Wind Energ. Sci., 11, 4205–4235, 2026
https://doi.org/10.5194/wes-11-4205-2026
© Author(s) 2026. This work is distributed under
the Creative Commons Attribution 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE

Machine learning framework for scour detection across multiple offshore wind farms

Mahmoud Abdelhak[1], Luke J. Prendergast[2], Jacques Tott-Buswell[2], Ramin Ghiasi[1], and Abdollah Malekjafarian[1]

[1]School of Civil Engineering, University College Dublin, Dublin, Ireland
[2]Department of Civil Engineering, University of Nottingham, Nottingham, United Kingdom

Correspondence: Mahmoud Abdelhak (mahmoud.abdelhak@ucdconnect.ie)

Received: 17 April 2026 – Discussion started: 23 April 2026
Revised: 22 September 2026 – Accepted: 26 September 2026 – Published: 11 November 2026

Abstract. This paper proposes a global scour detection framework for monopile foundations across offshore wind farms (OWFs) using vibration data acquired from a single accelerometer mounted on the turbine tower. The proposed framework is designed to address the real-world complexity and heterogeneity of offshore wind systems operating under diverse environmental, geotechnical, and structural conditions across multiple sites. To validate the framework, numerically generated acceleration data are obtained for various offshore wind turbines (OWTs) distributed across multiple OWFs using an integrated OpenFAST-bespoke soil–structure interaction (SSI) model. The generated dataset accounts for a wide array of operating conditions, including varying sea states, soil properties, turbine structural characteristics, and site-specific scour scenarios. For each OWT, acceleration responses are simulated using foundation stiffness values derived from the SSI model, reflecting the dynamic behaviour variations corresponding to site conditions, turbine characteristics and scour state. A multi-source-domain generalization (DG) strategy is then employed, in which a model is trained using a combined dataset containing one representative turbine per farm (referred to as source turbines) and subsequently evaluated on the remaining previously unseen turbines (referred to as target turbines) across geographically disparate wind farms. The results demonstrate that the proposed framework accurately identifies scour states across multiple target turbines without target-domain data during training, outperforming single-source and non-DG approaches. Model performance is strongly influenced by source-turbine selection, while increasing the number of source turbines improves prediction accuracy, particularly for early-stage scour. Furthermore, vibration measurements collected near the tower base substantially improve scour detection performance. The DG methods maintain high accuracy for intermediate and deep scour depths but struggle to detect shallow scour of 1 m.

1 Introduction

Monopile foundations support the majority of installed offshore wind turbines (OWTs); however, they are vulnerable to several issues, including corrosion, fatigue, boat collisions, and scour (Wang et al., 2018). Scour, in particular, occurs due to soil erosion around the foundation caused by wave–current interactions (Duan et al., 2025). It can have severe effects on the structural and dynamic behaviour of monopile foundations, leading to excessive lateral pile-head displacement (Mostafa, 2012) and increased bending moments (Lin et al., 2016). Additionally, scour causes a notable reduction in the natural frequency (Niu et al., 2023; Prendergast et al., 2015) and changes in the tower's mode shapes (Niu et al., 2023). Many studies have investigated scour detection using vibration-based methods (Dai et al., 2021; Malekjafarian et al., 2026; Niu et al., 2023; Prendergast et al., 2018; Weijtjens et al., 2017). These works primarily examined dynamic response characteristics, such as natural frequencies, mode shapes, and damping ratios, and investigated their usage as indicators for scour occurrence. A common finding is that the low probability of detecting shallow scour depths using frequency measurements arises from the difficulty in dis-

tinguishing minor frequency shifts caused by shallow scour from those induced by environmental conditions (Abdelhak et al., 2025). At moderate scour depths, the reduction in the first natural frequency is relatively small, with a 2.2 % decrease reported at $1.5D$ scour for monopile-supported NREL 5 MW OWTs, compared with 3.6 % for the second natural frequency (Rezaei et al., 2018). The reduction is more pronounced in loose sand than in dense sand (Menéndez-Vicente et al., 2023; Prendergast et al., 2015). In heterogeneous, layered soils, natural frequencies may remain similar to those in homogeneous soils; however, the removal of stiff layers can further reduce foundation stiffness and, consequently, the natural frequency (Niu et al., 2022). In addition to changes in the natural frequencies, scour erosion can modify the mode shapes of a monopile-supported OWT by reducing the effective soil–structure stiffness. In this context, Jawalageri et al. (2022) investigated this effect using an integrated OpenFAST–SESAM numerical model incorporating Winkler-type p–y springs and identified the first two mode shapes from acceleration responses at nine locations along the tower using frequency domain decomposition (FDD). Their results showed progressive changes in the lateral displacement distributions of both the first- and second-mode shapes with increasing scour depth for different soil densities. Generally, the second mode is more sensitive to scour than the first mode, particularly in medium-density and dense sand, which indicates that scour can change mode shapes in addition to shifting natural frequencies.

Recent developments in structural health monitoring (SHM) and machine learning (ML), particularly data-driven approaches, show strong potential to improve scour detection methods. In this context, several studies have applied data-driven approaches for scour detection, yielding promising results. For example, Jawalageri et al. (2024) proposed a data-driven approach using a Naïve Bayes model, trained on labelled acceleration data, to classify scour depths around monopile foundations supporting an NREL 5 MW OWT. Extending beyond acceleration data, Jeong et al. (2020) proposed a data fusion approach, which combined both acceleration and angular velocity responses to predict scour occurrence. Similarly, Jatoliya et al. (2024) applied various ML models to predict scour hole depths, while Lu et al. (2024) integrated scour physics and video processing techniques to improve scour detection capabilities. In addition, Najafzadeh et al. (2025) proposed ML approaches to estimate the severity of scour damage in protected monopile foundations. By evaluating five ML algorithms – random forest (RF), support vector machine (SVM), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), and categorical boosting (CatBoost) – they demonstrated that ML techniques can effectively assess damage severity in scour protection systems. Despite their success in detecting scour around single turbines, these approaches primarily rely on supervised learning models, which require access to labelled data – a process that is expensive and time-consuming. Abdelhak et al. (2025) addressed this limitation by developing an unsupervised ML framework based on a one-class support vector machine (OCSVM). By training on healthy-state acceleration data, the detection of scour around OWTs is conducted without requiring labelled scour data. Their results demonstrated strong potential to identify scour around monopile foundations under varying environmental conditions without access to scour-damage labels. Models trained on larger, more diverse datasets that incorporate environmental variability outperformed those developed under limited conditions. Despite these advances, previous studies have mainly focused on scour detection for individual turbines under site-specific conditions, requiring a separate model for each turbine. Additionally, these approaches rely on multiple data sensors distributed along the tower, which increases installation costs and requires comprehensive monitoring systems.

Traditional machine learning (TML) techniques used for scour detection are typically developed and evaluated using datasets collected from the same OWT, where the training and testing data are obtained from a single turbine with specific structural and environmental characteristics. However, applying these models to datasets from other turbines results in a significant performance drop, driven by shifts in the underlying data distribution (Teng et al., 2022). This limitation restricts practical applicability, as offshore wind farms (OWFs) comprise multiple turbines with diverse site-specific conditions and scour scenarios. These variations introduce significant distributional shifts between representative turbine domains, thereby increasing the complexity of large-scale scour detection across OWFs. A transfer learning framework developed by Abdelhak et al. (2026a) trained on a source turbine located at a specific offshore site was able to predict the scour state at target turbines installed at different offshore sites with varying environmental conditions. In addition, a scour-detection-based domain adaptation (DA) approach, proposed by Abdelhak et al. (2026b), shows strong potential for scour detection at the farm level. Primarily, DA addresses domain shift by aligning the distributions of the source- and target-turbine data. Despite being trained exclusively on numerically generated healthy-state acceleration data from a single turbine, the geodesic flow kernel (GFK) technique demonstrates high accuracy in predicting the foundation status (healthy or scoured) of multiple turbines across a hypothetical OWF. However, this technique requires access to the target turbine during the training process, necessitating the installation of a data acquisition system on both the source and target turbines. While practical for a single OWF when the target data are included in training, its applicability across multiple farms is limited when such acceleration data are unavailable. Consequently, data access during model development remains a critical challenge for the practical application of DA-based scour detection methods at the inter-farm level. This challenge is particularly pronounced in OWTs, where acquiring labelled scour data is inherently difficult. In addition, scour develops gradually over time, while underwa-

ter inspections are expensive, infrequently scheduled, and dependent on the availability of specialized inspection vessels and professional divers. Despite the availability of acceleration measurements from source or target turbines, the acquisition of representative labelled datasets remains costly and impractical under many operational conditions, since obtaining reliable damage labels requires direct knowledge of the structural condition of the OWT, which typically can only be established through physical inspection of the turbine. Consequently, only a limited number of turbines are expected to possess reliable labelled datasets, whereas the majority of the turbine fleet is likely to remain unlabelled. Furthermore, modern OWFs frequently comprise turbines operating under various structural configurations and site-specific environmental and geotechnical conditions, resulting in significant variations in the dynamic characteristics between individual turbines. Differences in water depth, soil properties, foundation characteristics, and environmental loading can further alter the structural dynamic response, thereby constraining the generalization capability of conventional data-driven ML models trained on data from a single turbine or site.

To address this challenge, this study proposes a multi-source-domain generalization (DG)-based scour detection framework that learns domain-invariant feature representations consistent across multiple OWFs. This is conducted by transferring the knowledge from a limited number of labelled turbines to previously unseen target turbines without requiring additional target-domain labels. The generalizability of scour detection across multiple wind farms is achieved through a global scour detection framework based on multi-source DG techniques, trained on a combined dataset containing one source turbine per wind farm and designed to detect scour in the remaining, previously unseen target turbines across these farms. The proposed approach relies on vibration measurements acquired from a single accelerometer installed on an OWT tower, with power spectral density (PSD) features derived from the measured acceleration signals serving as the input to the ML model. The proposed framework is implemented using a domain-adversarial neural network (DANN) technique, trained on data from a single reference turbine per farm, assumed to represent the early life stage, and including labelled healthy and early-stage scour states with 1 m scour depth. Despite relying on lightly instrumented source turbines, the framework demonstrates the capability to detect multiple scour levels across geographically varying OWFs. In addition, a comprehensive comparison is conducted between the proposed multi-source DG model, a baseline model without domain generalization, and a model trained using data from a single-source turbine to enable a systematic evaluation of the performance and the generalization capability of the proposed approach. Additionally, the study investigates the sensitivity of model performance to the location of sensor data along the tower and the volume of source data, specifically the number of source turbines. The effectiveness of alternative DG methods, such as causal disentanglement domain generalization (CCDG), is furthermore evaluated to further validate the proposed methodology.

2 Methodology

2.1 Proposed framework

This study proposes a generalized scour detection framework at the inter-wind-farm scale. In this framework, a model trained on labelled data from individual source turbines – each located at a specific OWF (e.g. Farm 1, Farm 2, and Farm 3) – is used to identify the scour state of previously unseen target turbines installed across these farms without requiring scour labels. An overview of the proposed methodology is briefly described in Fig. 1, where source turbines are indicated by red boxes and target turbines by green boxes. The source turbines are assumed to provide early-life acceleration data with healthy and shallow-scour state labels. For each source turbine, acceleration data are numerically generated using an integrated OpenFAST soil–structure interaction (SSI) model developed in MATLAB. The simulated tower acceleration responses are transformed into PSD representations, which characterize the distribution of vibration energy across the frequency spectrum and capture changes in the turbine dynamic behaviour caused by scour. The extracted PSD features are subsequently normalized and organized into batches to facilitate latent feature learning and serve as the input for model training. The proposed framework combines supervised deep learning with domain generalization to learn scour-related representations from labelled source turbines and transfer this knowledge to previously unseen target turbines without requiring target-domain data during training. More details on the DG methods being used are provided in Sect. 2.3 and 2.4, with the DANN and CCDG methods selected for model development. Their methodologies and network architectures are described in Sect. 2.2 and 2.3, respectively. The trained models are first evaluated on unseen target turbines to assess their ability to accurately predict scour conditions under unseen operating scenarios. To further enhance the models' generalization capability across multiple wind farms, various training and testing protocols, outlined in Sect. 2.5, are incorporated. The proposed framework is validated using a training dataset generated numerically, which accounts for variations in turbine structural properties, environmental conditions, soil characteristics, and scour-induced foundation degradation. Details of data generation, including foundation specifications, environmental and operational conditions, the integrated OpenFAST–SSI model, and scour implementation, are provided in Sect. 3.1, 3.2, and 3.3. Data preprocessing and distribution shifts between turbines across different OWFs are described in Sect. 3.4 and 3.5. The proposed methodology's effectiveness is further evaluated by comparing single-source and multi-source-turbine training with a non-DG baseline, assessing the sensitivity to sensor

location along the tower, and investigating the influence of the number of source turbines and the choice of DG model.

2.2 Domain generalization

DG can be implemented through several approaches, including adversarial training, metric learning, meta-learning, and integrated learning techniques, all aimed at extracting domain-invariant features across different operating conditions (Dayal et al., 2023). In the context of fault diagnosis in engineering systems, DG provides strong potential for robust damage detection (Cao et al., 2025). Accordingly, several studies applied DG models to transfer the fault knowledge from multiple source domains to unseen target domains. For instance, Li et al. (2020) proposed a domain generalization method based on adversarial training for fault diagnosis in rotating machinery, while Jia et al. (2023) developed a causality-based approach that decomposes fault features to achieve a more reliable domain-generalized diagnosis.

For a detailed description of the DG model for scour detection at the OWF level, the source turbines can be represented as $D_s = \{D_{s,1}, D_{s,2}, \ldots, D_{s,m}\}$, where m is the number of the source turbines (or domains). Each source-turbine dataset is defined as $D_{s,m} = \{(x_{i,m}, y_{i,m})\}_{i=1}^{N_m}$, where N_m is the number of labelled samples, $x_{i,m} \in X$ denotes the feature vector, and $y_{i,m} \in Y$ demonstrates the corresponding class label. Similarly, a target-turbine dataset can be defined as $D_t = \{(x_{j,t})\}_{j=1}^{N_t}$, where N_t is the number of samples in the target turbine. In this study, while the source turbines have labelled datasets, the target labels y^t are assumed to be unavailable during training. Each farm includes an individual reference turbine, characterized by unique site-specific conditions. Consequently, both marginal and conditional distributions vary across domains such that $P(X_{s,1}) \neq P(X_{s,2}) \neq \ldots \neq P(X_t)$, and $P(Y_{s,1}|X_{s,1}) \neq P(Y_{s,2}|X_{s,2}) \neq \ldots \neq P(Y_t|X_t)$, where $P(X_s)$ denotes the marginal distribution of the source features, and $P(Y_s|X_s)$ represents their conditional distribution. For each source turbine, the feature space X_s and label space Y_s are defined, forming the joint distribution $P(X_s, Y_s)$. Thus, the generalized joint distribution is defined as $P(X_m, Y_m)$. In this study, the source turbines include two class labels (healthy and scoured), with the source dataset containing 1440 samples, divided equally between the two classes. Given datasets from m source turbines, the objective of domain generalization is to learn a predictive function that minimizes the generalization error in unseen target turbines.

2.3 Domain-adversarial neural network (DANN)

DANN mainly incorporates an adversarial training strategy based on a domain discriminator to encourage indistinguishability between features extracted from different domains (Ju et al., 2025). Although DANN is commonly used for domain adaptation, this study adapted it for domain generalization by training the model to extract domain-invariant features. In this framework, the adversarial training process simultaneously optimizes two objectives: a feature extractor, which trains to generate domain-invariant features that enable accurate classification of scour states, and a domain discriminator, which attempts to identify the source domain for each sample. Figure 2 demonstrates the architecture of the DANN technique, consisting of three main components, as described by Ganin and Lempitsky (2015): a feature extractor network, a task-specific classifier, and a domain discriminator. It employs a multi-class domain discriminator that learns to distinguish between source domains based on the feature representations. Meanwhile, the feature extractor is optimized to both support the task-specific classifier and simultaneously confuse the domain discriminator, thereby encouraging the model to learn domain-invariant features. This adversarial interaction is employed through the gradient reversal layer (GRL), which reverses the gradient during backpropagation to the feature extractor. Therefore, the discriminator becomes less effective at determining the source domain of a given sample, while the extractor learns domain-invariant features that are valuable for scour classification.

The DANN model is designed to learn class-discriminative features and domain-invariant representations through joint adversarial training. Table 1 illustrates the architecture of the feature extractor and classifier. The feature extractor is implemented as a one-dimensional convolutional neural network (1D-CNN) that takes PSD features as input and comprises four convolutional layers, each followed by a max-pooling layer. Following the final convolutional block, the extracted feature signals are flattened to produce a 256-dimensional latent feature vector for each sample. These latent representations are then fed into two parallel branches: a label classifier and a domain classifier. The label classifier consists of a fully connected layer that maps the latent features to the two output classes (healthy and scoured) and is optimized using the cross-entropy classification loss to maximize discrimination between scour states in the source domains. In parallel, the adversarial domain classifier uses latent features to predict the source-turbine order (1, 2, or 3). A GRL is implemented between the feature extractor and the domain classifier to enable adversarial training. During optimization, the domain classifier minimizes the domain classification loss, whereas the GRL reverses the corresponding gradients before propagating them to the feature extractor. This encourages the feature extractor to learn latent representations that are simultaneously discriminative for scour detection and invariant to domain-specific characteristics, thereby improving generalization to previously unseen target turbines.

