Environmental Data Science (2026), 5: e9, 1–26
doi:10.1017/eds.2026.10036

CAMBRIDGE UNIVERSITY PRESS

APPLICATION PAPER

Data-driven discovery of meteotsunami patterns from sparse observations

Ardiansyah Fauzi[1], Emiliano Renzi[1], Frederic Dias[2,3], Daniel Santiago Pelaez-Zapata[3] and Tatjana Kokina[4]

[1]Mathematics of Complex and Nonlinear Phenomena (MCNP), School of Engineering, Physics and Mathematics, Northumbria University, UK
[2]School of Mathematics and Statistics, University College Dublin, Ireland
[3]Centre Borelli, ENS Paris-Saclay, France
[4]Met Eireann, Ireland
Corresponding author: Ardiansyah Fauzi; Email: ardiansyah.fauzi@northumbria.ac.uk

Received: 20 May 2025; **Revised:** 19 February 2026; **Accepted:** 06 March 2026

Keywords: coastal hazard monitoring; data-driven modeling; dynamic mode decomposition; meteotsunamis; sparse observations

Abstract

Meteotsunamis—tsunami-like sea level oscillations generated by atmospheric disturbances—pose underestimated risks to coastal regions worldwide. Despite growing evidence of their frequency and impact, limited offshore observations and forecasting capabilities hinder effective monitoring and early detection. Here, we present a data-driven framework for identifying and characterizing meteotsunami dynamics using sparse observational data. Leveraging dynamic mode decomposition and clustering techniques, we extract dominant spatiotemporal patterns and optimize the placement of offshore monitoring stations. We demonstrate the effectiveness of this approach using high-resolution simulations of the 2022 Ireland meteotsunami, a well-documented event exhibiting clear atmospheric forcing and sea-level response. Our results show that a minimal network of five strategically positioned sensors can accurately capture the essential dynamics of the event. This framework establishes a scalable methodology for designing cost-effective monitoring systems, enhancing our ability to detect and understand meteotsunamis under data-scarce conditions.

Impact Statement

Meteotsunamis are hazardous ocean waves triggered by atmospheric disturbances, yet they remain undermonitored due to sparse offshore observations. This study presents a data-driven framework that combines dynamic mode decomposition, numerical modeling, and data assimilation to optimize the detection of meteotsunamis in data-scarce regions. Using the 2022 Ireland meteotsunami as a case study, we demonstrate that a minimal, strategically positioned sensor network can effectively capture essential wave dynamics. Our methodology supports cost-effective offshore monitoring and is adaptable to other hazard-prone coastal regions. The approach is scalable, cost-efficient, and applicable to other vulnerable coastal areas. By enhancing early detection and forecasting capabilities, this work provides policymakers with valuable insights, ultimately improving coastal resilience to rapidly evolving oceanic hazards in a changing climate.

This Application Paper was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

1. Introduction

Meteotsunamis are tsunami-like waves generated by atmospheric disturbances. Despite their seemingly rare nature, meteotsunamis pose significant risks to coastal communities, yet remain inadequately monitored. Their sudden onset, coupled with limited observational infrastructure and forecasting tools, makes them particularly challenging to detect and predict. In this study, we present a rigorous approach that combines high-resolution numerical modeling, dynamic mode decomposition (DMD), and data assimilation (DA) techniques to better characterize meteotsunami dynamics and optimize offshore monitoring strategies. Using the 2022 Ireland meteotsunami as a case study, we demonstrate how data-driven insights can inform the strategic placement of observational stations, enhancing early warning systems and improving coastal resilience. Our framework not only advances the scientific understanding of meteotsunamis but also offers a scalable methodology for hazard-prone regions worldwide.

Meteotsunamis are wave oscillations caused by atmospheric disturbances that share characteristics similar to those of seismically generated tsunamis (Monserrat et al., 2006). The generation of meteotsunamis on open coasts is typically a multi-resonant process initiated by a sudden atmospheric pressure jump, usually only a few hPa, or by wind disturbances associated with moving weather systems such as storms, squalls, or frontal passages. This initial water level perturbation, primarily driven by the inverse barometer effect, is amplified as it propagates with the triggering disturbance through external resonance processes and nearshore wave transformations, such as shoaling, refraction, and wave reflection.

Proudman resonance (Proudman, 1929) plays a crucial role. It is typically the main phenomenon in the initial generation of meteotsunamis, where the maximum energy transfer occurs when the speed of the atmospheric pressure disturbance is close to the celerity of shallow water waves. Additionally, resonances such as the Greenspan resonance (Greenspan, 1956), whereby the atmospheric perturbation travels at a speed close to that of edge wave modes along the coast, combined with local bathymetry and harbor seiche matching (Denamiel et al., 2018), can significantly amplify wave heights, sometimes reaching destructive levels of several meters. In certain cases, the Greenspan resonance may play a more dominant role than the Proudman resonance (e.g., Kim et al., 2022).

Meteotsunamis have been reported in regions such as the Mediterranean, the US East Coast, and the Great Lakes, where well-documented events have enabled deeper scientific understanding. In contrast, the Celtic Sea remains largely understudied, with meteotsunamis in this region often perceived as rare and relatively benign. This geographic gap presents a valuable opportunity to advance knowledge in a poorly characterized yet potentially vulnerable coastal setting. Lewis et al. (2023) reported that meteotsunamis occur more frequently in the vicinity of the Celtic Sea than previously believed. With relatively flat bathymetry (~100 m), the region is susceptible to Proudman resonance when atmospheric disturbances travel at 20–50 m/s—yet events here remain under-detected and under-studied. One key reason is the limited offshore monitoring infrastructure: only two buoys (M3 and M5) are positioned near the study area, both with coarse, one-hour sampling intervals (Figure 1). Such low temporal resolution may miss meteotsunami signals with periods of just a few hours.

Unlike earthquake-prone regions—where dense offshore networks like Japan's DONET (Baba et al., 2014; Kaneda et al., 2015) and S-Net (Yamamoto et al., 2016) or the U.S. DART system (Titov et al., 2005) reduce uncertainty in tsunami source modeling—meteotsunami-prone areas lack comparable systems. Yet meteotsunami forecasting faces similarly high uncertainty, particularly in estimating the origin, shape, and propagation of atmospheric pressure disturbances. Most weather stations are land-based, offering little insight into oceanic forcing. Expanding offshore weather observations is therefore essential to resolving atmospheric inputs and enabling accurate transformation into coastal sea level responses via numerical modeling. This study addresses these challenges by proposing a data-driven methodology to optimize offshore sensor placement, using meteotsunami simulations and spatiotemporal decomposition to overcome observational sparsity and improve predictive capacity.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

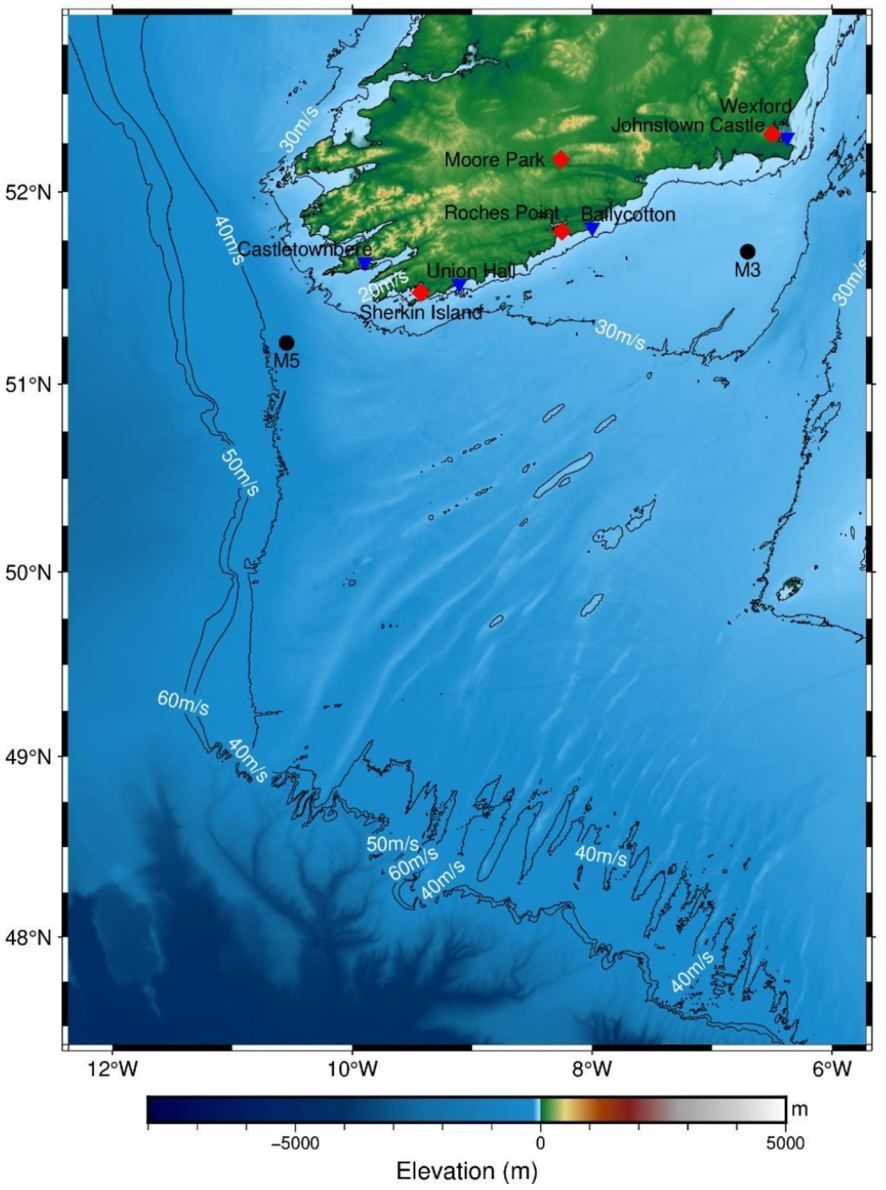


Figure 1. *Bathymetry and topography data with ~15 arcsec resolution. Blue inverted triangles and red diamonds indicate the location of tide gauges and AWSs, respectively. Black circles denote the location of the existing offshore buoys. The contour lines represent the shallow water wave celerity* $c = \sqrt{gd}$ *in terms of water depth* d*.*

In summary, this work introduces a data-driven framework for meteotsunami monitoring that unifies numerical forward simulations, DMD, and DA. The main findings and contributions are as follows:

- An integrated numerical simulation—DMD–DA framework combines nonlinear shallow-water modelling, mode decomposition, and data assimilation to extract dominant meteotsunami dynamics from sparse observations.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

- The approach represents one of the first applications of both DMD and DA to meteotsunamis, enabling the identification of coherent spatiotemporal patterns and supporting the optimization of offshore observational networks.
- Application to the 18 June 2022 Ireland meteotsunami shows that a minimal array of five strategically positioned sensors can reproduce the essential wave dynamics with accuracy comparable to a dense network.
- The framework provides a scalable and cost-efficient methodology that can be generalized to improve meteotsunami detection and forecasting in other data-sparse coastal regions.

