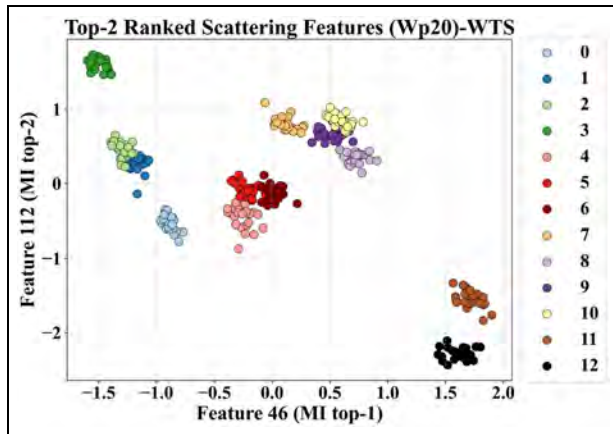


Figure 9. Scatter plot of the top-2 features ranked by mutual information, based on scattering transforms employed on Wp20; samples are colour-coded by labelled classes (Table 1).

per channel, normalised, temporally averaged, and concatenated. The top two features ranked by mutual information are visualised in Figure 9.

Figure 9 shows effective separation across classes. However, overlap persists between crack-related classes 5 and 6, and classes 9 and 10, likely from subtle crack characteristic differences (length, severity) harder to distinguish in 2D projections but linearly separable in full feature space.

To examine EOV impact on data distribution and assess WTS's domain-invariant feature capture, scatter analysis used three datasets (Wn15, Wp20, Wp40). The two most discriminative MI-ranked features appear in Figure 10(a). For comparison, WTS was substituted with a CNN comprising four one-dimensional

convolutional layers with batch normalisation, ReLU activations, global average pooling, and linear projection to 64-dimensional embedding space. The CNN was trained in a supervised manner on source-domain data. The same three datasets were processed, with top-ranked features in Figure 10(b). Features were ranked using Wp20 for consistency with Figure 9; five observations per class per domain were plotted for visibility.

In Figure 10(a) (WTS), class clusters exhibit clear domain-wise ordering: Wn15 samples consistently left, Wp20 centred, Wp40 right along the first MI feature, with this left-centre-right pattern repeating across classes while preserving intra-class compactness (red-dashed rectangular). Conversely, Figure 10(b) (CNNs) shows weaker domain regularity and greater intermingling of temperature samples within class groups despite comparable separability, suggesting the Wp20-trained CNN captures class-discriminative cues but with reduced domain awareness and poorer EOV alignment as will be discussed in the subsequent sections.

A one-dimensional ResNet-18 variant (ResNet1D) was additionally evaluated as a deeper learned backbone. The architecture comprised a convolutional stem (kernel size 7, stride 2, 64 channels, batch normalisation, ReLU, max-pooling with kernel size 3 and stride 2), followed by four residual stages of two BasicBlock1D modules each, with output channels of 64, 128, 256 and 512 and stride-2 downsampling at the first block of stages 2 through 4. Skip connections in

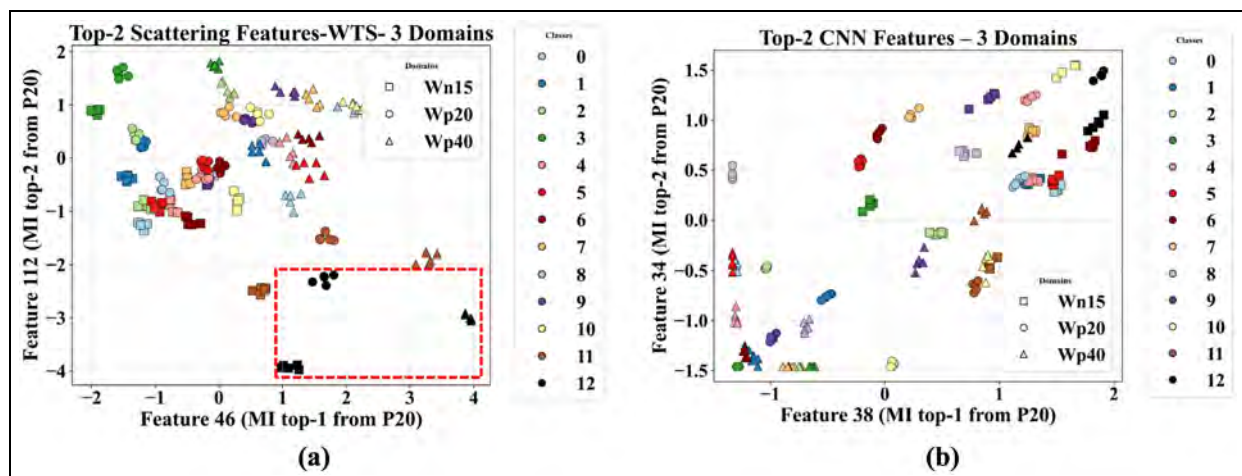


Figure 10. Scatter plot of the top-2 features ranked using Wp20 data, visualised across three temperature conditions (Wn15, Wp20, Wp40); for (a) WTS-based and (b) CNNs-based feature extraction. WTS: wavelet time scattering; CNN: convolutional neural network.

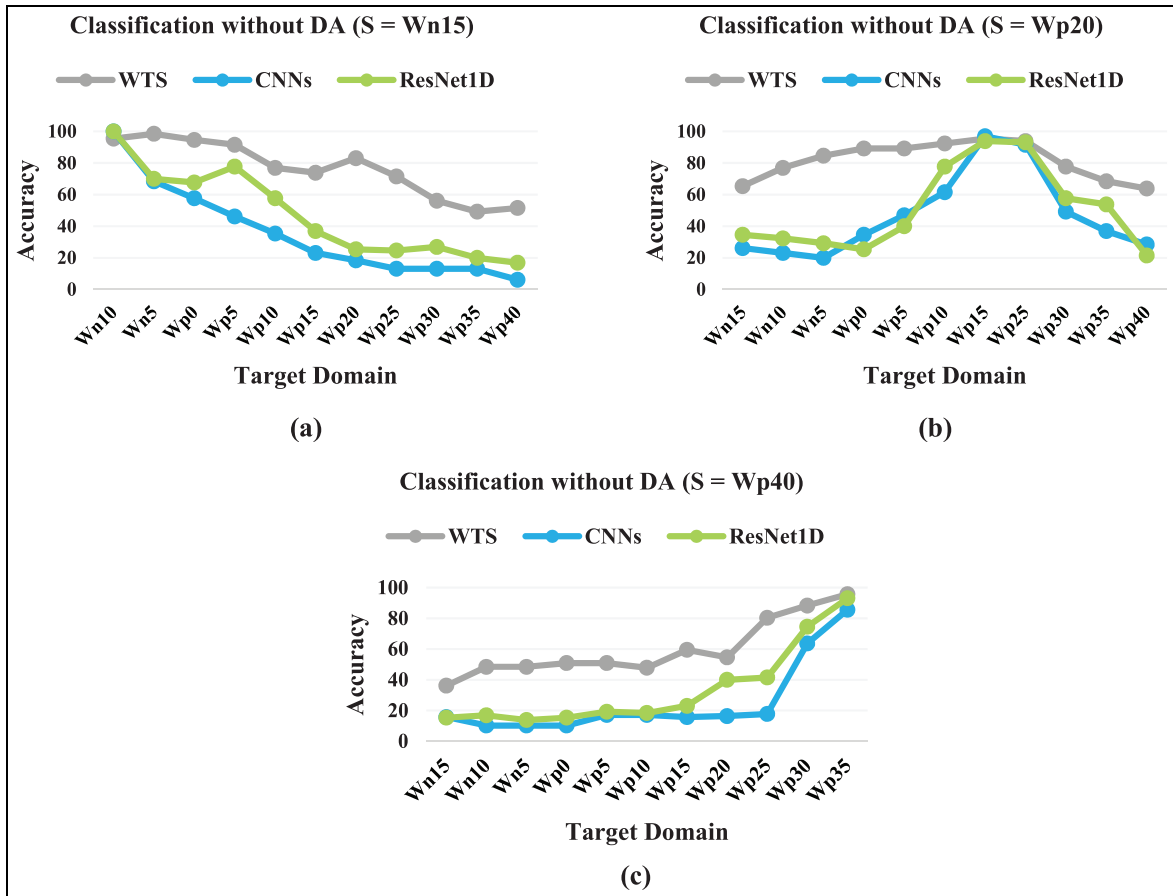


Figure 11. Damage detection utilising WTS, CNNs and ResNet1D without DA when (a) Wn15, (b) Wp20 and (c) Wp40 was assigned as the source domain. WTS: wavelet time scattering; CNN: convolutional neural network; DA: domain adaptation.

downsampled blocks used a 1 by 1 convolution for dimension matching. Global average pooling and a linear projection with dropout (0.2) produced a 64-dimensional embedding. The network was trained in a supervised manner on source-domain data using the same protocol as the CNN baseline. The effects of this feature extraction technique on the damage detection also will be discussed on WTB-VibClimate as well as CONCEPT in the following sections.

Damage detection without DA. Damage detection was conducted assuming no DA stage; 3 temperatures (Wn15, Wp20, Wp40) were independently designated as source domains, with remaining datasets as targets. DA loss weights (λ_{tgt} , λ_{adv}) were set to zero (Equation (21)). Grid search across 1000 feature spaces identified optimal classification accuracy. Figure 11(a) to (c) shows damage detection results without DA for Wn15, Wp20 and Wp40 as sources, respectively, following WTS feature extraction, classifier training on labelled source data and evaluation on target test sets. The

CNN framework described previously was applied as an alternative to WTS under the same source assumptions, with results presented in the same figures.

Figure 11 reveals clear temperature proximity effects. Using WTS, average accuracies across targets were 76.57% (Wn15 source), 81.54% (Wp20) and 60.0% (Wp40). Corresponding CNN averages were markedly lower: 35.87, 46.85 and 25.24%. Wp20 consistently provided strongest generalisation for both methods.

WTS degraded gracefully as source-target temperature gaps widened, while both learned backbones were markedly more sensitive to mismatch. With Wp20 as source, average accuracies across targets were 81.54% (WTS), 50.84% (ResNet1D) and 46.85% (CNN). With Wn15 as source, corresponding averages were 76.57% (WTS), 47.62% (ResNet1D) and 35.87% (CNN). With Wp40 as source, averages were 60.0% (WTS), 33.78% (ResNet1D) and 25.24% (CNN). ResNet1D consistently outperformed the shallower CNN but remained substantially below WTS across all source settings, indicating that increased network depth alone does not

Table 3. Hyperparameter candidates and values in SPADA.

Component	Value	Count
Epochs per trial	300, 400	2
Feature width	64, 128	2
Learning rate	1.5×10^{-5} , 2×10^{-5} , 5×10^{-5}	3
Domain loss weight	0.5, 1	2
Prototype weighting strength	0.05, 0.1, 0.2	3
Entropy gate	0.5, 0.6, 0.7	3
RMSprop momentum	0.5, 0.9	2
RMSprop weight decay	1×10^{-6} , 1×10^{-5}	2
Label smoothing	0.1, 0.2	2
Prototype-alignment loss weight	0.15, 0.2	2
Confidence threshold	0.7, 0.8	2
Prototype blend	0.8, 0.95	2
Weight on target pseudo-label loss	0.02, 0.03, 0.05, 0.07, 0.1, 0.12	6
Softmax temperature during adaptation	0.9, 1.0, 1.2, 1.5, 2, 2.5, 3, 3.5	8
Schedule steepness	5, 10, 12, 15, 20	5
Grid batch size	4, 8, 16, 24, 32	5
Global seed (per-trial)	1, 2, 3, 4, 5	5
Maximum scale (in WTS)	8, 9, 10, 15	4
Quality factor (in WTS)	1, 2	2
Maximum scattering order (in WTS)	1, 2	2
Log epsilon (in WTS)	1×10^{-6} , 1×10^{-5}	2
Batch size (in WTS)	24, 64	2

SPADA: scattering-based prototype-aligned domain adaptation; WTS: wavelet time scattering.

compensate for the absence of the translation-invariance and deformation-stability properties that WTS provides.