2.4 Conditional contrastive domain generalization (CCDG)

To assess the effectiveness of the proposed methodology and its sensitivity to the choice of DG model, conditional con-

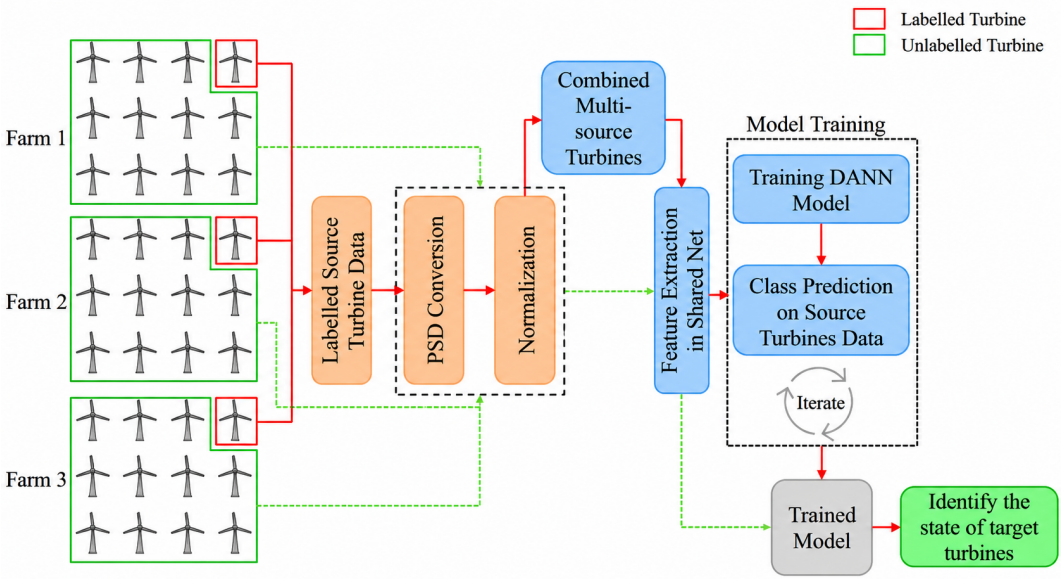


Figure 1. Conceptual schematic of the proposed multi-farm scour detection framework.

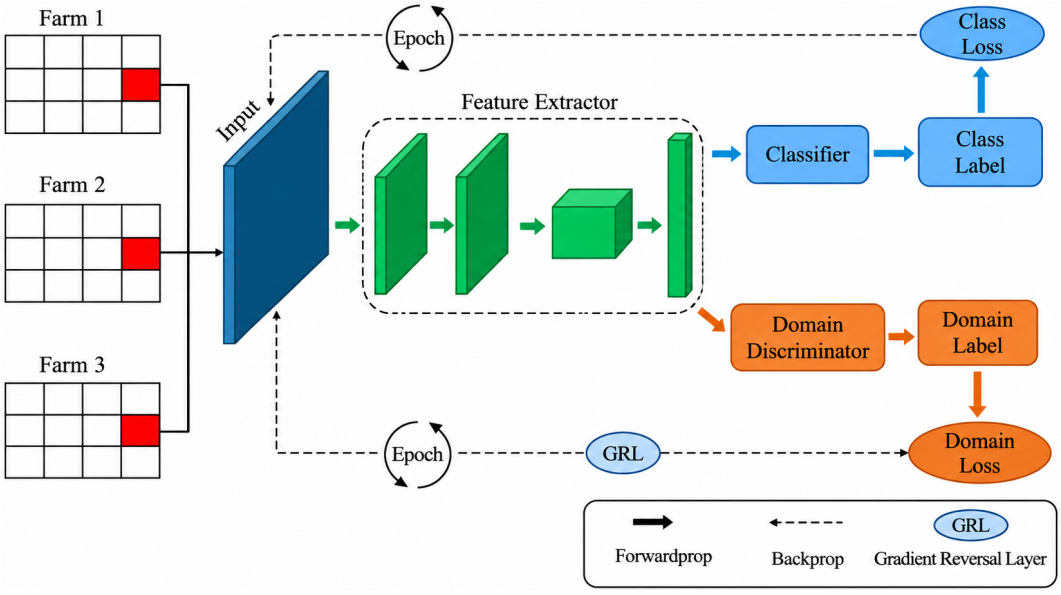


Figure 2. Architecture of the DANN method for DG.

trastive domain generalization (CCDG) is employed alongside DANN. CCDG is a contrastive learning-based approach that aims to improve generalization across domains by incorporating domain information into the learned feature representations (Ragab et al., 2022). Similar to DANN, CCDG reduces the influence of domain-specific variations on the learned representations, thereby improving generalization to unseen domains. However, unlike conventional feature alignment approaches, CCDG incorporates a contrastive learning objective to enhance the extraction of domain-invariant information in the latent feature space. In this study, CCDG is employed as a benchmark DG method to further evaluate the robustness of the proposed framework under varying source- and target-turbine conditions. CCDG enhances the class-level feature alignment by maximizing the mutual information between samples belonging to the same class across different domains while minimizing the mutual information between samples from different classes (Ragab et al., 2022). This enables the model to preserve discriminative features associated with scour conditions while reducing sensitivity to domain-related variations. The CCDG model is trained using a composite objective function consisting of a classification

Table 1. Architecture of feature extractor and classifier.

Module	Layer type	Kernel number	Kernel size	Stride	Activation	Normalization
Feature extractor (CNN)	Convolution	16	64×1	1	ReLU	Batch normalization
	Max-pooling	–	2×1	2	–	–
	Convolution	32	16×1	1	ReLU	Batch normalization
	Max-pooling	–	2×1	2	–	–
	Convolution	64	5×1	1	ReLU	Batch normalization
	Max-pooling	–	2×1	2	–	–
	Convolution	64	5×1	1	ReLU	Batch normalization
	Max-pooling	–	2×1	2	–	–
Classifier	Fully connected	–	Class number (2)	–	Softmax	–

loss and a contrastive loss. The classification loss ensures accurate identification of scour states, whereas the contrastive loss encourages domain-invariant feature learning by reducing the distance between same-class samples and increasing the separation between different-class samples in the latent feature space.

2.5 Training and evaluation protocols

Scour primarily affects the dynamic behaviour of OWTs by changing the foundation stiffness as the seabed soil around the monopile is removed. In time-domain representations, the acceleration signals show negligible impact from scour due to their low-frequency bandwidth. However, changes in the dynamic performance of OWTs due to scour are more pronounced in the frequency domain, making frequency-domain features more suitable for scour detection (Abdelhak et al., 2026b; Prendergast et al., 2016). Accordingly, the logarithmic PSD of acceleration signals is used as input to the DG model, with the analysis limited to the 0–8 Hz range, where the signal contains substantial energy. Since the proposed framework relies on acceleration measurements up to 8 Hz, sensor and data acquisition quality, particularly frequency response and signal-to-noise ratio, are important for practical deployment. Previous work by Abdelhak et al. (2026a) showed that measurement noise has a limited effect on the identified first four natural frequencies located within 0–8 Hz, while the transfer learning model successfully detected scour with noise levels of up to 5 % of the original signal root mean square (RMS). These findings indicate that scour-induced changes remain identifiable under moderate measurement noise. Nevertheless, reliable acceleration measurements across the 0–8 Hz frequency range should be ensured and further validated using field measurements.

In this study, multiple training protocols (similar to tasks in the results section) are designed to evaluate the robustness of the proposed approach across varying acceleration-data-availability scenarios, including training on representative source turbines located at offshore sites with shallow, intermediate, and deep water depths, with low, medium, and high soil densities within each OWF. In Protocol 1, the model is trained using labelled healthy and scoured data from representative source turbines located at shallow-water sites with low soil density (i.e. Turbine 1 from each OWF). Protocol 2 employs source turbines located at sites with intermediate water level and medium soil density (i.e. Turbine 6 across different farms). In Protocol 3, the source turbines are selected from the deepest water sites, characterized by the highest soil density. In addition to these protocols, which use source turbines with consistent site characteristics across all wind farms, an additional protocol is proposed to investigate scenarios in which the source turbines are selected from varying water depths. Moreover, an extended training protocol is introduced to reduce domain mismatch through incorporating data from six representative turbines spanning both shallow- and intermediate-water-depth conditions. This protocol also employs a target normalization scheme, in which each target-turbine dataset is normalized with respect to the closest reference-source turbine to further improve cross-domain generalization. Across different training protocols, a consistent evaluation is applied, where the trained model is systematically evaluated on the remaining previously unseen target turbines within the wind farms, regardless of their location or relative distance from the source turbines. Importantly, each target turbine has unique site specifications and is subjected to different levels of scour, ranging from shallow scour depths ($S = 1$–3 m, $S/D \approx 0.5$), intermediate scour depths ($S = 4$–6,m, $S/D \approx 1.0$), and deep scour depths ($S = 7$–9 m, $S/D \approx 1.5$), where S represents the scour depth, while D represents the monopile diameter. The scour depth is represented as a fraction of the monopile diameter to enable better interpretation of the results and account for variations in turbine and foundation dimensions among different wind farms. Notably, the scour in this study is mainly characterized by scour depth, given its significant impact on the system's dynamic characteristics, while other scour hole parameters, such as width and slope, are neglected due to their minor impact. A simple accuracy metric, defined as the ratio of correctly predicted samples to the total number of target samples, is used to evaluate the model. For example, 50 %

Table 2. NREL 5 MW OWT properties (Jonkman et al., 2009).

Description	Value
Rating	5 MW
Configuration	3 blades
Rotor, hub diameter	126, 3 m
Hub height	90 m
Rated rotor speed	12.1 rpm
Rotor mass	110 000 kg
Nacelle mass	240 000 kg
Tower mass	347 500 kg
Monopile (MP) diameter	6.0 m
MP wall thicknesses	0.060 m (above mudline)
Tower top diameter, wall thickness	3.87, 0.019 m
Tower base diameter, wall thickness	6.0, 0.027 m
Support structure steel density	8500 kg m^{-3}
Steel Young modulus	210 GPa
Steel shear modulus	80.8 GPa

accuracy indicates that the model fails to distinguish between classes and is biased toward a single state (either all healthy or all scoured), whereas 100 % accuracy reflects correct predictions for all samples.

3 Data generation

This study adopts the NREL 5 MW OWT (Jonkman et al., 2009) as the reference model for data generation. A comprehensive overview of the structural parameters is provided in Table 2.

3.1 Foundation specifications across wind farms

This study uses numerically generated vibration datasets representing 3 medium-sized wind farms, each consisting of 11 turbine units. The farms are assumed to be located in sites with dense, medium-dense, and loose sand and are subsequently referred to as Farm 1, Farm 2, and Farm 3, respectively. Each farm has monopile-supported OWTs designed according to site-specific soil characteristics. Turbines within each farm are installed in water depths ranging from 20–25 m, with variations in soil properties to account for local soil heterogeneity and its influence on the structural response. The monopile dimensions are also adjusted between turbine locations to fulfil the required safety and design criteria (Arany et al., 2017). The foundation specifications for Farm 1 are provided in Appendix A (Table A1). Farm 1 comprises monopiles with a consistent diameter of 6 m and a length of 30 m, with minor variations in wall thickness to achieve the design requirements. Farm 2, assumed to be founded in medium-dense sand, is detailed in Table A2. The pile diameter varies from 6.0 to 6.2 m, while the wall thickness remains constant at 90 mm. An embedment depth of 40 m is adopted according to (Arany et al., 2017). The specifications for Farm 3 are presented in Table A3. This farm uses varying pile diameters with a constant wall thickness and an embedment depth of 50 m for all turbines.

3.2 Environmental conditions

Due to variation in installation locations, each wind farm is assumed to experience distinct environmental conditions, which aligns with realistic representations of real-world situations. Each farm is simulated under 12 load cases (LCs) encompassing variations in wind speed, wave height, and wave period. All farms are assumed to be located in the Celtic Sea, where measured environmental data are available from a single buoy (M5, near Wicklow) (Huang and Yang, 2021). To represent farm-specific environmental variability, measurements from different years are assigned to each farm: Farm 1 uses 2022 records, Farm 2 uses 2023 records, and Farm 3 uses 2024 records. To focus on the most frequently occurring environmental conditions and exclude extreme outliers, the selected data are restricted to the 5th–95th percentile range. A linear regression is then performed to establish the relationship between wind speed and wave height. The selected wind speed range represents different turbine operating conditions, while the corresponding sea state records provide the associated wave height and period. Due to the limited availability of current measurements, the DNV-GL current model (DNV, 2007) is used to estimate the sea surface current speed as a function of the wind speed at the 10 m level above the sea level. It should be noted that the water depth at the buoy location may differ from that assumed for the wind farms because of limited availability of environmental measurements at the required depths. The environmental conditions for Farms 1–3 are presented in Tables A4–A6, respectively.

3.3 Numerical acceleration data generation

In this study, numerical acceleration signals are generated using an integrated OpenFAST–SSI model developed in MATLAB. The dynamic response of each OWT is simulated using the aero-hydro-servo-elastic solver OpenFAST for the NREL 5 MW reference-turbine model under combined aerodynamic and hydrodynamic loading. Turbulent wind fields are generated using TurbSim based on the IEC Kaimal turbulence model (IECKAI) (Cheynet et al., 2017), as recommended by the IEC 61400-1 standard for simulating atmospheric turbulence acting on wind turbines. A normal turbulence model (NTM) with IEC turbulence class B and a turbulence intensity of 12 % was adopted to represent typical offshore operating conditions. The mean wind speed at the reference height of 90 m was specified based on the defined environmental conditions in Sect. 3.2, while the vertical wind profile was modelled using the IEC power-law formulation. The resulting stochastic wind field captures the spatial and temporal fluctuations in atmospheric turbulence across the rotor plane, providing realistic aerodynamic loading for

the OpenFAST simulations. The OpenFAST Aerodyne module included the wake effect using blade element momentum (BEM) theory. In addition, the JONSWAP spectrum in OpenFAST represents irregular wave conditions, with the corresponding significant wave height and peak wave period defined according to the environmental conditions considered for each OWF, as indicated in Sect. 3.2. Aerodynamic loads are computed using BEM theory (Zhao et al., 2019), whereas hydrodynamic loads acting on the support structure are evaluated using Morison's equation (Wolfram and Naghipour, 1999). The sea current is incorporated using the normal current model. It is important to note that this study does not take yaw misalignment into account, which enables the analysis to concentrate only on changes in the dynamic response occurring due to the scour effect.

A fixed random seed is used across the simulations to maintain consistent stochastic wind and wave realizations and thereby isolate the influence of scour on the structural response. Each simulation is performed for 3700 s, and the first 100 s are discarded to remove start-up transients associated with model initialization, resulting in 3600 s of response data for subsequent analysis. The simulation implemented a time step of 0.001 s, while the response outputs at every 0.01 s. The relatively long simulation duration is selected to improve the statistical convergence of the response features used for scour detection. While this controlled approach facilitates a direct comparison between different scour conditions, the use of a fixed seed does not explicitly account for variability between independent stochastic realizations. Using a fixed random seed ensures identical turbulent wind realizations are generated, allowing changes in the structural response to be attributed solely to the investigated scour conditions rather than variations in the environmental loading. This assumption is considered appropriate for evaluating the domain generalization capability of the proposed framework, where the objective is to capture variations arising from different turbine and foundation characteristics rather than farm-level aerodynamic interactions. The robustness of the proposed framework to different random seeds is therefore identified as an important aspect for further investigation.

As OpenFAST assumes a rigid foundation and does not explicitly model SSI, a bespoke SSI model is developed in MATLAB and incorporated in OpenFAST by generating equivalent stiffness for zero-dimensional (0D) ground springs. The monopile foundation is represented using a modified Winkler formulation approach (Prendergast and Gavin, 2016), comprising distributed nonlinear lateral p–y springs together with rotational ($m-\theta$) and pile-tip springs to capture the flexibility of the pile–soil system (Tott-Buswell et al., 2024). The nonlinear soil resistance is derived from cone penetration test (CPT)-based geotechnical properties using a modified p–y approach, while the rotational resistance accounts for shaft-friction mobilization generated during pile rotation (Byrne et al., 2015). The nonlinear load–displacement and moment–rotation responses are first evaluated in MATLAB under the mudline loading obtained from an equivalent rigid-foundation OpenFAST simulation. Equivalent lateral and rotational secant stiffnesses are then calculated and implemented in OpenFAST as 0D ground springs at the mudline. Consequently, although the underlying soil model is nonlinear, it is incorporated into OpenFAST through equivalent secant stiffness coefficients implemented as 0D ground springs at the mudline, resulting in a computationally efficient decoupled representation of foundation flexibility. To simulate the scoured conditions, the soil layer equivalent to the scour depth is removed, and the stiffness of the soil springs located within the scoured region is assigned to zero, effectively reducing the pile embedment and updating the foundation stiffness matrix. For simplicity, the springs below the scour depth are assumed to remain unaffected, as no established methodology currently exists to quantify the influence of reduced overburden on multi-spring soil–pile interaction at depth, and this is beyond the scope of this paper. The updated equivalent stiffness matrix is then implemented in OpenFAST to simulate the dynamic response of the turbine under each scour condition. In this study, scour is characterized by scour depth, which represents the primary parameter influencing the loss of foundation stiffness and the resulting changes in turbine dynamic response. This study covers nine scour states, each represented by a specific scour depth, with scour depths ranging from 1 m to 9 m (corresponding to $S = 1, 2, \ldots, 9$ m, with $S/D = 0.167 : 1.5$) considered. The decoupled sub-structuring approach implemented herein has previously been validated against monotonic pile pushover experiments (Tott-Buswell et al., 2024) and has been successfully applied in previous OpenFAST-based scour modelling studies (Abdelhak et al., 2025, 2026a, b).