The remainder of this paper is organized as follows. Section 2 presents the related works and background concepts that support the proposed framework. Section 3 describes the numerical model setup and atmospheric forcing representation. Section 4 explains the implementation of DMD and DA. Section 5 presents and discusses the results, while Section 6 provides the concluding remarks.

2. Background and related work

2.1. Related works

Previous studies have investigated meteotsunamis in different regions of the world, revealing both their common generation mechanisms and their distinct local features. Early works identified the principal resonant mechanisms responsible for amplification—Proudman resonance (Proudman, 1929), Greenspan resonance (Greenspan, 1956), and harbor resonance (Denamiel et al., 2018). Numerical and observational analyses, such as those by Hibiya and Kajiura (1982) in Nagasaki Bay and Vilibić et al. (2008) in the Adriatic, it showed how bathymetry and atmospheric forcing jointly shape meteotsunami behavior. More recent studies have explored meteotsunami dynamics in various basins, including the Adriatic Sea (Denamiel et al., 2019a, 2019b), the Yellow Sea (Kim et al., 2022), the coast of Portugal (Kim and Omira, 2021), and the United Kingdom (Thompson et al., 2020; Lewis et al., 2023). Together, these works demonstrate that meteotsunamis are more frequent and globally distributed than once believed (Gusiakov, 2021).

Numerical models based on the nonlinear shallow-water equations (NLSWE) remain the standard tool for simulating tsunami and meteotsunami propagation. Recent developments have also highlighted the potential of combining physical models with data-driven techniques to improve computational efficiency and interpretability (Fauzi and Mizutani, 2020a; Andraud, 2025).

From the observational perspective, dense offshore networks such as DONET and S-Net in Japan (Baba et al., 2014; Kaneda et al., 2015; Kubota et al., 2021) and the DART system (Titov et al., 2005) have significantly improved tsunami detection and early warning. However, their high cost and maintenance demands (Bernard and Titov, 2015) limit their feasibility for meteotsunami monitoring. This limitation motivates research into optimal sensor placement methods that can maximize information from a minimal number of stations.

In recent years, data-driven techniques have increasingly been applied to tsunami science. DA has been used to integrate model predictions with observations in real time (Kalnay, 2002; Maeda et al., 2015; Byrne et al., 2023). Similar hybrid approaches combining DA with machine-learning frameworks have also been explored for rapid tsunami wavefield reconstruction (Fauzi and Mizutani, 2020b). For optimizing sensor networks, most tsunami studies have relied on Empirical Orthogonal Function (EOF) analysis to identify spatial patterns that capture the largest variance (Mulia et al., 2017, 2019; Ren et al., 2022). However, EOF assumes stationary linear modes, which may not fully describe the nonstationary and time-dependent characteristics of meteotsunamis forced by evolving atmospheric disturbances. To address this, DMD has emerged as a promising approach. DMD captures time-dependent structures and dominant frequencies within complex flow fields (H. Tu et al., 2014) and has recently been applied to atmospheric and oceanic datasets (Gugole and Franzke, 2020; Li et al., 2023). Building upon these developments, the present study integrates DMD and DA to identify dynamically significant regions for offshore monitoring and to reconstruct meteotsunami wavefields from sparse observations.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

2.2. Background concept

The numerical modelling of meteotsunamis is commonly based on the NLSWE with atmospheric pressure as an external forcing term. These equations describe the depth-averaged conservation of mass and momentum, allowing simulation of long-wave propagation in coastal and shelf regions. The sea-surface displacement is governed by hydrostatic balance, responding to spatial and temporal variations in air pressure through the inverse-barometer effect. When the propagation speed of the atmospheric disturbance approaches the wave celerity $c = \sqrt{gd}$, where d is the local ocean depth, Proudman resonance occurs, producing resonant energy transfer from air to sea. Additional amplification mechanisms, such as Greenspan and harbor resonances, depend on coastal geometry and bathymetry (Denamiel et al., 2018).

Atmospheric pressure forcing is typically represented using idealized Gaussian-shaped perturbations (Anarde et al., 2021) or by interpolating observational data (Kim and Omira, 2021; Kim et al., 2022). These formulations enable the reconstruction of realistic atmospheric inputs for numerical simulations and are widely used for assessing meteotsunami generation and propagation.

To extract dominant spatiotemporal patterns from the simulated fields, DMD is employed as the primary decomposition method within the framework. Coupling DMD with sequential DA (Kalnay, 2002; Maeda et al., 2015) allows model states to be updated using sparse observations, enhancing both reconstruction accuracy and monitoring efficiency.

These physical and methodological concepts form the foundation of the present framework, which combines NLSWE-based simulation, DMD analysis, and DA to improve meteotsunami detection and forecasting in data-sparse coastal regions. In the following, we describe the 2022 Ireland meteotsunami and associated modelling choices.

3. The 2022 Ireland meteotsunami

On June 18, 2022, a sudden change in sea surface elevation was reported by locals along the southern coast of Ireland. A preliminary study by Renzi et al. (2023) identified the presence of a high-intensity jet stream in the region on the day of the event. Additionally, high relative humidity was observed along the jet stream's border—conditions that closely resemble those of previously studied meteotsunami events. To better understand the characteristics and underlying causes of this event, this study analyzes key observational datasets. The two primary datasets used are tidal observations and air pressure measurements.

Tidal data were obtained from the Marine Institute of Ireland (www.marine.ie) and recorded at four tide gauges along the southern coast: Castletownbere, Ballycotton, Union Hall, and Wexford, each with a sampling interval of 5 minutes (Figure 1). Although the M3 and M5 buoys also provided air pressure data, they were excluded from the analysis due to their low sampling interval (1 h), which is insufficient for capturing the shorter-period atmospheric disturbances. The previous study applied a single high-pass filter with a 4-h cutoff period (~0.000069 Hz), which removes long-period tidal components but may still retain higher-frequency variability. In contrast, our study used a band-pass Butterworth filter with a high-pass cutoff of 0.00008 Hz (~3.47 h) and a low-pass cutoff of 0.003 Hz (~0.93 h) to better isolate the meteotsunami band. The additional low-pass cutoff suppresses short-period noise, resulting in a smoother signal focused on the 15–60 min frequency range typical of meteotsunamis. Since determining the optimal cutoff frequency is sensitive to both tide-gauge and pressure data, we performed iterative forward simulations to identify the range that best reproduced the observed waveforms. Additionally, wavelet spectrum analysis was applied to confirm that the recorded waves exhibited long-period characteristics, which are a defining feature of meteotsunamis. We use the Morlet continuous wavelet spectrum (Torrence and Compo, 1998), which has also been used in many previous studies (Kim and Omira, 2021; Kim et al., 2022; Renzi et al., 2023). The original, filtered, and wavelet spectrum analyses are shown in Figure 2.

From the filtered waveforms, the first small wave was recorded at Castletownbere around 12:25 UTC, followed by a larger wave at 15:00 UTC, reaching a maximum crest-to-trough wave height of 0.25 m. Similar wave patterns were subsequently observed at other stations, with maximum crest-to-trough wave

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

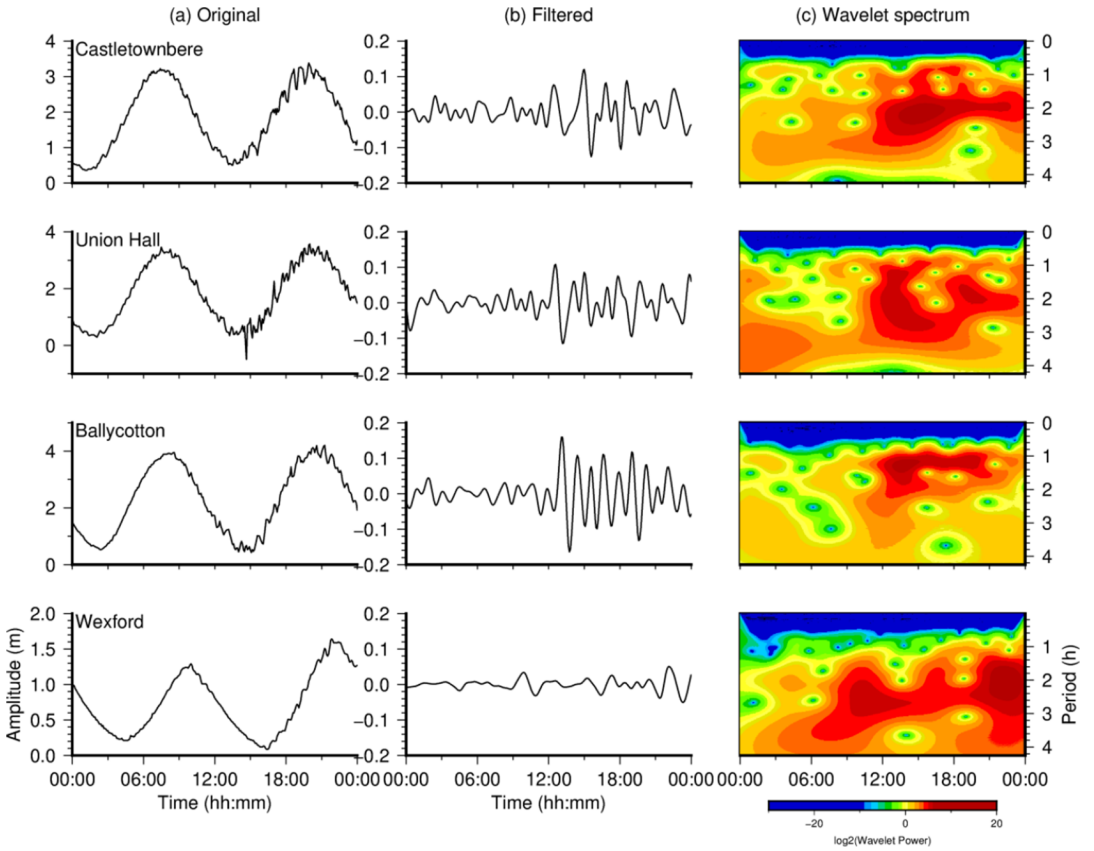


Figure 2. *Waveforms recorded at four tide gauges on 18 June 2022. (a) The original waveforms. (b) the filtered waveforms. (c) Morlet wavelet spectra.*

heights of 0.22 m at 12:35 UTC in Union Hall, 0.32 m at 13:05 UTC in Ballycotton, and 0.10 m at 22:10 UTC in Wexford. Tsunami waves typically exhibit wave periods between a few minutes and nearly 3 h (Rabinovich, 2020). The wavelet spectrum analysis in Figure 2c provides clear evidence that the recorded waves of the 2022 Ireland event are within the tsunami frequency range, where the wave periods at Castletownbere, Union Hall, and Wexford range between 1–3 h, while 1–2 h at Ballycotton.