Persistently lower accuracies at larger temperature differences indicate systematic generalisation gaps, strongly justifying DA incorporation to mitigate temperature-induced covariate shift and stabilise performance across dissimilar operating conditions.

Damage detection with DA. To address detection gaps under large temperature differences, full SPADA with DA was deployed. Hyperparameters were tuned via random search over 1000 configurations without replacement. For each configuration, 15 independent seeds were run. All 15 seeds were executed for every sampled configuration, not only for the final selected configuration. This ensures that the mean unsupervised score used for configuration selection reflects the full seed-level variability of each candidate, rather than being estimated from a single run. Epochs were selected using the unsupervised two-part hold-out score on unlabelled target validation subsets A and B, as described in the training summary. Models were retrained with the chosen configuration under the same 15 seeds; test performance was reported based on the mean accuracy on held-out target test sets. Reproducibility was ensured by controlling all randomness sources: seeds were applied consistently to

Python, NumPy, PyTorch and scikit-learn, with deterministic data loading, shuffling and augmentation. Target-domain data used in DA was fixed at 60% class-balanced proportion; consequently, 10 observations per class were allocated to testing.

In WTS, maximum scale was constrained to 2^{15} ($J = 15$), satisfying $2^{15} \leq T$ where $T = 39,200$ (augmented dataset signal length). Table 3 summarises hyperparameters; random subsets were sampled without replacement from the full grid.

The unsupervised selection score used fixed combination weights $\omega_{\text{comp}} = 0.5$ and $\omega_{\text{cov}} = 0.1$. These weights were not tuned but set to ensure that diversity and confidence dominate the score while compactness and coverage act as secondary regularisers.

Following the without-DA scenario, three source-to-target settings were considered. Figure 12(a) to (c) presents highest target accuracies, with CNN and DA results plotted for comparison.

With Wp20 as source (Figure 12(b)), accuracies were uniformly high, averaging 97.58%. With Wn15 as source (Figure 12(a)), perfect results occurred for nearby targets, but performance decreased for hottest targets (Wp30 66.92%, Wp35 61.54%), averaging 86.29%. With Wp40 as source (Figure 12(c)), coldest targets were most challenging (Wn10 56.92%, Wn15 59.23%), while warm targets remained strong, averaging 79.65%. Comparing to Figure 11 (without DA), averages increased from 35.87, 46.85 and 25.24% to

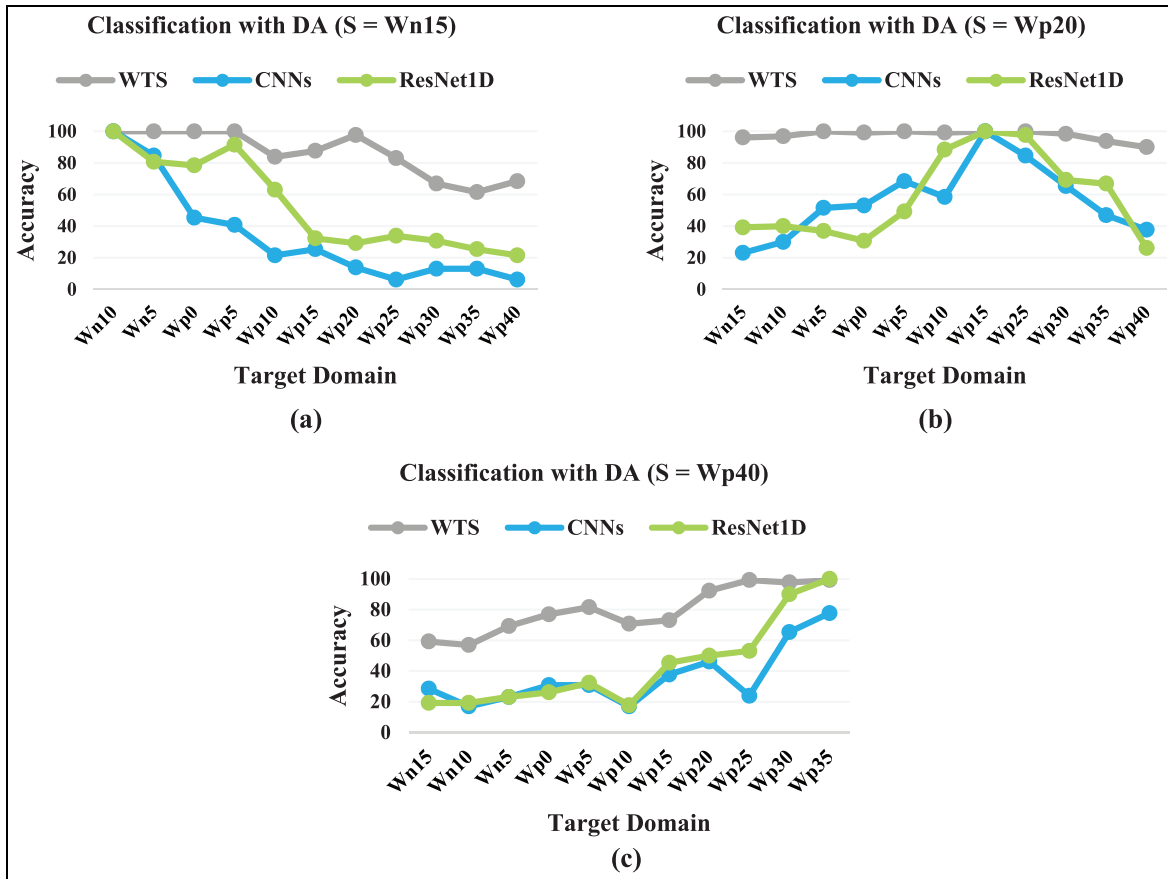


Figure 12. Damage detection utilising WTS, CNNs and ResNet1D with DA when (a) Wn15, (b) Wp20 and (c) Wp40 was assigned as the source domain. WTS: wavelet time scattering; CNN: convolutional neural network; DA: domain adaptation.

86.29, 97.62 and 79.65%, indicating earlier low-accuracy gaps were largely closed. For instance, Wn15 to Wp35 increased from 13.08 to 61.54%.

Both learned backbones with DA achieved lower accuracies than WTS with DA. ResNet1D with DA improved over its without-DA baseline (averages rising from 50.84 to 58.60% with Wp20, from 47.62 to 53.36% with Wn15 and from 33.78 to 43.29% with Wp40), confirming that the adaptation mechanism provides benefit even with deeper learned features. CNN with DA showed similar but smaller gains, and in some transfers, performance fell below without-DA settings, reflecting negative transfer. Nonetheless, both learned backbones with DA remained substantially below WTS with DA (97.58, 86.29 and 79.65% for the three source settings), reinforcing that the stability properties of WTS features facilitate more effective domain alignment than representations optimised purely for source-domain discrimination.

To understand which classes SPADA (with DA) struggles to classify, confusion matrices are presented for 4 target domains (Wn10, Wp0, Wp30, Wp35) in Figure 13(a) to (d), respectively, assuming Wp20 as source. These targets were selected because SPADA did not achieve full accuracy, allowing detailed limitation examination.

Figure 13 shows SPADA's residual errors concentrate almost entirely in pairwise confusions between labels 5 and 6; misclassifications are symmetric and limited, indicating tight decision boundaries rather than widespread class drift. These labels correspond to closely related crack configurations (two vs three 5 cm cracks), inducing similar stiffness reductions and mode-shape perturbations. Under temperature shift, spectral signatures become more alike due to thermal softening and peak broadening, narrowing margins between class prototypes. Effects are amplified by (i) sensor-limited operation (three channels), reducing spatial sensitivity

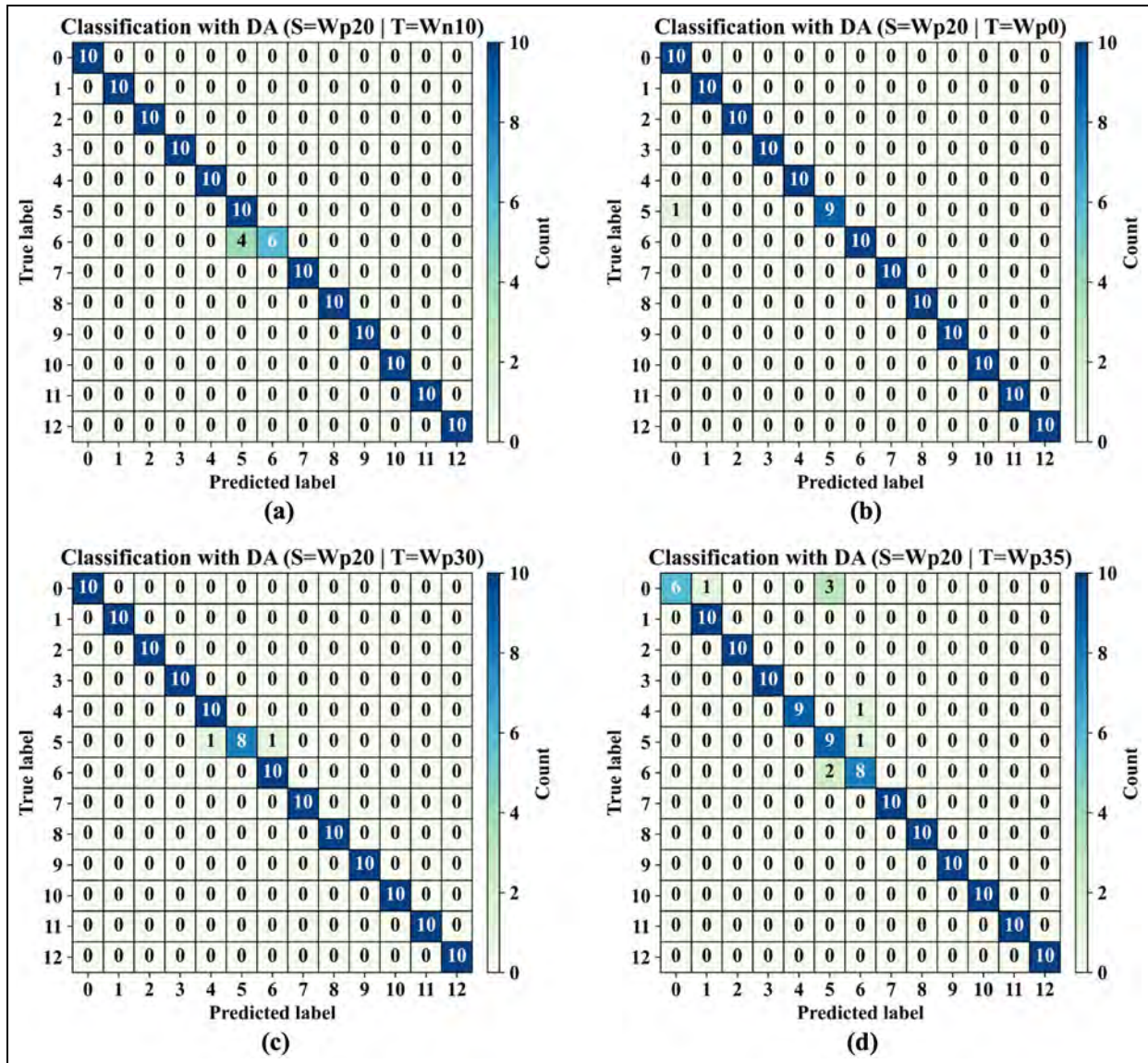


Figure 13. Confusion matrices of SPADA with Wp20 as the source and (a) Wn10, (b) Wp0, (c) Wp30 and (d) Wp35 as target domains. SPADA: scattering-based prototype-aligned domain adaptation.

to crack multiplicity; (ii) windowed segmentation, preserving local transients but weakening global geometry cues; and (iii) conservative pseudo-labelling during adaptation, slightly relaxing class margins for near-neighbour classes.