It should be noted that the variations in soil–structure interaction behaviour arising from different soil properties and scour depths are represented through corresponding modifications of the foundation stiffness. In this study, simplified homogeneous soil profiles are adopted to isolate the influence of global soil stiffness on the dynamic response of OWTs and to evaluate the proposed framework under controlled geotechnical conditions. Although layered soil stratification can influence scour development and foundation dynamics, its effect on modal characteristics is generally smaller than that resulting from substantial changes in overall soil stiffness (e.g. dense versus loose sand) (Zefzouf et al., 2025). Therefore, the proposed framework is expected to remain effective under layered soil conditions, although its performance may be influenced by the degree of soil heterogeneity and layer configuration and therefore requires further investigation. Two sensor locations are considered to generate the acceleration signals within the considered farms: Sensor 1, located at the tower base, 10 m above sea level, and Sensor 2, located at the tower top. Figure 3 presents representative examples of the generated acceleration signals for the healthy and 9 m scour conditions, obtained from Sensors 1 and 2. The

numerical model accounts for the structural damping; however, other types of damping, such as hydrodynamic damping associated with the submerged components, are not considered. A damping ratio of 1 % is adopted for the first and second tower fore–aft and side–side modes in the numerical simulations.

Accelerometers are commonly installed on modern OWTs for structural health and condition monitoring purposes. Therefore, the motivation for the proposed reference-turbine approach is not the unavailability of acceleration measurements, but the limited availability of suitable and accessible labelled data from the target turbine, particularly for different scour conditions. In practice, measurements associated with known scour levels are difficult to obtain because scour development is site-specific, and continuous ground-truth information is rarely available. In contrast, data from healthy or reference turbines can be more readily collected to establish baseline dynamic characteristics. The proposed framework therefore investigates the potential to exploit data from reference turbines and transfer the learned scour-related characteristics to other turbines, reducing the dependence on extensive labelled data from the target turbine. To assess this approach under realistic conditions, the numerical model described above is used to generate datasets that incorporate the variability expected between turbines and offshore sites.

The numerical model described above is used to generate the training and testing datasets under variations in operational and site conditions, thereby representing a range of realistic OWT operating scenarios. Variations in water depth, soil properties, turbine characteristics, environmental conditions, and scour depth are incorporated to capture the heterogeneous conditions encountered across OWFs. Their interaction with SSI and turbine dynamic characteristics exposes the ML model to a broad range of dynamic responses associated with different turbine and foundation configurations. Accordingly, the scenarios and tasks defined in Sect. 2.5 are used to train, evaluate, and validate the capability of the proposed framework to detect scour across multiple OWF environments and operating conditions. Although the dynamic characteristics of OWTs are influenced by various factors – including environmental variability; SSI conditions; and other degradation mechanisms, such as corrosion, marine growth, settlement, and grout deterioration – this study focuses exclusively on scour-induced changes resulting from reductions in foundation stiffness and modifications to boundary conditions. Accordingly, the proposed framework is intended to detect scour-related variations in the turbine's bending modal characteristics rather than serve as a universal diagnostic tool for all degradation mechanisms. Other operational uncertainties, including controller effects, rotor imbalance, sensor installation errors, and measurement noise, are not explicitly considered in the present numerical framework. In addition to these modelling assumptions, the operating regime considered in the simulations also affects the dynamic response and the associated excitation characteristics.

The simulations primarily represent above-rated operating conditions, under which the rotor speed is maintained close to the rated value by the turbine controller. Therefore, the 3P excitation frequency remains relatively consistent among the investigated turbines. However, the proposed scour detection framework does not rely on rotor harmonic components as the primary indicator of scour; instead, it exploits changes in the global dynamic response of the OWT system associated with variations in SSI performance and foundation stiffness. Above-rated operation represents an important condition for OWT monitoring, as turbines frequently operate under these conditions during their service life. Nevertheless, extending the dataset to incorporate below-rated and transient operating conditions, where rotor speed and aerodynamic excitation vary significantly, represents an important direction for future research to evaluate the robustness and applicability of the proposed framework.

3.4 Data preprocessing

The acceleration data generated from the integrated OpenFAST–SSI model have 3600 s per sensor and are sampled at 100 Hz, resulting in 360 000 time instants. Further division of acceleration signals into fixed-width windows of 6000 time instants was conducted, resulting in 60 windows per LC and a total of 720 data windows per turbine across the 12 LCs for each scour state. The PSD of acceleration signals at each window is estimated using Welch's method (Jwo et al., 2021), with a Hamming window of 1000 points and 50 % overlap. The resulting PSDs are first converted to a logarithmic scale and then trimmed to the first 80 frequency bins to capture the low-frequency range where scour has the most pronounced effect. The extracted PSD features are subsequently normalized using the minimum and maximum values of the healthy state of the source turbine and employed to train the DG model. The selected frequency range of 0–8 Hz captures the dominant structural responses of the turbine, including the first four global bending frequency modes affected by variations in foundation stiffness, as well as the dynamic response associated with scour development within the given wind farms. Although the rotor excitation frequency 1P is located at a much lower frequency range, the structural modes within the 0–8 Hz range contain important information regarding changes in the system dynamic behaviour. Therefore, the extracted spectral features within this frequency range provide physically meaningful indicators for identifying scour-induced degradation.

Figure 4 presents the t-distributed stochastic neighbour embedding (t-SNE) visualization (Cieslak et al., 2020) of the healthy and 2 m scoured state datasets for Turbines T1.6, T2.6, and T3.6 in reduced dimension. t-SNE is a nonlinear dimensionality reduction technique that projects high-dimensional data into a low-dimensional space while pre-

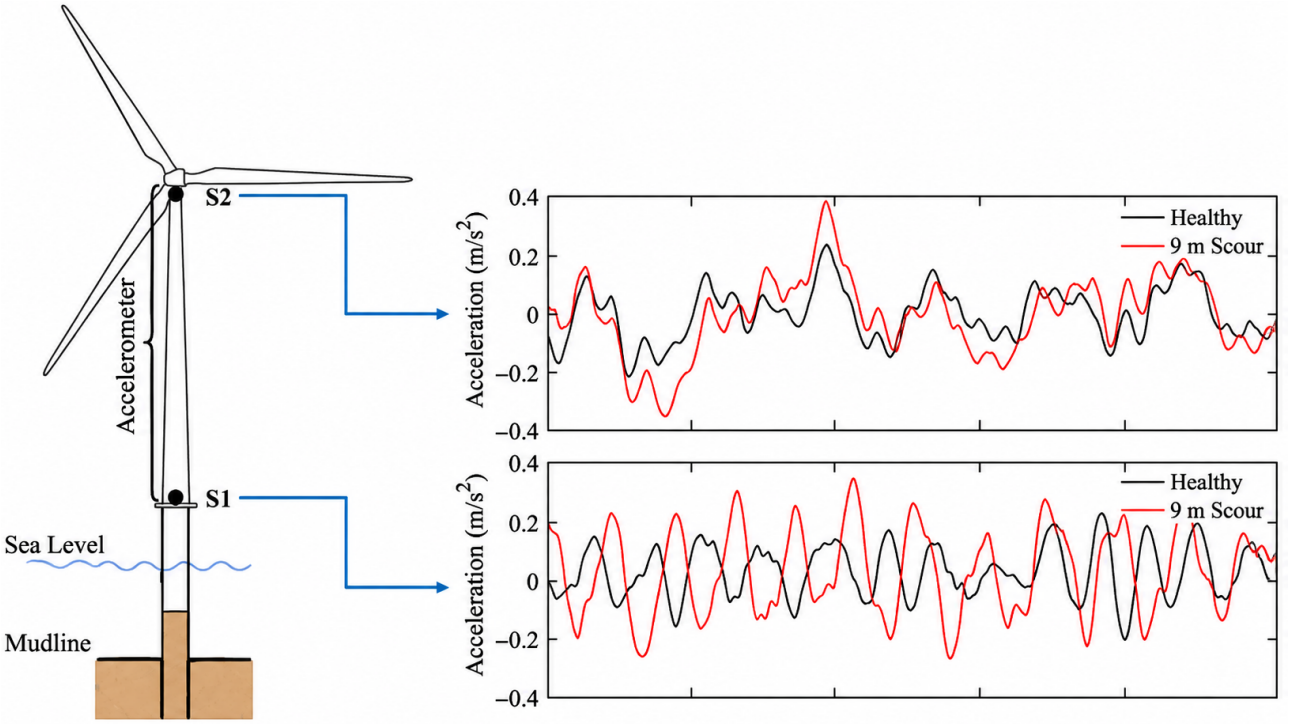


Figure 3. Acceleration response obtained from Sensors 1 and 2 at healthy and 9 m scour states.

serving local similarities between samples. The *t*-SNE visualization shown in Fig. 4 is performed based on the raw input data without the application of the DG model. Therefore, the visualization illustrates the distribution of the reduced-dimensional raw PSD features, not the latent features extracted from the DG model. The results indicate the challenge of distinguishing between healthy and scoured states from the available data, highlighting the system's complexity and the difficulty of scour detection at the inter-farm scale.

3.5 Similarity study between OWT units across farms

A similarity study is conducted between the source and target turbines within each farm using the maximum mean discrepancy (MMD) metric (Ghojogh and Ghodsi, 2024). MMD is a statistical metric primarily used to estimate the difference between two probability distributions, in this case, the PSD feature distributions of the source and target turbines. It evaluates the domain similarity by comparing their mean embeddings in a reproducing kernel Hilbert space (RKHS) (Ghojogh et al., 2021). The MMD distances between the healthy-state data of the reference turbines and the remaining turbines under different scour conditions are presented in Appendix A. Specifically, Table A7 reports the results for Turbine T1.1 in Farm 1, Table A8 presents the corresponding results for Turbine T2.1 in Farm 2, and Table A9 provides the equivalent results for Turbine T3.1 in Farm 3. The results reflect a significant difference between source- and target-turbine data within each farm, primarily due to variations in soil types, water depth, and structural configuration, all of which alter the system's dynamic performance. These factors contribute to considerable overlap between the healthy and scoured states across turbines within a farm. The intersection in the turbines' data distributions across different states demonstrates the complexity of farm-scale scour detection, where each turbine is site-specific and affected by changes in the offshore environment. The MMD distances between Turbine T3.1 and the remaining turbines across all three farms under healthy conditions are presented in Table A10. The results indicate that the inter-farm shift is larger than the intra-farm shift; for example, in Farm 3, the intra-farm shift increases progressively from Turbines T3.1 to T3.11 due to variations in water depth and localized soil changes. Beyond these intra-farm shifts, a clear inter-farm shift is observed even between turbines installed at comparable water depths (e.g. T3.1, T2.1, and T1.1), demonstrating that differences in soil characteristics strongly influence the discrepancy in data distributions. The combined effect of inter-farm and intra-farm shifts highlights a significant challenge in developing a fully generalized scour detection model capable of handling variations across environmental and structural conditions. The influence of scour-induced modal characteristic changes is assessed by investigating the data distribution for OWT in both healthy and scoured states using the MMD. The results show that the MMD distance between healthy and scoured conditions increases with scour depth, reflect-

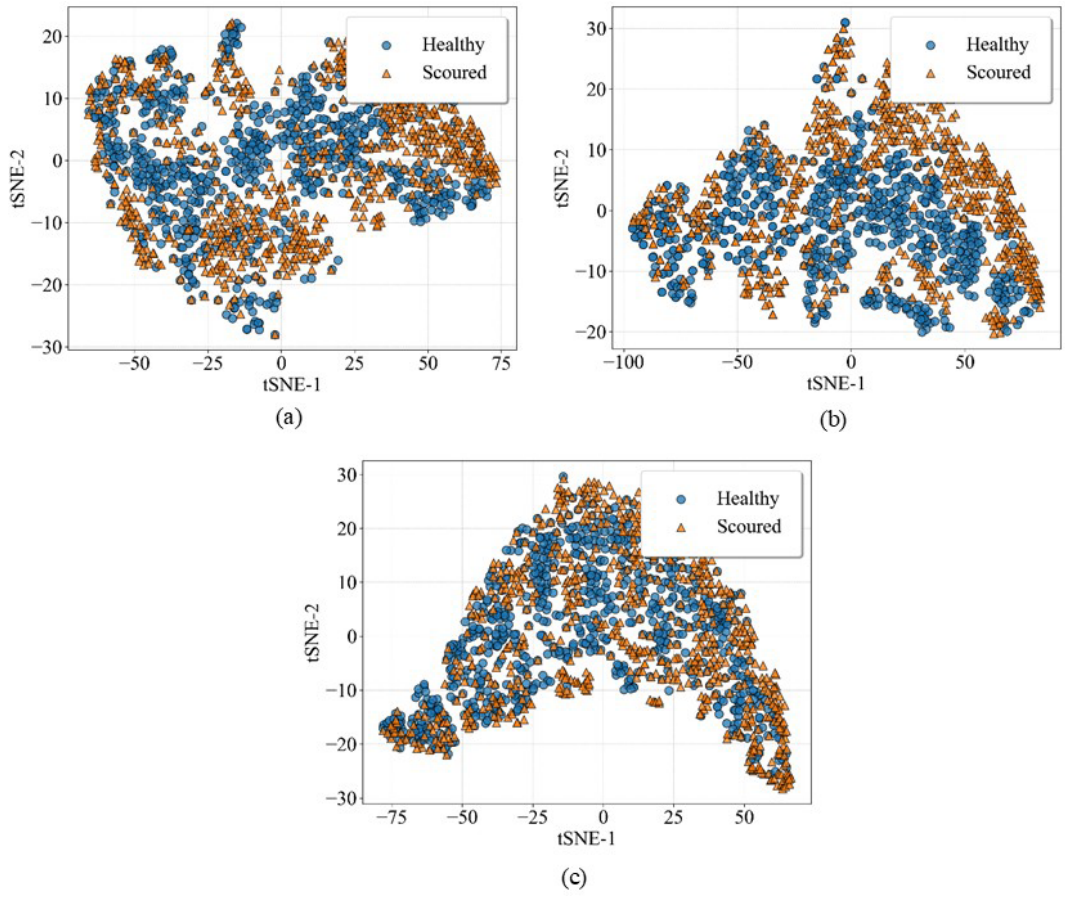


Figure 4. *t*-SNE distribution for raw PSD features for turbines **(a)** T1.6, **(b)** T2.6, and **(c)** T3.6 under 2 m scour depth state.

ing the increasing deviation in the dynamic response caused by foundation stiffness degradation. In contrast, variations in soil density between healthy turbines introduce significant domain shifts, as observed from the increased MMD values between turbines located at different farms. These shifts become more pronounced when variations in soil conditions are combined with changes in water depth, further increasing the discrepancy between source- and target-turbine responses. These results highlight the challenge of scour detection under realistic offshore conditions. Small scour-induced frequency shifts may be difficult to distinguish from normal operational variability and site-to-site differences, whereas larger domain shifts can result in healthy-state responses from one turbine or farm overlapping with scoured-state responses from another. Therefore, effective scour detection requires a DG framework capable of identifying scour-related dynamic changes while reducing sensitivity to non-damage-related variations.

To further investigate the physical influence of scour on the turbine dynamic response, the natural frequencies of the healthy and scoured foundation conditions are estimated from the simulated acceleration responses using the FDD method. Table 3 shows the natural frequencies for the first four bending modes as well as the rate of variation/reduction under different scour states ($S = 3$, 6, and 9 m scour depth; $S/D = 0.5$, 1, and 1.5) based on healthy-state measurements. It represents the natural frequencies for different turbines, characterized by various water depths and soil densities (T1, T6, and T11) across different wind farms. The results demonstrate that increasing scour depth leads to a reduction in the effective foundation stiffness due to the progressive loss of soil support, resulting in a decrease in the tower natural frequencies. The first bending natural frequency shows only a negligible change with increasing scour severity, whereas the second bending frequency decreases from 1.422 Hz under healthy conditions to 1.081 Hz at a scour depth of 9 m, corresponding to a reduction of approximately 24 %. This frequency shift increases consistently with scour severity, confirming that scour-induced degradation of the soil–foundation system directly modifies the global dynamic characteristics of the OWT. To further investigate the extent to which these scour-induced changes vary across different turbine and site conditions, the effect of increasing scour severity on the dynamic response of multiple turbines across different wind farms is subsequently assessed, as shown in Table 3. Soil density has a pronounced effect

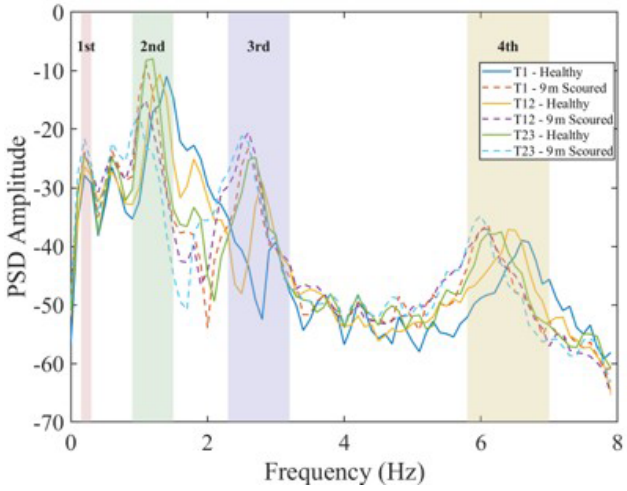


Figure 5. PSD for turbines (T1.1, T2.1, T3.1) in both healthy and scoured states in the 0–8 Hz range.

on natural frequencies. When the soil type is varied while maintaining a constant water depth, the second natural frequency decreases significantly, from 1.422 Hz for T1.1 to 1.244 Hz for T2.1 and 1.148 Hz for T3.1. A similar trend is observed for higher-order natural frequencies, indicating that variations in soil density substantially influence the dynamic characteristics of the OWT compared to water depth. At a constant water depth of 20 m, reducing soil density at Turbines T1.1, T2.1, and T3.1 decreases the lateral foundation stiffness, increasing the flexibility of the soil–foundation system and consequently reducing the natural frequencies. In contrast, greater water depths increase the unsupported tower height and global structural flexibility, with more pronounced effects observed at higher-order modes.

