

Monitoring Offshore Ocean Wave Heights from Land Seismic Networks: Multi-Location ANN Estimations and Insights from Microseism Wavefield Simulations

Samaneh Baranbooei ^{1,2}, Christopher J. Bean ¹, Meysam Rezaeifar ^{1,3}, Athira Vijayan ¹, Florian Le Pape ⁴

¹Geophysics Section, School of Cosmic Physics, Dublin Institute for Advanced Studies, Dublin, Ireland, ²School of Natural Sciences, Earth and Ocean Sciences, University of Galway, Ireland, ³Flood Forecasting Division, Met Éireann, Dublin, Ireland, ⁴Geo-Ocean, UMR6538 Univ Brest, CNRS, Ifremer, Plouzané, France

Author contributions: *Conceptualization:* Samaneh Baranbooei, Christopher J. Bean. *Methodology:* Samaneh Baranbooei, Athira Vijayan, Florian Le Pape, Christopher J. Bean. *Software:* Samaneh Baranbooei, Athira Vijayan, Florian Le Pape. *Validation:* Samaneh Baranbooei, Christopher J. Bean. *Formal Analysis:* Samaneh Baranbooei, Meysam Rezaeifar, Athira Vijayan, Florian Le Pape. *Writing - Original draft:* Samaneh Baranbooei. *Writing - Review & Editing:* Christopher J. Bean, Meysam Rezaeifar, Athira Vijayan. *Visualization:* Samaneh Baranbooei, Meysam Rezaeifar, Athira Vijayan, Florian Le Pape. *Supervision:* Christopher J. Bean.

Abstract This study investigates the potential of using land based seismic observations of secondary microseisms to estimate oceanic significant wave height (SWH) in the Northeast Atlantic. A feedforward artificial neural network (ANN) was trained using physical buoy measurements, hindcast wave data and seismic amplitudes recorded at multiple broadband stations across Ireland. The model successfully reproduces temporal variability in SWH at different locations and demonstrates strong predictive capability across both real and virtual offshore sites.

Spatial analysis reveals that model performance is not uniform, with higher accuracy and stronger signal coherence observed at northern and northwestern locations compared to sites further south. To understand this variability, numerical simulations of microseism propagation were performed for different source regions. The simulations show that microseism wavefield propagation from northern sources is less distorted than for southern sources, where wavefields interact strongly with the shelf.

These results suggest that the spatial variability in ANN performance is consistent with and may be partly influenced by the physical characteristics of microseism wavefield propagation, particularly the degree of scattering along source to receiver paths. The agreement between ANN predictions and simulated wavefields demonstrates that the model captures physically meaningful relationships between ocean wave activity and seismic response. This work highlights the potential of terrestrial seismic networks as a complementary tool for offshore wave monitoring and provides new insight into the limitations imposed by spatial variations in microseism wavefield complexity.

Production Editor:

Laura Petrescu

Handling Editor:

Andrea Llenos

Copy & Layout Editor:

Mengjie Zheng

Received:

June 24, 2026

Accepted:

September 26, 2026

Published:

October 26, 2026

1 Introduction

Accurate and continuous monitoring of offshore wave conditions is critical for a variety of applications, including climate studies, coastal hazard assessment, maritime navigation, and the development of renewable energy infrastructure. Among wave parameters, the Significant Wave Height (SWH) is a key metric widely used in both operational forecasting and long term climatological research. Traditional SWH monitoring methods such as ocean buoys, satellite altimetry, and numerical wave models each present limitations. Buoys are spatially sparse and vulnerable to failure in extreme conditions, satellite observations can be weather dependent and temporally intermittent, and numerical models require ongoing validation against ground truth data such as hindcast wind.

An emerging complementary approach uses the solid

Earth as a natural sensor. Ocean surface gravity waves, driven primarily by wind speed, duration, and fetch, interact with the seafloor and with opposing wave systems to generate microseisms. In particular, "secondary microseisms", produced by non-linear interactions between counter propagating waves, are continuously recorded by global broadband seismic networks (Hasselmann, 1963; Longuet-Higgins, 1950; Ardhuin et al., 2011). These seismic signals offer a weather independent, low maintenance, and spatially extensive proxy for offshore wave activity (Bromirski et al., 1999; Chevrot et al., 2007; Ardhuin et al., 2015). Building on this, seismic based techniques have been developed to estimate SWH, initially using empirical transfer functions derived from buoy comparisons (Bromirski, 2001), and more recently through data driven methods. In particular, Artificial Neural Networks (ANNs) have shown promise in capturing complex, non-linear relationships between secondary microseism amplitudes and wave conditions, allowing the inference of ocean

*Corresponding author: samane@cp.dias.ie

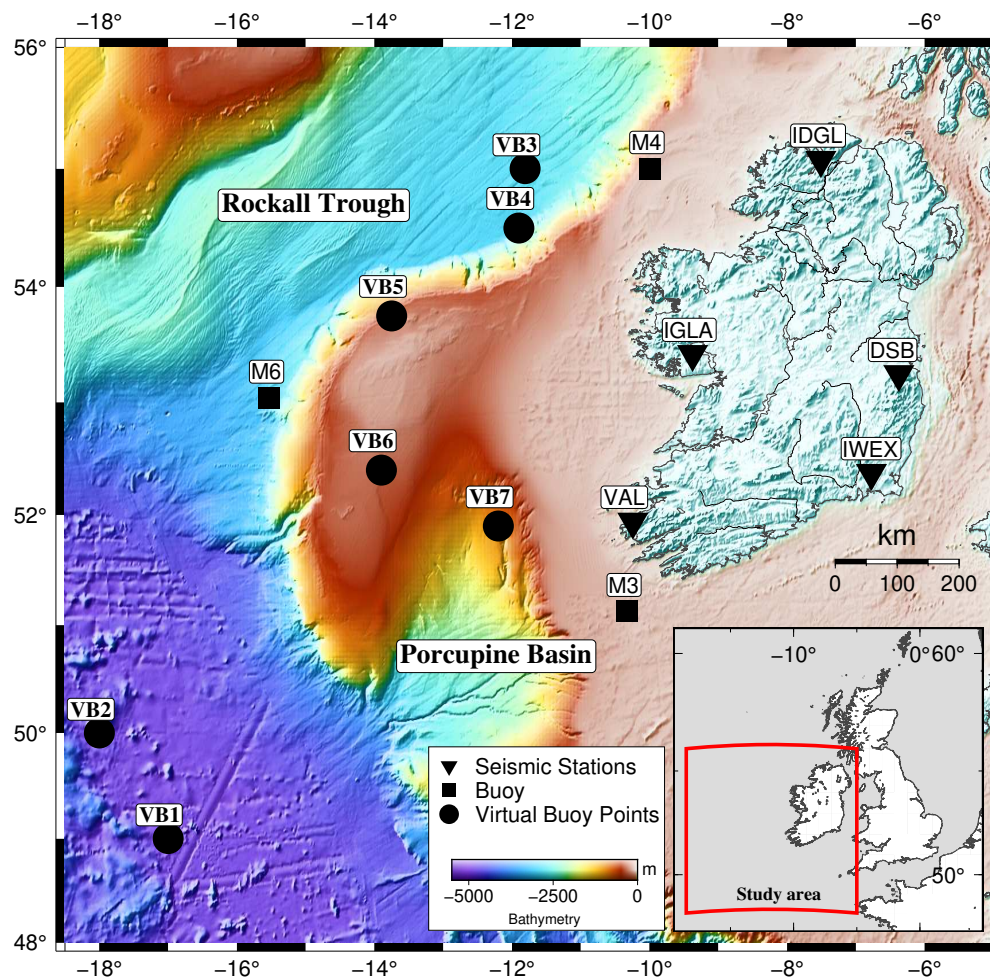


Figure 1 Study area showing locations of seismic stations (inverted triangles), offshore physical buoys (squares), and virtual hindcast points (Virtual Buoys (VB); circles). The area encompasses the energetic Northeast Atlantic, west of Ireland.