To address safety-relevant diagnostic performance, per-class recall was computed for all transfer pairs with Wp20 and Wn15 as source domains. Figure 14 presents these values as heatmaps, where rows correspond to target transfers and columns to health-scenario classes. Since the false negative rate is the complement of recall (False Negative Rate = 1 - recall), only the recall heatmaps are shown. With Wn15 as source

(Figure 14(a)), a clear gradient emerges: nearby targets retain high recall across all classes, while distant targets (Wp25 to Wp40) exhibit reduced recall for mid-severity crack classes (4-9), reflecting the compounded difficulty of distinguishing closely related damage configurations under large thermal shifts. With Wp20 as source (Figure 14(b)), recall remains at or near unity across the majority of transfers and classes; the only visible degradation concentrates on classes 5 and 6 at the widest temperature gaps (Wp35 and Wp40), consistent with the pairwise crack-configuration confusion identified in the confusion matrices. Because all health scenarios contain equal numbers of observations, the

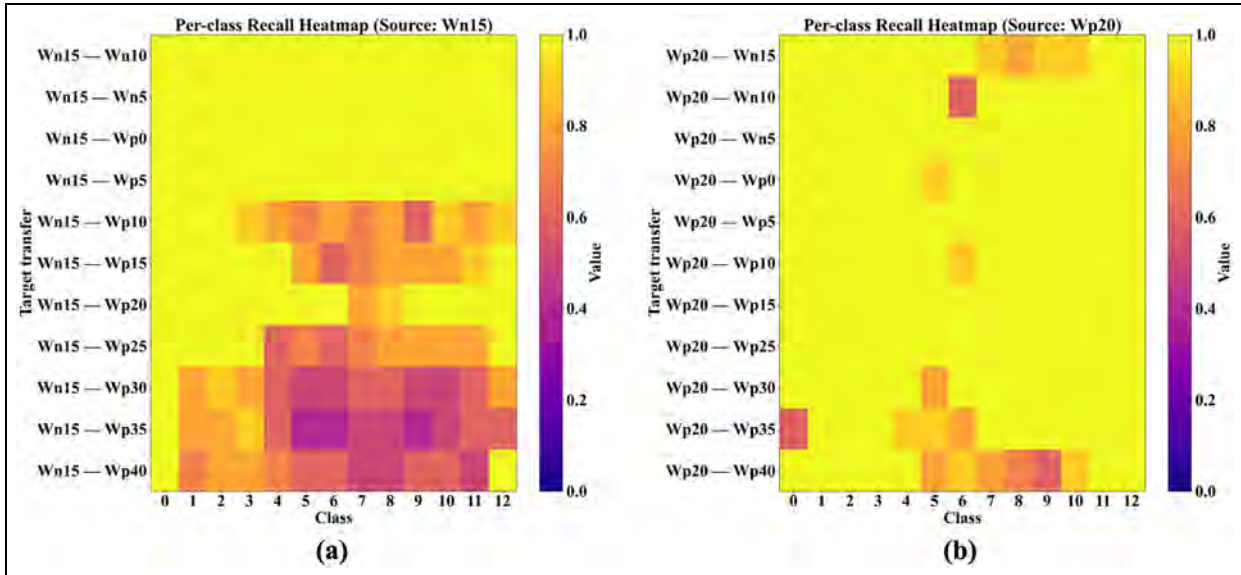


Figure 14. Per-class recall heatmaps for WTB-VibClimate with DA: (a) Wn15 and (b) Wp20 as source domain. WTB-VibClimate: small-scale wind turbine blade under varying climate conditions; DA: domain adaptation.

macro-averaged F1 score is numerically close to the overall accuracy for each transfer pair and is therefore not reported separately.

Concept. Before presenting the CONCEPT adaptation results, it is important to note that this case study constitutes a constructed evaluation setting: healthy signals are experimental across all seven temperatures, while damaged signals at non-baseline temperatures are synthetic, generated by the Conv-CVAE described in previous sections. Consequently, performance figures for CONCEPT transfers reflect the combined effect of the adaptation mechanism and the generator fidelity, and should not be interpreted as evidence of robustness against experimentally measured damaged responses at varying temperatures. Direct experimental support for cross-temperature robustness is provided by the WTB-VibClimate results, where real damaged data exist at all temperature conditions.

SPADA was evaluated on composite plate damage detection using the CONCEPT dataset augmented with synthetic damaged signals generated by the convolutional conditional variational autoencoder. To avoid DA bias, the baseline temperature dataset (Cp30), which contains the only experimentally measured damaged signals, was excluded and used only for scatter plots. Only Cp0, Cp10, Cp20, Cp40, Cp50 and Cp60 were considered. Two scenarios were examined: (1) Cp0 (lowest temperature) as source with remaining temperatures as targets and (2) Cp60 (highest temperature) as source with others as targets.

During DA, 50% of target data were allocated for training, 50% for testing (class-balanced, randomly selected). Consequently, 15 observations per class were classified in testing.

Feature extraction. The same WTS block from WTB-VibClimate was employed, with maximum scale limited to 2^9 due to shorter signal length (1000 points per channel). To assess feature extraction efficacy, scattering transforms were applied to three-channel guided waves at Cp30 (maximum scale 2^7). Features from each channel were normalised, temporally averaged, and concatenated. The top two MI-ranked features appear in Figure 15.

Figure 15 demonstrates successful WTS class separation. However, meaningful class hierarchy must also be maintained under temperature variation. Figure 16(a) and (b) present 2D scatter plots of MI-ranked features for three domains (Cp0, Cp30, Cp60) using WTS and CNNs, respectively. CNNs used the same supervised framework as WTB-VibClimate; five observations per health scenario per domain were plotted.

Figure 16 shows CNNs achieved reasonable domain cluster distinguishability with visible class-wise separations (apparent overlaps result from figure density). WTS demonstrated strong domain and class-level separability. The domain separation pattern from WTB-VibClimate (Figure 10(a), red box) recurs here but differently positioned. In Figure 16(a), Cp30 clusters (baseline temperature) occupy the lower left rather than class-area centres. Figure 16(b) indicates CNN-

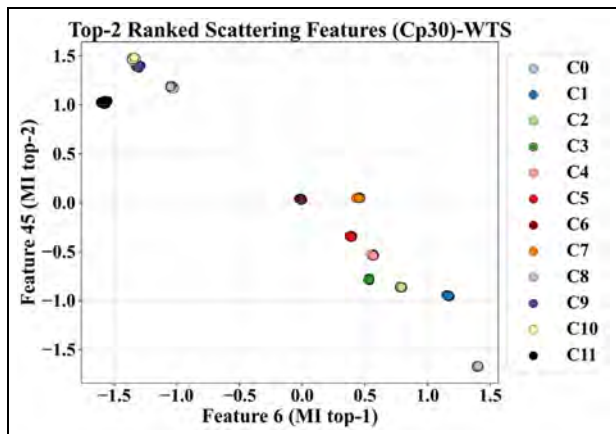


Figure 15. Scatter plot of the top-2 features ranked by mutual information, based on scattering transforms employed on Cp30; samples are colour-coded by labelled classes (Table 2).

extracted features exhibit large domain gaps, challenging subsequent UDA adjustment.

Damage detection without DA. DA was initially omitted to examine WTS’s domain-gap compensation. Cp0 and Cp60 were designated independent sources, with remaining datasets as targets (λ_{tgt} , $\lambda_{adv} = 0$). Grid search across 1000 feature spaces determined optimal accuracy. Figure 17(a) and (b) present results without DA for Cp0 and Cp60 sources, respectively, following WTS extraction, classifier training, and target test evaluation. The same CNN framework was employed as an alternative; results appear in the same figures.

Figure 17 shows WTS significantly outperformed both learned backbones. With Cp0 as source, average accuracies were 73.89% (WTS), 29.67% (ResNet1D) and 30.40% (CNN). With Cp60 as source, averages were 71.89% (WTS), 42.99% (ResNet1D) and 51.33% (CNN). ResNet1D and CNN performed comparably, both exhibiting sharp accuracy drops at wider temperature gaps, while WTS maintained substantially higher performance across all transfers.

Damage detection with DA. Full SPADA was employed with the hyperparameters in Table 3, searching 1000 random configurations with maximum scale constrained to 2^9 ($J = 6, 7, 8, 9$). Each target split into two class-balanced subsets: 11 samples per class for DA, 4 for the validation and 15 for testing. As in the WTB-VibClimate study, model selection for each configuration used the unsupervised two-part A/B hold-out score on unlabelled target validation data, and the configuration with the highest mean unsupervised score across seeds was retained. Figure 18(a) and (b) shows results with Cp0 and Cp60 sources.

Figure 18 shows DA stage increases accuracy, with the largest gains at wider temperature gaps. With Cp0 source and Cp40 target, accuracy rose from 48.33 to 94.44%; with Cp0 source and Cp60 target, accuracy increased from 45 to 93.33%. With Cp60 source, accuracy increased from 41.67%, 51.67%, 66.11% to 91.67%, 94.44%, 100% for Cp0, Cp10, Cp20 targets, respectively. Small gaps (Cp0 source with Cp10 target) maintained 100%. Room for improvement remains at widest gaps (e.g., Cp60 source with Cp0 target: 91.67%). ResNet1D with DA improved over its without-DA baseline (averages rising from 29.67 to 47.34%)

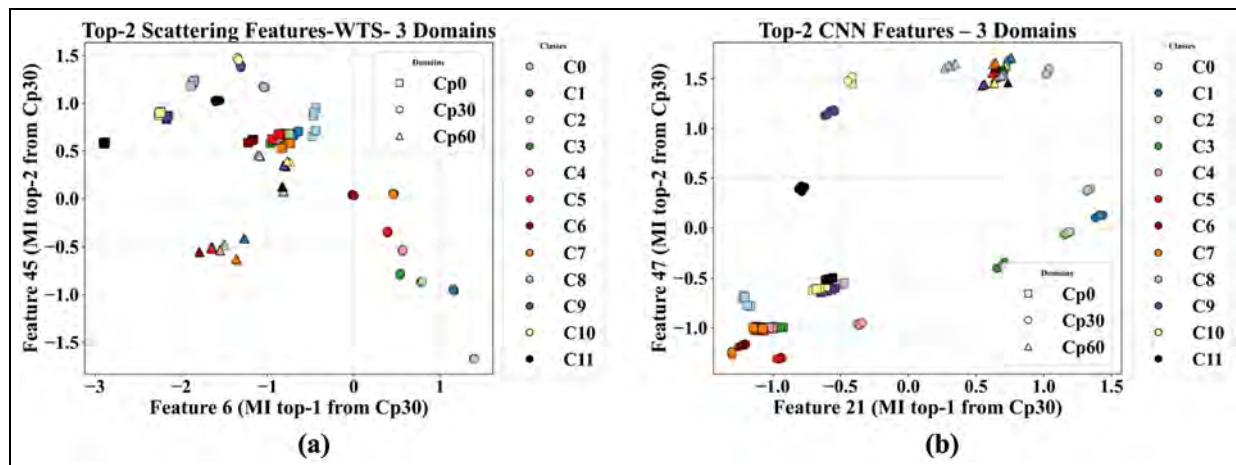


Figure 16. Scatter plot of the top-2 features ranked using Cp30 data, visualised across three temperature conditions (Cp0, Cp30, Cp60); for (a) WTS-based and (b) CNNs-based feature extraction. WTS: wavelet time scattering; CNN: convolutional neural network.