Figure 5 represents the PSD for turbines (T1.1, T2.1, T3.1) in both healthy and scoured states in the 0–8 Hz range, with the frequency ranges corresponding to the first, second, third, and fourth modes highlighted. As shown, the first-mode natural frequency exhibits the smallest variation with changes in scour condition and turbine location. In contrast, higher modes, particularly the second mode, exhibit greater sensitivity to changes in scour condition, along with additional variations associated with turbine location. A similar trend is expected for the third and fourth modes.

4 Results and discussion

In this section, the methodology explained in Sect. 2.1 is implemented and validated using the datasets generated in Sect. 3.3. First, the proposed framework is validated through different training and testing protocols to assess its generalization capability. A sensitivity analysis is then conducted to investigate the effects of DG versus a conventional supervised learning baseline model, as well as using single-source versus multi-source turbines and the impact of sensor locations at the source turbines on the framework's performance. The sensitivity to the number of source turbines and their spatial locations within the wind farm is subsequently assessed. Finally, the influence of the selected DG model on the detection performance is investigated.

4.1 Global scour detection across different farms

For multi-source-domain generalization, the feature sets from the three source turbines are first combined into a unified training dataset. Each sample is assigned two labels: a structural condition label indicating the scour state (healthy or scoured) and a domain label identifying its source turbine. The three source turbines are assigned distinct domain labels (1, 2, and 3), and the corresponding feature vectors, condition labels, and domain labels are concatenated to form the training dataset. The combined data are then passed to a data loader, which organizes the samples into batches of 720. Each batch contains data from all three source domains, thereby providing the model with diverse turbine-specific responses during training and encouraging the learning of both scour-discriminative and domain-invariant features. In each training iteration, the data are processed through the shared feature extractor to obtain latent representations. These representations are subsequently processed by two branches: a classification branch that predicts the structural condition and a domain-discrimination branch that identifies the source turbine. For the DANN framework, the training objective comprises two complementary components: a classification loss and a domain classification loss. The classification loss promotes accurate identification of the scour state from the labelled source-domain data, while the domain classification loss encourages the feature extractor to learn representations that are less dependent on the source domain. Through adversarial training, the feature extractor is optimized to improve scour classification while simultaneously making the extracted representations increasingly difficult for the domain discriminator to distinguish between source domains. This process suppresses domain-specific characteristics associated with differences in turbines, sites, and operating conditions while retaining features relevant to scour detection, resulting in more discriminative and transferable representations. The classifier and domain discriminator are optimized using learning rates of 0.00075 and 0.001, respectively. A gradually adjusted learning rate is applied during training to improve optimization stability.

At each iteration, the classification and domain losses are calculated from the corresponding branches and combined to update the network parameters according to the DANN objective. Model performance is monitored throughout the training process, with classification accuracy evaluated on the source-domain data at each iteration. Early stopping is applied to limit overfitting, with training terminated when the classification accuracy on the source turbines decreases

Table 3. First four bending mode natural frequencies of healthy and scoured turbines across different farms.

Turbine	Mode	Damage condition				Percentage of change		
		Healthy (Hz)	3 m (Hz)	6 m (Hz)	9 m (Hz)	3 m (%)	6 m (%)	9 m (%)
T1.1	1st mode	0.244	0.244	0.244	0.244	0.0	0.0	0.0
	2nd mode	1.422	1.337	1.210	1.081	6.0	14.9	24.0
	3rd mode	2.909	2.844	2.705	2.611	2.2	7.0	10.2
	4th mode	6.595	6.479	6.273	6.116	1.8	4.9	7.3
T1.6	1st mode	0.246	0.246	0.244	0.244	0.0	0.8	0.8
	2nd mode	1.399	1.337	1.210	1.131	4.4	13.5	19.2
	3rd mode	2.992	2.916	2.781	2.659	2.5	7.1	11.1
	4th mode	6.755	6.568	6.411	6.097	2.8	5.1	9.7
T1.11	1st mode	0.244	0.244	0.244	0.244	0.0	0.0	0.0
	2nd mode	1.422	1.348	1.210	1.148	5.2	14.9	19.3
	3rd mode	3.095	3.019	2.844	2.746	2.5	8.1	11.3
	4th mode	6.866	6.758	6.500	6.273	1.6	5.3	8.6
T2.1	1st mode	0.248	0.248	0.244	0.244	0.0	1.6	1.6
	2nd mode	1.244	1.210	1.140	1.100	2.7	8.4	11.6
	3rd mode	2.792	2.732	2.652	2.568	2.1	5.0	8.0
	4th mode	6.416	6.292	6.150	6.039	1.9	4.1	5.9
T2.6	1st mode	0.248	0.248	0.248	0.244	0.0	0.0	1.6
	2nd mode	1.228	1.159	1.064	1.100	5.6	13.4	10.4
	3rd mode	2.791	2.752	2.650	2.568	1.4	5.1	8.0
	4th mode	6.405	6.260	6.098	6.039	2.3	4.8	5.7
T2.11	1st mode	0.244	0.244	0.244	0.244	0.0	0.0	0.0
	2nd mode	1.324	1.210	1.164	1.092	8.6	12.1	17.5
	3rd mode	2.945	2.881	2.801	2.698	2.2	4.9	8.4
	4th mode	6.561	6.518	6.308	6.163	0.7	3.9	6.1
T3.1	1st mode	0.244	0.244	0.244	0.244	0.0	0.0	0.0
	2nd mode	1.148	1.099	1.047	0.968	4.3	8.8	15.7
	3rd mode	2.661	2.617	2.542	2.450	1.7	4.5	7.9
	4th mode	6.212	6.108	6.029	5.972	1.7	2.9	3.9
T3.6	1st mode	0.244	0.244	0.244	0.244	0.0	0.0	0.0
	2nd mode	1.210	1.132	1.074	1.011	6.4	11.2	16.4
	3rd mode	2.737	2.687	2.642	2.542	1.8	3.5	7.1
	4th mode	6.293	6.212	6.044	6.029	1.3	4.0	4.2
T3.11	1st mode	0.248	0.248	0.248	0.235	0.0	0.0	5.2
	2nd mode	1.210	1.159	1.098	1.024	4.2	9.3	15.4
	3rd mode	2.820	2.790	2.714	2.628	1.1	3.8	6.8
	4th mode	6.365	6.283	6.162	6.070	1.3	3.2	4.6

by more than 3 %. In addition, model performance is evaluated and logged on the target-domain data at regular intervals to monitor the generalization behaviour throughout training. During the evaluation phase, the unseen target dataset is normalized using the corresponding reference turbine at the same farm and employed to assess model accuracy at the final iteration. The model's capability for detecting various scour levels at the remaining previously unseen target turbines across the given farms is thoroughly assessed. This study employs multiple training and evaluation protocols, as described in Sect. 2.5, to assess the performance of the proposed framework across various practical scenarios. Table 4 presents the training protocols/tasks used for model development. In Task 1, the model is trained on data from the reference turbine, located in shallow water with the lowest soil density within each farm (T1.1, T2.1, and T3.1). Task 2 and Task 3 extend this by training the model on turbines located at intermediate water depths with medium soil density (T1.6,

T2.6, and T3.6) and in deep water with the highest soil density (T1.11, T2.11, and T3.11), respectively. Additionally, Task 4 simulates challenging scenarios with mixed-source turbines installed at varying water depths (T1.1, T2.11, and T3.6).

Figure 6a–c illustrate the performance of the proposed model in identifying the scour state at target turbines in Farms 1, 2, and 3, respectively. The robustness of the proposed framework is assessed by evaluating its capability to detect the healthy condition, which verifies the model's capacity to distinguish normal operating conditions from scour-induced foundation degradation. The results show that the proposed DG framework accurately identifies healthy foundation conditions while maintaining reliable prediction performance for the different scour depths across previously unseen target turbines. For example, the proposed model effectively detected the healthy state for all target turbines across different wind farms when trained using Task 1. In addition, the results indicate that in Task 1, the model effectively detects scour states with scour depths ($S > 3$ m, with $S/D > 0.5$) around monopile foundations in Farm 1 and Farm 2. Its performance is reduced at Farm 3, particularly for the deep-water units (T3.8–T3.11). Although performance decreases for shallow scour depths ($S = 1$–3 m, with $S/D = 0.167$–0.5), the model reliably detects intermediate scour ($S = 4$–6 m, with $S/D = 0.67$–1.0) and deep scour levels ($S = 7$–9 m, with $S/D > 1$) across the studied farms. Overall, Task 1 demonstrates high robustness in consistently detecting scour depths across various turbines, regardless of their site-specific conditions. The classification performance shown in Fig. 6 represents the prediction accuracy for each scour condition and does not imply a direct linear relationship between scour depth and detection performance. Increasing scour depth leads to greater modification of the SSI system and consequently larger changes in the turbine dynamic response. As scour depth increases, the foundation stiffness decreases, resulting in a more flexible support system. Initially, the structural response is dominated by bending behaviour; however, as scour progresses, the response may gradually transition towards a more rigid-body rotational mode. Consequently, the incremental changes in the dynamic characteristics become smaller with increasing scour depth, potentially making the detection of deeper scour levels more challenging. This trend is also reflected in the diminishing rate of increase in the natural period (or corresponding reduction in natural frequency) as scour depth continues to increase.

Figure 7a–c illustrate the model's performance when trained on Task 2 and evaluated on the remaining target turbines across the given OWFs. As shown, the model achieves satisfactory performance for detecting the healthy state at turbines located in deeper water depth and with higher soil density when trained using Task 2. However, the model fails to detect the healthy state for turbines located in shallower water depth and with lower soil density. This demonstrates that the learned representations are associated with scour-induced changes in structural dynamics rather than turbine-specific characteristics arising from differences in foundation properties, structural configuration, or environmental conditions. In Task 2, the model exhibits low generalization capability across wind farm units, with satisfactory accuracy limited to turbines closely matching the source conditions, particularly in terms of water depth and soil density. A similar trend is observed across all three wind farms, where turbines located at greater water depths than the source turbine are accurately identified. In contrast, the scour detection performance degrades for turbines in shallower water due to the distributional shift in data caused by differences in water depth and soil stiffness. These shifts lead to misclassification, with the model often incorrectly identifying a healthy state as scoured due to similarities between the source turbines' healthy state and the target turbines' scoured state. Despite its limited performance for shallow-water turbines (T1.1–T1.5, T2.1–T2.5, and T3.1–T3.5), the model effectively detects scour in deep-water turbines, achieving accuracies above 90 %.

Figure 8a–c indicate the model accuracy when trained on Task 3 and evaluated on the remaining turbines across all three farms. When trained using Task 3, the model demonstrates satisfactory capability in identifying the healthy state for turbines operating in deeper water depths and characterized by higher soil densities. However, its performance deteriorates for turbines located in shallower water depths and softer soil conditions, where the healthy state is not reliably detected. The results show that model performance is strongly influenced by the source-turbine location within each farm. Notably, scour detection at Farm 1 yields the most consistent results compared with Farms 2 and 3. The model achieves acceptable accuracy for the two turbines closest to the source (T1.9, T1.10); however, its accuracy drops sharply for turbines in shallower water, indicating limited generalization across target turbines operating under varying conditions. This reduction in accuracy is primarily due to differences between the invariant features learned during training and those observed during evaluation. A comparison across the different training tasks further reveals that using shallow-water turbines with lower soil stiffness as source domains consistently outperforms training on intermediate- or deep-water turbines with higher soil stiffness, indicating the importance of carefully selecting source turbines to maximize scour detection performance at the wind farm scale.

Figure 9a–c represent the results for the model trained on Task 4 and assessed across multiple target turbines. As shown, the model demonstrates strong capability in identifying the healthy state for Farm 1; however, its performance deteriorates when applied to Farms 2 and 3. In these cases, satisfactory detection is achieved only for turbines located in deeper waters, while the accuracy decreases significantly for turbines operating in shallower water conditions. The results demonstrate that models trained on combined source data achieve higher scour detection accuracy when

Table 4. Training task specifications used for the model development.

Task	Source turbine	Target turbine	Task description
Task 1	{T1.1, T2.1, T3.1}	Corresponding to the remaining turbines across the three farms, including Farm 1 {T1.1–T1.11}, Farm 2 {T2.1–T2.11}, and Farm 3 {T3.1–T3.11}	Evaluation of cross-farm generalization using shallow-water turbines with relatively low foundation stiffness as source domains
Task 2	{T1.6, T2.6, T3.6}		Evaluation using intermediate-water turbines and intermediate-foundation-stiffness conditions
Task 3	{T1.11, T2.11, T3.11}		Evaluation using deep-water turbines with relatively high-foundation-stiffness conditions
Task 4	{T1.1, T2.11, T3.6}		Evaluation using heterogeneous source turbines representing diverse foundation and environmental conditions

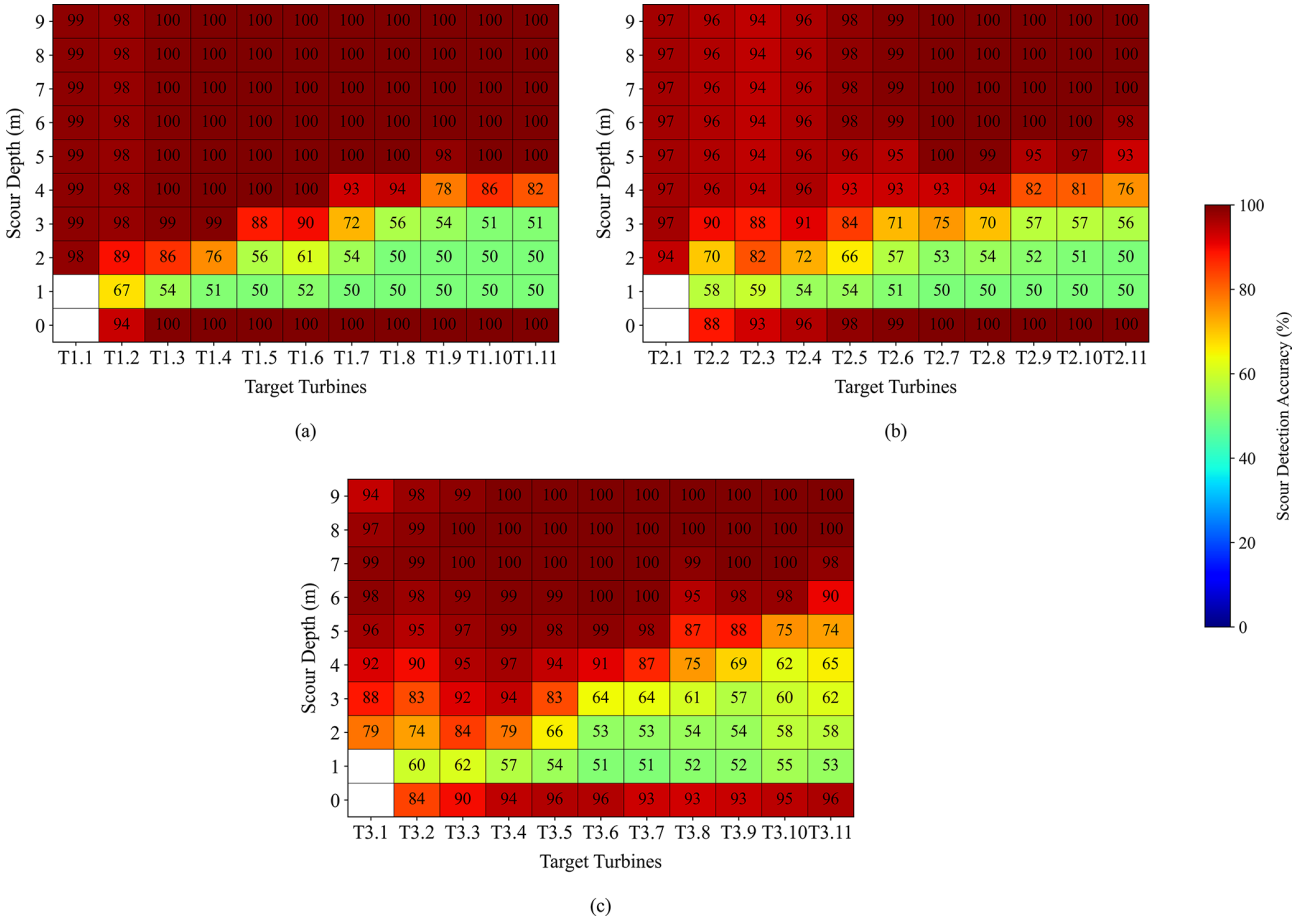


Figure 6. The proposed model performance when trained on Task 1 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