We also gathered high-resolution 1-min sampling frequency atmospheric pressure data from the Irish Meteorological Service (www.met.ie) at four AWSs, including Sherkin Island, Roches Point, Moore Park, and Johnstown Castle. These AWSs are situated between 21 and 62 m above mean sea level and located across the south and eastern regions of Ireland (see Figure 1). To localize the sudden pressure changes during the event, we applied a high-pass filter to eliminate the low-frequency components of the pressure. Using a fifth-order Butterworth high-pass filter, we carefully choose a cutoff period of 0.001 Hz (~0.278 h). Recorded air pressure variations provide strong evidence for the presence and propagation of pressure jumps during the event. The comparison of the original, filtered, and their associated wavelet spectra is shown in Figure 3.

Analysis of the filtered pressure data reveals a significant pressure change first observed at Sherkin Island AWS between approximately 12:00 and 13:00 UTC. The pressure dropped from −0.59 hPa at 11:59 UTC and suddenly increased to 1.92 hPa at 12:52 UTC, resulting in a maximum pressure jump of 2.52 hPa within an hour. A similar pattern was subsequently recorded at Roches Point, Moore Park, and Johnstown Castle AWS. At Roches Point, the maximum pressure jump was 3.33 hPa, increasing from −1.28 hPa at 12:30 UTC to 2.05 hPa at 13:11 UTC. At Moore Park, pressure rose from −0.56 hPa at 12:39

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

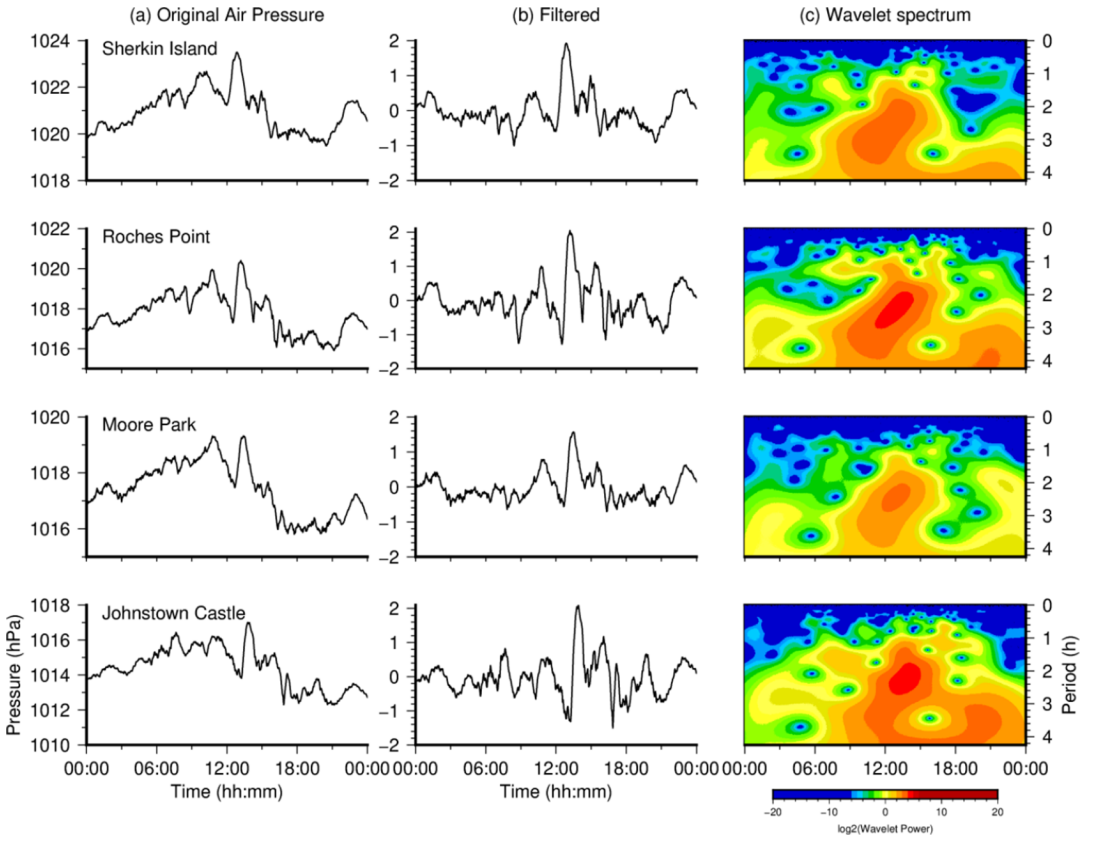


Figure 3. *Air pressure disturbances recorded at four AWSs on 18 June 2022. (a) The original air pressure. (b) the filtered air pressure. (c) Morlet wavelet spectra.*

UTC to 1.57 hPa at 13:29 UTC, yielding a maximum jump of 2.13 hPa. Meanwhile, at Johnstown Castle, the pressure increased from −1.32 hPa at 13:14 UTC to 2.09 hPa at 13:55 UTC, marking the highest observed jump of 3.40 hPa. Further wavelet spectrum analysis confirms the long-wave characteristics, with wavelet power indicating wave periods between 1 and 4 h.

The arrival time of the first wave crest at Castletownbere and Union Hall falls within the time range of the pressure jump recorded at Sherkin Island AWS, the closest location. This temporal coincidence strongly suggests that air pressure fluctuations played a key role in driving the sea surface changes. A similar pattern was observed at Ballycotton, which correlated well with pressure variations at Roches Point AWS. However, Wexford displayed a different response, showing a distinct correlation with Johnstown Castle AWS. These findings indicate that atmospheric pressure disturbances were likely a primary forcing mechanism behind the 18 June event, with regional variations in pressure gradients influencing the water level at different tide gauge stations.

3.1. Numerical model description: governing equations

We use a numerical model based on a non-linear shallow water equation (NLSWE) with atmospheric pressure as the source model. NLSWE-based models have been widely used in meteotsunami studies. For example, Vilibić et al. (2008), Šepić et al. (2018a), and Kim and Omira (2021) employed depth-averaged NLSWE models with idealized atmospheric pressure forcing to simulate meteotsunami dynamics, validating their results against field observations. We consider a spherical coordinate system (R, θ, λ) with the origin at the Earth's center. The Earth is assumed to be a sphere, hence the radius R is constant and

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

equal to the Earth's radius, whereas θ is latitude (measured southwards) and λ is longitude (measured eastwards). The governing equations of NLSWE are given by the following expressions:

$$\frac{\partial D}{\partial t}+\frac{1}{R\sin\theta}\frac{\partial(uD)}{\partial\lambda}+\frac{1}{R}\frac{\partial(vD)}{\partial\theta}=0, \tag{1}$$

$$\frac{\partial u}{\partial t}+\frac{u}{R\sin\theta}\frac{\partial u}{\partial\lambda}+\frac{v}{R}\frac{\partial u}{\partial\theta}+\frac{g}{R\sin\theta}\frac{\partial h}{\partial\lambda}+C_f\frac{u\sqrt{u^2+v^2}}{D}+\frac{1}{\rho R\sin\theta}\frac{\partial P}{\partial\lambda}=0, \tag{2}$$

$$\frac{\partial v}{\partial t}+\frac{u}{R\sin\theta}\frac{\partial v}{\partial\lambda}+\frac{v}{R}\frac{\partial v}{\partial\theta}+\frac{g}{R}\frac{\partial h}{\partial\theta}+C_f\frac{v\sqrt{u^2+v^2}}{D}+\frac{1}{\rho R}\frac{\partial P}{\partial\theta}=0, \tag{3}$$

$$C_f=\frac{gn^2}{D^{⅓}}. \tag{4}$$

Manning's roughness coefficient n is used to calculate the non-dimensional frictional coefficient C_f. A constant value of $n=0.025\ \mathrm{m^{1/3}s^{-1}}$ is applied homogeneously in the whole computational domain. This value is widely used in ocean wave propagation modeling (Imamura et al., 2006). In equations (1)–(4), u and v are depth-averaged horizontal velocity fluxes along the latitude and longitude axes, respectively, P is air pressure, t is time, g is the gravitational constant, and D is the total water depth, which is a summation of water depth d and sea surface amplitude h. We assume that seawater density $\rho \sim 1025\ \mathrm{kg/m^3}$. The edge of the computational domain is used as an open boundary and the Coriolis effect is neglected in this study. The computational time step should satisfy the Courant–Friedrichs–Lewy (CFL) condition, whereby the time step Δt must be equal to or smaller than the time required by a wave to travel on the spatial grid size.

3.2. Air pressure forcing

Air pressure modeling is a critical variable in meteotsunami simulations and represents a key distinction from seismic tsunami modeling. In seismically generated tsunamis, once the hypocenter coordinates are known, the source model is typically estimated using waveform inversion of either oceanic or seismic waves (Satake, 1989; Gusman and Tanioka, 2013). This is possible not only due to the availability of observing sensors but also because seismic waves travel through the Earth's layers and can be detected globally (Obara, 2007; Gee and Leith, 2011; Di Giacomo et al., 2014). However, estimating a tsunamigenic atmospheric perturbation, such as an air pressure jump, is challenging due to the limited spatial coverage of observational stations. As a consequence, there is significant uncertainty regarding the shape, propagation direction, and speed of the pressure perturbation. Gaussian shapes are commonly used to represent the shape of the pressure jump (Anarde et al., 2021). A more realistic source model can be constructed by interpolating between observed pressure values (Kim and Omira, 2021; Kim et al., 2022).