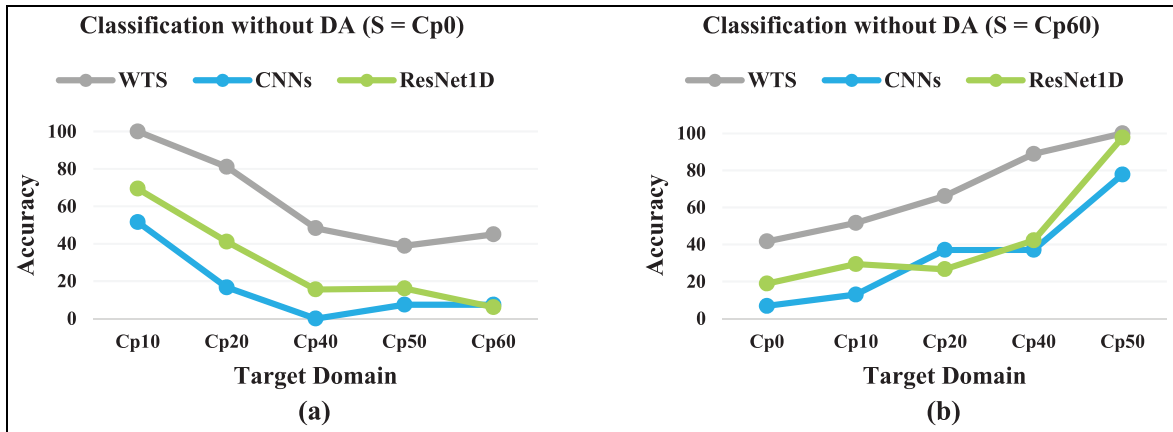


Figure 17. Damage detection utilising WTS, CNNs and ResNet1D without DA when (a) Cp0 and (b) Cp60 was assigned as the source domain. WTS: wavelet time scattering; CNN: convolutional neural network; DA: domain adaptation.

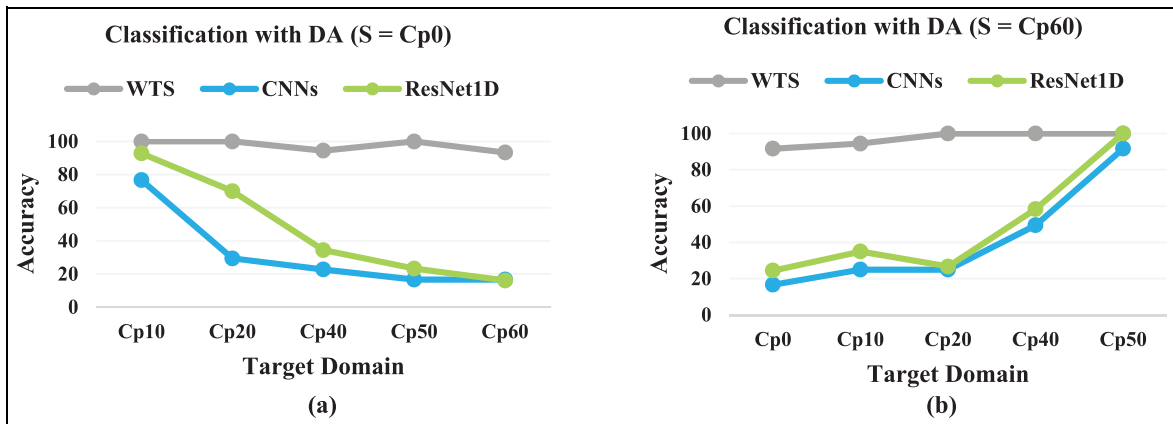


Figure 18. Damage detection utilising WTS, CNNs and ResNet1D with DA when (a) Cp0 and (b) Cp60 were assigned as the source domain. WTS: wavelet time scattering; CNN: convolutional neural network; DA: domain adaptation.

with Cp0 and from 42.99 to 48.89% with Cp60), confirming that DA provides partial benefit with deeper learned features. However, these figures remained far below WTS with DA (96.11 and 95.56% for the same source settings), indicating that the structured, bounded feature-space displacements produced by WTS are substantially more amenable to prototype-based alignment than the less constrained representations learned by ResNet1D.

To determine residual misclassification patterns where accuracy was below 100%, Figure 19(a) to (d) presents confusion matrices for four settings: Cp0 as source with Cp60 as target, Cp0 as source with Cp40 as target, Cp60 as source with Cp0 as target and Cp60 as source with Cp10 as target.

When Cp0 was used as source and Cp60 as target (Figure 19(a)), matrices are mostly diagonal with errors concentrated between adjacent severity classes. Largest

exchanges occur between C2 and C3, with smaller leakages from C4 to C3, C1 to C2 and C7 to C8. When Cp0 was used as source and Cp40 as target (Figure 19(b)), almost all classes are correct; main deviations are C9 misclassified as C10 and minor C3 misclassified as C4. These patterns align with CONCEPT physics: Lamb waves from neighbouring severities produce highly similar dispersion and attenuation signatures, especially under larger temperature separations.

With Cp60 as source and Cp0 as target (Figure 19(c)), residual errors remain local, concentrated within C2–C4, indicating limited margins between adjacent prototypes under largest shifts. When Cp60 was used as source and Cp10 as target (Figure 19(d)), diagonals tighten further. Largest residuals occur for C2 (two samples predicted as C1, three as C3), minor errors for C3 (two samples predicted as C4, one as C5) and C8 (two samples predicted as C9); all others

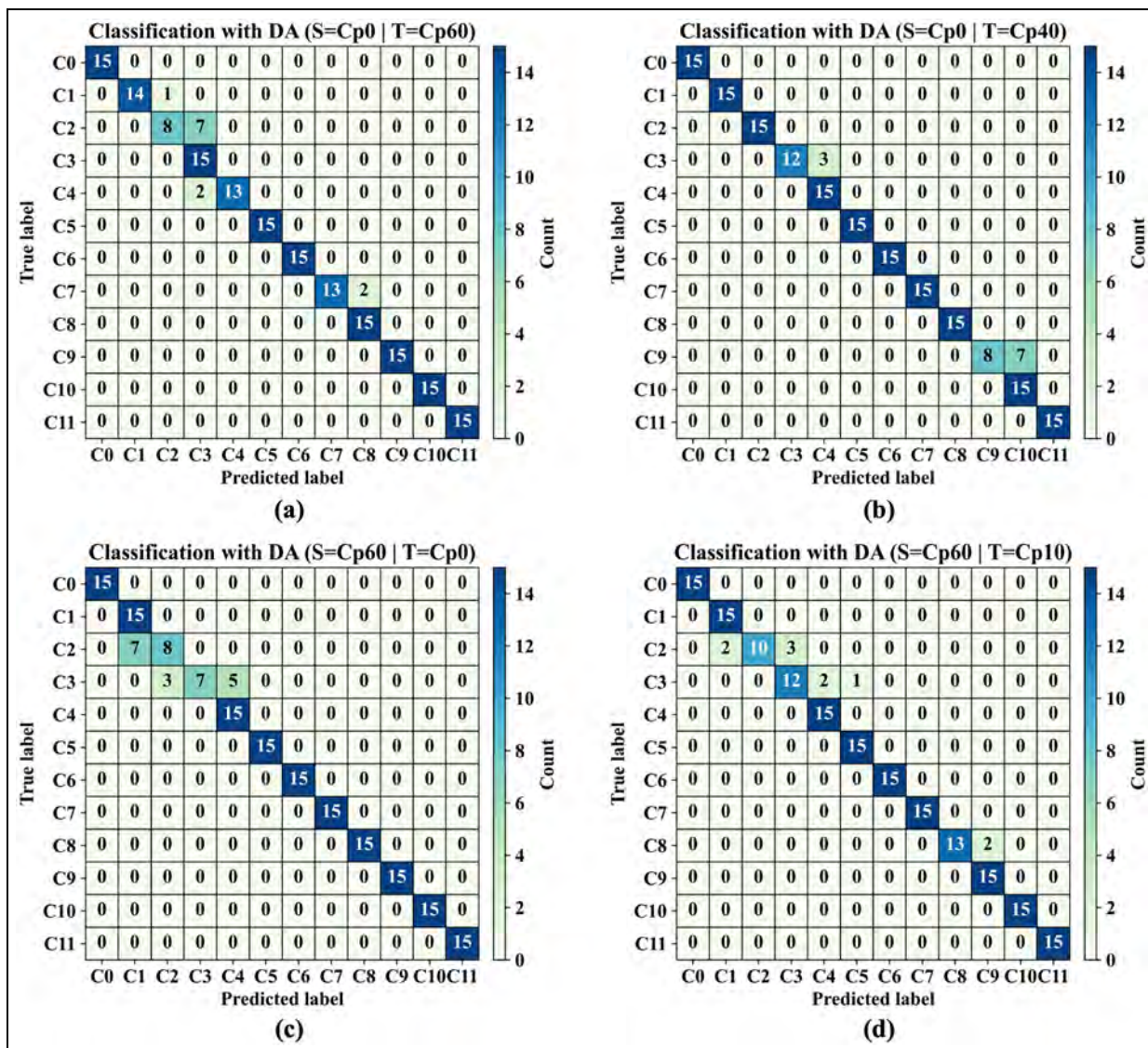


Figure 19. Confusion matrices for SPADA on CONCEPT for (a) source Cp0 and target Cp60, (b) source Cp0 and target Cp40, (c) source Cp60 and target Cp0 and (d) source Cp60 and target Cp10. SPADA: scattering-based prototype-aligned domain adaptation; CONCEPT: carbon–epoxy composite plate.

correct. Interpreting classes as healthy C0 and increasing delamination severities C1–C11 from incremental putty coverage, local swaps are physically plausible because adjacent severities perturb wavefields similarly between actuators and receivers. Results suggest value in class-conditional alignment or explicit pairwise margin penalties rather than global alignment alone.

Internal mechanism visualisation and interpretability analysis

To enhance transparency, SPADA logged internal states every 10 training epochs during DA: prototype vectors, instance features, pseudo-labels, entropy values

and attention weights. Four t-SNE-based visualisations reveal progressive domain alignment while preserving class separability. Analysis uses two CONCEPT scenarios: Cp0 source with Cp40 target (94.44% accuracy) and Cp0 source with Cp50 target (100% accuracy). Epoch numbers on axes represent logged snapshots; multiply by 10 for actual training epochs. Entropy threshold was $\tau = 0.6$.