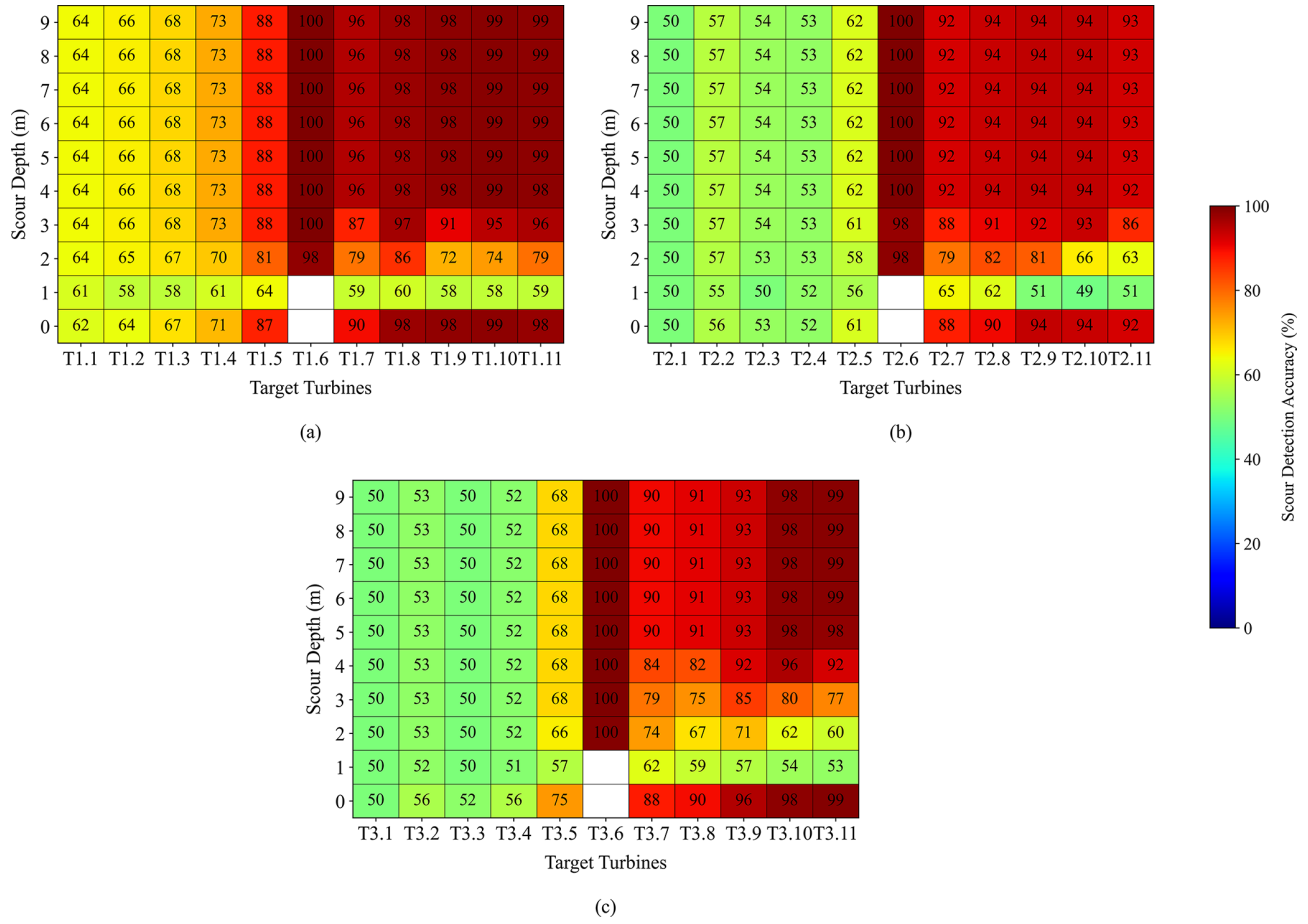


Figure 7. The proposed model performance when trained on Task 2 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

shallow-water turbines with lower soil density are employed as source domains. In contrast, performance decreases when intermediate- or deep-water turbines with increased soil density are selected. This reduction in accuracy is primarily attributed to the model's limited capability to identify the healthy state, which is misclassified as scour due to distributional shifts. Overall, the model reliably detected scour for turbines installed at greater water depths, particularly those with higher soil stiffness than the training conditions. Consequently, the proposed framework exhibits improved generalization capability for practical deployment across heterogeneous OWFs, where both healthy and scoured foundation conditions are expected to coexist. Despite its enhanced performance in detecting the healthy state when only healthy data exist at the unseen turbine, its performance is reduced sharply for detecting 1–2 m scour depth because the reduced impact on the dynamic performance leads to small changes due to scour, which are masked by operational and environmental variation.

To assess the benefit of applying DG strategies to generalize scour detection at the inter-farm level, a conventional supervised learning baseline model developed using the empirical risk minimization (ERM) method is also implemented. The ERM model employs the same network architecture, training data, and optimization strategy as the proposed DANN framework, with the only difference being the absence of the domain-adversarial component. Consequently, the ERM model learns solely by minimizing the classification loss on the source-domain data without enforcing domain-invariant feature representations. Comparing the ERM baseline with DANN enables the assessment of the performance improvement obtained due to applying DG, along with its isolation from the influence of the source-turbine selection. Any performance difference between the ERM baseline and the DG approaches can be directly attributed to the ability of DG methods to learn feature representations that are robust to the distribution shifts between source and target turbines. The capability of the DANN-based scour detection method is further assessed to mitigate the domain shift arising from variations in site-specific conditions, turbine characteristics, and environmental conditions across different OWFs. Figures A1–A4 demonstrate the model perfor-

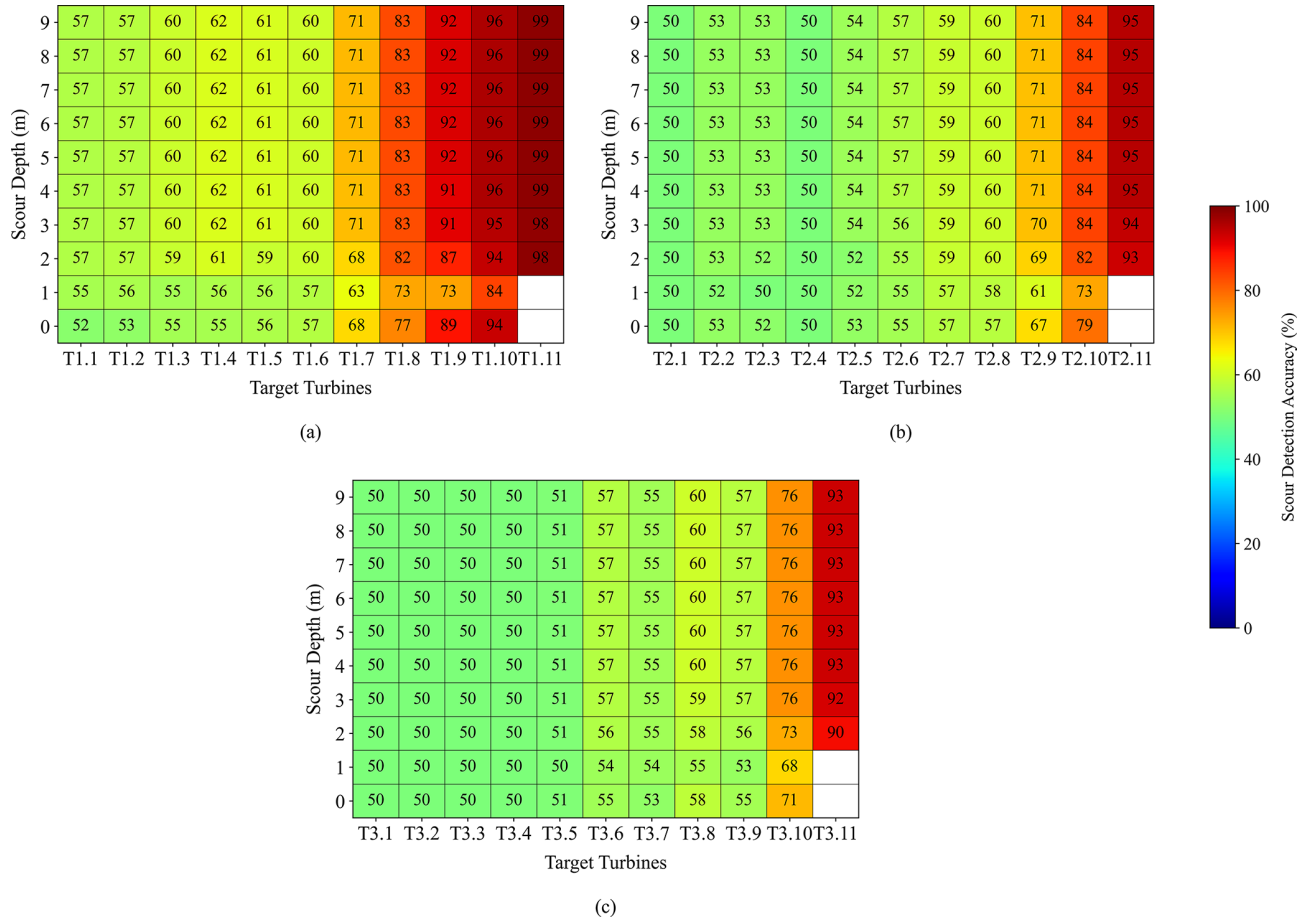


Figure 8. The proposed model performance when trained on Task 3 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

mance for detecting scour using different training and testing tasks (Tasks 1–4) and are presented in Appendix A. The results show that low to moderate gains in performance are obtained due to the application of DG. For example, the application of DG leads to an enhancement in scour detection at shallow-water turbines (T1.1, T1.2, T2.1, T2.2, T3.1, and T3.2) across different wind farms, reaching a 23 % enhancement for detecting 6 m scour at T3.1. Under Task 1 training, the application of DG shows the highest impact in detecting scour at Farm 3. Similar improvement is obtained for training Tasks 3 and 4. Despite enhancing the detectability for shallow scour depths at deep Turbines T6–T11 across different turbines at Task 2, its application induced an adverse effect at shallow-water turbines. The observed improvement is not solely attributable to the selection of the source turbines, but also to the domain-invariant feature learning enabled by the adversarial training process.

To further evaluate the effectiveness of the proposed DG-based training strategy, an additional single-source baseline is considered. In this case, the model is trained exclusively using data from a single-source turbine, e.g. Turbine T1.1, which represents the shallowest water depth and lowest foundation stiffness among the investigated OWTs within Farm 1. Figure 10 represents the performance of the single-source DG model, developed using the PSD features extracted from the acceleration responses of Turbine T1.1 under healthy and 1 m scour conditions. The model is evaluated for scour-state detection across multiple target turbines located in Farms 1, 2, and 3. Compared with the proposed DG-based framework using the same unseen target turbines, the single-source model achieves satisfactory results for detecting the scour state at the same farm (Farm 1). However, in terms of its performance in detecting scour at Farm 2, its capability to detect scour deteriorates significantly at other farms (Farms 2 and 3), failing to detect any scour scenario at any target turbines within these farms. Although the single-source model achieves satisfactory performance for turbines with similar dynamic characteristics, its prediction accuracy decreases as the discrepancy between the source and target turbines increases. In contrast, the DG-based framework consistently demonstrates improved prediction accuracy and robustness across the diverse target turbines. These findings indicate

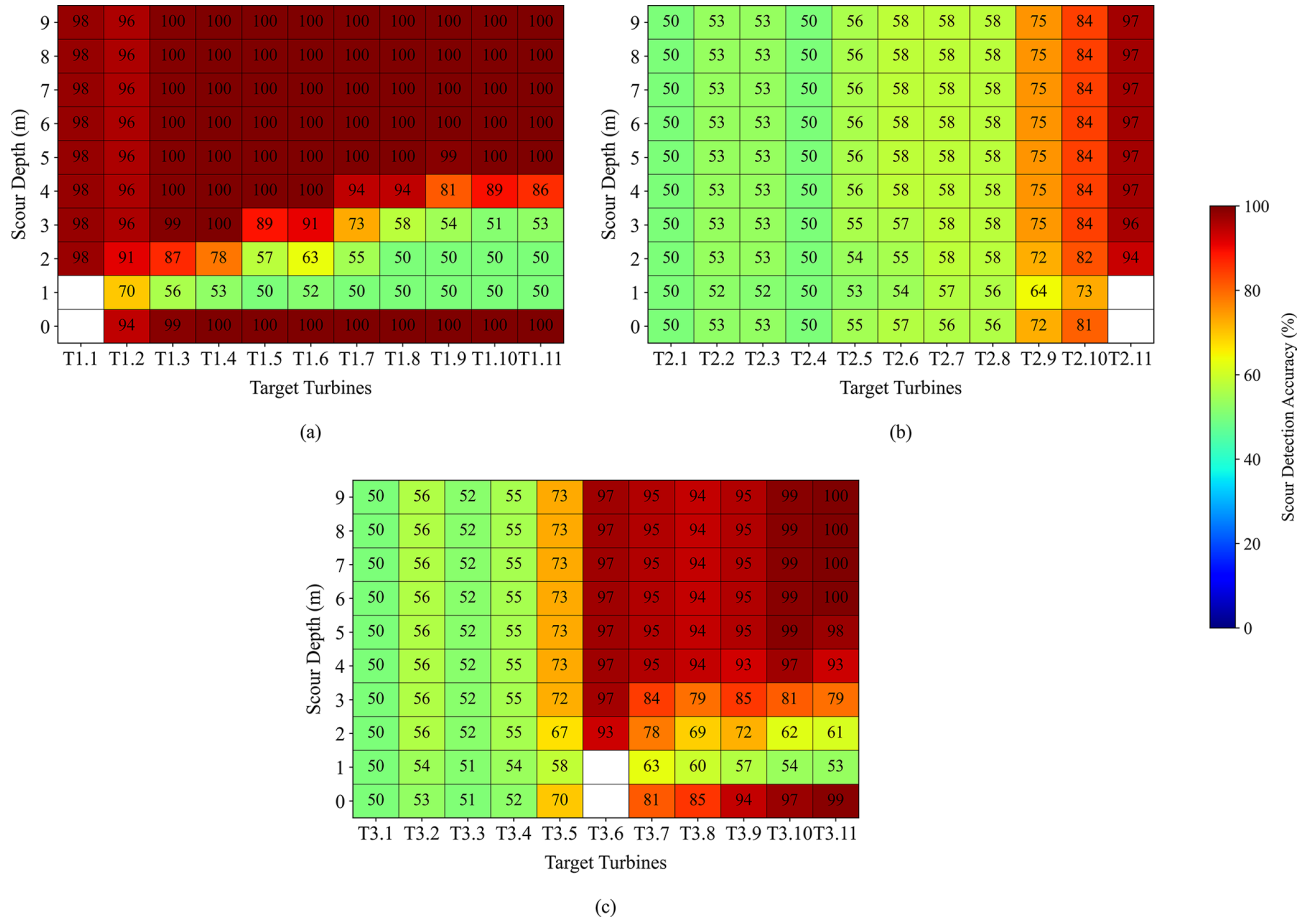


Figure 9. The proposed model performance when trained on Task 4 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

that the enhanced performance of the proposed framework is primarily attributed to the increased diversity of the source-domain data, which enables the model to learn more generalized scour-related features, rather than relying on the characteristics of a single favourable source turbine.

4.2 Model sensitivity to tower sensor positioning

To enhance the applicability of the proposed framework and evaluate the effectiveness of acceleration sensors at different locations, this section investigates the impact of sensor placement on model performance. Specifically, the model is re-evaluated using acceleration data collected at the tower top rather than at the sensor located at the tower base. Consistent with earlier analyses, the model is trained on Task 1, with the architecture and hyperparameters kept constant. Table 5 represents the impact of sensor location on model performance, measured as the average scour detection accuracy across all target turbines in all given farms (T1.2–T1.11, T2.1–T2.11, and T3.1–T3.11) over a range of scour depths. However, the performance may vary between individual OWTs; the average accuracy reflects the overall trend. The results show that, regardless of sensor location, the model performs well for larger scour depths ($S > 6$ m, with $S/D > 1$), demonstrating its strong potential for farm-level scour detection. However, the model trained using data collected at the tower base achieves higher accuracy for shallow and intermediate scour depths, particularly for in-range $S = 3$–6 m, with $S/D = 0.5$–1, highlighting that the tower base is a more effective sensor location. Primarily, the base-mounted sensor provides more stable performance across farms, with lower variability in detection accuracy than the tower-top sensor, particularly for turbines farther from the source conditions. Regardless of the sensor location, the model struggles to detect very shallow scour depths ($S = 1$–2 m, with $S/D =$ 0.167-0.33), likely due to the minimal changes in acceleration response at early scour stages, which exhibit small scour depths compared to the pile length ($L = 30$, 40, and 50 m for Farms 1, 2, and 3, respectively). The superior scour detection performance achieved using the tower base accelerometer highlights the importance of sensor location in capturing scour-induced dynamic variations. The foundation–soil stiffness degradation caused by scour primarily affects the