SWH (Donne et al., 2014; Cannata et al., 2020; Iafolla et al., 2022; Baranbooei et al., 2025; Bailey et al., 2025).

A key limitation of the approaches taken thus far is their reliance on a single buoy as the source of sea state information. The spatial separation between the buoy and the seismic stations can influence the reliability of the data, particularly due to local variations in bathymetry and seismic wave propagation (Bailey et al., 2025; Le Pape et al., 2021; Ying et al., 2014). A critical gap therefore remains in understanding how seismic based SWH estimation performs across multiple offshore locations, and how well these estimates can be contextualized by physical models of microseism propagation.

In this paper, we partly address these gaps by extending the previous single buoy framework presented by Baranbooei et al. (2025) in two important ways. First, we incorporate observations from three offshore buoy locations and seven numerical hindcast virtual locations, substantially expanding the spatial coverage of seismic based SWH estimation. Second, we conduct numerical simulations of microseism source regions to assess source–receiver relationships and to better quantify the physical basis for our ANN results. Together, these approaches allow us to evaluate both the robustness and generalization of seismic proxies for wave height mon-

itoring.

2 Data and Methods

2.1 Study Area and Observational Framework

The study area (Fig. 1) is located along the Northeast Atlantic margin, a region characterized by energetic wave climates and frequent storm systems that generate persistent microseism activity. Strong North Atlantic swells interact with the coastline of Ireland, producing seaward propagating ocean waves that interact with incoming waves, generating continuous microseism recorded onshore (Moni et al., 2013; Möllhoff and Bean, 2016; Le Pape et al., 2021). Secondary microseisms can also arise from interactions between opposing swell systems in the open ocean, independent of coastal reflections, where the interacting ocean wavefields are not necessarily produced by reflected waves but may originate from separate storm systems (Moni et al., 2013; Bensen et al., 2007; Bromirski et al., 2005; Bromirski, 2001). This natural coupling between ocean wavefields and the solid Earth makes the region an ideal natural laboratory for investigating seismic proxies of ocean wave dynam-

ics.

Seismic data were obtained from five broadband stations of the Irish National Seismic Network (INSN; INSN, 1993), which provides stable, long term coverage across Ireland. These stations are equipped with broadband sensors capable of accurately recording ground motions in the 2–10 s secondary microseism period range, ensuring high quality observations suitable for quantitative amplitude analysis. The array geometry (Fig. 1) captures azimuthally diverse source directions, enabling assessment of how regional wavefields influence inland seismic amplitudes.

2.2 Integrated Datasets

This study combines three complementary datasets to explore the relationship between land recorded microseisms and offshore wave conditions:

1. **In situ observations:** Hourly Significant Wave Height (SWH) data from three offshore buoys operated by the Irish Met Office (Met Éireann; Met Éireann, 2024). Each buoy computes the SWH as four times the Root Mean Square (RMS) of surface elevation over the final 17.5 minutes of each hour.
2. **Numerical hindcast data:** Simulated SWH values derived from the WaveWatch III model (Cheung, 2010; Arduin et al., 2011), extracted at both the above mentioned buoy coordinates and seven additional Virtual Buoy (VB) locations (Fig. 1) chosen to extend spatial coverage over both deeper and poorly instrumented areas of the Northeast Atlantic.
3. **Seismic observations:** Continuous broadband ground motion time series from five INSN stations distributed across Ireland, providing inland records of microseism amplitude variations linked to oceanic wave activity.

Integrating these datasets within a machine learning framework allows us to demonstrate the potential of terrestrial seismic data for geographically disparate estimations of SWH, potentially in real time.

2.3 Seismic Data Processing

Our seismic pre-processing followed established ambient noise workflows (Bensen et al., 2007; Baranbooei et al., 2025) with adaptations to ensure uniform treatment across multiple years and stations. The steps are summarized below:

1. **Instrument correction:** Raw velocity recordings were corrected for ground displacement using published sensor response functions.
2. **Resampling and detrending:** Data were downsampled from 100 samples per second (sps) to 5 sps, demeaned, and detrended to remove long term offsets.
3. **Filtering:** A bandpass filter (2–10 s period) was applied to isolate the secondary microseism frequency band.

4. **Window extraction:** The final 17.5 minutes of each hour were extracted to align with the buoy sampling intervals.
5. **Amplitude metric:** For each hourly window, the Significant Microseism Amplitude (SMA) was calculated as four times the RMS of the filtered displacement. This approach parallels the oceanographic definition of SWH.
6. **Transformation:** Because seismic amplitudes scale approximately with the square of the generating ocean wave amplitudes (Hasselmann, 1963; Longuet-Higgins, 1950), we applied a square root transformation ($RSMA = \sqrt{SMA}$) to linearize the relationship prior to network training.
7. **Quality control:** Time windows affected by regional or teleseismic earthquakes were identified using the INSN event catalogue and removed to avoid bias.

These standardized steps yield a consistent RSMA time series for each station, ensuring that temporal variations predominantly reflect oceanic rather than tectonic processes.

2.4 Wave and Seismic Data Integration

The buoy and hindcast datasets were synchronized with the seismic RSMA time series to create combined matrices for model training. Only periods with complete data coverage across all relevant seismic stations and ocean wave sources were retained, preventing the ANN from learning spurious apparent correlations due to data gaps. The merged dataset spans January 2011 to June 2023, encompassing approximately 42,000 hourly samples after all quality control procedures. Data from 2011–June 2022 were used for ANN model development, whereas the July 2022–June 2023 observations presented in Section 3 were excluded from model development and retained as an independent temporal evaluation dataset. This chronological separation provides a more stringent assessment of the ability of the trained ANN to reproduce significant wave height during a later, previously unseen period and reduces the potential influence of temporal data leakage associated with randomly partitioning autocorrelated hourly observations.

This integrated database provides an internally consistent framework for exploring how microseism amplitudes recorded onshore relate to spatially distributed offshore SWH variations. The diverse network geometry of three buoy locations, seven virtual hindcast locations, and five inland seismic stations provides a range of azimuthal source–receiver paths. This allows us to examine the robustness of seismic based SWH estimation across various ocean wave and microseism propagation regimes.

2.5 Artificial Neural Network Framework

To quantify the relationship between land recorded microseisms and offshore wave heights, we employed a supervised machine learning regression model based on

an Artificial Neural Network (ANN). The network was designed to learn the complex, non-linear mapping between seismic amplitude features and Significant Wave Height (SWH) values, accounting implicitly for variations in microseism source strength and propagation path effects.

The ANN follows a feedforward architecture similar to that described in [Baranbooei et al. \(2025\)](#), consisting of an input layer receiving pre-processed seismic features (RSMA from five stations), five hidden layers with five neurons each, and an output layer containing ten neurons, corresponding to the three physical buoy locations and seven virtual hindcast locations. The ordering of the output neurons is fixed according to the ordering of the corresponding locations in the target matrix used during training. Hidden layers use a hyperbolic tangent sigmoid activation function to capture non-linear relationships. The network was implemented and trained using the MATLAB Neural Network fitting environment ([The MathWorks Inc., 2022](#)) with Bayesian regularization. This approach provides a probabilistic framework that simultaneously minimizes the mean squared prediction error and the magnitude of the network weights, thereby controlling model complexity and reducing overfitting. Training was performed using the Levenberg–Marquardt optimization algorithm over ten epochs. The 2011–June 2022 model-development dataset was partitioned into 80% for training and 20% for internal testing, with an internal validation subset used during training for hyperparameter adjustment. This internal partition was used for model development and diagnostic assessment and is distinct from the independent July 2022–June 2023 evaluation period. Because hourly seismic and ocean-wave observations are temporally autocorrelated, performance on a randomly partitioned internal test subset may provide an optimistic estimate of temporal generalisation. The independent 2022–2023 evaluation period was therefore kept entirely separate from model development and provides the primary assessment of model performance on temporally unseen data.