Prototype trajectory evolution. Prototype trajectories were projected into 2D t-SNE space fitted on concatenated source and target prototypes across logged epochs. Trajectories sampled every 10th epoch: source

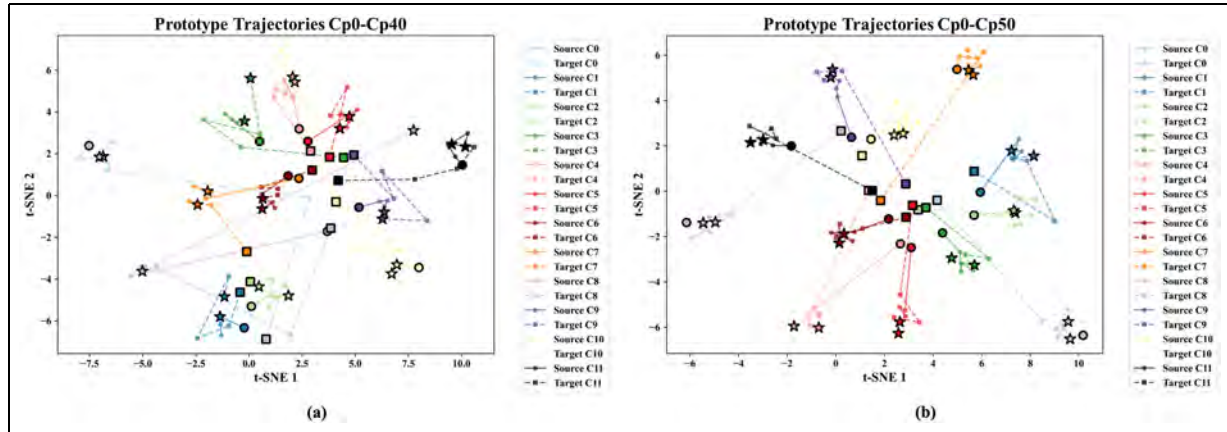


Figure 20. Prototype trajectory evolution through t-SNE embedding for (a) Cp0-to-Cp40 transfer and (b) Cp0-to-Cp50 transfer. t-SNE: t-distributed stochastic neighbour embedding.

prototypes as solid lines with circles, target prototypes as dashed lines with squares. Starting positions have larger filled markers with black edges; ending positions have star symbols. Figure 20(a) and (b) shows trajectories.

For Cp0 source with Cp50 target (Figure 20(b), 100% accuracy), source and target prototypes converge to nearly coincident positions, indicating complete alignment. For Cp0 source with Cp40 target (Figure 20(a), 94.44% accuracy), class C8 prototypes diverged substantially, ending at opposite regions, likely contributing to residual errors. Classes C3 and C4 show partial convergence, while C6 and C7 achieve satisfactory alignment.

Instance-prototype cosine similarity. Heatmaps display cosine similarities between instances and prototypes at final epoch. Rows represent instances (source upper, target lower); columns represent prototypes (source left, target right). White lines delineate domains. Figure 21(a) and (b) shows similarity matrices.

Both graphs exhibit pronounced diagonal structures with elevated values along main diagonals, confirming instances align with corresponding class and domain prototypes. Cp0 source with Cp50 target (Figure 21(b)) demonstrates sharper diagonals with minimal off-diagonal activations (100% accuracy). Cp0 source with Cp40 target (Figure 21(a)) shows weaker diagonal intensity and moderate off-diagonal values, reflecting 5.56% error.

Prototype attention weight dynamics. Three panels track attention dynamics. Left: median source attention with IQR shading. Middle: median target attention with IQR shading. Right: pseudo-label coverage (solid line) and median target weight (dashed line). Vertical dashed

line marks selected epoch. Figure 22(a) and (b) presents dynamics.

Source weights remain at unity. For Cp0 source with Cp50 target (Figure 22(b)), best epoch at 12, target weights stabilise around 0.3, coverage rises to 0.5 and plateaus. For Cp0 source with Cp40 target (Figure 22(a)), best epoch at 29, target weights fluctuate between 0.2 and 0.5 with wider IQR, coverage peaks near 0.6 at epoch 33 then declines. Earlier stabilisation in Cp0 source with Cp50 target corroborates superior accuracy.

Decision boundaries in feature space. Instances and prototypes projected into 2D t-SNE at best epoch. Source instances: circles; target instances: squares; source prototypes (Proto): X markers; target prototypes: filled plus symbols. Figure 23(a) and (b) shows boundaries.

Cp0 source with Cp50 target (Figure 23(b)) shows tight, well-separated clusters by class with prototypes coincident or proximate, confirming complete convergence. Cp0 source with Cp40 target (Figure 23(a)) exhibits looser clustering with noticeable prototype gaps, visually corroborating 5.56% accuracy deficit.

Native-space quantitative metrics. To complement the t-SNE visualisations with quantitative measures computed directly in the 64-dimensional latent space, two metrics were tracked over training epochs for the Cp0 to Cp40 and Cp0 to Cp50 transfers (Figure 24). The prototype separation ratio (PSR), defined as the mean inter-class prototype Euclidean distance divided by the mean intra-class instance-to-prototype distance, quantifies how well-separated the class prototypes are relative to the dispersion of instances around them. The silhouette score, computed on instance features with class labels (true labels for source, pseudo-labels for

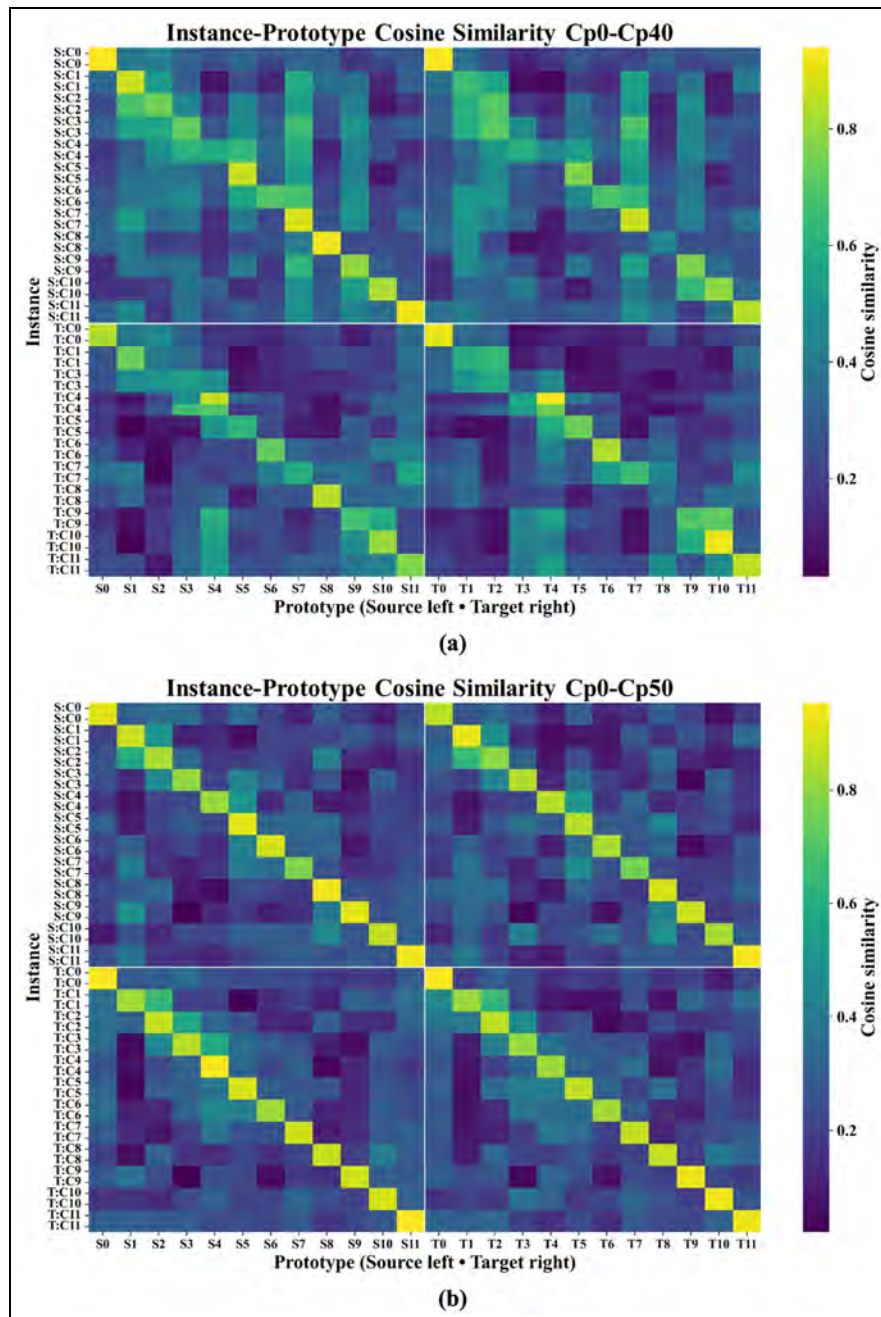


Figure 21. Instance-prototype cosine similarity heatmap for (a) Cp0-to-Cp40 transfer and (b) Cp0-to-Cp50 transfer.

target), provides a complementary global measure of cluster quality.

Figure 24 represents that for both transfers, the source-domain PSR and silhouette remained high and stable throughout training (PSR approximately 2, silhouette above 0.9), confirming that source class structure was preserved during adaptation. The target-domain metrics exhibited markedly different

trajectories: PSR rose from 1.03 to 1.77 (Cp40) and from 0.98 to 2.28 (Cp50), while silhouette increased from 0.42 to 0.82 (Cp40) and from 0.07 to 0.94 (Cp50). The convergence of target metrics towards source-domain values provides quantitative evidence, independent of t-SNE projection, that the adaptation mechanism progressively organises target representations into class-discriminative clusters consistent with the source-