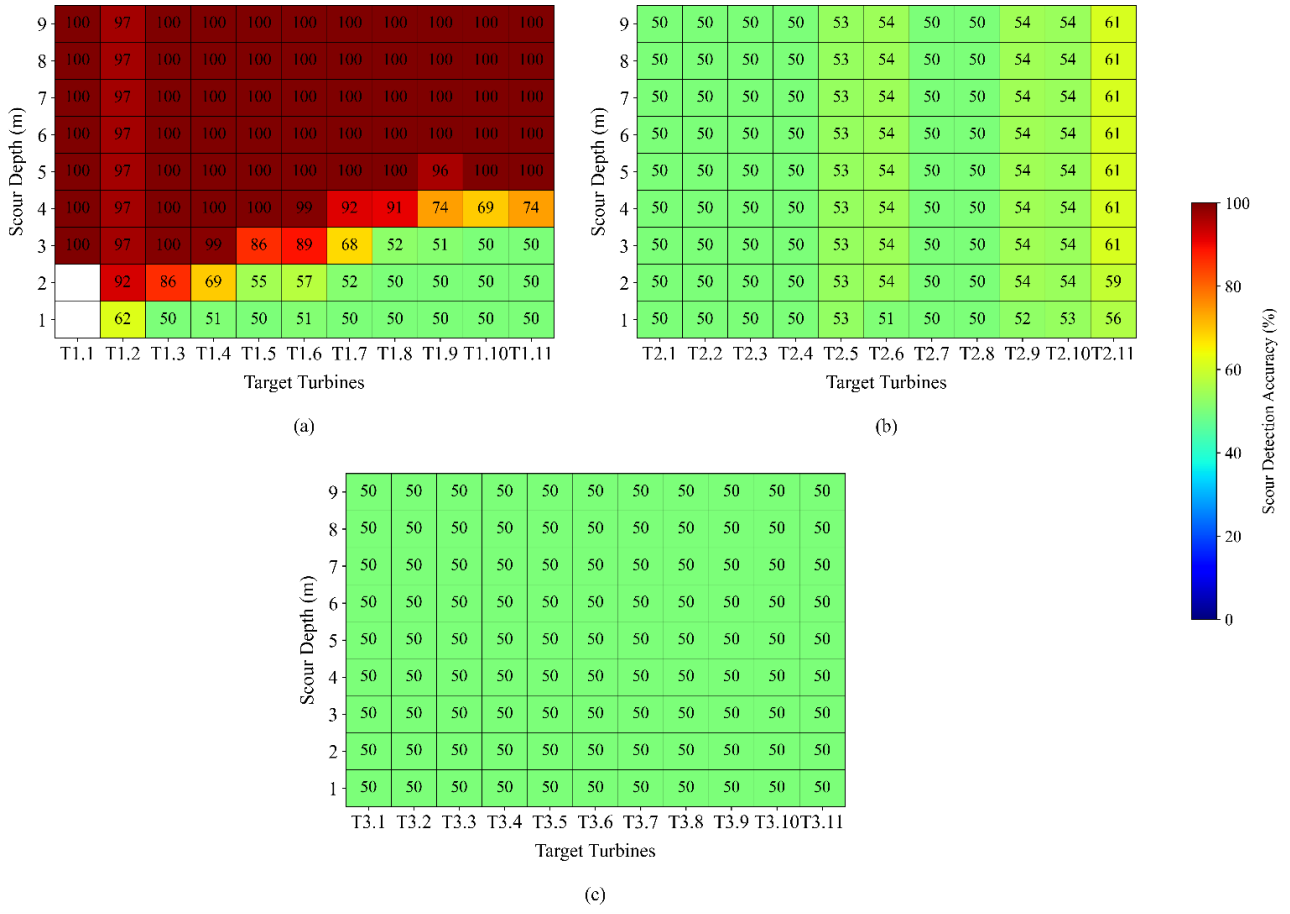


Figure 10. The performance of a single model when trained on Turbine T1.1 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

SSI system and consequently modifies the global vibration characteristics of the turbine, particularly the second mode, which involves significant deformation near the tower base and foundation region. The second structural mode is expected to exhibit increased sensitivity to foundation stiffness variations because its deformation pattern includes a greater contribution from the lower tower and foundation response. Measurements obtained near the tower base are therefore more representative of changes in the foundation dynamic behaviour and provide enhanced sensitivity to scour-induced variations. Conversely, tower top measurements are influenced by the response of the entire turbine structure and may include additional contributions from aerodynamic loading and higher-order structural effects, reducing the relative contribution of foundation-related changes. These observations demonstrate that appropriate sensor placement can significantly enhance vibration-based scour detection performance by improving the sensitivity to the structural modes affected by foundation degradation.

4.3 Sensitivity of model performance to the number of source turbines

This section examines the influence of source data volume and diversity on model accuracy and introduces an alternative approach to improve the prediction performance for turbines located far from the source turbines. The effect of increasing the number of source turbines used for training the DG model is thoroughly investigated by expanding the source set from three to six turbines (T1.1, T1.6, T2.1, T2.6, T3.1, and T3.6). In this section, each target turbine is normalized with respect to its closest source turbine. Figure 11 presents the impact of the number of source turbines on model performance. The average accuracy across all turbines in the given OWF is reported at different scour depths and compared with the results from Sect. 4.2, where only three source turbines with sensors located at the tower base were employed. The results show that increasing the diversity of the source dataset significantly enhances model accuracy, particularly for shallow scour depths ($S = 2$–4 m), with limited impact at deeper scour depths. These findings highlight the importance of source data volume and diversity, suggest-

Table 5. The impact of sensor location at the source turbines on the scour detection accuracy at the target turbines across the farms.

Farm	Sensor location		Average scour detection accuracy across the farm under different scour states (%)									
			Healthy	1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m
Farm 1	Sensor 1	Avg.	99.3	56.4	65.6	77.9	93.4	99.5	99.6	99.6	99.6	99.6
		SD	±1.9	±14.0	±18.3	±21.3	±8.00	±0.80	±0.80	±0.80	±0.80	±0.80
	Sensor 2	Avg.	99.3	54.9	61.7	72.2	82.5	95.9	98.9	99.0	99.0	99.0
		SD	±1.7	±11.0	±16.9	±21.2	±19.4	±4.40	±2.20	±2.30	±2.30	±2.30
Farm 2	Sensor 1	Avg.	97.4	56.2	63.7	76.2	90.4	96.2	97.9	98.2	98.2	98.2
		SD	±3.8	±12.2	±14.6	±14.8	±7.20	±2.00	±2.10	±2.20	±2.20	±2.20
	Sensor 2	Avg.	97.4	54.2	60.9	68.2	76.9	88.4	96.3	98.0	98.3	98.3
		SD	±3.8	±7.90	±14.9	±19.6	±18.7	±13.1	±3.70	±2.80	±2.80	±2.80
Farm 3	Sensor 1	Avg.	92.9	56.3	64.6	73.4	83.4	91.6	97.6	99.5	99.6	99.1
		SD	±3.5	±6.80	±12.2	±14.5	±13.0	±9.20	±2.70	±0.60	±0.90	±1.90
	Sensor 2	Avg.	98.5	53.2	59.4	67.4	76.1	84.4	91.6	96.4	97.8	98.1
		SD	±1.5	±5.00	±11.7	±17.9	±18.0	±16.3	±9.80	±3.60	±2.80	±2.90

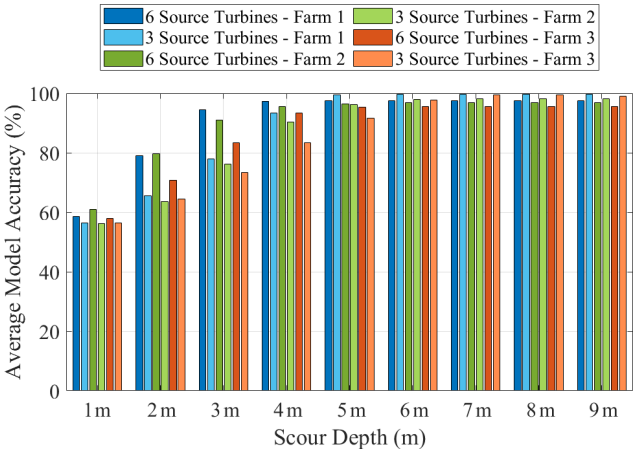


Figure 11. The impact of the number of source turbines on the model performance.

ing that increasing access to a wide set of source turbines enhances overall scour detection performance.

4.4 Influence of the DG method on model behaviour

This section evaluates the impact of the chosen DG method on overall model performance by comparing the CCDG approach with the proposed DANN framework for scour detection. The CCDG method is designed to learn domain-invariant features that generalize to previously unseen target domains by combining classification loss with conditional contrastive loss. This formulation encourages similarity among features of the same class across different domains while minimizing similarity between different classes. The training strategy for CCDG combines supervised scour classification with a contrastive learning objective to improve the robustness of the extracted feature representations. The overall training loss consists of two complementary components: the classification loss and the contrastive loss. The classification loss is calculated using the negative log-likelihood loss and guides the model to correctly predict the scour state from the labelled source-domain samples to ensure the learned features retain the information required for accurate scour identification. The contrastive loss is applied to the latent feature representations extracted by the shared network. By using domain labels, the contrastive objective encourages samples from different domains to learn more consistent feature representations while reducing domain-specific variations. The two losses are combined to form the total training objective, where the classification loss promotes discriminative features for scour detection, and the contrastive loss improves feature alignment across domains. During each iteration, the model parameters are updated by minimizing this combined loss, resulting in feature representations that are both sensitive to scour-related changes and more robust against variations in turbine characteristics, environmental conditions, and operational scenarios. This approach enhances the generalization capability of the model when applied to previously unseen turbines or offshore wind farm conditions. For the feature extractor, a 1D CNN is employed, using an architecture consistent with that of the DANN model. To ensure a fair comparison, the source turbines, network architecture, and hyperparameters are kept identical for both methods, and the average performance of the CCDG vs. DANN method for scour detection at Farms 1, 2, and 3 is compared as shown in Fig. 12. The results reveal that both DG models achieve high detection accuracy, with only minor differences in performance. At Farm 1, both models exhibit comparable results; however, CCDG has better accuracy in detecting shallow scour depths

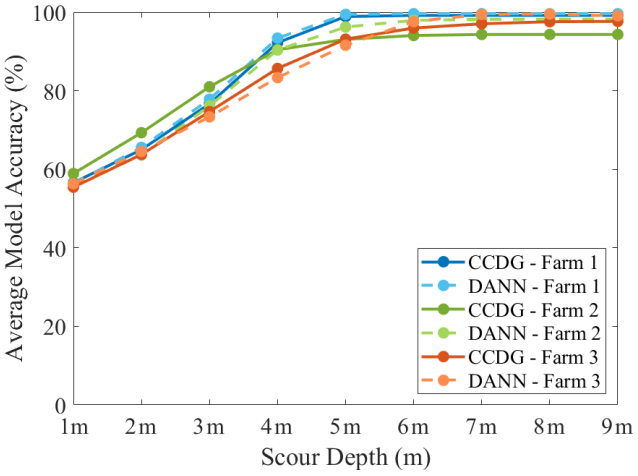


Figure 12. The impact of the selected DG method on the model performance.

at Farms 2 and 3, whereas DANN performs better in detecting larger scour depths.

The present study is limited to hypothetical wind farms, in which site conditions are assumed to vary uniformly across turbines within a specified range of scour depths. It is further assumed that the soil density changes with water depth, which is restricted to a range of 20–25 m for the considered wind farms. Due to data availability constraints, the validation of the proposed framework is based on numerically generated datasets obtained using a coupled OpenFAST–SSI model; therefore, variations in model performance may occur when the proposed approach is applied to real-world measurements. Additionally, the analysis is limited to the predefined operational and environmental conditions using an NREL 5 MW turbine model. These conditions replicate typical operational conditions, with sea currents modelled theoretically rather than measured. Therefore, further changes in model performance may be anticipated when the model is applied to field data. Although the proposed framework demonstrates effective cross-farm scour detection using labelled data from representative source turbines, future research should investigate learning strategies that further reduce the reliance on labelled data. In particular, self-supervised and semi-supervised learning approaches offer considerable potential for exploiting the large volumes of unlabelled operational monitoring data routinely collected from OWTs. Furthermore, extending the proposed framework to heterogeneous offshore wind fleets comprising different turbine ratings, manufacturers, support structures, and foundation types represents an important step towards scalable fleet-wide structural health monitoring and practical industrial deployment.

5 Conclusion

This study proposes a general framework for scour detection at the inter-offshore-wind-farm level. The framework primarily relies on the domain generalization model, developed using early installation data from a single sensor on an individual reference turbine at each farm. Training on the combined source dataset, obtained from three farms, evaluates the model's capacity for scour detection at the remaining previously unseen target turbines, yielding the following results.

- The model trained on source turbines located in shallow water with lower soil densities demonstrates strong potential for generalized scour detection at different target turbines with higher soil densities across the three farms. However, models trained on intermediate- or deep-water turbines with higher soil densities show reduced generalization capability.

- The model excels for scour detection across farms when developed on datasets obtained from the shallowest and lowest-stiffness foundations. However, training the model on scenarios with increased water depth and higher-stiffness foundations resulted in reduced performance.

- The proposed multi-source DG model outperforms both the baseline model without DG and the single-source DG model in detecting scour across different turbines located at multiple wind farms.

- Models developed using tower base sensor data outperform those based on a tower top sensor, highlighting that the tower bottom is the optimal sensor placement for effective scour detection.

- Increasing the number of source turbines has a notable impact on enhancing the model performance, specifically at shallow scour depths.

- Both DANN and CCDG domain generalization methods maintain high performance for generalized scour detection, successfully detecting intermediate and deep scour depths but failing to detect shallow scour depths of 1 m.

This study is limited to 3 hypothetical wind farms, each comprising 11 OWT units. The effect of increasing the number of farms and unit locations on the model performance necessitates further investigation. The robustness of the proposed framework to different random initializations also requires further investigation. In addition, considering other degradation mechanisms or operational uncertainties combined with scour, such as controller effects, rotor imbalance, sensor installation errors, and measurement noise, necessitates further assessment. Furthermore, the applicability of the framework under below-rated and transient conditions, where rotor

speed and aerodynamic excitation vary more significantly, remains to be investigated. Additionally, the impact of incorporating alternative feature types on model performance requires further assessment. Future work should also explore the use of other domain generalization methods to mitigate data shifts, as well as the potential of training and testing across different farms to better evaluate the model's generalization capacity.

Appendix A

Table A1. Key specifications of OWT foundations at Farm 1.

Unit name	Water depth (m)	Soil relative density	Soil unit weight (kN m^{-3})	Pile diameter (m)	Wall thickness (mm)	Embedment depth (m)
T1.1	20.0	0.70	19.0	6	85	30
T1.2	20.5	0.72	19.1	6	85	30
T1.3	21.0	0.74	19.2	6	85	30
T1.4	21.5	0.76	19.3	6	90	30
T1.5	22.0	0.78	19.4	6	90	30
T1.6	22.5	0.80	19.5	6	90	30
T1.7	23.0	0.82	19.6	6	90	30
T1.8	23.5	0.84	19.7	6	95	30
T1.9	24.0	0.86	19.8	6	95	30
T1.10	24.5	0.88	19.9	6	95	30
T1.11	25.0	0.90	20.0	6	95	30

Table A2. Key specifications of OWT foundations at Farm 2.

Unit name	Water depth (m)	Soil relative density	Soil unit weight (kN m^{-3})	Pile diameter (m)	Wall thickness (mm)	Embedment depth (m)
T2.1	20.0	0.40	18.0	6.0	90	40
T2.2	20.5	0.42	18.1	6.0	90	40
T2.3	21.0	0.44	18.2	6.0	90	40
T2.4	21.5	0.46	18.3	6.0	90	40
T2.5	22.0	0.48	18.4	6.0	90	40
T2.6	22.5	0.50	18.5	6.1	90	40
T2.7	23.0	0.52	18.6	6.1	90	40
T2.8	23.5	0.54	18.7	6.1	90	40
T2.9	24.0	0.56	18.8	6.1	90	40
T2.10	24.5	0.58	18.9	6.2	90	40
T2.11	25.0	0.60	19.0	6.2	90	40

Table A3. Key specifications of OWT foundations at Farm 3.

Unit name	Water depth (m)	Soil relative density	Soil unit weight (kN m^{-3})	Pile diameter (m)	Wall thickness (mm)	Embedment depth (m)
T3.1	20.0	0.20	17.0	6.0	90	50
T3.2	20.5	0.22	17.1	6.0	90	50
T3.3	21.0	0.24	17.2	6.0	90	50
T3.4	21.5	0.26	17.3	6.0	90	50
T3.5	22.0	0.28	17.4	6.0	90	50
T3.6	22.5	0.30	17.5	6.1	90	50
T3.7	23.0	0.32	17.6	6.1	90	50
T3.8	23.5	0.34	17.7	6.1	90	50
T3.9	24.0	0.36	17.8	6.1	90	50
T3.10	24.5	0.38	17.9	6.2	90	50
T3.11	25.0	0.40	18.0	6.2	90	50

Table A4. Environmental conditions at Farm 1.

Load case	Wind speed ($m\,s^{-1}$)	Significant wave height (m)	Wave peak period (second)	Current speed ($m\,s^{-1}$)
LC 1.1	10.0	1.61	5.14	0.26
LC 1.2	11.0	1.73	5.23	0.29
LC 1.3	11.5	1.79	5.28	0.30
LC 1.4	12.0	1.85	5.32	0.31
LC 1.5	12.5	1.91	5.37	0.32
LC 1.6	13.0	1.97	5.41	0.34
LC 1.7	13.5	2.03	5.46	0.35
LC 1.8	14.0	2.09	5.50	0.36
LC 1.9	14.5	2.15	5.55	0.38
LC 1.10	15.0	2.21	5.59	0.39
LC 1.11	15.5	2.27	5.64	0.40
LC 1.12	16.0	2.33	5.68	0.41

Table A5. Environmental conditions at Farm 2.

Load case	Wind speed ($m\,s^{-1}$)	Significant wave height (m)	Wave peak period (second)	Current speed ($m\,s^{-1}$)
LC 2.1	9.60	1.64	5.16	0.25
LC 2.2	10.2	1.73	5.23	0.26
LC 2.3	10.8	1.81	5.30	0.28
LC 2.4	11.4	1.90	5.37	0.30
LC 2.5	12.0	1.98	5.44	0.31
LC 2.6	12.6	2.07	5.51	0.33
LC 2.7	13.2	2.15	5.58	0.34
LC 2.8	13.8	2.24	5.65	0.36
LC 2.9	14.4	2.32	5.72	0.37
LC 2.10	15.0	2.41	5.79	0.39
LC 2.11	15.6	2.49	5.86	0.40
LC 2.12	16.2	2.58	5.94	0.42

Table A6. Environmental conditions at Farm 3.