Given that ANNs extrapolate poorly beyond their training range, the model was trained on a comprehensive dataset (2011–June 2022) to include a full spectrum of oceanic conditions, from calm to storm events. Inputs and targets were normalized prior to training to improve numerical stability and convergence.

For all buoy and hindcast virtual locations, a single multi-site, multi-output ANN was trained using coincident seismic amplitude data. For each set of five seismic inputs, the network simultaneously estimates SWH at the three physical buoy and seven virtual locations through its ten output neurons. This multi-site approach extends previous single-location analyses by allowing SWH at multiple offshore targets to be estimated within a common network framework. The trained models were subsequently applied to the unseen and independent July 2022–June 2023 evaluation dataset, which was not used during model development. Predictive performance on this temporally unseen dataset was quantified using standard statistical metrics, including the Root Mean Square Error (RMSE),

Mean Absolute Error (MAE), and coefficient of determination (R^2). It should be noted that the multi-site ANN results are comparable to those obtained from individual ANNs trained separately for each location. Consequently, the multi-site approach offers a practical advantage by providing reliable SWH estimates for multiple locations within a single training and inference framework. This reduces computational effort, simplifies model deployment, and demonstrates that a common relationship between microseism amplitudes and ocean wave conditions can be learned across a broad geographic region. The ability of a single network to generalise across multiple offshore locations further highlights the robustness of the relationship between seismic wave and ocean SWH captured by the ANN.

A schematic overview of the full processing workflow, from seismic signal conditioning to the ANN prediction and validation, is presented in [Baranbooei et al. \(2025\)](#).

3 Results

3.1 ANN Predictions at Buoy Locations

ANN models trained using coincident seismic amplitude and buoy SWH data reproduce the observed variability at all three buoy sites with high fidelity. The estimated SWH time series track both the seasonal cycle and short term storm peaks, confirming that the models capture the dominant temporal dynamics of offshore wave conditions (Fig. 2, and Figs. S1 and S2 available in the Supplementary Materials). Statistical performance metrics (RMSE, MAE, and R^2) indicate strong agreement between predicted and measured values, consistent with previous single station applications ([Baranbooei et al., 2025](#)).

Across the three buoy locations, R^2 values typically exceed 0.8, with RMSE values below 1 m without systematic bias between the observed and predicted SWH. The models effectively reproduce both moderate and high energy events with SWH of up to 8 m, demonstrating the robustness of the trained networks even when applied to previously unseen data. Figures 2, S1, and S2 summarize the ANN performance for each buoy, comparing the estimated SWH with the buoy recorded values. These results confirm that terrestrial seismic data contain sufficient information to infer offshore SWH at multiple buoy locations, providing a reliable complementary proxy when direct measurements are unavailable from the buoys.

3.2 ANN Predictions at the Virtual Buoy (VB) Locations

The ANN framework was subsequently evaluated at seven additional offshore locations where only hindcast SWH data were available. These virtual locations provided an independent means of assessing model performance across a broader range of oceanographic conditions and source–receiver geometries. The hindcast SWH data, previously shown to closely reproduce buoy observations where co-located measurements are available ([Baranbooei et al., 2025](#)), served as target values

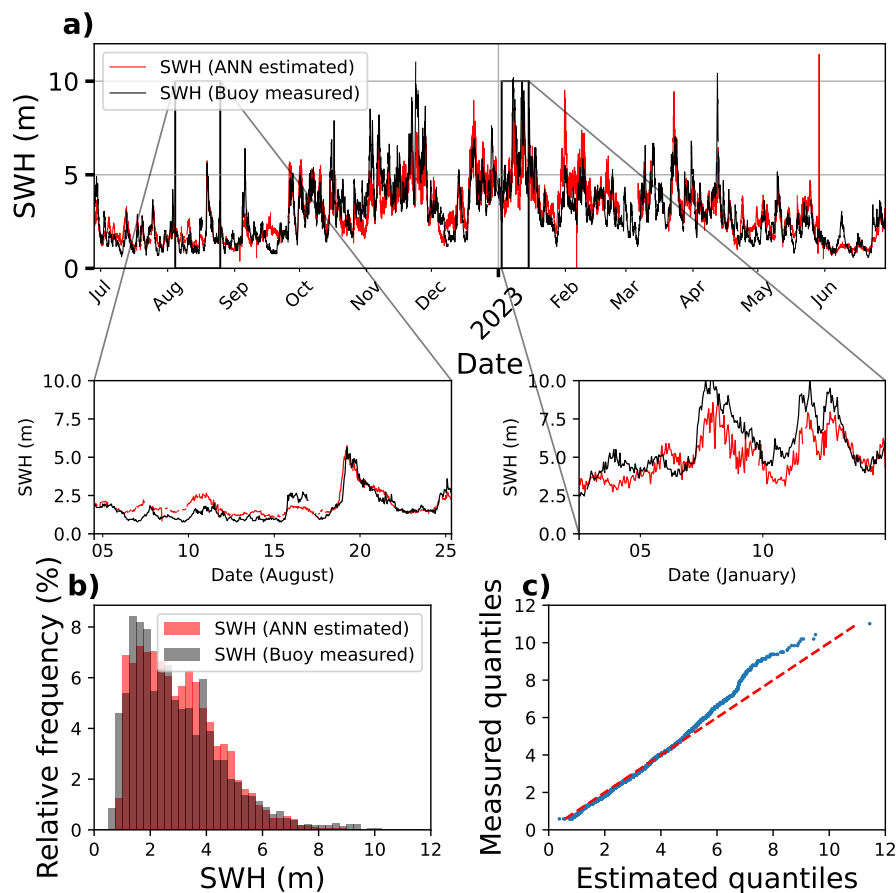


Figure 2 Examples of ANN results for M3 location (See Fig. 1 for location). For ANN results of M4 and M6 locations please refer to the Supplementary Materials (Figs. S1 and S2). Each panel shows (a) the full time series and two enlarged periods of ANN estimated and buoy measured SWH, (b) corresponding histograms, and (c) Q–Q comparison between predicted and measured values.

during training. These virtual locations represent positions of potential operational or scientific interest beyond the limited physical buoy network. The trained model was subsequently evaluated at each of these seven locations using unseen hindcast SWH data at that location. Figures 3 and S3 to S8 present the ANN estimated SWH at these locations and compare them with the corresponding, previously unseen WaveWatch III hindcast values.

Overall, the ANN estimated SWH time series is consistent with the hindcast SWH data across all virtual buoy locations, effectively capturing both seasonal amplitude modulation and short term variability (e.g., storm sequences between December and March; Fig. 3 and Figs. S3–S8 in Supplementary Materials). Histograms and quantile–quantile (Q–Q) plots reveal that the statistical distributions of estimated and reference SWH are a close match, with some deviations at the largest wave heights (e.g. $\text{SWH} > 6$ m).

Quantitatively, the ANN models achieve mean R^2 values of 0.82–0.91 across the virtual buoys, with RMSE values typically between 0.6 and 1.1 m. Locations in deeper water (e.g., VB4–VB6) show slightly im-

proved correspondence, likely due to more stable hindcast–seismic relationships. In contrast, nearshore locations (e.g., VB1–VB3) exhibit marginally larger scatter, possibly due to a more complex ocean wavefield, or more scattering in microseism wave propagation or both.