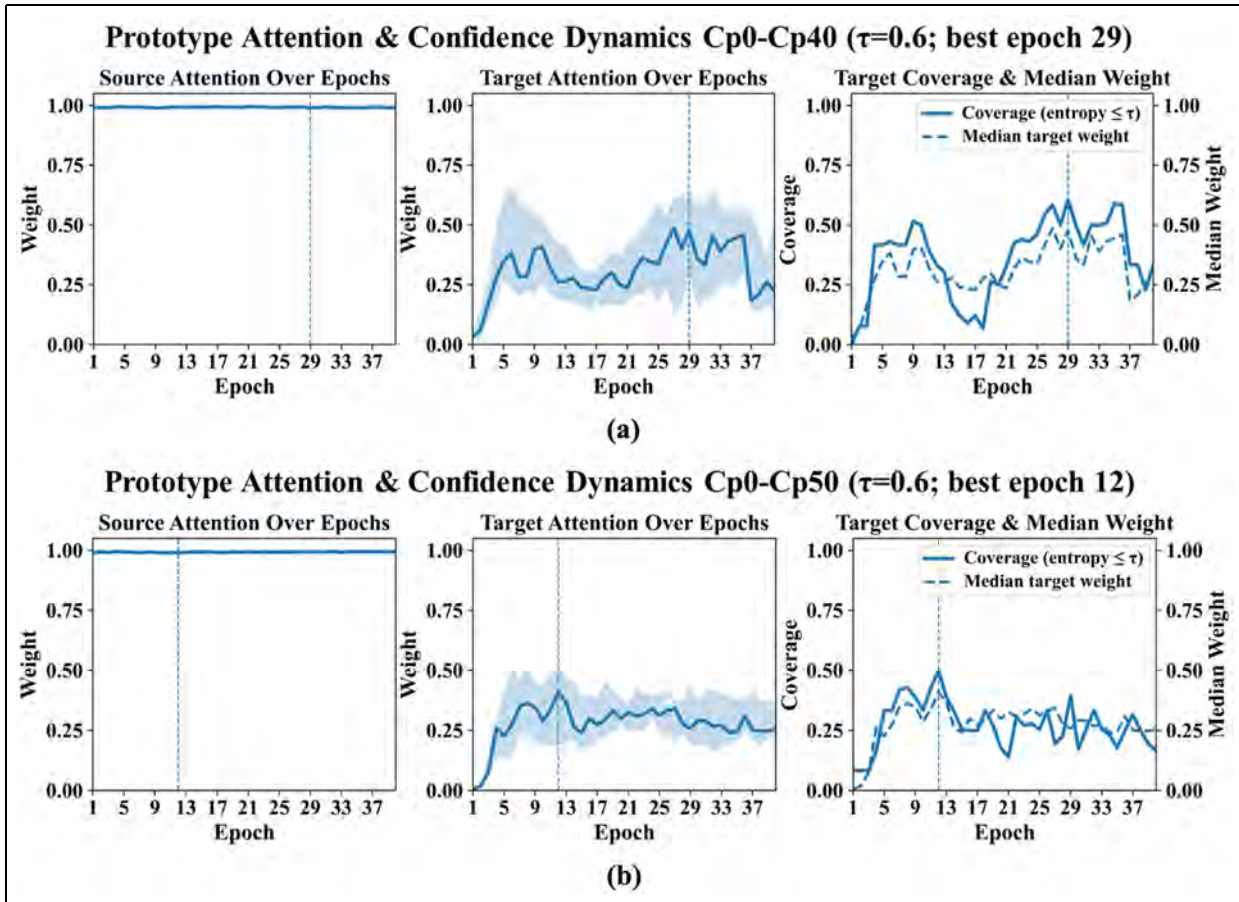


Figure 22. Prototype attention and confidence for (a) Cp0-to-Cp40 transfer and (b) Cp0-to-Cp50 transfer.

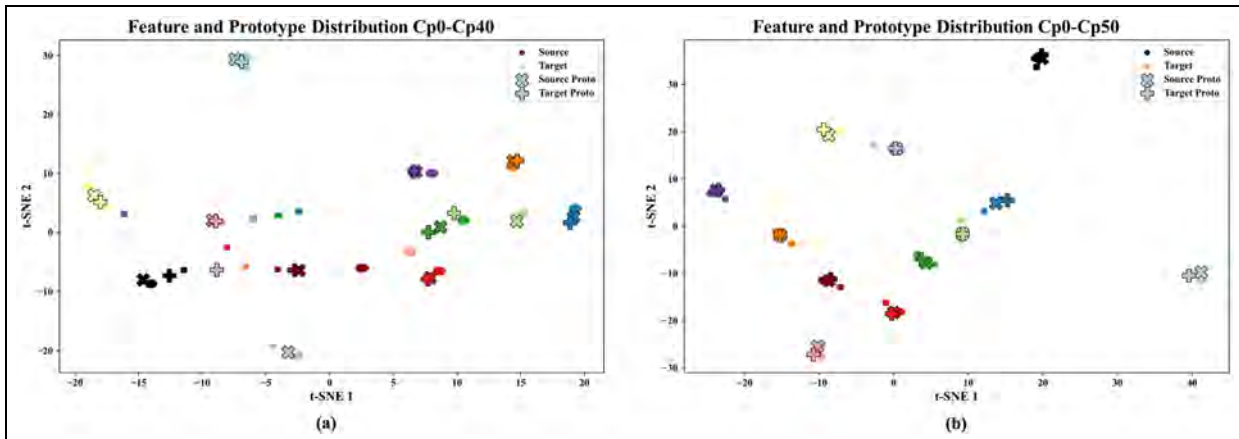


Figure 23. Decision boundaries in t-SNE embedding space for (a) Cp0-to-Cp40 transfer and (b) Cp0-to-Cp50 transfer. t-SNE: t-distributed stochastic neighbour embedding.

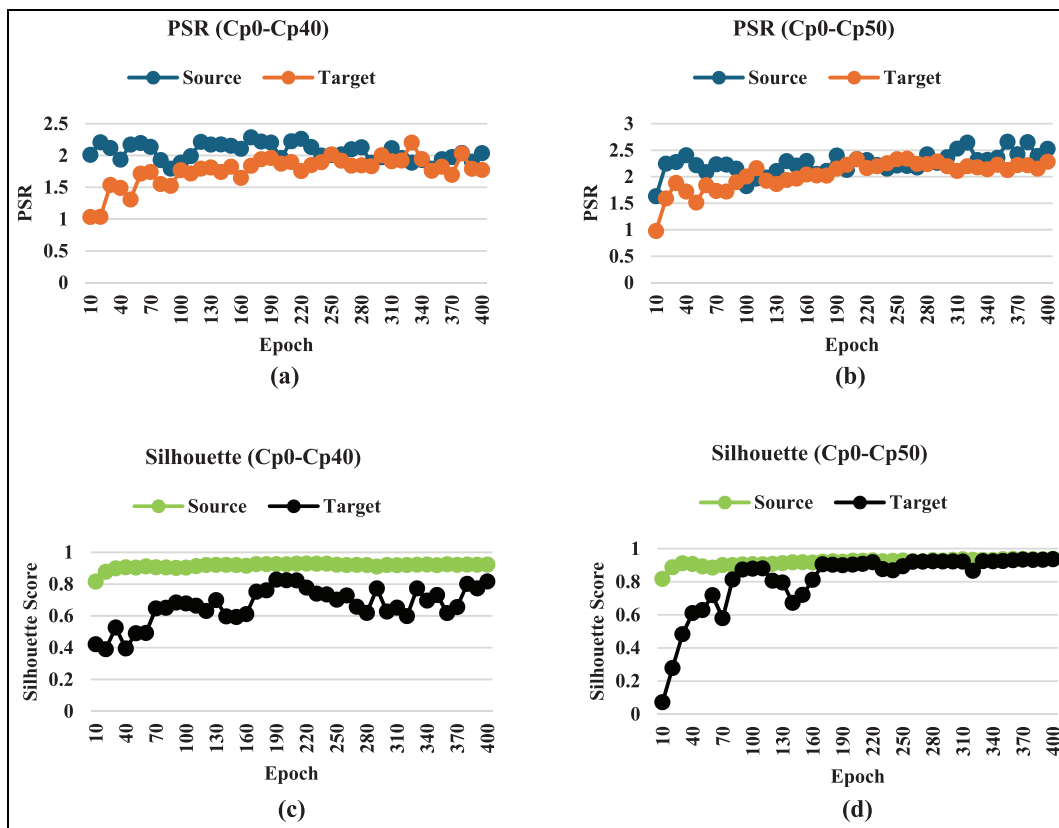


Figure 24. Native-space quantitative metrics over training epochs for CONCEPT: PSR for (a) Cp0-to-Cp40 transfer and (b) Cp0-to-Cp50 transfer, and silhouette score for (c) Cp0-to-Cp40 transfer and (d) Cp0-to-Cp50 transfer. CONCEPT: carbon–epoxy composite plate; PSR: prototype separation ratio.

domain structure. The stronger improvement for Cp50 is consistent with its higher final classification accuracy (100 vs 94.44% for Cp40).

Comparison study

SPADA framework was evaluated against six reference methods by employing WTS for feature extraction with different UDA modules: adversarial DA with prototypes (ADA-Proto),³⁹ multi-domain adversarial DA with prototype attention (MADA-Proto),⁴⁰ generative adversarial network with prototype weighting (GAN-Proto),⁴¹ central moment discrepancy (CMD),⁴² correlation alignment (CORAL)⁴³ and neighbour refinement consistency with virtual adversarial training (NRC-VAT).⁴⁴ Experiments used augmented WTB-VibClimate and CONCEPT datasets. Challenging shifts were considered: Wp20 source with Wn15 and Wp40 targets for WTB-VibClimate; Cp0 source with Cp60 target and Cp60 source with Cp10 target for CONCEPT.

For each method, 1000 configurations were sampled without replacement from the corresponding

hyperparameter ranges and evaluated across 15 independent seeds. All methods followed the same search protocol; the sole difference was the selection criterion: SPADA retained the configuration with the highest mean unsupervised A/B score, while the baselines retained the configuration with the highest mean target-domain validation accuracy. SPADA therefore operates under a more constrained selection regime, as no target labels are used at any stage.

To validate that the unsupervised A/B score provides a meaningful proxy for true target-domain performance, Spearman rank correlations were computed between the mean unsupervised score and mean target accuracy across 1000 sampled configurations (15 seeds each) for two representative CONCEPT transfers. For Cp0 to Cp10 (small temperature gap), $\rho = 0.317$ ($p < 5 \times 10^{-6}$); for Cp0 to Cp40 (large temperature gap), $\rho = 0.315$ ($p < 6 \times 10^{-6}$). Both correlations are statistically significant and confirm a positive monotonic association between the unsupervised score and true accuracy. The unsupervised-selected configuration achieved 100% for Cp0 to Cp10 (matching the oracle) and 94.44% for Cp0 to Cp40 (oracle: 95.39%),

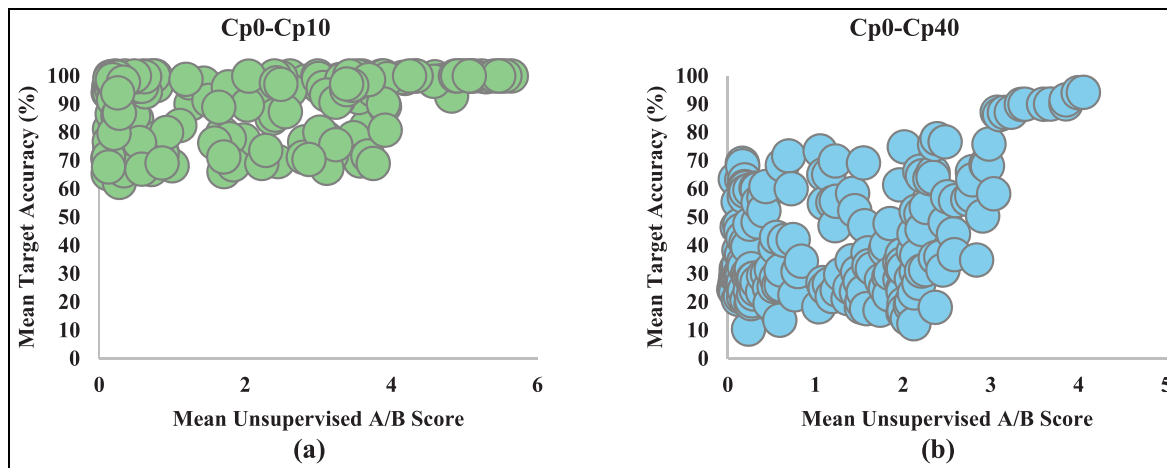


Figure 25. Scatter plots of mean unsupervised A/B score versus mean target accuracy for (a) Cp0-to-Cp10 transfer and (b) Cp0-to-Cp40 transfer.

resulting in a selection penalty of 0 and 1.22%, respectively. These results confirm that the proposed score reliably identifies near-optimal configurations without access to target labels. Figure 25 presents the corresponding scatter plots.