Due to the limited information on the spatial distribution of air pressure, this study employs a Gaussian-shaped profile to estimate the characteristics of air–pressure disturbances.

$$P_{i,j,t}=\sum_{k=1}^{N}P_k\,exp\left[\left(-\frac{x_{i,j}sin\phi+y_{i,j}cos\phi-V_o t}{\sigma_{k/2}}\right)^2\right], \tag{5}$$

where P is air pressure on a two-dimensional grid, P_k is the initial amplitude of the kth pressure component (with a total of N components), x and y are longitude and latitude coordinates, ϕ is the propagation direction of the pressure disturbance measured counterclockwise from the x-axis, V_o is the propagation speed, t is the numerical time step, and σ is the spread of the kth component, which controls the width of the Gaussian distribution. Similar to Anarde et al. (2021), we create a synthesized representation of the physical pressure perturbation by combining individual components into a smooth, continuous two-dimensional distribution. A total of four crests and troughs, each with distinct P_0 and σ values, are identified from the pressure records. The spatial distribution of the AWSs allows for a quantitative estimation of the direction and propagation speed of the air-pressure disturbance. Based on the locations

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

of the AWSs (Figure 1) and the clear sequential pattern of maximum pressure jumps over time (Figure 3), we conclude that the pressure disturbances propagated in a northeastward direction, in agreement with Renzi et al. (2023). We assume that the pressure disturbance propagates at a constant speed and angle. This assumption is widely used in the previous study (Šepić et al., 2018b; Anarde et al., 2021; Kim and Omira, 2021).To determine the propagation speed and direction, the process typically begins with hypothetical values, followed by multiple trial-and-error iterations to best fit the observations. In this study, we first identified the time of maximum air-pressure perturbation recorded at each AWS along the track of the synoptic disturbance. The average propagation speed was then calculated as the mean distance between Sherkin Island and the other AWSs divided by the corresponding time lag between their peak-pressure records, yielding an average speed of 55.4 m/s. The propagation direction (44° from the west) was estimated from the linear alignment of these stations and the spatiotemporal sequence of the pressure peaks. The parameter estimation workflow is summarized in Figure 4, illustrating the stepwise process from AWS observations to the synthesized pressure field.

To illustrate the resonance condition, Figure 1 includes contours of the local long-wave phase speed $c = \sqrt{gd}$. The region where the atmospheric disturbance speed matches the oceanic wave celerity indicates where Proudman resonance is expected to occur—along the 50–60 m/s contour southwest of Ireland. Figure 5 presents the comparison between simulated and observed air pressure at the four AWS locations. To quantify the agreement between observed and simulated air pressure, we use the normalized cross-correlation coefficient (NCC), where values closer to 1 indicate a strong correlation. Additionally, we evaluate the root mean square error (RMSE). Overall, the simulated air pressure demonstrates good agreement with observations, yielding an NCC value of 0.83 and an RMSE of 0.38 hPa across all AWS locations.

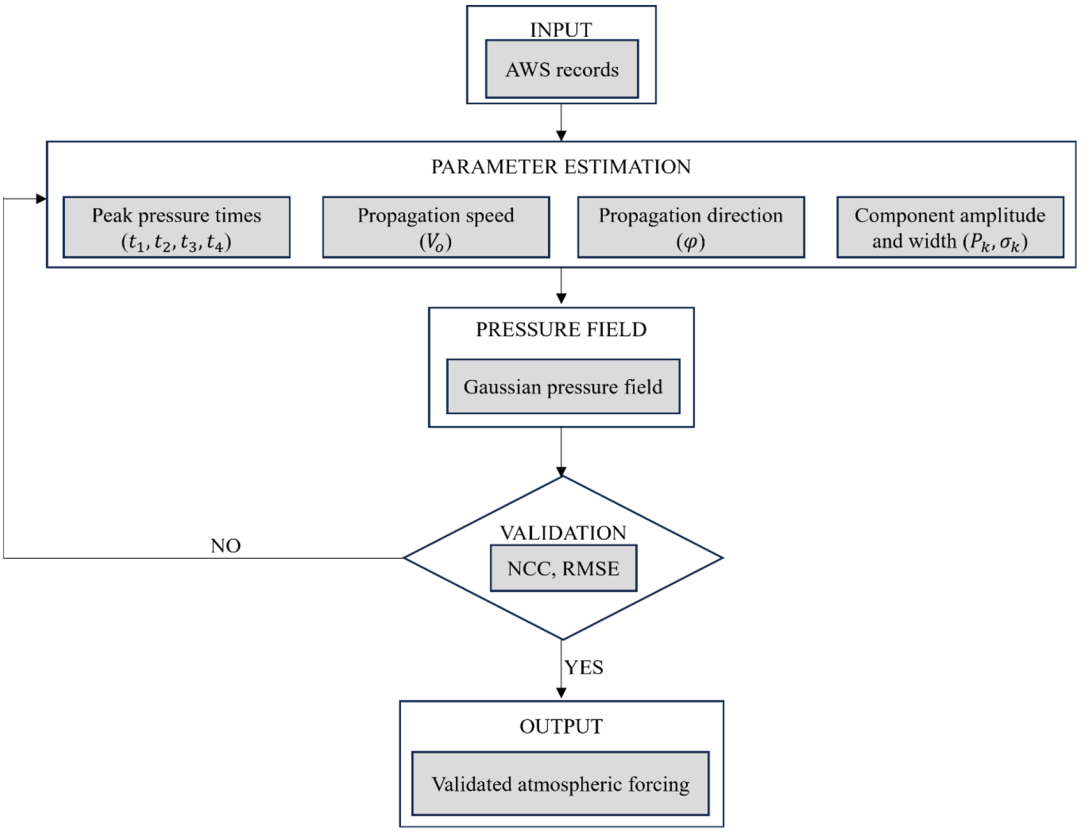


Figure 4. *Atmospheric pressure parameter estimation workflow.*

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

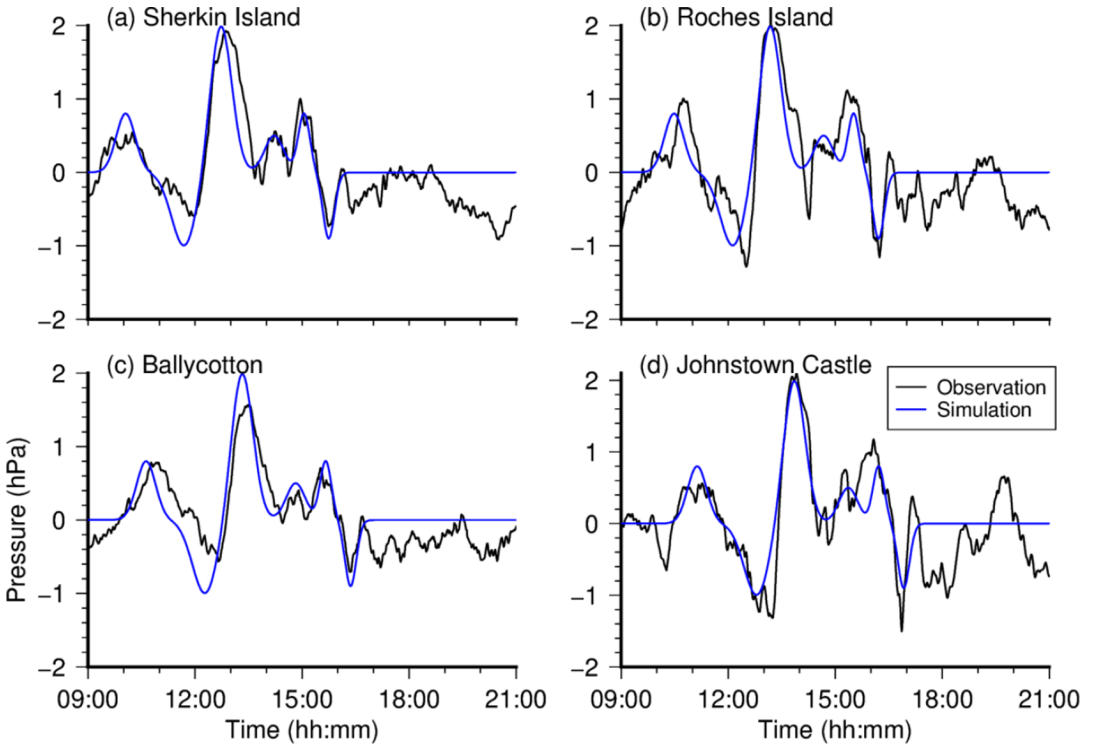


Figure 5. *Comparisons between simulated and observed air pressure at four AWSs.*

3.3. Meteotsunami generation and propagation

At the beginning of the numerical simulation (at $t = 0$), the air pressure disturbance is transformed into a two-dimensional grid to create an initial sea surface using hydrostatic approximation,

$$h = -\frac{\Delta P}{\rho g}. \tag{6}$$

The topography and bathymetry dataset are obtained from the GEBCO 2024 grid with a resolution of 15 arcsec (~450 m), see Figure 1. To satisfy the CFL condition, the numerical time step Δt is set to 1 s. We run the 12 h simulation time using simulated air pressure from 18 June 2022, 09:00 to 21:00 UTC to observe the dynamics of meteotsunami during the generation and propagation phases.

The snapshots of pressure and wave propagation over the computational domain are shown in Figure 6, whereas the comparison between observed and simulated waves at four coastal gauges is presented in Figure 7. Figure 6 shows that meteotsunami propagation is directly guided by pressure forcing along its propagation path. Our numerical model also captures wave amplification due to local bathymetry. Initially, the model generates only a few centimeters of sea surface displacement in the offshore region, which then propagates toward the coast, increasing in magnitude as a result of nearshore amplification processes. As shown in Figure 7, the simulated waveforms exhibit good agreement with observations at Castletownbere, although the model slightly overestimates both the maximum and minimum wave amplitudes. The simulated maximum amplitude is 0.17 m, overestimating observations only by 0.05 m, and the overall waveform pattern closely aligns with the observed data. Similarly, comparisons at Union Hall and Ballycotton show a strong correlation, despite a minor temporal shift in the timing of the maximum wave amplitude. The maximum wave amplitudes are 0.12 m at Union Hall and 0.15 m at Ballycotton, with overestimations of 0.015 m and underestimations of 0.01 m, respectively. At Wexford, however, the model overestimates the observed wave heights, suggesting potential discrepancies in model