3.3 Summary of ANN Performance Metrics

To consolidate ANN results across all study locations, Figure 4 summarizes the main performance metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2) for the three physical and seven virtual buoy locations. Each bar represents the value derived from the hourly comparisons between predicted and referenced SWH.

In general, the ANN exhibits good performance throughout the study locations. RMSE values range between 0.74 and 1.1 m, while MAE values remain below 0.72 m, indicating relatively small prediction errors compared to the overall variability of SWH across the study area. The R^2 coefficients cluster between 0.58 and 0.81, confirming a good and stable correlation between seismic based estimates and reference data at both ob-

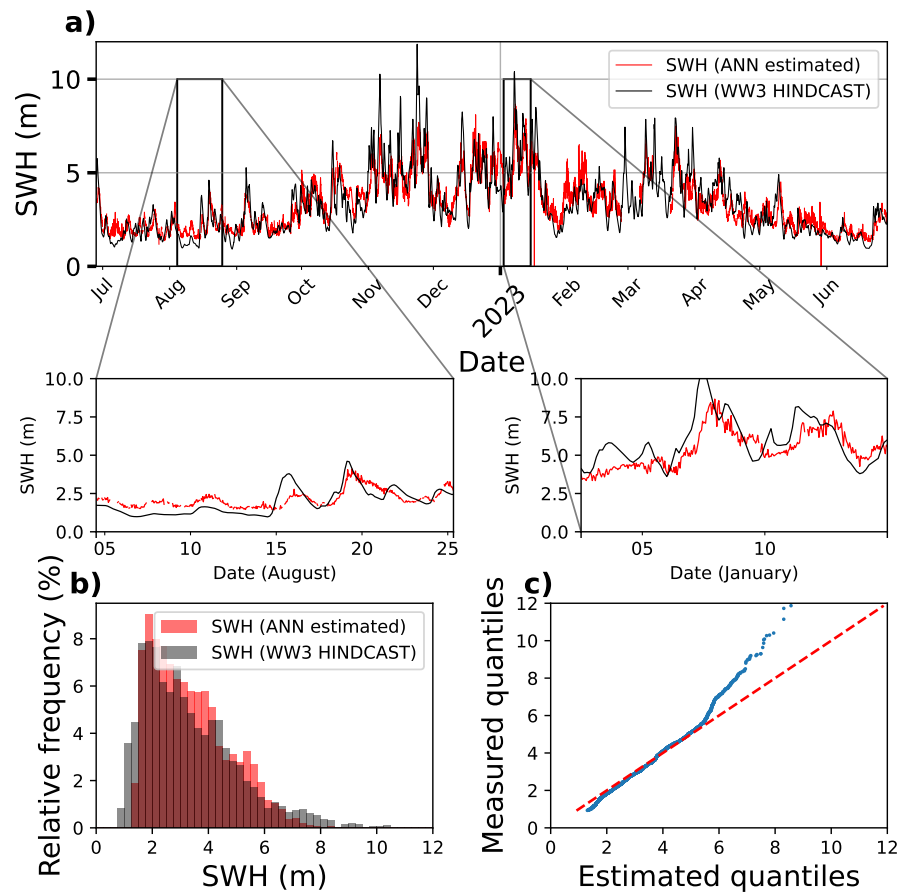


Figure 3 ANN estimated versus modeled SWH using hindcast numerical data at Virtual Buoy location VB1 (See Fig. 1 for location) For ANN results of other virtual buoy locations please refer to the Supplementary Materials (Figs. S3-S8). Each panel shows (a) the full time series and two enlarged periods of ANN estimated and hindcast SWH, (b) corresponding histograms, and (c) Q-Q comparison between predicted and hindcast values.

served and virtual points (Figs. 4 and S9).

At the buoy sites, ANN performance is best at M3, followed by M4, while M6 shows the weakest buoy performance, consistent with the lower errors and higher R^2 at M3 and M4 relative to M6. For virtual buoy locations, ANN performance is poorest at VB1-VB4, remains moderate at VB5, and improves markedly at VB6-VB7, where R^2 increases and both RMSE and MAE decrease. Spatially, this indicates that ANN performance is poorer at the more distal western and southwestern locations. Overall, the results show a clear location dependent gradient in ANN performance across both real and virtual buoy locations.

3.4 Location dependency of ANN Estimated SWH

Although ANN results for individual locations (Figs. 2 and 3 and Supplementary Materials Figures. S1 to S8) demonstrate predictive performance, the error time series, in agreement with the ANN performance metrics, demonstrate that ANN performance is strongly location dependent and not uniform across the study locations (Figs. 5 and 6).

Overall, the ANN consistently exhibits lower and more stable errors at M4 and M3 buoy location compared to M6 location, indicating that these locations are inherently better captured by the model (Fig. 5). This contrast is further supported by the virtual buoy results, where error magnitudes and variability differ between sites rather than following a simple trend as a function of distance (Fig. 6). During energetic periods, all stations show increased residuals, however, the amplitude and coherence of these errors vary between locations. These observations highlight that ANN performance is not spatially uniform, but instead depends on the specific characteristics of each site and illustrates relatively better ANN performance at the northern locations than at those located in the southwest. In particular, the two most southwesterly virtual buoys consistently produce the poorest fits relative to the other locations. Although the limited number of sites prevents definitive attribution, this pattern suggests that location dependent factors may influence the results. Consequently, numerical simulations were conducted to further investigate the mechanisms that could explain the observed variability.

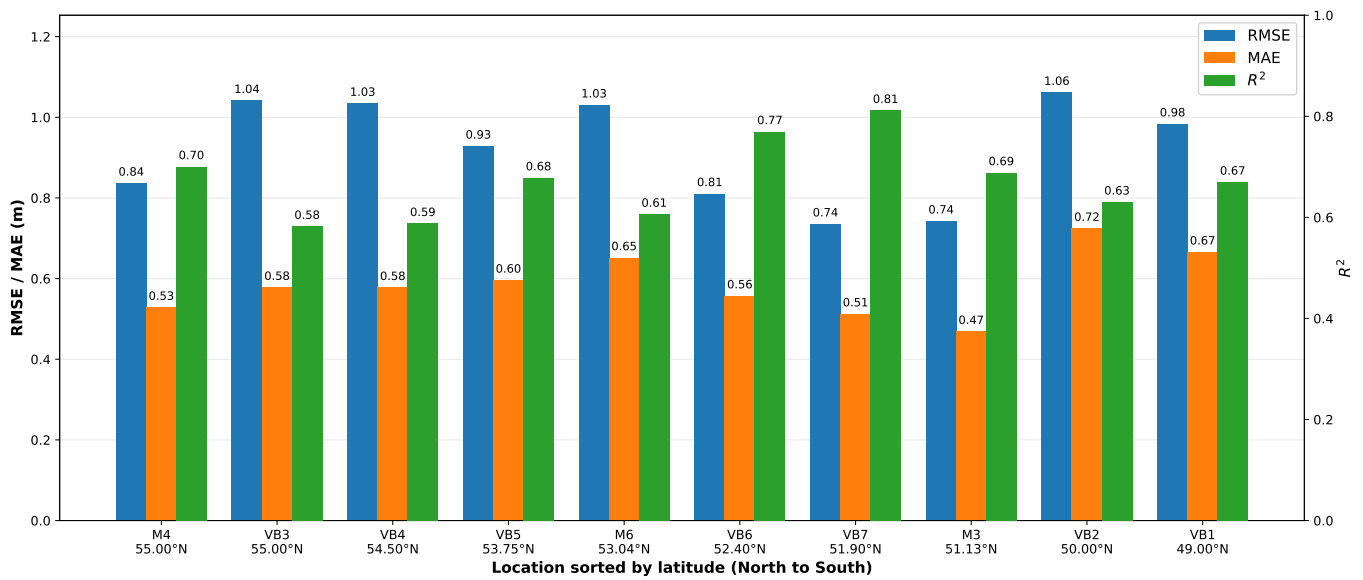


Figure 4 Summary of ANN performance metrics (RMSE, MAE, and R^2) across all physical and virtual buoy locations sorted by latitude. Each set of three bars corresponds to one location. For performance metrics sorted by the relative distance between the location and nearest seismic station please refer to Figure S9 available in supplementary material.