Load case	Wind speed ($m\,s^{-1}$)	Significant wave height (m)	Wave peak period (second)	Current speed ($m\,s^{-1}$)
LC 3.1	9.30	1.60	5.13	0.24
LC3.2	9.90	1.69	5.20	0.26
LC 3.3	10.5	1.77	5.27	0.27
LC 3.4	11.1	1.86	5.34	0.29
LC 3.5	11.7	1.94	5.41	0.30
LC 3.6	12.3	2.03	5.48	0.32
LC 3.7	12.9	2.11	5.55	0.33
LC 3.8	13.5	2.19	5.62	0.35
LC 3.9	14.1	2.28	5.69	0.37
LC 3.10	14.7	2.36	5.76	0.38
LC 3.11	15.3	2.45	5.83	0.40
LC 3.12	15.9	2.53	5.90	0.41

Table A7. MMD value between the healthy-state data of Turbine T1.1 and the remaining turbines at Farm 1 under both healthy and scoured states.

Scour depth	Target turbines										
	T1.1	T1.2	T1.3	T1.4	T1.5	T1.6	T1.7	T1.8	T1.9	T1.10	T1.11
0 m (healthy)	0.00	0.02	0.12	0.20	0.32	0.28	0.36	0.44	0.48	0.52	0.52
1 m	0.05	0.02	0.03	0.07	0.22	0.18	0.30	0.39	0.42	0.45	0.45
2 m	0.25	0.16	0.06	0.04	0.07	0.08	0.19	0.31	0.36	0.41	0.42
3 m	0.52	0.42	0.29	0.26	0.12	0.15	0.13	0.17	0.24	0.31	0.33
4 m	0.72	0.64	0.55	0.53	0.36	0.37	0.24	0.21	0.20	0.25	0.27
5 m	0.88	0.81	0.75	0.74	0.61	0.61	0.47	0.43	0.34	0.41	0.41
6 m	1.00	0.95	0.90	0.91	0.81	0.80	0.69	0.68	0.59	0.62	0.65
7 m	1.12	1.07	1.03	1.03	0.95	0.94	0.87	0.86	0.79	0.80	0.82
8 m	1.21	1.17	1.15	1.16	1.08	1.06	1.00	1.00	0.94	0.95	0.98
9 m	1.29	1.25	1.22	1.22	1.17	1.15	1.10	1.12	1.07	1.09	1.11

Table A8. MMD value between the healthy-state data of Turbine T2.1 and the remaining turbines at Farm 2 under both healthy and scoured states.

Scour depth	Target turbines										
	T2.1	T2.2	T2.3	T2.4	T2.5	T2.6	T2.7	T2.8	T2.9	T2.10	T2.11
0 m (healthy)	0.00	0.04	0.03	0.12	0.21	0.38	0.45	0.47	0.57	0.61	0.62
1 m	0.03	0.01	0.01	0.05	0.12	0.28	0.36	0.38	0.50	0.55	0.56
2 m	0.14	0.05	0.07	0.03	0.05	0.15	0.21	0.24	0.38	0.44	0.45
3 m	0.32	0.19	0.21	0.13	0.08	0.07	0.10	0.13	0.24	0.30	0.33
4 m	0.50	0.36	0.37	0.32	0.20	0.11	0.12	0.12	0.15	0.19	0.21
5 m	0.64	0.52	0.53	0.51	0.36	0.26	0.27	0.27	0.17	0.19	0.20
6 m	0.76	0.66	0.66	0.67	0.53	0.44	0.48	0.47	0.34	0.35	0.31
7 m	0.85	0.78	0.78	0.79	0.67	0.59	0.64	0.64	0.52	0.54	0.48
8 m	0.90	0.89	0.89	0.90	0.78	0.73	0.78	0.77	0.68	0.69	0.63
9 m	0.90	1.00	1.01	1.01	0.90	0.84	0.89	0.88	0.80	0.81	0.77

Table A9. MMD value between the healthy-state data of Turbine T3.1 and the remaining turbines at Farm 3 under both healthy and scoured states.

Scour depth	Target turbines										
	T3.1	T3.2	T3.3	T3.4	T3.5	T3.6	T3.7	T3.8	T3.9	T3.10	T3.11
0 m (healthy)	0.00	0.04	0.05	0.11	0.23	0.38	0.42	0.47	0.55	0.64	0.67
1 m	0.02	0.01	0.01	0.05	0.16	0.29	0.33	0.40	0.48	0.58	0.62
2 m	0.11	0.02	0.02	0.02	0.07	0.17	0.21	0.29	0.37	0.48	0.53
3 m	0.26	0.11	0.13	0.08	0.04	0.07	0.10	0.17	0.24	0.35	0.42
4 m	0.39	0.26	0.29	0.23	0.11	0.07	0.09	0.11	0.14	0.22	0.29
5 m	0.53	0.41	0.42	0.37	0.25	0.19	0.19	0.12	0.13	0.16	0.20
6 m	0.65	0.55	0.57	0.52	0.40	0.32	0.31	0.21	0.21	0.20	0.19
7 m	0.77	0.68	0.70	0.66	0.55	0.46	0.46	0.34	0.34	0.30	0.25
8 m	0.86	0.79	0.79	0.78	0.69	0.61	0.62	0.49	0.50	0.46	0.39
9 m	0.93	0.87	0.89	0.86	0.79	0.74	0.75	0.63	0.65	0.61	0.55

Table A10. MMD distance between Turbine T3.1 and the remaining turbines across different farms in the healthy state (X is 1–3 for Farms 1–3).

	TX.1	TX.2	TX.3	TX.4	TX.5	TX.6	TX.7	TX.8	TX.9	TX.10	TX.11
Farm 3	0.00	0.04	0.05	0.11	0.23	0.38	0.42	0.47	0.55	0.64	0.67
Farm 2	0.71	0.73	0.73	0.80	0.82	0.90	0.96	0.96	0.98	1.02	1.00
Farm 1	1.15	1.15	1.17	1.21	1.23	1.19	1.21	1.26	1.27	1.29	1.28

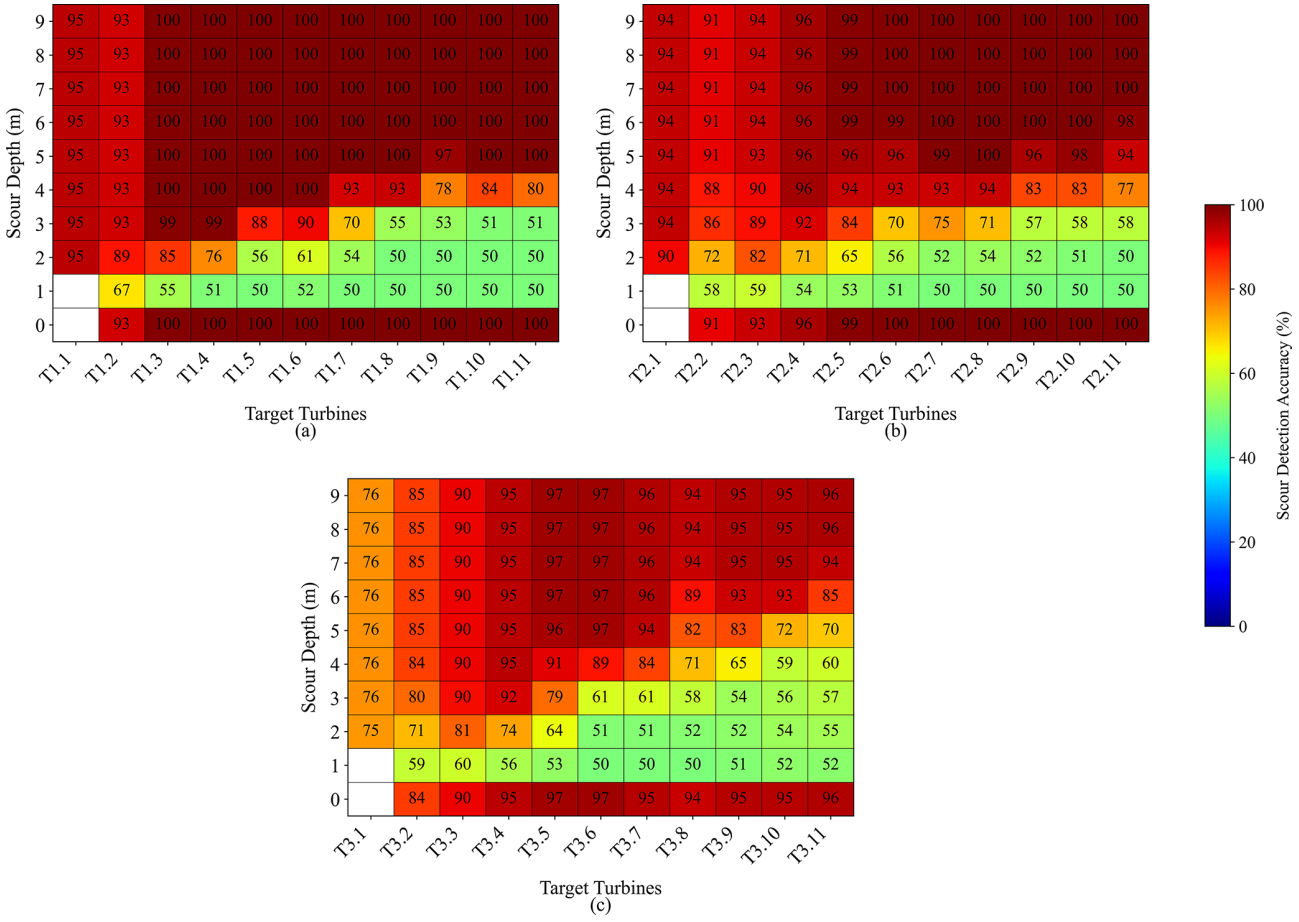


Figure A1. The baseline model performance when trained on Task 1 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

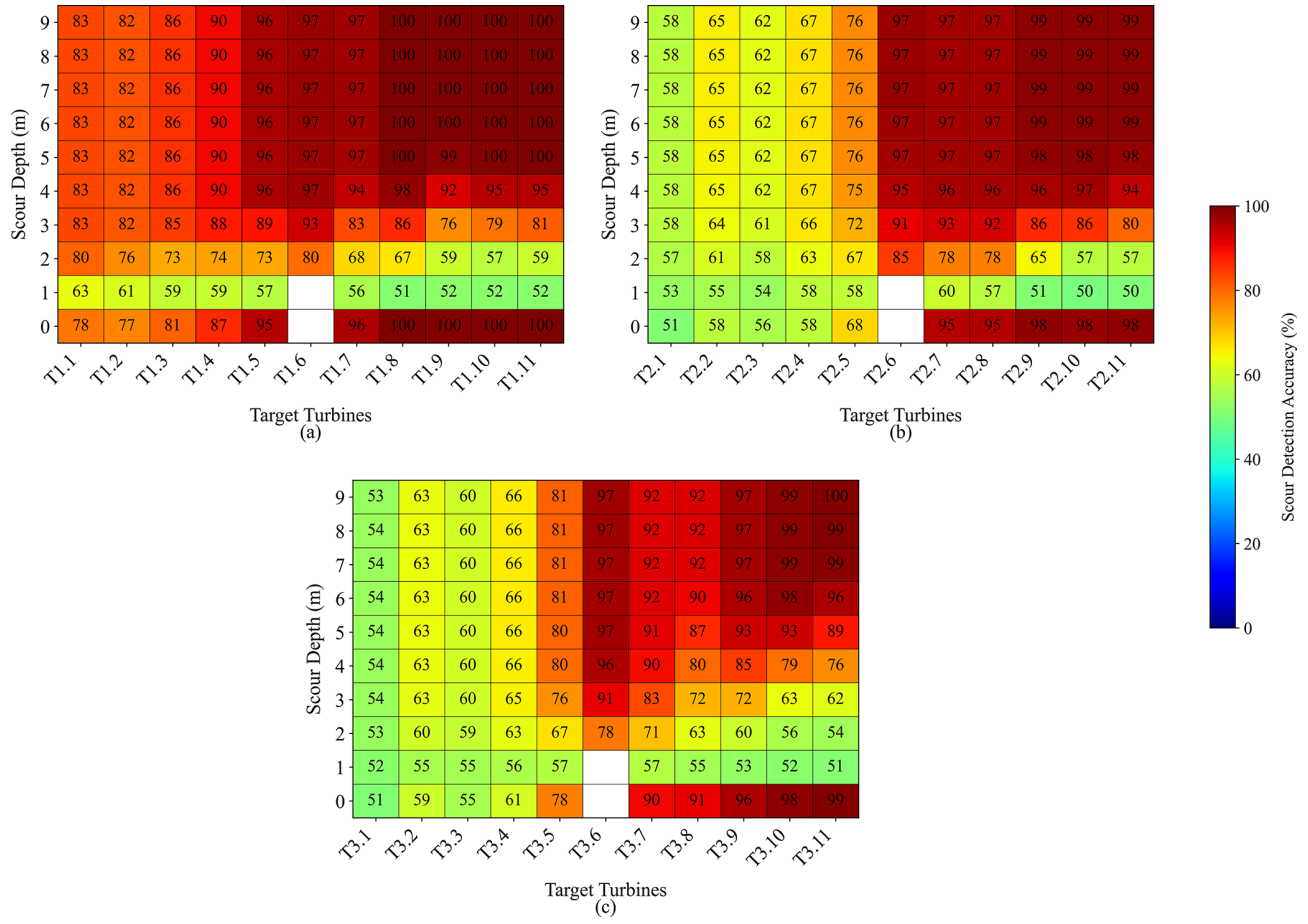


Figure A2. The baseline model performance when trained on Task 2 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

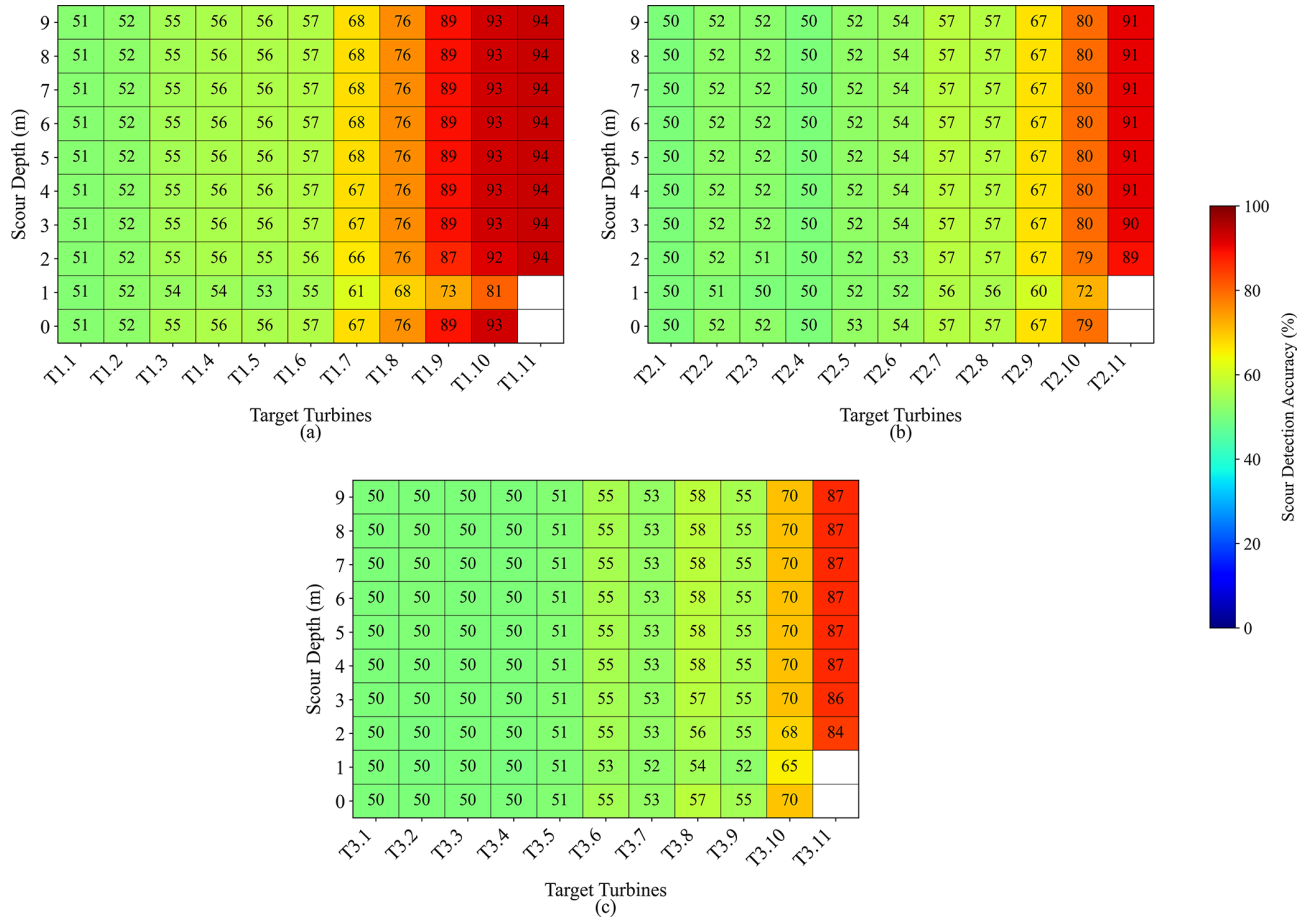


Figure A3. The baseline model performance when trained on Task 3 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