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

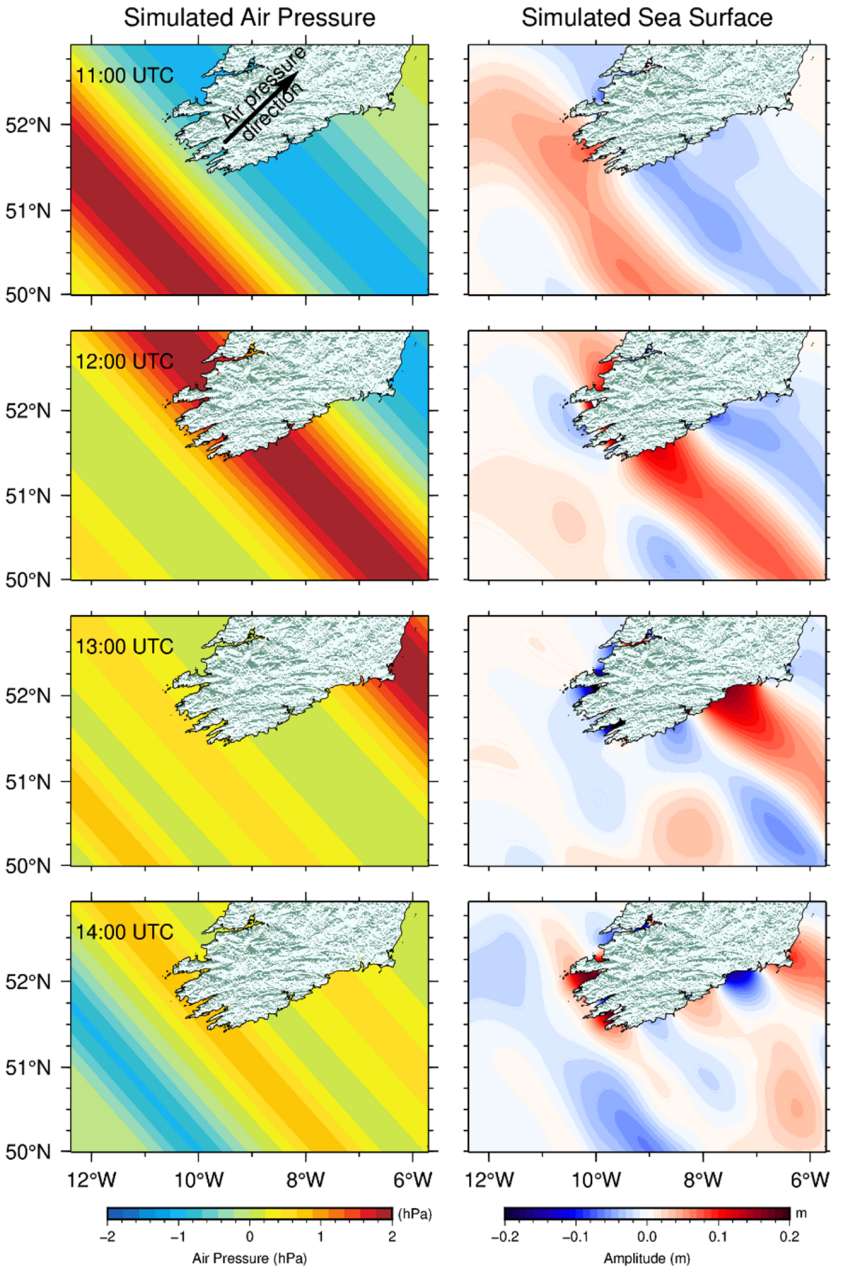


Figure 6. *Snapshots of simulated air pressure and the influenced sea surface over computational domain.*

accuracy at this location, which will be discussed in Section 4. The maximum simulated amplitude at Wexford is 0.1 m, compared to 0.03 m in observations, indicating an overestimation of 0.07 m.

3.4. Computational setup

The numerical simulations were performed using Fortran for the nonlinear shallow-water solver and Python for the post-processing analyses, including DMD, K-means clustering, and DA. The Fortran code was compiled and executed on the Linux-based HPC cluster at Northumbria University. Each run used a

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

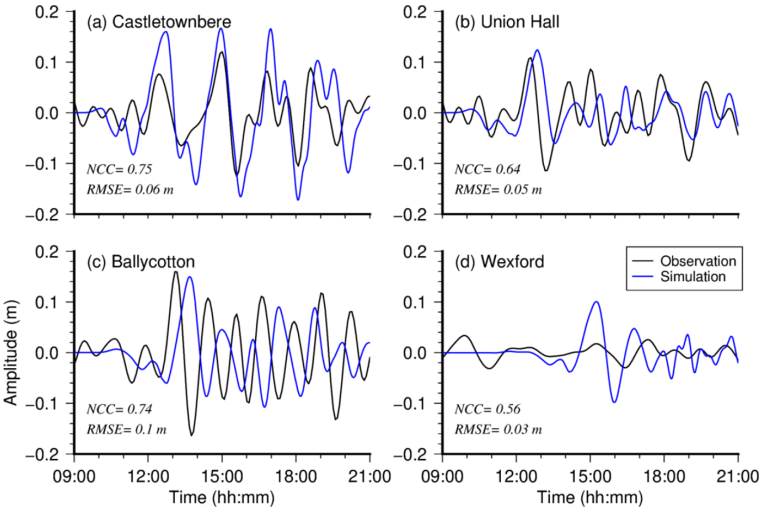


Figure 7. *Comparisons between simulated and observed waveforms at four tide gauges.*

single compute node (Intel Xeon 2.2 GHz CPU), and a 12-hour simulation of the 18 June 2022 event required about three hours of wall-clock time.

The computational complexity of the numerical solver scales linearly with the number of spatial grid points N expressed as $O(N)$, while the DMD analysis scales with both the number of spatial points and temporal snapshots M, approximately $O(N \times M)$. Data assimilation using Optimal Interpolation operates at $O(N)$ per assimilation cycle. The numerical solver structure is summarized in Figure 8, and all simulation parameters are provided in Table 1 to ensure reproducibility. Detailed computational specifications—including hardware configuration, compiler version, and optimization flags, and software environment—are also reported in Table 1.

4. Optimization of offshore monitoring stations

Lewis et al. (2023) investigated the occurrence of meteotsunamis in the UK and identified key hotspots, including the southwest region, where most historical events have been recorded. These meteotsunamis exhibit distinct seasonal patterns, with winter events clustering around the southwest UK, particularly along the Celtic Sea boundary. This seasonal pattern aligns with the passage of mid-latitude depressions, in which convective elements are embedded within cold fronts and low-pressure troughs. However, the detection and analysis of these events are often complicated by storm-driven water level fluctuations, which can obscure the meteotsunami signals (Thompson et al., 2020). Because of their rapid onset and limited predictability, individual meteotsunamis can cause abrupt changes in water levels and flow velocities, threatening human safety, coastal infrastructure, and marine activities. However, the limited availability of offshore observational stations remains a critical challenge for early warning (Denamiel et al., 2019a, 2019b), as most existing monitoring networks are primarily designed for earthquake-generated tsunamis. As a result, meteotsunamis have often been considered rare events. However, recent studies (e.g., Lewis et al., 2023) have shown that meteotsunamis occur more frequently than previously believed, challenging this long-held perception. Building on this issue, we evaluate the potential deployment of offshore observational stations along the southeastern coast of Ireland, using the 2022 Ireland meteotsunami as a case study. We show that expanding offshore observational networks with higher-frequency pressure sensors enhances detection capabilities, improves early warning systems, and refines predictive models.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

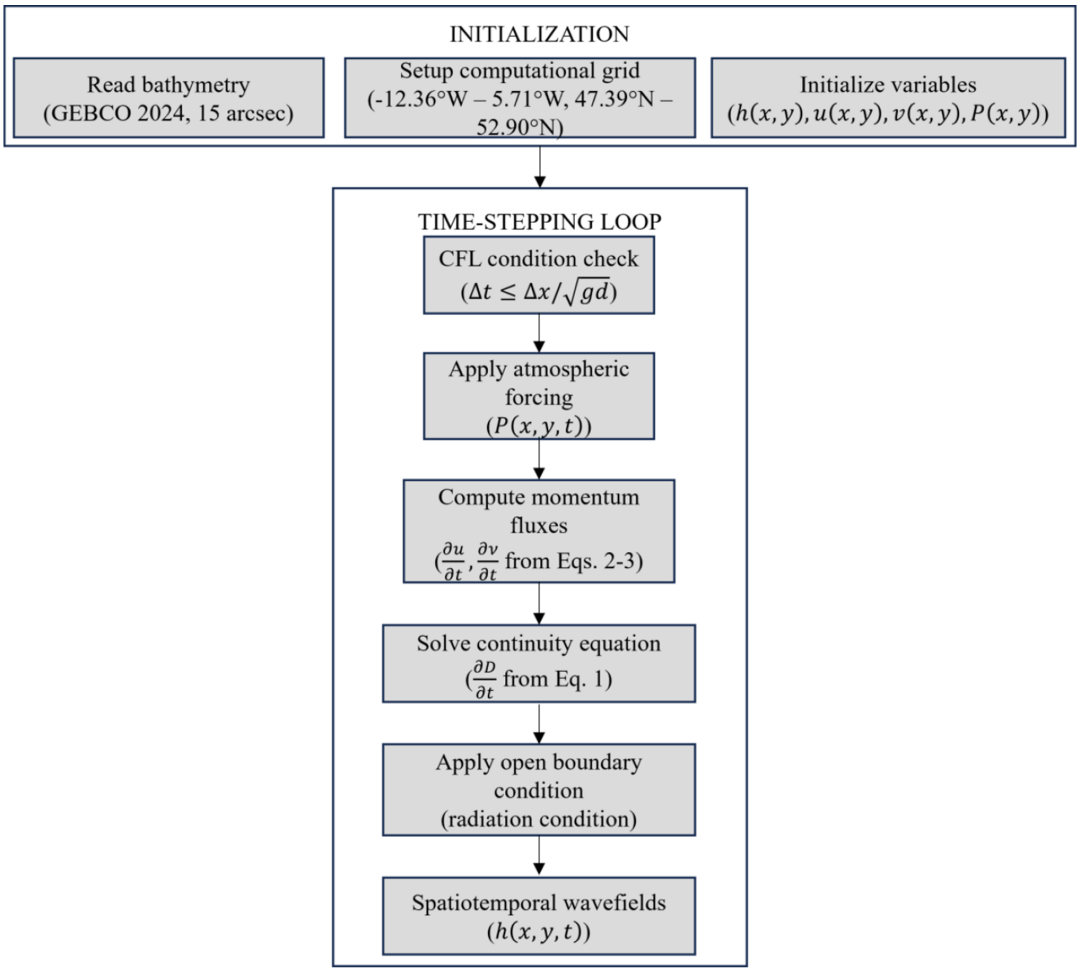


Figure 8. *Nonlinear shallow-water equation solver structure.*

4.1. Dynamic mode decomposition and sensitivity analysis

We use dynamic mode decomposition (DMD) for analyzing spatiotemporal patterns, and here we apply it to the 2022 Ireland meteotsunami case, where high-quality data offer a valuable opportunity to demonstrate its strengths. Currently, on the Irish side of the Celtic Sea, only two offshore buoys (M3 and M5) are located near the study area, with a 1-h sampling interval and positioned relatively far apart. This spatial and temporal sparsity is insufficient to capture meteotsunami dynamics. To improve detection and monitoring, observational stations should be strategically located in regions where the most energetic dynamics occur.