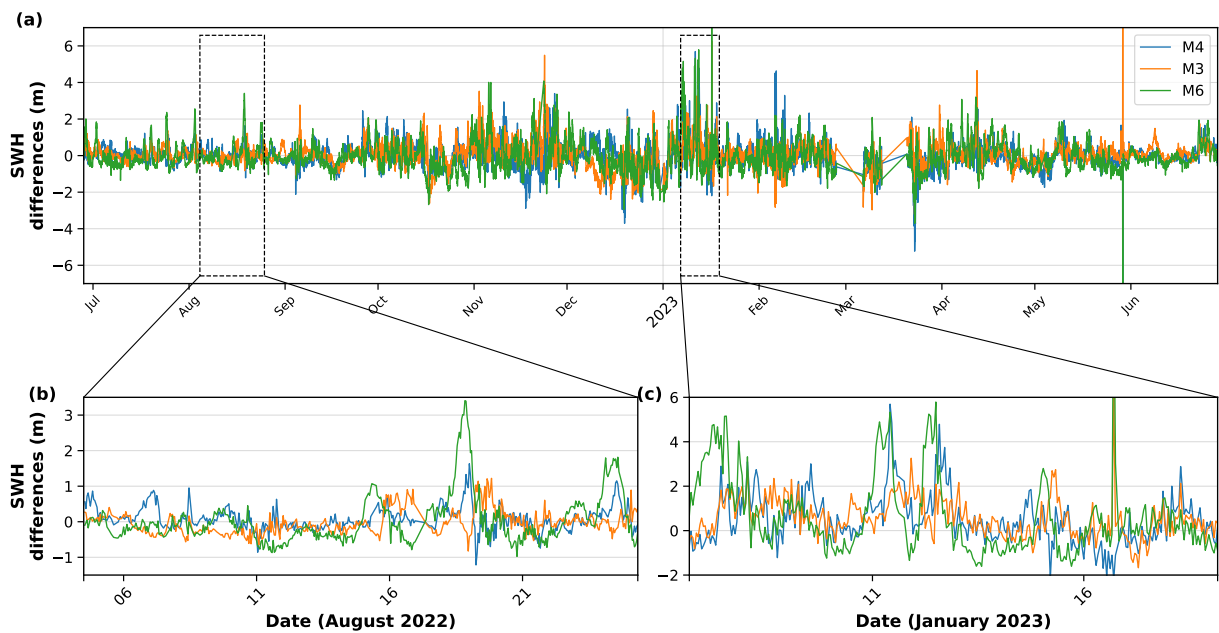


Figure 5 Error time series (hourly measured–estimated) of ANN estimated SWH at physical buoy locations (M3, M4, M6). (a) Annual comparison; (b–c) magnified summer and winter periods showing similar temporal patterns and variability in prediction errors among locations. These estimates represent worst cases, as phase delays between predicted and observed data are not taken into account in these calculations.

3.5 Numerical Simulation of Microseism Propagation

We suspect that spatial variations in microseism wave-field propagation might play a role in the spatial variations observed in the ANN estimates of SWH. Hence we conducted 3D seismo-acoustic numerical simulations of secondary microseism propagation. The numerical simulations were performed using SPECFEM3D cartesian software (Komatitsch and Tromp, 1999; Peter et al., 2011). The 3D volume model used here is from Le Pape et al. (2021). The model area spans 910 km x 1000 km x 60 km with water, sediments, and crustal layers. To

effectively replicate the secondary microseism source without modelling the full source physics (Stutzmann et al., 2012), a vertical pressure Ricker wavelet source of dominant frequency 0.2 Hz is placed 15 m below the water surface. The simulations represent two contrasting source scenarios: 1) a dominant northern source located in the Central Rockall Trough and 2) a southern source in the Porcupine Basin. For additional details on the numerical model formulation, source characterization, and validation of simulation performance, please refer to Le Pape et al. (2021). It should be noted that these single-source simulations are intended to iso-

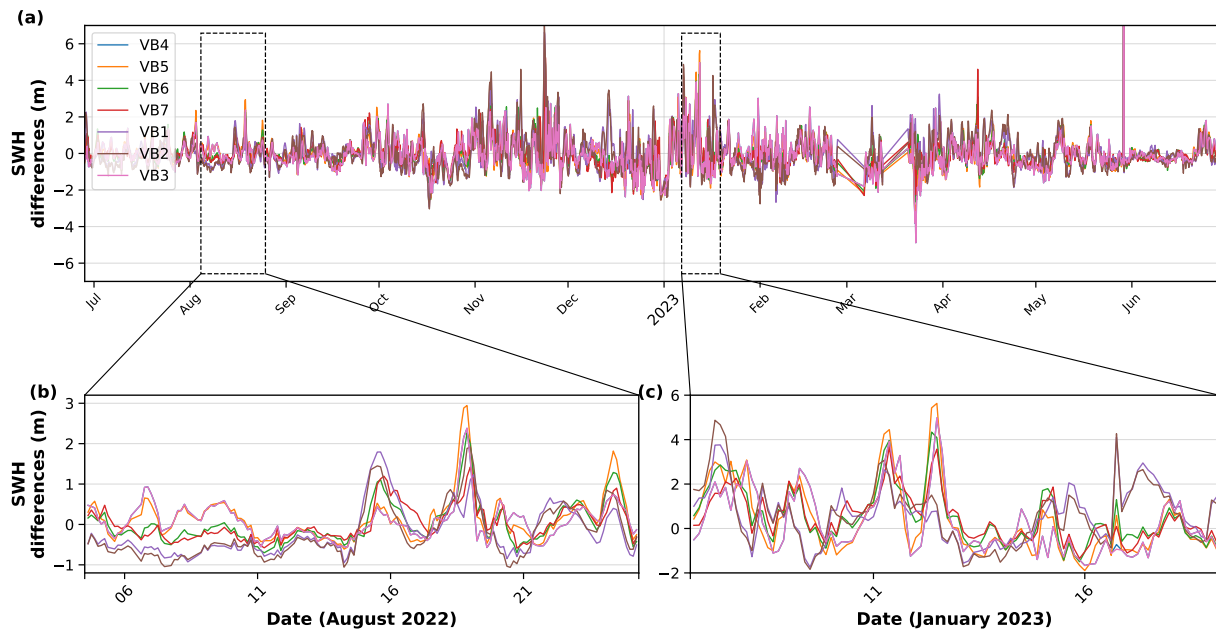


Figure 6 Error time series (3-hourly measured–estimated) of ANN estimated SWH at seven virtual buoy locations (VB1–VB7). (a) Annual comparison; (b–c) magnified summer and winter periods showing similar temporal patterns and variability in prediction errors among locations. These estimates represent worst cases, as phase delays between predicted and observed data are not taken into account in these calculations. Since the VB3 and VB4 predictions are highly similar and largely overlap, the VB4 curve is mainly obscured by the VB3 curve.

late and compare the propagation characteristics associated with contrasting source regions rather than to reproduce the complete observed microseism wavefield, which may contain contributions from multiple simultaneously active and spatially distributed sources.