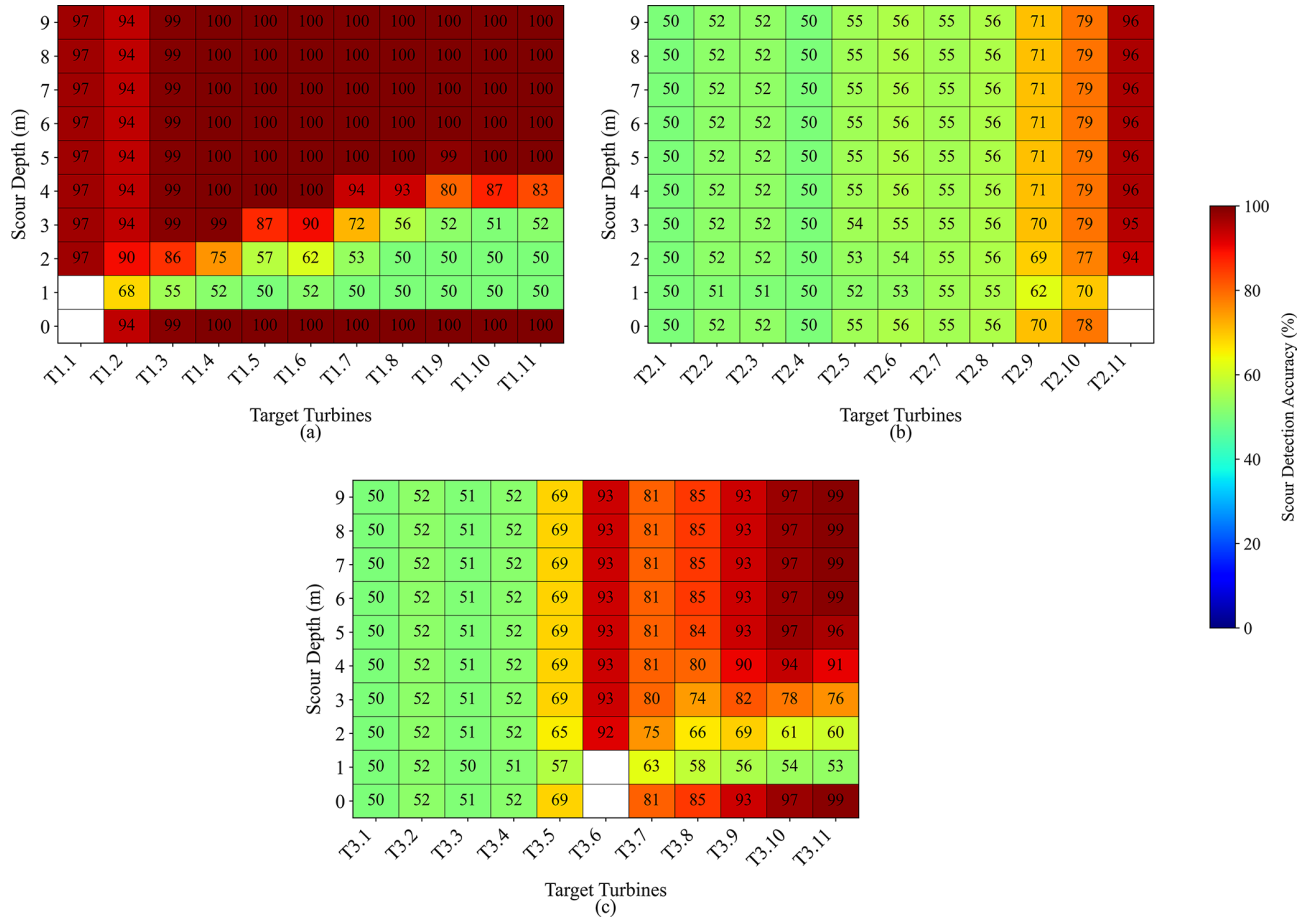


Figure A4. The baseline model performance when trained on Task 4 and evaluated on target OWTs located at **(a)** Farm 1, **(b)** Farm 2, and **(c)** Farm 3.

Data availability. The dataset generated and analysed during the current study is exceptionally large and comprises numerous raw and processed files. Public deposition is impractical due to the data volume, storage requirements, and the need to provide associated metadata and documentation for correct interpretation. The data will therefore be made available from the corresponding author upon reasonable request.

Author contributions. MA: writing (original draft), visualization, validation, software, methodology, formal analysis, data curation, conceptualization. LJP: writing (review and editing), project administration, methodology, funding acquisition, conceptualization. JTB: software, methodology, data curation. RG: writing (review and editing), validation, supervision, methodology. AM: writing (review and editing), visualization, supervision, project administration, methodology, funding acquisition, conceptualization.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. The authors at the University of Nottingham are grateful for financial support from UKRI through Innovate UK and the KTP initiative.

Financial support. This research was supported by UK Research and Innovation (UKRI) through Innovate UK.

Review statement. This paper was edited by Erin Bachynski-Polić and reviewed by two anonymous referees.

References

Abdelhak, M., Tott-Buswell, J., Prendergast, L. J., Ghiasi, R., and Malekjafarian, A.: Scour identification around offshore wind turbine monopile foundations using an unsupervised anomaly detection approach, Ocean Eng., 340, 122365, https://doi.org/10.1016/j.oceaneng.2025.122365, 2025.

Abdelhak, M., Prendergast, L. J., Tott-Buswell, J., Ghiasi, R., and Malekjafarian, A.: A transfer learning-based framework for scour detection around monopile-supported offshore wind turbines with limited data, Ocean Eng., 365, 127315, https://doi.org/10.1016/j.oceaneng.2026.127315, 2026a.

Abdelhak, M., Prendergast, L. J., Tott-Buswell, J., Ghiasi, R., and Malekjafarian, A.: Wind farm-level scour detection around offshore monopile foundations using unsupervised domain adaptation, Mech. Syst. Signal Pr., 251, 114186, https://doi.org/10.1016/j.ymssp.2026.114186, 2026b.

Arany, L. S., Bhattacharya, S., Macdonald, J. H. G., and Hogan, S. J.: Design of monopiles for offshore wind turbines in 10 steps, Soil Dyn. Earthq. Eng., 92, 126–152, https://doi.org/10.1016/j.soildyn.2016.09.024, 2017.

Byrne, B., Mcadam, R., Burd, H., Houlsby, G., Martin, C., Zdravkovic, L., Taborda, D. M. G., Potts, D., Jardine, R., Sideri, M., Schroeder, F., Gavin, K., Doherty, P., Igoe, D., Wood, A., Kallehave, D., and Gretlund, J.: New design methods for large diameter piles under lateral loading for offshore wind applications, Proceedings of the Third International Symposium on Frontiers in Offshore Geotechnics, ISFOG 2015, Oslo, Norway, 2015.

Cao, W., Meng, Z., Li, J., Guan, Y., Fan, J., He, H., and Fan, F.: A bearing fault diagnosis method for unknown operating conditions based on differentiated feature extraction, ISA Transactions, 156, 468–478, https://doi.org/10.1016/j.isatra.2024.10.024, 2025.

Cheynet, E., Jakobsen, J. B., and Obhrai, C.: Spectral characteristics of surface-layer turbulence in the North Sea, Energy Proced., 137, 414–427, https://doi.org/10.1016/j.egypro.2017.10.366, 2017.

Cieslak, M. C., Castelfranco, A. M., Roncalli, V., Lenz, P. H., and Hartline, D. K.: *t*-Distributed Stochastic Neighbor Embedding (*t*-SNE): A tool for eco-physiological transcriptomic analysis, Marine Genom., 51, 100723, https://doi.org/10.1016/j.margen.2019.100723, 2020.

Dai, S., Han, B., Wang, B., Luo, J., and He, B.: Influence of soil scour on lateral behavior of large-diameter offshore wind-turbine monopile and corresponding scour monitoring method, Ocean Eng., 239, 109809, https://doi.org/10.1016/j.oceaneng.2021.109809, 2021.

Dayal, A., Vimal, K. B., Cenkeramaddi, L. R., Mohan, C., Kumar, A., and Balasubramanian, V. N.: MADG: margin-based adversarial learning for domain generalization, Adv. Neur. In., 36, 58938–58952, https://doi.org/10.52202/075280-2572, 2023.

DNV: DNV-OS-J101: Design of Offshore Wind Turbines, Det Norske Veritas, October 2007.

Duan, B., Wang, D., Qin, C., and Duan, L.: Local Scour Around Marine Structures: A Comprehensive Review of Influencing Factors, Prediction Methods, and Future Directions, Buildings, 15, 2125, https://doi.org/10.3390/buildings15122125, 2025.

Ganin, Y. and Lempitsky, V. S.: Unsupervised Domain Adaptation by Backpropagation. International Conference on Machine Learning, Proceedings of the 32nd International Conference on Machine Learning, PMLR, 37, 1180–1189, 2015.

Ghojogh, B. and Ghodsi, A.: Maximum Mean Discrepancy and Generative Moment Matching Networks: Tutorial and Survey, OSF [preprint], https://doi.org/10.31219/osf.io/34d9u, 2024.

Ghojogh, B., Ghodsi, A., Karray, F., and Crowley, M.: Reproducing Kernel Hilbert Space, Mercer's Theorem, Eigenfunctions, Nyström Method, and Use of Kernels in Machine Learning: Tutorial and Survey, arXiv [preprint], https://doi.org/10.48550/arXiv.2106.08443, 2021.

Huang, W. H. and Yang, R. Y.: Water depth variation influence on the mooring line design for fowt within shallow water region, J. Marine Sci. Eng., 9, https://doi.org/10.3390/jmse9040409, 2021.

Jatoliya, A., Bhattacharya, D., Manna, B., Bento, A. M., and Fazeres Ferradosa, T.: Physics-based and machine-learning models for accurate scour depth prediction, Philos. T. R. Soc. A, 382, 20220403, https://doi.org/10.1098/rsta.2022.0403, 2024.

Jawalageri, S., Prendergast, L. J., Jalilvand, S., and Malekjafarian, A.: Effect of scour erosion on mode shapes of a 5 MW monopile-supported offshore wind turbine, Ocean Eng., 266, 113131, https://doi.org/10.1016/j.oceaneng.2022.113131, 2022.

Jawalageri, S., Ghiasi, R., Jalilvand, S., Prendergast, L. J., and Malekjafarian, A.: A data-driven approach for scour detection around monopile-supported offshore wind turbines using Naive Bayes classification, Marine Struct., 95, 103565, https://doi.org/10.1016/j.marstruc.2023.103565, 2024.

Jeong, S., Kim, E. J., Shin, D. H., Park, J. W., and Sim, S. H.: Data fusion-based damage identification for a monopile offshore wind turbine structure using wireless smart sensors, Ocean Eng., 195, 106728, https://doi.org/10.1016/j.oceaneng.2019.106728, 2020.

Jia, S., Li, Y., Wang, X., Sun, D., and Deng, Z.: Deep causal factorization network: A novel domain generalization method for cross-machine bearing fault diagnosis, Mech. Syst. Signal Pr., 192, 110228, https://doi.org/10.1016/j.ymssp.2023.110228, 2023.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development, National Renewable Energy Laboratory (NREL), https://doi.org/10.2172/947422, 2009.

Ju, X., Wu, X., Dai, S., Li, M., and Hu, D.: Domain adversarial learning with multiple adversarial tasks for EEG emotion recognition, Expert Systems with Applications, 266, 126028, https://doi.org/10.1016/j.eswa.2024.126028, 2025.

Jwo, D.-J., Chang, W.-Y., and Wu, I. H.: Windowing Techniques, the Welch Method for Improvement of Power Spectrum Estimation. Computers, Materials and Continua, 67, 3983–4003, https://doi.org/10.32604/cmc.2021.014752, 2021.

Li, X., Zhang, W., Ma, H., Luo, Z., and Li, X.: Domain generalization in rotating machinery fault diagnostics using deep neural networks, Neurocomputing, 403, 409–420, https://doi.org/10.1016/j.neucom.2020.05.014, 2020.

Lin, C., Han, J., Bennett, C., and Parsons, R. L.: Analysis of laterally loaded piles in soft clay considering scour-hole dimensions, Ocean Eng., 111, 461–470, https://doi.org/10.1016/j.oceaneng.2015.11.029, 2016.

Lu, B., Zuo, J., Shahhosseini, M., Wang, H., Liu, H., Zhang, M., and Yu, G.: Visual deep learning with physics constraints for local

scour evolution prediction at monopiles, J. Ocean Eng. Sci., 10, 342–352, https://doi.org/10.1016/j.joes.2024.04.001, 2024.

Malekjafarian, A., Riad, M., Chen, J., Ghiasi, R., Da-Silva-Silvestre, G., and Golpayegani, F.: Cooperative UAVs/-Drones for Inspection of Civil Infrastructure: A Review of Current Challenges and Future Directions, Structural Control and Health Monitoring, 2026, 5775327, https://doi.org/10.1155/stc/5775327, 2026.

Menéndez-Vicente, C., López-Querol, S., Bhattacharya, S., and Simons, R.: Numerical study on the effects of scour on monopile foundations for Offshore Wind Turbines: The case of Robin Rigg wind farm, Soil Dyn. Earthq. Eng., 167, 107803, https://doi.org/10.1016/j.soildyn.2023.107803, 2023.

Mostafa, Y. E.: Effect of Local and Global Scour on Lateral Response of Single Piles in Different Soil Conditions, Engineering, 4, 297–306, https://doi.org/10.4236/eng.2012.46039, 2012.

Najafzadeh, M., Bento, A. M., Basirian, S., and Fazeres-Ferradosa, T.: Dynamic damage functions for scour protection at monopile foundations: Application of ensemble machine learning models, Ocean Eng., 323, 120590, https://doi.org/10.1016/j.oceaneng.2025.120590, 2025.

Niu, X., Chen, X., Qiu, Z., Wang, H., and Xie, A.: Effects of scour on the first natural frequency of monopile in different layered foundations, Ocean Eng., 265, 112580, https://doi.org/10.1016/j.oceaneng.2022.112580, 2022.

Niu, X., Chen, X., Du, W., Xie, A., and Wang, H.: Method for monitoring scour depth of monopile under different relative densities of soil, Ships Offshore Struc., 18, 1438–1447, https://doi.org/10.1080/17445302.2022.2120298, 2023.

Prendergast, L. J. and Gavin, K.: A comparison of initial stiffness formulations for small-strain soil–pile dynamic Winkler modelling, Soil Dyn. Earthq. Eng., 81, 27–41, https://doi.org/10.1016/j.soildyn.2015.11.006, 2016.

Prendergast, L. J., Gavin, K., and Doherty, P.: An investigation into the effect of scour on the natural frequency of an offshore wind turbine, Ocean Eng., 101, 1–11, https://doi.org/10.1016/j.oceaneng.2015.04.017, 2015.

Prendergast, L. J., Gavin, K., and Reale, C.: Sensitivity Studies on Scour Detection Using Vibration-based Systems, Transp. Res. Proc., 14, 3982–3989, https://doi.org/10.1016/j.trpro.2016.05.495, 2016.

Prendergast, L. J., Reale, C., and Gavin, K.: Probabilistic examination of the change in eigenfrequencies of an offshore wind turbine under progressive scour incorporating soil spatial variability, Marine Struct., 57, 87–104, https://doi.org/10.1016/j.marstruc.2017.09.009, 2018.

Ragab, M., Chen, Z., Zhang, W., Eldele, E., Wu, M., Kwoh, C. K., and Li, X.: Conditional Contrastive Domain Generalization for Fault Diagnosis, IEEE T. Instrum. Meas., 71, 1–12, https://doi.org/10.1109/TIM.2022.3154000, 2022.

Rezaei, R., Duffour, P., and Fromme, P.: Scour influence on the fatigue life of operational monopile-supported offshore wind turbines, Wind Energy, 21, 683–696, https://doi.org/10.1002/we.2187, 2018.

Teng, S., Chen, G., Yan, Z., Cheng, L., and Bassir, D.: Vibration-based structural damage detection using 1-D convolutional neural network and transfer learning, Struct. Health Monit., 22, 2888–2909, 2022.

Tott-Buswell, J., Prendergast, L. J., and Gavin, K.: A CPT-based multi-spring model for lateral monopile analysis under SLS conditions in sand, Ocean Eng., 293, 116642, https://doi.org/10.1016/j.oceaneng.2023.116642, 2024.

Wang, X., Zeng, X., Li, J., Yang, X., and Wang, H.: A review on recent advancements of substructures for offshore wind turbines, Energ. Convers. Manage., 158, 103–119, https://doi.org/10.1016/j.enconman.2017.12.061, 2018.

Weijtjens, W., Verbelen, T., Capello, E., and Devriendt, C.: Vibration based structural health monitoring of the substructures of five offshore wind turbines, Proced. Eng., 199, 2294–2299, https://doi.org/10.1016/j.proeng.2017.09.187, 2017.

Wolfram, J. and Naghipour, M.: On the estimation of Morison force coefficients and their predictive accuracy for very rough circular cylinders, Appl. Ocean Res., 21, 311–328, https://doi.org/10.1016/S0141-1187(99)00018-8, 1999.

Zefzouf, M., Chioukh, N., Castro, M., Miranda, F., Leal Sousa, R., Carvalho, E., Rosa-Santos, P., Taveira-Pinto, F., and Fazeres-Ferradosa, T.: Laboratory study on scour development around a monopile in layered soils, Ocean Eng., 342, 122968, https://doi.org/10.1016/j.oceaneng.2025.122968, 2025.

Zhao, D., Han, N., Goh, E., Cater, J. E., and Reinecke, A.: Offshore wind turbine aerodynamics modelling and measurements, Wind Turbines and Aerodynamics Energy Harvesters, Academic Press, 373–400, https://doi.org/10.1016/B978-0-12-817135-6.00005-3, 2019.