Previous studies have utilized empirical orthogonal function (EOF) analysis to determine the most energetic locations for seismically generated tsunamis (Mulia et al., 2017, 2019; Ren et al., 2022). EOF, essentially a principal component analysis, identifies static spatial patterns that explain the highest variance in the dataset. However, it assumes linear modes, which may not always adequately represent the evolution of wave propagation. In seismically generated tsunamis, wave propagation is primarily driven by water height gradients, which influence velocity changes. However, in meteotsunamis, waves propagate due to both water height gradients and atmospheric pressure forcing (Eqs. [2] and [3]). As a result, meteotsunamis exhibit distinct spatiotemporal variations compared to seismically generated tsunamis, which require a different approach.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

Table 1. *Numerical simulation parameters used in the meteotsunami modeling*

Parameters	Values	Description
Grid parameters		
Spatial resolution	15 arcsec (~450 m)	GEBCO 2024
Grid dimensions (N)	1,825,920 cells	Total spatial grid points
Longitude range	−12.36°W – 5.71°W	Computational domain
Latitude range	47.39°N – 52.90°N	Computational domain
Physical parameters		
Gravity (g)	9.81 m s^{-2}	Gravitational constant
Water density (ρ)	1025 kg m^{-3}	Seawater density
Manning coefficient (n)	0.025 $m^{1/3}$ s^{-1}	Bottom friction (homogeneous)
Temporal parameters		
Time step (Δt)	1 s	Chosen to satisfy CFL condition
Simulation duration	12 hours	09:00–21:00 UTC, 18 June 2022
Output interval	5 min ($M = 48$)	Snapshots for DMD analysis
Pressure forcing parameters		
Propagation speed (V_o)	55.4 m s^{-1}	Estimated from AWS data
Propagation direction (φ)	44° (from west)	Estimated from AWS alignment
Number of components	4	Gaussian components
Computational resources		
Hardware	Intel Xeon 2.2 GHz	Single node HPC cluster
RAM	62 GB	Total system memory
Wall-clock time	~3 hours	For 12-hour simulation
Compiler	gcc 5.2.0	Compiled as gfortran
Operating system	CentOS Linux 7 (Core)	Kernel 3.10.0–327.36.3.el7.x86_64

To address this limitation, we employ DMD, which captures time-dependent patterns and frequencies within the dataset (Tu et al., 2014). Unlike EOF, DMD extracts modes that evolve dynamically over time, offering a more accurate representation of wave evolution. Here, we derive the DMD modes of meteotsunami amplitudes resulting from numerical simulations. For the DMD computation, tsunami amplitudes within the model domain are stored at 5-minute intervals. We use the first 4 h of the meteotsunami simulation, yielding a total $m = 48$ snapshots. The DMD algorithm decomposes the spatiotemporal meteotsunami data into dynamic modes, each associated with specific frequencies and growth rates. Similar mode decomposition approaches have been used to isolate physically significant instabilities in nonlinear wave simulations, such as the dipole-shaped superharmonic unstable modes linked to wave breaking (Mansar et al., 2025).

First, the simulated sea surface dataset $\{h_1, \ldots, h_m\}$ is arranged into two sequential matrices,

$$X_1 = [h_1, \ldots, h_{m-1}], \tag{7}$$

$$X_2 = [h_1, \ldots, h_m], \tag{8}$$

where each ith column represents a sea surface elevation at a given time interval. The goal of DMD is to find a best-fit linear operator A such that,

$$X_2 \approx AX_1. \tag{9}$$

Since A is typically high-dimensional, a singular value decomposition is applied to X, yielding a reduced-order representation. The matrix X is factorized as

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

$$X = U\Sigma V^T, \tag{10}$$

where U and V are orthonormal matrices, and Σ is a diagonal matrix containing singular values. To obtain a low-rank approximation of A, we compute

$$\tilde{A} = U^T X_2 V \Sigma^{-1}, \tag{11}$$

where $\tilde{A}$ is a reduced-order representation that captures the dominant system dynamics while mitigating computational complexity. Next, we solve the eigenvalue problem,

$$\tilde{A}W = W\Lambda, \tag{12}$$

where Λ contains the DMD eigenvalues, which define the oscillation frequencies and growth rates, and W contains the corresponding eigenvectors, representing temporal evolution. Finally, the true DMD modes, which describe the spatial structures of wave dynamics, are reconstructed using

$$\varphi = X_2 V \Sigma^{-1} W. \tag{13}$$

These DMD modes reveal the dominant spatiotemporal patterns of meteotsunami evolution, providing a more dynamic characterization than traditional EOF analysis. By employing DMD, we capture the complex interactions between pressure forcing and ocean response over time, offering a more comprehensive approach for identifying key regions where offshore observational stations should be deployed.

Furthermore, since the DMD results consist of multiple spatial points, many of which may be redundant, we apply K-means clustering (Lloyd, 1982) as an optimization step to group similar modes and remove redundancies. K-means clustering optimizes the selection of dominant modes by minimizing the intra-cluster variance, which is given by the objective function:

$$J = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2, \tag{14}$$

where k is the number of clusters, C_i represents the set of data points in cluster i, μ_i is the centroid of cluster i, and $\|x - \mu_i\|^2$ denotes the squared Euclidean distance between a data point x and its cluster centroid. By iteratively updating centroids and reassigning points, K-means partitions the DMD modes into a set of representative clusters, reducing noise and redundancy in the final analysis. This process ensures that only the most significant and distinct spatiotemporal patterns are retained, improving the interpretability and efficiency of meteotsunami dynamic analysis.

The results of the first two spatial modes and their corresponding extrema are presented in Figure 9a,b. We focus on these two leading modes, as they capture the majority of the variance and represent the most energetic meteotsunami dynamics in the study area, primarily associated with meteotsunami propagation. While the first mode likely represents the largest-scale, slowest-changing features in the dataset, the second mode captures secondary variations, such as reflected waves, smaller oscillations, and bathymetric influences. The energy variance for these modes is similar, with the first mode representing 25.04% and the second mode for 22.68% of the total variance. Although subsequent modes also contribute to the overall variance (e.g., 15.96% for the third mode and 13.58% for the fourth mode), their contributions are significantly smaller, and higher modes primarily represent less energetic processes. For comparison, Figure 9d,e display the spatial modes and extrema derived from the traditional EOF analysis.

To identify optimal observation locations, we evaluated the extrema of modes 1 and 2 through an iterative, greedy optimization based on DA sensitivity testing. In each iteration, one candidate extremum was removed, and the DA performance was recalculated. The extrema whose exclusion caused the largest decrease in the overall assimilation score (i.e., the most dynamically significant locations) were retained, while those whose removal had minimal impact on the score were discarded. The performance metric was quantified using a scalar skill score that combines improvements in both amplitude accuracy and waveform coherence:

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

$$Score = w\left(1 - \frac{RMSE}{RMSE_f}\right) + (1-w)\left(\frac{NCC - NCC_f}{1 - NCC_f}\right), \tag{15}$$

where $RMSE_f$ and NCC_f denote the forward modeling baselines, and $w = 0.5$ assigns equal weight to amplitude and phase coherence. Higher values of the score indicate better assimilation performance.

The evolution of this optimization process for both DMD and EOF extrema is shown in Figure 9h, where the score changes with the successive removal of candidate stations. This greedy elimination continued until all non-significant extrema were removed. In addition, extrema located less than 30 km from the coastline (D1, E1, and E13) were excluded. The remaining candidate stations should have a space of approximately 85–100 km apart, ensuring a balanced spatial distribution that enhances meteotsunami monitoring and forecasting (Heidarzadeh et al., 2019; Ren et al., 2022).

The final DMD-based configuration initially identified the last seven candidate stations (D2, D3, D11, D12, D15, D16, and D20), where the assimilation score generally decreased as stations were removed. However, stations D11 and D12 were manually excluded to maintain adequate spacing between observation points. The resulting five optimized stations—D2, D3, D15, D16, and D20—correspond to the final five iterations in Figure 9h (see Figure 9c for their spatial locations).

Similarly, the EOF-based configuration retains three key stations—E2, E5, and E23—as shown in Figure 9f. Although the score shows a minor increase when E2 was excluded (Figure 9h), this likely reflects the sensitivity of the greedy procedure rather than a physically meaningful improvement. Station E2 is spatially well separated from the other EOF extrema and extends the observational coverage toward the eastern shelf, ensuring that the optimized network samples a broader portion of the meteotsunami propagation path. Therefore, E2 was retained to maintain spatial balance and avoid excessive clustering of observation points in the southern region.

4.2. Meteotsunami data assimilation

To model the meteotsunami monitoring system, we apply data assimilation using optimized station locations derived from DMD and K-means clustering to assess their performance. Sequential data assimilation (DA) techniques have been widely used in weather simulation (Kalnay, 2002). Maeda et al. (2015) introduced a DA approach for real-time tsunami wavefield simulation using the optimal interpolation method. Byrne et al. (2023) applied a similar technique to transform tide gauge records into a regional storm surge model. Optimal interpolation is computationally less demanding than more advanced methods, such as the ensemble Kalman filter (Yang et al., 2019), provided that the system remains linear. Despite its simplicity, this DA approach has demonstrated excellent agreement with both observed tsunami data and synthetic cases. To the best of our knowledge, this is the first study to implement data assimilation for meteotsunamis, with the goal of identifying potential offshore observational station locations.