The simulated wavefield amplitudes display clear directional contrasts between the two configurations. The northern source case (Fig. 7–left panel) generates a focused and coherent energy lobe propagating toward northwest Ireland, where the majority of stations receive clear arrivals from the source region. In contrast, the southern source case (Fig. 7–right panel) produces a more diffuse propagation pattern over the whole region. In this second case, onland arrivals emanate from a highly chaotic offshore microseism wavefield and therefore do not necessarily propagate directly from the actual source region to the terrestrial seismic stations.

To assess how these spatial differences manifest at individual stations, we extracted time series of simulated amplitudes at representative INSN sites, IDGL and VAL (Fig. 8; see Fig. 1 for station locations). The northern source scenario exhibits a more coherent signal across the stations, with waveforms at IDGL and VAL appearing quite similar, whereas the southern source case shows more complex interference patterns and less coherent wavefronts arriving at the terrestrial stations. These results are consistent with the ANN performance, which is generally better for ocean wavefields in the north than in the south, likely reflecting stronger microseism path effects for microseisms generated further south.

4 Discussion

4.1 Performance and Generalization of the ANN Models

The ANN results demonstrate that land based seismic recordings can be used to reconstruct offshore SWH across a broad region. The consistent metrics across physical and virtual buoy locations (Fig. 4) confirm that the models generalize well spatially. This is a further step from single buoy studies, (Bromirski et al., 1999; Donne et al., 2014; Cannata et al., 2020; Iafolla et al., 2022; Baranbooei et al., 2025), which typically lacked cross site validation or spatial information.

The robust performance of the ANN across diverse azimuths and water depths suggests that the network successfully captures physical relationships between microseism amplitudes and offshore wavefields. Variations in ANN prediction between different locations may reflect the combined influences of source–receiver distance and geometry, water depth and bathymetric setting, and propagation path effects that modify the recorded microseism signal. Although these propagation effects are implicitly embedded within the seismic amplitudes used to train the ANN, and are therefore incorporated into the learned statistical relationships, they are not explicitly represented within the model. Consequently, systematic differences in wavefield coherence and scattering between source regions may still influence predictive performance. The observed spatial variability therefore suggests that propagation effects are sufficiently strong to leave a measurable imprint on the relationship between seismic wave and ocean SWH, even after the ANN has learned the dominant patterns present in the training data.

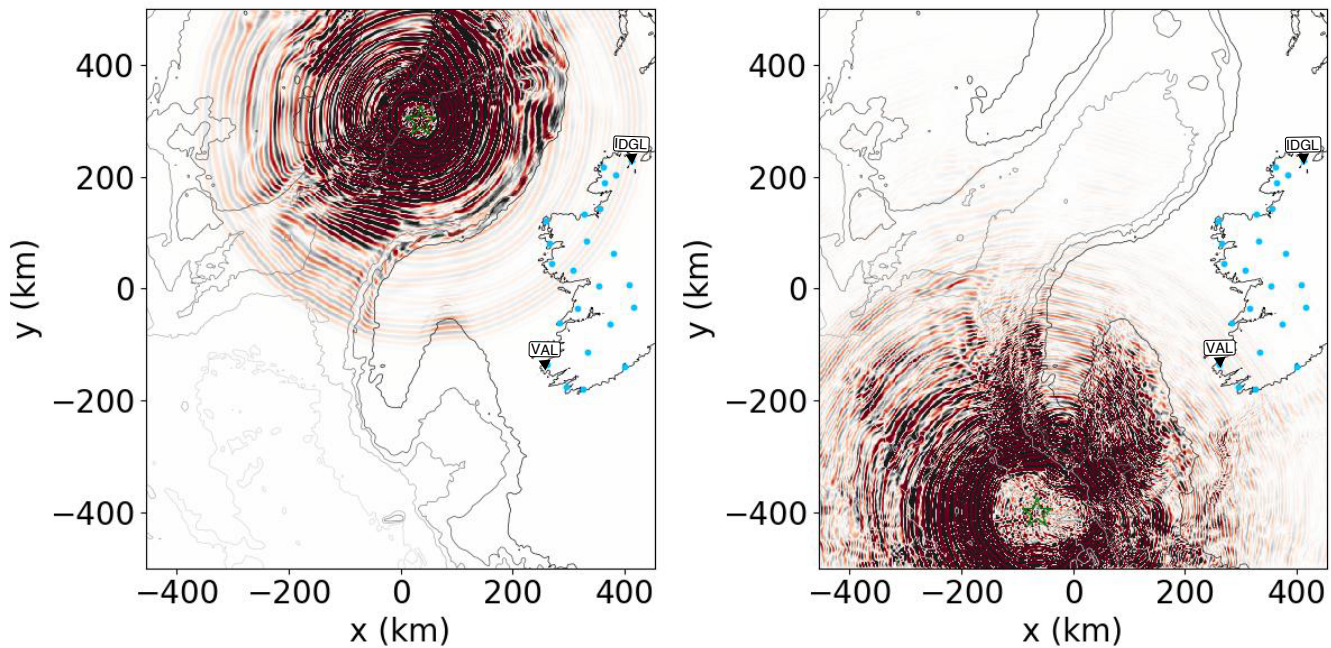


Figure 7 Snapshots of simulated secondary microseism wavefields. (Left) Northern source simulation after 62.2 s showing high energy lobes directed toward northwest Ireland; (Right) Southern source simulation after 125 s showing stronger wavefield distortion due to trapped waves in the Porcupine Basin and interactions with the continental shelf. Inverted triangles show the locations of the two synthetic INSN stations, used in Fig. 8. The light blue dots represent additional seismic receiver locations included in the numerical simulation to record the simulated wavefield. Additional snapshots for different times are available in Figure S10 in the supplementary material.

Seismic waves generated by ocean wave interactions propagate through heterogeneous crustal structures and varying bathymetry, introducing amplitude distortions and scattering effects before reaching on-land stations. These propagation effects can alter the seismic response independently of the source strength itself and may contribute to the slightly larger variability observed at some locations. The generally stable ANN performance therefore suggests that the network is able to learn and compensate for a substantial part of these complex propagation signatures, supporting the use of regional seismic networks as effective proxies for offshore ocean wave monitoring even in the absence of physical instrumentation in the ocean.

The Q-Q plots (Figs. 2 and 3 and Supplementary Figs. S1-S8) indicate that the ANN tends to underestimate the highest SWH values, particularly for extreme wave conditions above approximately 6–8 m. This behaviour is consistent with the tendency of regression models to favour the more densely represented range of the training distribution, while extreme wave events constitute a relatively small fraction of the available observations. Consequently, although the ANN reliably captures the timing and general evolution of high energy wave events, the predicted magnitude of the largest SWH peaks should be interpreted with caution. This limitation is particularly relevant for applications involving extreme weather monitoring, where accurate estimation of peak wave height is important. Increasing the representation of extreme events in the training dataset, or applying training strategies that place greater weight on high SWH observations, may improve

prediction of the upper tail of the SWH distribution. The present model should therefore be considered a complementary monitoring approach rather than a replacement for direct observations or operational wave forecasting systems when accurate quantification of extreme wave heights is required.

4.2 Location Variability and Wavefield Interpretation

The ANN performance varies between the investigated offshore locations (Figs. 5–6), although the spatial pattern does not indicate a simple dependence on source direction alone. For example, comparable performance at some locations with different azimuths, together with variations among the virtual locations, suggests that several geographically correlated factors may influence predictive skill. These include source–receiver distance, water depth and bathymetric setting, the position of the source relative to the continental shelf and shelf break, and propagation through heterogeneous crustal structure. Given the limited number of offshore locations considered here and the covariance between several of these parameters, their individual contributions cannot be uniquely separated from the ANN results alone.