Figure 26(a) to (d) shows accuracy for each source–target combination.

Figure 26 reveals SPADA delivers the highest accuracy across all transfers, with largest advantage on Wp20 to Wp40. For WTB-VibClimate, SPADA attains 96.15% (Wp20 to Wn15), exceeding ADA-Proto by 1.06 percentage points, and 90% (Wp20 to Wp40), outperforming next best by 11.47 percentage points. Prototype-based adversarial baselines generally outperform CMD, CORAL, and NRC-VAT but trail SPADA. For CONCEPT, SPADA achieves 91.67% (Cp0 to Cp60), ahead of GAN-Proto by 1.67 percentage points, and 94.44% (Cp60 to Cp10), ahead of ADA-Proto by 1.11 percentage points. Across all four shifts, SPADA shows smallest performance spread at 6.15%, indicating both improved peak performance and stability across distinct domain shifts.

To assess how sensitive the final performance is to random seed variation in the frameworks discussed above, Table 4 reports the standard deviations (across 15 seeds) of the accuracies for SPADA, ADA-Proto, MADA-Proto, GAN-Proto, CMD, CORAL and NRC-VAT on the four representative source–target domain crosses described above.

Table 4 indicates that SPADA’s performance is weakly sensitive to random seed changes, with standard deviations ranging from ± 0.47 to $\pm 0.62\%$ across the four representative transfers. This low dispersion suggests that the reported accuracies are not driven by a few favourable initialisations or stochastic training effects, but are reproducible across independent runs.

In contrast, a number of baselines show noticeably higher variability in at least one transfer (e.g., CMD up to $\pm 1.11\%$, GAN-Proto up to $\pm 1.06\%$, NRC-VAT up to $\pm 1.05\%$), indicating less stable optimisation or greater sensitivity to stochasticity under the same evaluation protocol. Consistent with this, the manuscript reports final results as means across 15 independent seeds, with all major randomness sources controlled, supporting reliable comparison between methods under identical experimental conditions.

Ablation study

To isolate the contribution of each SPADA component, three ablation variants were evaluated on the Cp0 to Cp40 transfer, using the unsupervised-selected configuration across 15 seeds. Removing adversarial alignment ($\lambda_{\text{adv}} = 0$) reduced accuracy from 94.44 to 63.89%, indicating that domain-level alignment is the strongest contributor to adaptation in this setting. Removing prototype compactness ($\lambda_{\text{proto}} = 0$) diminished accuracy to 70%, showing that preservation of class-level structure also contributes substantially. Removing pseudo-labelling ($\lambda_{\text{tgt}} = 0$) produced 95% accuracy, which is slightly above the full model. This pattern is consistent with the limited target DA set size in CONCEPT, with six observations per class, where entropy-gated pseudo-labels on very small target sets may introduce noise that offsets potential gains. Together with the WTS comparison against CNN and ResNet1D, these ablations indicate that adversarial and prototype terms are essential for this transfer, the WTS backbone provides a major feature-quality advantage, and the effect of pseudo-labelling is data-regime-dependent rather than uniformly beneficial.

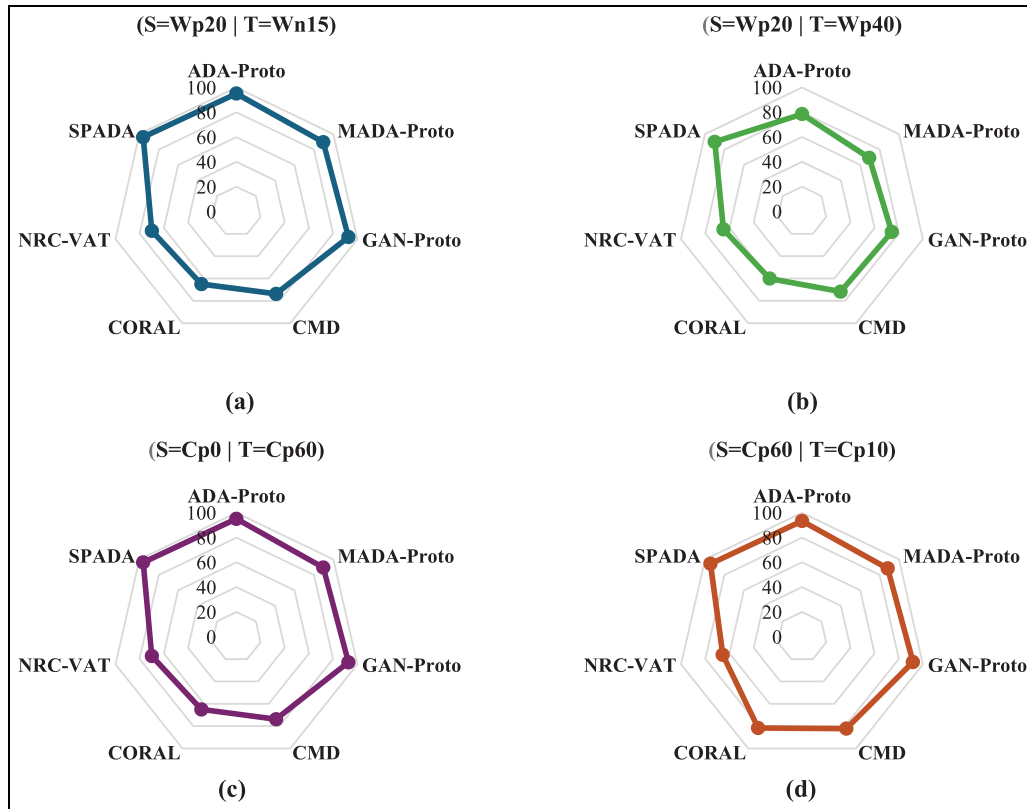


Figure 26. Damage detection results of UDA benchmarks against SPADA for (a) WTB-VibClimate source Wp20 and target Wn15, (b) WTB-VibClimate source Wp20 and target Wp40, (c) CONCEPT source Cp0 and target Cp60 and (d) CONCEPT source Cp60 and target Cp10. UDA: unsupervised domain adaptation; SPADA: scattering-based prototype-aligned domain adaptation; WTB-VibClimate: small-scale wind turbine blade under varying climate conditions; CONCEPT: carbon-epoxy composite plate.

Table 4. Standard deviations of accuracy across 15 seeds for the four representative source-target crosses.

Pipeline	Scenario			
	Wp20-Wn15	Wp20-Wp40	Cp0-Cp60	Cp60-Cp10
SPADA	± 0.51 (%)	± 0.47 (%)	± 0.50 (%)	± 0.62 (%)
ADA-Proto	± 0.94 (%)	± 0.46 (%)	± 0.68 (%)	± 0.52 (%)
MADA-Proto	± 0.29 (%)	± 0.38 (%)	± 0.24 (%)	± 0.52 (%)
GAN-Proto	± 1.06 (%)	± 0.64 (%)	± 0.74 (%)	± 0.67 (%)
CMD	± 1.04 (%)	± 1.11 (%)	± 0.56 (%)	± 0.37 (%)
CORAL	± 0.89 (%)	± 0.90 (%)	± 0.63 (%)	± 0.62 (%)
NRC-VAT	± 0.62 (%)	± 0.45 (%)	± 1.05 (%)	± 0.71 (%)

SPADA: scattering-based prototype-aligned domain adaptation; ADA-Proto: adversarial domain adaptation with prototypes; MADA-Proto: multi-domain adversarial DA with prototype attention; GAN-Proto: generative adversarial network with prototype weighting; CMD: central moment discrepancy; CORAL: correlation alignment; NRC-VAT: neighbour refinement consistency with virtual adversarial training.

Computation efficiency

This study placed particular emphasis on the computational efficiency and practical implementation of the SPADA framework. All training, evaluation and the grid search were performed in Python 3.10 using

Jupyter Notebook on a workstation running Microsoft Windows 11 Pro for Workstations (Build 26200). The experiments were executed on CPU only, using an Intel(R) Xeon(R) Gold 6248R @ 3.00 GHz with 24 cores and 48 logical processors, 192 GB of physical

memory and 250 GB of virtual memory. Configurations were executed in parallel across the available cores. The core software stack comprised PyTorch 2.0.1, Scikit-learn, NumPy and Matplotlib. The wall-clock time per hyperparameter configuration (single seed, including training and unsupervised validation) was 11.88 and 16.14 s for CONCEPT and WTB-VibClimate, respectively. Each configuration was evaluated across 15 seeds, and the full grid search over 1000 sampled configurations took about 102 and 138 min for CONCEPT and WTB-VibClimate, respectively, with parallelisation across available cores. Feature extraction via WTS was cached across configurations sharing the same scattering parameters and constituted a negligible fraction of total computation. Per-sample inference latency at test time (feature extraction plus forward pass) was below 1 ms on the same hardware. From a computational standpoint, the inference stage of SPADA is compatible with online monitoring requirements on standard server-class hardware; the grid search represents an off-line training cost that would not recur during operational deployment.

Conclusion

This work addressed the challenge of deploying damage detection models under changing environmental conditions in SHM by developing an interpretable UDA framework (SPADA). The framework combines WTS for physics-guided feature extraction with prototype-based transfer learning and dedicated interpretability modules, providing clear evidence of how damage-sensitive patterns behave during environmental transitions while maintaining strong classification performance. The key findings are summarised below:

- (1) *Practicality*: WTS is used to obtain deformation-stable representations that suppress selected environmental variations while preserving damage-related modulations in guided wave and vibration signals. These scattering coefficients are processed by a prototype-based adaptation mechanism that maintains explicit correspondence between source and target exemplars, while instance to prototype similarities and low-dimensional visualisations reveal how prototypes move in latent space, how decision regions emerge and where misclassifications tend to concentrate.
- (2) *Effectiveness*: Validation on an experimental composite plate under temperature variation, using temperature-conditioned synthetic

damaged responses generated by the Conv-CVAE, and on the WTB benchmark dataset of vibration signals, demonstrated successful knowledge transfer without requiring labelled target data. Prototype tracking showed that damage classes largely preserved their structural characteristics during temperature-induced transfer, with trajectories and confusion patterns following expected trends. The interpretability diagnostics indicated that performance degradations mainly occurred between neighbouring damage severities and that the class structure of source and target data remained well separated, suggesting that adaptation preserved damage-sensitive information rather than collapsing classes due to environmental effects. These findings are drawn from two controlled laboratory benchmarks, one of which (CONCEPT) employs synthetic damaged signals at non-baseline temperatures generated by the Conv-CVAE. Accordingly, the robustness claims for CONCEPT transfers should be interpreted within the context of this constructed evaluation setting, while the WTB-VibClimate results provide direct experimental support across real temperature conditions.