In the numerical simulation, the meteotsunami wavefield at the nth time step is represented as $\boldsymbol{x_n}(\eta(n\Delta t,x,y),M(n\Delta t,x,y),N(n\Delta t,x,y))$, where η is the meteotsunami height, M and N are the depth-integrated meteotsunami fluxes in the x and y directions and Δt is the time step. The DA method can be expressed as

$$\boldsymbol{x}_n^f = \boldsymbol{F}\boldsymbol{x}_{n-1}^a, \tag{16}$$

$$\boldsymbol{x}_n^a = \boldsymbol{x}_n^f + \boldsymbol{W}\left(\boldsymbol{y}_n - \boldsymbol{H}\boldsymbol{x}_n^b\right). \tag{17}$$

At each time step, the forecasted meteotsunami wavefield, $\boldsymbol{x}_n^f$, is obtained by solving the non-linear long wave equation using the assimilated wavefield from the previous time step, $\boldsymbol{x}_{n-1}^a$. Here, $\boldsymbol{F}$ is the propagation matrix corresponding to the two-dimensional NLSWE model. The residual—defined as the difference between the observed meteotsunami amplitude at the observational stations ($\boldsymbol{y}_n$) and the simulation output ($\boldsymbol{H}\boldsymbol{x}_n^b$)—is calculated, where $\boldsymbol{H}$ is a vector with entries of 1 at the observational stations and 0 elsewhere, used to extract the forecasted meteotsunami height. These residuals are then multiplied

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

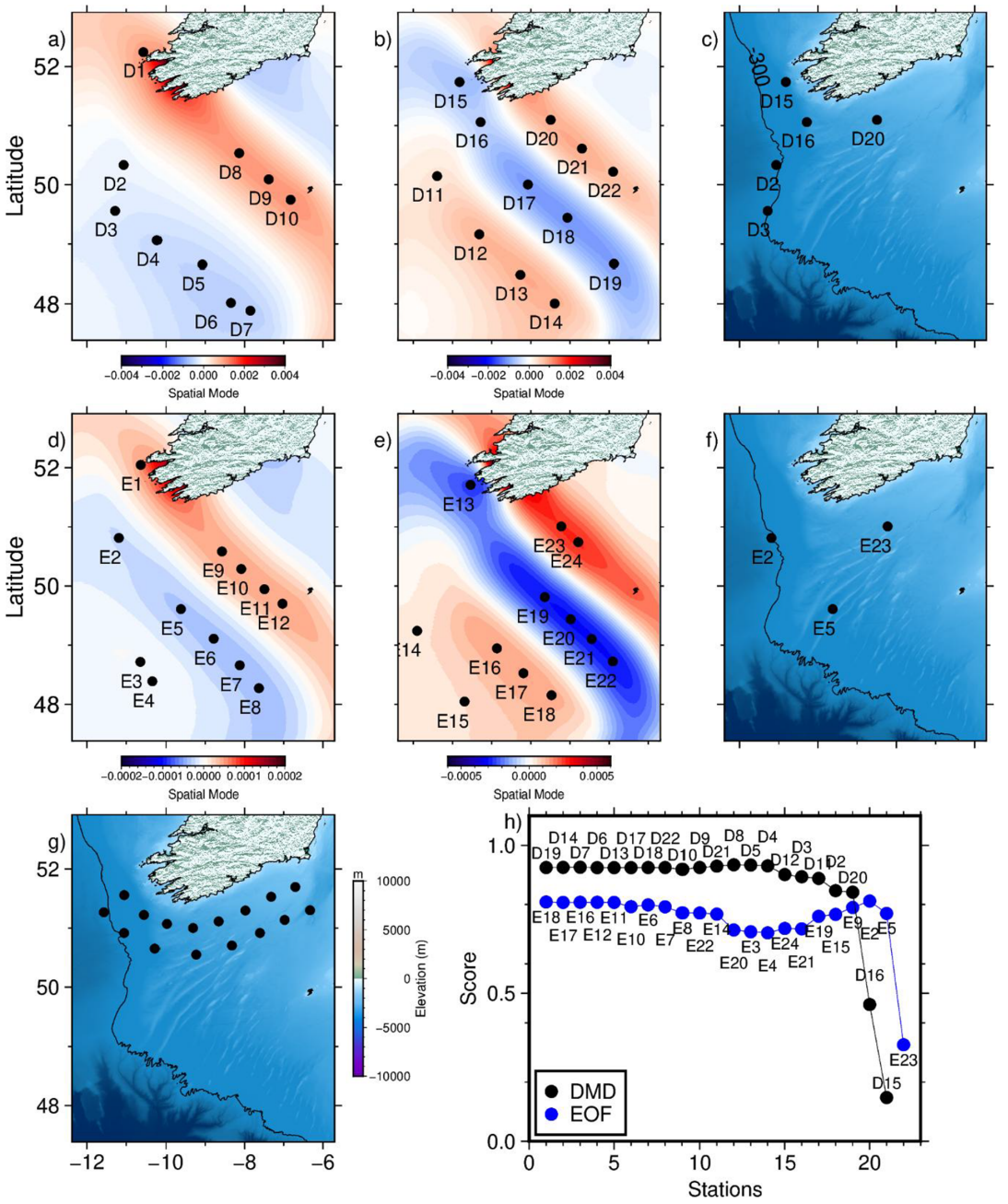


Figure 9. *Spatial distributions and station selections based on DMD and EOF analysis. (a) First DMD mode and (b) second DMD mode, with extrema indicated by black circles. (c) Selected extrema from DMD modes identified as potential observational stations (black circles). (d) First EOF mode and (e) second EOF mode, with extrema indicated by black circles. (f) Selected extrema from EOF modes used as potential observational stations (black circles). (g) Predefined station locations resembling the S-Net array configuration, with approximately 50 km spacing. (h) Evolution of the greedy optimization process for DMD- and EOF-based station selection. Numbers above and below the markers indicate the stations removed at each iteration.*

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

by a smoothing matrix, $\boldsymbol{W}$, which is crucial for bringing the assimilated wavefield closer to the observed meteotsunami field. The smoothing matrix is determined by solving the following equation:

$$\boldsymbol{W} = \boldsymbol{PH^T} + \left(\boldsymbol{HPH^T} + \boldsymbol{R}\right)^{-1}, \tag{18}$$

where $\boldsymbol{P} = \langle \varepsilon^f \varepsilon^{fT} \rangle$ and $\boldsymbol{R} = \langle \varepsilon^O \varepsilon^{OT} \rangle$ are the error covariance matrices of the forward simulation and the observations, respectively. ε^f and ε^O are the errors of the forward simulation and observations, while ε^{fT} and ε^{OT} are the corresponding transpose matrices. In practice, $\boldsymbol{P}$ is modelled using an isotropic Gaussian correlation $P_{ij} = \sigma_p^2 exp\left[-\frac{d_{ij}^2}{L^2}\right]$, where d_{ij}^2 is the distance between grid points i and j, σ_p^2 is the background-error variance, and L is the characteristic distance controlling the spatial influence of each observation. Observation errors are assumed to be uncorrelated, so $R = \sigma_r^2 I$. Following Maeda et al. (2015), the relative error ratio is expressed by the nondimensional parameter ρ, defined as the root-mean-square observation error normalized by the forecast error ($R/P = \rho_d^2$). Similarly, we adopted $\rho_d = 1$ implying equal confidence in model and observations. Both covariance matrices assume Gaussian-distributed errors with a characteristic distance of 20 km (Maeda et al., 2015; Wang et al., 2018). The assimilation cadence was identical to the numerical time step (Δt) used in the forward model. By iteratively applying Eqs. (16) and (17), the tsunami wavefield is effectively assimilated, enabling the generation of forecasted meteotsunami waveforms at any location within the model domain during and after the assimilation process. The assimilation was performed sequentially at each model time step corresponding to the available observation interval, with an assimilation cadence of 1 min. The complete data assimilation procedure, including the optimal interpolation scheme based on Maeda et al. (2015), is visualized in Figure 10. This sequential algorithm updates the wavefield at each time step using observations from the optimized station network.

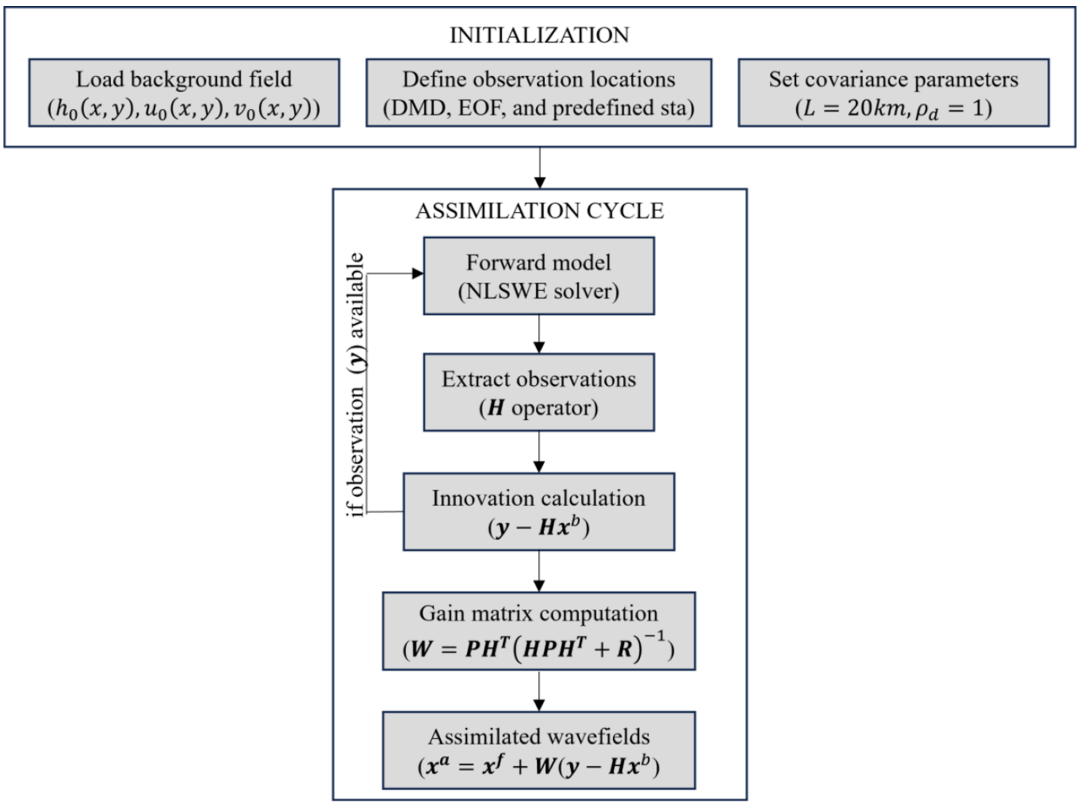


Figure 10. *Flowchart of the sequential data assimilation procedure using optimal interpolation.*

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

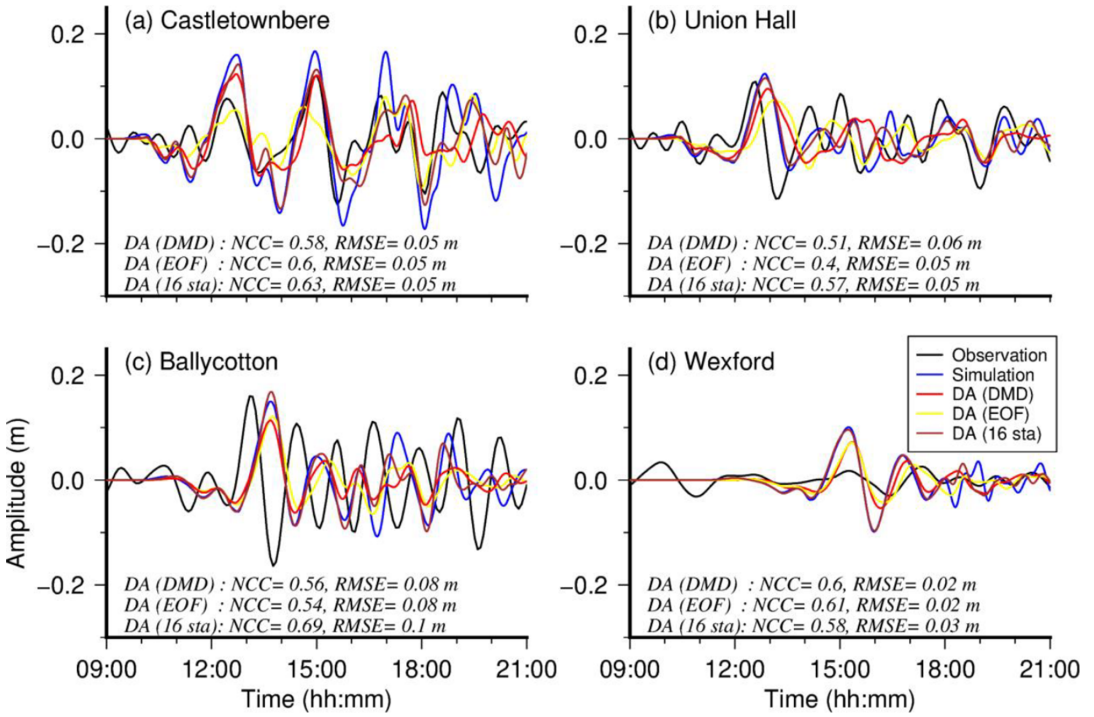


Figure 11. *Comparisons between simulated, observed, and assimilated waveforms at four tide gauges.*