To investigate the possible contribution of propagation effects of this variability, we performed numerical simulations of microseism propagation (Fig. 7). These simulations allow us to examine how source location influences the characteristics of the seismic wavefield recorded on land. The results show a clear contrast between northern and southern source scenarios. For

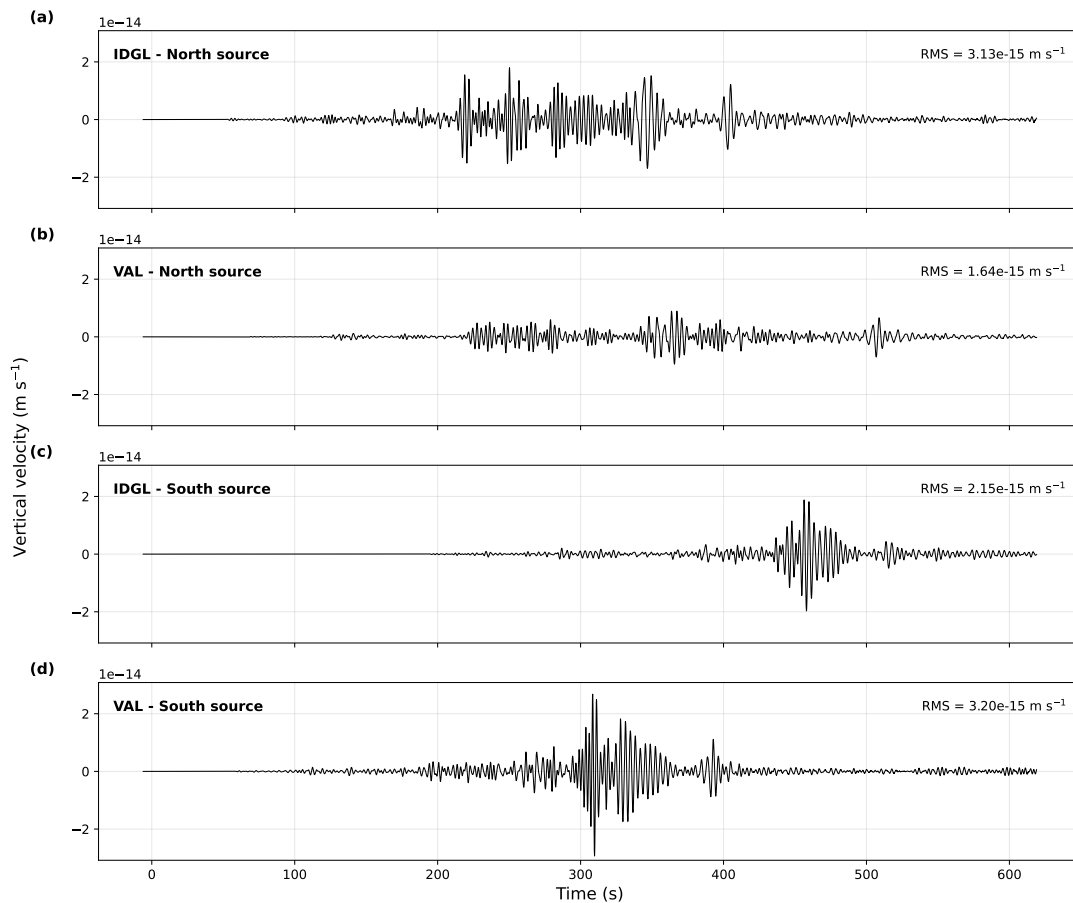


Figure 8 Synthetic vertical velocity seismograms at stations IDGL and VAL for the northern and southern microseism source scenarios. Panels show (a) IDGL and (b) VAL for the northern source, and (c) IDGL and (d) VAL for the southern source. The RMS velocity amplitude and source–receiver distance are indicated in each panel. Please refer to Fig. 1 for real station locations.

northern sources, the simulated wavefield remains relatively coherent, with focused energy propagating toward Ireland. In contrast, southern sources produce a more diffuse and scattered wavefield, with less coherent energy reaching the recording stations (Fig. 7). The snapshots further indicate that the incoming wavefield from the northern source remains comparatively coherent as it approaches Ireland, whereas the southern source produces a more fragmented wavefield with local portions of the wavefront exhibiting back azimuths that deviate from the source direction. The simulations also show substantial modification of the wavefield around the continental margin and shelf break, indicating that bathymetric and structural transitions along the propagation path may contribute to the observed differences.

The reduced coherence and increased scattering in the southern simulation are consistent with the lower ANN performance observed for the southern locations, suggesting that propagation effects may contribute to the spatial variability in the seismic–SWH relationship. However for computational reasons given the limited number of source scenarios considered here, the simulations do not allow the relative contributions of source–receiver distance, shelf and shelf-break geometry, bathymetry, and crustal heterogeneity to be uniquely separated. These factors may act together to modify the amplitude and coherence of the microseism

wavefield before it reaches the onshore seismic network. The simulations should therefore be interpreted as evidence that source–receiver propagation differs between the northern and southern scenarios, rather than as uniquely attributing the observed differences in ANN performance to a particular geological or bathymetric structure.

It should also be noted that the simulations represent idealised single-source scenarios, whereas the observed microseism wavefield may comprise the superposition of energy generated simultaneously across multiple source regions. Interference between these contributions may further modify the amplitudes and coherence observed at individual stations. Consequently, the simulations are intended primarily to isolate and illustrate differences in source-to-receiver propagation and should not be interpreted as reproducing the full amplitude characteristics or complexity of the observed microseism wavefield.

4.3 Implications for Seismological Ocean Monitoring

An additional complexity in seismic-based ocean monitoring arises from the non-linear nature of secondary microseism generation. Secondary microseisms result from interactions between ocean wave components of similar frequency travelling in approximately opposing

directions (Longuet-Higgins, 1950; Hasselmann, 1963). Their amplitude therefore depends on opposing ocean wave heights, but also on the directional wave spectrum and the strength of opposing wave interactions (Ardhuin et al., 2011, 2015). Consequently interacting ocean wave sets may produce different microseism excitation, depending on their interaction angles. The seismic amplitude subsequently recorded on land is further modified by the source location, bathymetry, propagation path, and local site response. Consequently, a given SWH does not correspond to a unique land recorded microseism amplitude and ANNs are well suited to this problem because they can learn this nonlinear, multi-variate relationship directly from observations at multiple seismic stations.

Beyond SWH estimation, the framework presented here could also contribute to the broader development of seismological ocean monitoring. Combining prediction residuals and spatial variations in ANN performance with numerical wavefield simulations may help identify regions or conditions where the observed seismic response departs from that expected from ocean-wave forcing alone. Future developments incorporating directional wave information and frequency dependent seismic attributes could further help separate variations associated with microseism generation from those introduced during propagation. In this way, terrestrial seismic networks may provide not only a complementary means of estimating offshore wave conditions, but also an observational framework for investigating ocean–solid Earth coupling and its spatial and temporal variability.

5 Conclusions

This study demonstrates that continuous land based seismic data can be used to estimate ocean significant wave height (SWH) using a data driven ANN framework. The combined use of multi location ANN models and numerical simulations of microseism generation confirms the strong link between secondary microseisms and ocean wave activity, and highlights the potential of terrestrial seismic networks for monitoring offshore wave conditions in the Northeast Atlantic.

Additionally, the results show that model performance varies spatially, with northern and northwestern locations exhibiting stronger agreement than sites further south. Numerical simulations indicate that differences in microseism wavefield propagation may contribute to this contrast, with southern sources producing more scattered and less coherent seismic energy. Despite strong microseism propagation effects the ANN can still yield SWHs that are a good fit to real observations and/or hindcast numerical ocean wave models.