- (3) *Limitations and future work*: Key limitations warrant acknowledgement. The framework has been examined exclusively under temperature variations and has not yet been tested for other EOVs. The current formulation addresses classification rather than novelty detection, and the interpretability analysis relies on expert inspection rather than quantitative metrics of physical plausibility. The interpretability modules enhance diagnostic transparency but do not provide formal guarantees of physical correctness; they are designed as inspection instruments that support expert judgement rather than as certification mechanisms. Both case studies employ relatively simple structural geometries with regularly spaced sensors under controlled laboratory conditions. In more complex configurations, such as stiffened panels or curved shells, multipath reflections and mode conversions at geometric discontinuities would increase feature variability beyond temperature effects alone. While the WTS stability bound remains valid regardless of signal complexity, the prototype-based alignment may require greater adaptation capacity to accommodate richer feature distributions. Sparse or irregular sensor layouts would further reduce the

discriminability of scattering coefficients for closely spaced damage classes, and the Conv-CVAE fixed-offset synthesis strategy would be less likely to generalise where damage-induced waveform changes depend strongly on actuator-damage-sensor geometry. Future work will extend the framework to multi-source DA, integrate physics-based constraints derived from wave propagation models, and progress from classification to damage localisation and severity estimation within digital twin workflows. Addressing more complex structural geometries, irregular sensor configurations and additional environmental variabilities constitutes a further priority for investigation.



Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research study was carried out in the framework of the project “TU-LEARN—sTrUctural Life Extension enhanced by aRtificial iNtelligence,” funded by Unione europea—Next Generation EU, as part of Prin 2022 PNRR—D.D. n. 1409 del 14-09-2022 program.

ORCID iDs

Nima Rezazadeh  <https://orcid.org/0000-0002-1310-3012>
 Alessandro De Luca  <https://orcid.org/0000-0002-3760-0144>
 Giuseppe Lamanna  <https://orcid.org/0000-0002-6282-2224>
 Fawaz Annaz  <https://orcid.org/0000-0002-9236-9387>
 Mario de Oliveira  <https://orcid.org/0000-0002-3619-3989>

Data availability Statement

The data that support the findings of this study are available upon reasonable request from the corresponding author.

References

- Rezazadeh N, De Luca A, Perfetto D, et al. Domain-adaptive graph attention semi-supervised network for temperature-resilient SHM of composite plates. *Sensors* 2025; 25: 6847.
- Qu K, Logan A, Miller E, et al. Multi-phase adaptive methodology for mitigating environmental and operational variability in slowly changing time-variant engineering structures. *Mech Syst Signal Process* 2025; 229: 112494.
- Kaewniam P, Cao M, Alkayem NF, et al. Recent advances in damage detection of wind turbine blades: a state-of-the-art review. *Renew Sustain Energy Rev* 2022; 167: 112723.
- Cha Y-J, Ali R, Lewis J, et al. Deep learning-based structural health monitoring. *Autom Constr* 2024; 161: 105328.
- Perfetto D, Rezazadeh N, Aversano A, et al. Composite panel damage classification based on guided waves and machine learning: an experimental approach. *Appl Sci* 2023; 13: 10017.
- Azad MM and Kim HS. Noise robust damage detection of laminated composites using multichannel wavelet-enhanced deep learning model. *Eng Struct* 2025; 322: 119192.
- Ojha S, Jangid N, Shelke A, et al. Probabilistic impact localization in composites using wavelet scattering transform and multi-output Gaussian process regression. *Measurement* 2024; 236: 115078.
- Rezazadeh N, De Luca A and Perfetto D. Unbalanced, cracked, and misaligned rotating machines: a comparison between classification procedures throughout the steady-state operation. *J Braz Soc Mech Sci Eng*. Epub ahead of print 2022. DOI: 10.1007/s40430-022-03750-1.
- Ma S, Cheng B, Shang Z, et al. Scattering transform and LSPTSVM based fault diagnosis of rotating machinery. *Mech Syst Signal Process* 2018; 104: 155–170.
- Bull LA, Gardner PA, Gosliga J, et al. Foundations of population-based SHM, part I: homogeneous populations and forms. *Mech Syst Signal Process* 2021; 148: 107141.
- Bull LA, Gardner PA, Dervilis N, et al. On the transfer of damage detectors between structures: an experimental case study. *J Sound Vib* 2021; 501: 116072.
- Wang L, Liu G, Zhang C, et al. FEM simulation-based adversarial domain adaptation for fatigue crack detection using lamb wave. *Sensors* 2023; 23: 1943.
- da Silva S, Yano MO and Gonzalez-Bueno CG. Transfer component analysis for compensation of temperature effects on the impedance-based structural health monitoring. *J Nondestr Eval* 2021; 40: 64.
- Yang J, Gan Z, Wang T, et al. Multi-source dynamic adaptive domain generalization network for crack detection under unknown temperature environment. *Measurement* 2025; 240: 115588.
- Ge L and Sadhu A. Domain adaptation for structural health monitoring via physics-informed and self-attention-enhanced generative adversarial learning. *Mech Syst Signal Process* 2024; 211: 111236.
- Pan C, Shang Z, Tang L, et al. Open-set domain adaptive fault diagnosis based on supervised contrastive learning and a complementary weighted dual adversarial network. *Mech Syst Signal Process* 2025; 222: 111780.
- Fang H, Khodaei ZS and Aliabadi FMH. A novel data-driven K-SVD transferrable baseline method for multi-damage identification for composite fuselage panels. *Mech Syst Signal Process* 2025; 232: 112702.
- Salmanpour MS, Khodaei ZS and Ferri Aliabadi MH. Impact damage localisation with piezoelectric sensors

- under operational and environmental conditions. *Sensors*. Epub ahead of print 2017. DOI: 10.3390/s17051178.
19. Salmanpour MS, Sharif Khodaei Z and Aliabadi MH. Instantaneous baseline damage localization using sensor mapping. *IEEE Sens J*. Epub ahead of print 2017. DOI: 10.1109/JSEN.2016.2629279.
 20. Amer A and Kopsaftopoulos F. Gaussian process regression for active sensing probabilistic structural health monitoring: experimental assessment across multiple damage and loading scenarios. *Struct Health Monit*. Epub ahead of print 2023. DOI: 10.1177/14759217221098715.
 21. Yue N, Khodaei ZS and Aliabadi MH. Damage detection in large composite stiffened panels based on a novel SHM building block philosophy. *Smart Mater Struct*. Epub ahead of print 2021. DOI: 10.1088/1361-665X/abe4b4.
 22. Kang M-S and An Y-K. Explainable artificial intelligence-based flexural rigidity matrix estimation of bridges using spatially distributed inclinometers. *Eng Struct* 2026; 346: 121633.
 23. Kim K and Kim YS. Vibration spectrogram analysis for bearing fault diagnosis based on grad-cam for feature selection and statistical approach. *J Mech Sci Technol* 2024; 38: 5885–5898.
 24. Yan T, Xing X, Xia T, et al. Relation between fault characteristic frequencies and local interpretability Shapley additive explanations for continuous machine health monitoring. *Eng Appl Artif Intell* 2024; 136: 109046.
 25. Hanchate A, Bukkapatnam STS, Lee KH, et al. Explainable AI (XAI)-driven vibration sensing scheme for surface quality monitoring in a smart surface grinding process. *J Manuf Process* 2023; 99: 184–194.
 26. Chen G and Dong G. Temporal logic inference for interpretable fault diagnosis of bearings via sparse and structured neural attention. *ISA Trans* 2025; 158: 256–271.
 27. Li Y, Zhou Z, Sun C, et al. Variational attention-based interpretable transformer network for rotary machine fault diagnosis. *IEEE Trans Neural Netw Learn Syst* 2024; 35: 6180–6193.
 28. Rezazadeh N, Perfetto D, de Oliveira M, et al. A fine-tuning deep learning framework to palliate data distribution shift effects in rotary machine fault detection. *Struct Health Monit*. Epub ahead of print 30 November 2024. DOI: 10.1177/14759217241295951.
 29. Fan H, Liu S, Cao X, et al. A prototype-guided federated learning based fault diagnosis method of mechanical transmission system under label distribution skew. *Neurocomputing* 2025; 656: 131532.
 30. Rezazadeh N, de Oliveira M, Perfetto D, et al. Classification of unbalanced and bowed rotors under uncertainty using wavelet time scattering, LSTM, and SVM. *Appl Sci*. Epub ahead of print 2023. DOI: 10.3390/app13126861.
 31. Mallat S. Group invariant scattering. *Commun Pure Appl Math* 2012; 65: 1331–1398.
 32. Ou Y, Tassis KE, Dertimanis VK, et al. Vibration-based monitoring of a small-scale wind turbine blade under varying climate conditions. Part I: an experimental benchmark. *Struct Control Health Monit*. Epub ahead of print 2021. DOI: 10.1002/stc.2660.
 33. Ferreira L de PS, Teloli R de O, da Silva S, et al. Bayesian calibration for Lamb wave propagation on a composite plate using a machine learning surrogate model. *Mech Syst Signal Process*. Epub ahead of print 2024. DOI: 10.1016/j.ymssp.2023.111011.
 34. Lee U and Shin J. A frequency response function-based structural damage identification method. *Comput Struct* 2002; 80: 117–132.
 35. Chen J, Zhu C, Pu Y, et al. A systematic review of Coda Wave Interferometry technique for evaluating rock behavior properties: from single to multiple perturbations. *Earth Energy Sci* 2025; 1: 180–192.
 36. Senin P. Dynamic time warping algorithm review. *Science* 1979; 2007.
 37. Habermehl S, Schlesinger C and Prill D. Comparison and evaluation of pair distribution functions, using a similarity measure based on cross-correlation functions. *J Appl Crystallogr*. Epub ahead of print 2021. DOI: 10.1107/S1600576721001722.
 38. Łuczak D. Mechanical vibrations analysis in direct drive using CWT with complex Morlet wavelet. *Power Elect Drives* 2023; 8: 65–73.
 39. Fang Y, Chen C, Zhang W, et al. Prototype learning for adversarial domain adaptation. *Pattern Recognit* 2024; 155: 110653.
 40. Huang M, Xie Z, Sun B, et al. Multi-source unsupervised domain adaptation with prototype aggregation. *Mathematics* 2025; 13: 579.
 41. Wang D, Chen Z and Zhao H. Prototype transfer generative adversarial network for unsupervised breast cancer histology image classification. *Biomed Signal Process Control* 2021; 68: 102713.
 42. Li X, Hu Y, Zheng J, et al. Central moment discrepancy based domain adaptation for intelligent bearing fault diagnosis. *Neurocomputing* 2021; 429: 12–24.
 43. Sun B and Saenko K. Deep CORAL: Correlation alignment for deep domain adaptation. In: *Lect Notes Comput*. Epub ahead of print 2016. DOI: 10.1007/978-3-319-49409-8_35.
 44. Yang Z, Feng M, Huang T, et al. Gain from neighbors: boosting model robustness in the wild via adversarial perturbations toward neighboring classes. In: *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, 11–15 June 2025, pp. 25497–25507. Los Alamitos, CA: IEEE.

Notation

ADA-Proto	adversarial domain adaptation with prototypes
CC	cross correlation

CMD	central moment discrepancy	MADA-Proto	multi-domain adversarial domain adaptation with prototype attention
CNNs	convolutional neural networks	NRC-VAT	neighbour refinement consistency with virtual adversarial training
CONCEPT	carbon-epoxy composite plate	PCA	principal component analysis
CORAL	correlation alignment	PZT	lead zirconate titanate
CWI	coda wave interferometry	SHM	structural health monitoring
DA	domain adaptation	PSR	prototype separation ratio
DTW	dynamic time warping	t-SNE	t-distributed stochastic neighbour embedding
EMA	exponential moving average	UDA	unsupervised domain adaptation
EOVs	environmental and operational variabilities	WTB-VibClimate	small-scale wind turbine blade under varying climate conditions
FRFs	frequency-response functions	WTS	wavelet time scattering
GAN-Proto	generative adversarial network with prototype weighting		