These findings demonstrate that spatial variations in ANN performance are not simply a limitation of the modelling approach, but reflect fundamental differences in how microseism energy propagates from oceanic source regions to land based receivers. This highlights the importance of considering wavefield propagation effects when interpreting or applying seismic based SWH estimates, particularly in regions in-

fluenced by complex source geometries or propagation pathways. It should be noted that given the sensitivity to specific local microseism propagation effects, the trained ANN is likely not transferable to another geographical location, where retraining would be required.

Acknowledgements

Thanks to the Irish National Seismic Network (INSN), the Irish Met office, Ifremer, and the Laboratoire d'Océanographie Physique et Spatiale (LOPS) for providing us with the data used in this study. The authors sincerely acknowledge the official supervision and support of Martin White at the University of Galway. This work was funded by the Dublin Institute for Advanced Studies (DIAS) through a PhD stipend. The authors thank Laura Petrescu and two anonymous reviewers for their valuable suggestions that helped improve the manuscript.

Data and code availability

Seismic data are available at INSN data online (INSN, 1993), ocean buoy data by data requests to Irish Met office (Met Éireann), <https://www.met.ie/forecasts/marine-inland-lakes/buoys>, and ocean wave hindcast data are available at LOP-SWave Data <https://www.umar-lops.fr/Donnees/Vagues>. Codes and scripts used in this study are available at <https://github.com/BSamaneh/ANN-SWH.git>, and/or <https://doi.org/10.5281/zenodo.21428471>.

Competing interests

The authors declare no conflicts of interest. The sponsors had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Ardhuin, F., Stutzmann, E., Schimmel, M., and Mangeney, A. Ocean wave sources of seismic noise. *Journal of Geophysical Research: Oceans*, 116(C9), 2011. doi: 10.1029/2011jc006952.
- Ardhuin, F., Gualtieri, L., and Stutzmann, E. How ocean waves rock the Earth: Two mechanisms explain microseisms with periods 3 to 300 s. *Geophysical Research Letters*, 42(3):765–772, 2015. doi: 10.1002/2014gl062782.
- Bailey, E. S., Guillaumier, K., and Gauci, A. Enhancing Ocean Monitoring for Coastal Communities Using AI. *Applied Sciences*, 15 (19):10490, 2025. doi: 10.20944/preprints202508.0037.v1.
- Baranbooei, S., Bean, C. J., Rezaeifar, M., and Donne, S. E. Determining Offshore Ocean Significant Wave Height (SWH) Using Continuous Land-Recorded Seismic Data: An Example from the Northeast Atlantic. *Journal of Marine Science and Engineering*, 13(4):807, 2025. doi: 10.3390/jmse13040807.
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., Shapiro, N., and Yang, Y. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169 (3):1239–1260, 2007. doi: 10.1111/j.1365-246x.2007.03374.x.

- Bromirski, P. D. Vibrations from the “perfect storm”. *Geochemistry, Geophysics, Geosystems*, 2(7), 2001. doi: 10.1029/2000gc000119.
- Bromirski, P. D., Flick, R. E., and Graham, N. Ocean wave height determined from inland seismometer data: Implications for investigating wave climate changes in the NE Pacific. *Journal of Geophysical Research: Oceans*, 104(C9):20753–20766, 1999. doi: 10.1029/1999jc900156.
- Bromirski, P. D., Duennebieer, F. K., and Stephen, R. A. Mid-ocean microseisms. *Geochemistry, Geophysics, Geosystems*, 6(4), 2005. doi: 10.1029/2004gc000768.
- Cannata, A., Cannavò, F., Moschella, S., Di Grazia, G., Nardone, G., Orasi, A., Picone, M., Ferla, M., and Gresta, S. Unravelling the relationship between microseisms and spatial distribution of sea wave height by statistical and machine learning approaches. *Remote Sensing*, 12(5):761, 2020. doi: 10.3390/rs12050761.
- Cheung, K. WaveWatch III (WW3) Global Wave Model [ATNE], updated 2021. *Distributed by the Pacific Islands Ocean Observing System (PacIOOS)*, accessed June 2024, 2010. http://pacioos.org/metadata/ww3_global.html.
- Chevrot, S., Sylvander, M., Benahmed, S., Ponsolles, C., Lefèvre, J., and Paradis, D. Source locations of secondary microseisms in western Europe: Evidence for both coastal and pelagic sources. *Journal of Geophysical Research: Solid Earth*, 112(B11), 2007. doi: 10.1029/2007jb005059.
- Donne, S., Nicolau, M., Bean, C., and O’Neill, M. Wave height quantification using land based seismic data with grammatical evolution. In *2014 IEEE Congress on Evolutionary Computation (CEC)*, pages 2909–2916. IEEE, 2014. doi: 10.1109/cec.2014.6900563.
- Hasselmann, K. A statistical analysis of the generation of microseisms. *Reviews of Geophysics*, 1(2):177–210, 1963. doi: 10.1029/rg001i002p00177.
- Iafolla, L., Fiorenza, E., Chiappini, M., Carmisciano, C., and Iafolla, V. A. Sea wave data reconstruction using micro-seismic measurements and machine learning methods. *Frontiers in Marine Science*, 9:798167, 2022. doi: 10.3389/fmars.2022.798167.
- INSN. Irish National Seismic Network (INSN) Operated by the Dublin Institute for Advanced Studies and Supported by the Geological Survey Ireland. In *International Federation of Digital seismograph networks*, 1993. doi: 10.7914/SN/EI.
- Komatitsch, D. and Tromp, J. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, 139(3):806–822, 1999. doi: 10.1046/j.1365-246x.1999.00967.x.
- Le Pape, F., Craig, D., and Bean, C. J. How deep ocean-land coupling controls the generation of secondary microseism Love waves. *Nature Communications*, 12(1):2332, 2021. doi: 10.1038/s41467-021-22591-5.
- Longuet-Higgins, M. S. A theory of the origin of microseisms. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 243(857):1–35, 1950. doi: 10.1098/rsta.1950.0012.
- Met Éireann. Latest Buoy reports for M3, M4, and M6. In <https://www.met.ie/forecasts/marine-inland-lakes/buoys>, 2024.
- Möllerhoff, M. and Bean, C. J. Seismic noise characterization in proximity to strong microseism sources in the Northeast Atlantic. *Bulletin of the Seismological Society of America*, 106(2):464–477, 2016. doi: 10.1785/0120150204.
- Moni, A., Craig, D., and Bean, C. J. Separation and location of microseism sources. *Geophysical Research Letters*, 40(12): 3118–3122, 2013. doi: 10.1002/grl.50566.
- Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., et al. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes. *Geophysical Journal International*, 186(2):721–739, 2011. doi: 10.1111/j.1365-246x.2011.05044.x.
- Stutzmann, E., Ardhuin, F., Schimmel, M., Mangeney, A., and Patau, G. Modelling long-term seismic noise in various environments. *Geophysical Journal International*, 191(2):707–722, 2012. doi: 10.1111/j.1365-246x.2012.05638.x.
- The MathWorks Inc. Optimization Toolbox version: 9.4 (R2022b). <https://www.mathworks.com>, 2022. Natick, Massachusetts, United States.
- Ying, Y., Bean, C. J., and Bromirski, P. D. Propagation of microseisms from the deep ocean to land. *Geophysical Research Letters*, 41(18):6374–6379, 2014. doi: 10.1002/2014gl060979.

The article *Monitoring Offshore Ocean Wave Heights from Land Seismic Networks: Multi-Location ANN Estimations and Insights from Microseism Wavefield Simulations* © 2026 by Samaneh Baranbooei is licensed under CC BY 4.0.