For a clear performance comparison, we run DA simulations using three sets of observational station configurations: DMD-generated, EOF-generated, and predefined stations. The predefined stations, shown in Figure 9g, represent a dense network of observational stations with a spatial interval of approximately 50 km as well as for the approximate closest distance to the coast, resembling the dense S-Net network. We then compare the resulting waveforms at tide gauges and the maximum meteotsunami amplitudes to evaluate the performance of each approach. As shown in Figure 11, the assimilated waveforms from the DMD-generated, EOF-generated, and predefined stations exhibit good agreement with forward modeling results and observations. Even though EOF-generated stations exhibit similar statistical performance to DMD-generated, EOF tends to underestimate wave amplitudes. For instance, at Castletownbere, the maximum wave amplitudes are better captured by DMD-generated stations. The NCC values at this gauge are 0.58, 0.6, and 0.63 for the DMD, EOF-optimized, and predefined station scenarios, respectively, with a similar RMSE of 0.05 m. Similarly, the assimilated waveforms from DMD-generated stations at Union Hall, Ballycotton, and Wexford demonstrate good agreement with observations, with NCC values of 0.51, 0.56, and 0.60 and RMSE values of 0.06 m, 0.08 m, and 0.02 m, respectively. Detailed information including DA configurations are summarized in Table 2.

We also compared the assimilated maximum wave amplitudes to those from the meteotsunami forward simulation, since survey data for maximum wave height along the southern coast of Ireland is not available. The predefined station network shows good agreement with the simulated results, likely due to its wide spatial coverage, which enables better capture of meteotsunami dynamics. Comparison between assimilation and simulation near the coastline (Figure 12) further confirms that the predefined scenario closely aligns with the forward modeling. Similarly, the DMD-generated stations also show comparable maximum wave amplitudes, although some underestimation is observed in the longitude range between −7.5°W and − 6.5°W, which is likely due to a lower density of observation points in that area. Nevertheless, this underestimation is minimal and does not significantly impact the overall accuracy of the assimilation results. In contrast, the EOF-based stations exhibit more pronounced underestimation in

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

Table 2. *Data assimilation parameters and performance metrics for all station configurations*

Parameter	Value	Description
Assimilation settings		
Assimilation interval	1 min	
Background error decorrelation length (L)	20 km	Isotropic Gaussian correlation
Observation error parameter (ρ_d)	1	Equal confidence in model and obs.
Observation network	5 stations (DMD), 3 stations (EOF), 16 stations (predefined)	Station configurations compared
Predefined station network		
Spatial spacing	~50 km	Approximate inter-station distance
Minimum coastal distance	~30 km	Closest station to coast
Performance metrics by station	NCC	RMSE (m)
Castletownbere		
DMD-optimized	0.58	0.05
EOF-optimized	0.60	0.05
Predefined (16 sta)	0.63	0.05
Union Hall		
DMD-optimized	0.51	0.06
EOF-optimized	0.40	0.05
Predefined (16 sta)	0.57	0.05
Ballycotton		
DMD-optimized	0.56	0.08
EOF-optimized	0.54	0.08
Predefined (16 sta)	0.69	0.01
Wexford		
DMD-optimized	0.60	0.02
EOF-optimized	0.61	0.02
Predefined (16 sta)	0.58	0.03

both the −7.5°W to −6.5°W and the southwestern region between 51°N to 52°N and − 11°W to −9°W, which is attributed to the absence of EOF-generated stations in these areas.

5. Discussion

A comprehensive analysis of observational datasets, incorporating sea level and atmospheric pressure measurements, identifies meteorological forcing as the primary mechanism behind the tsunami-like waves observed along the southern coast of Ireland on 18 June 2022. The high-resolution air pressure records enable a more realistic representation of atmospheric forcing in meteotsunami genesis. Our numerical model plays a crucial role in reconstructing the hydrodynamics of the meteotsunami, from its generation to propagation.

Overall, the comparison between simulated and observed waveforms at tide gauges demonstrates good agreement in terms of wave amplitude and arrival times, and represents an improvement over previous studies that employed a sine-based source model for atmospheric pressure forcing (Renzi et al., 2023). However, at the northernmost tide gauge in Wexford, the simulation overestimates wave heights,

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

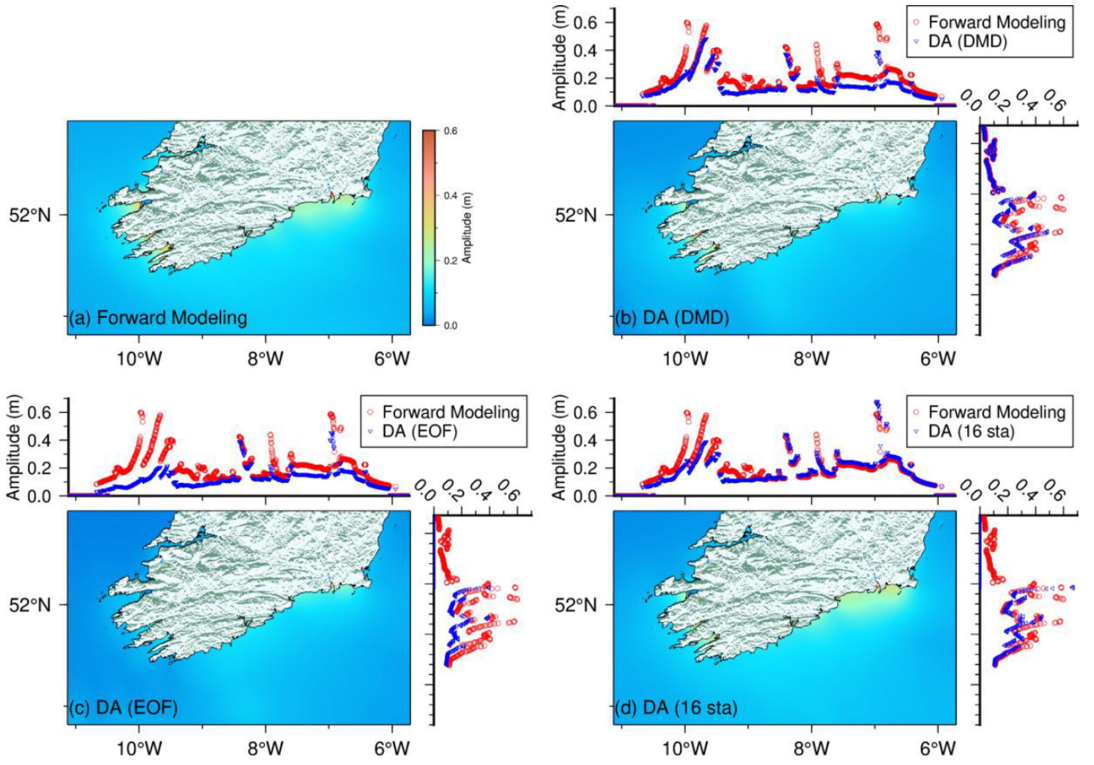


Figure 12. *Comparison of maximum meteotsunami amplitudes along the southern coast of Ireland from (a) the forward modeling simulation, (b) data assimilation (DA) using five DMD-optimized observational stations, (c) DA using three EOF-optimized observational stations, and (d) DA using 16 predefined observational stations. Insets above and to the right of (b), (c), and (d) show maximum wave amplitudes at coastal locations, highlighting differences in spatial accuracy.*

indicating a discrepancy between model predictions and observations. Several factors may contribute to this overestimation, including uncertainties in atmospheric forcing, limitations in bathymetric resolution, and complex wave interaction processes.

In this case, the uncertainties in atmospheric forcing appear to be the primary cause. Supporting this hypothesis, an analysis of filtered waveforms data and air pressure (Figures 2 and 3) reveals a positive correlation between Castletownbere and Ballycotton with Sherkin Island AWS, the nearest weather station. A similar pattern is observed at Union Hall, which correlates well with the air pressure recorded at Roches Point AWS. In contrast, Wexford does not exhibit a similar response to air pressure variations recorded at Johnstown Castle AWS, despite their proximity of just 6 km. This mismatch is also reported by Renzi et al. (2023), suggesting that localized atmospheric conditions or small-scale meteorological processes not fully captured in the dataset may have influenced the observed wave dynamics.

One possible explanation for this discrepancy is that the idealized Gaussian pressure field used in the numerical model may overestimate the actual spatial structure of the atmospheric pressure disturbance. Another potential explanation relates to the local geography of the Wexford tide gauge, which is situated within a harbor enclosed by a breakwater and further protected by a bay partially sheltered by a sandy barrier spit. The overestimation of wave amplitude may be due to unresolved local dissipation and complex wave transformations occurring within these nested, sheltered environments—processes that are not captured in the regional-scale numerical model. This is supported by the Wexford and Burrow Modelling Study (RPS, 2021), which underscores the need for high-resolution modeling to account for

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

localized features like breakwaters and barrier spits. Similar conclusions by Costa et al. (2023) highlight how small-scale coastal morphology can affect water level responses.

The EOF-based optimization curve (Figure 9h) exhibits a non-monotonic trend, with an initial decline, a brief recovery, and a final sharp drop in skill score. This behavior arises because many EOF extrema are spatially clustered—particularly in the central shelf region (e.g., E16–E18 and E10–E12)—resulting in redundant and highly correlated observations. The inclusion of these closely spaced extrema renders the gain matrix $\left(\boldsymbol{HPH^T}+\boldsymbol{R}\right)^{-1}$ increasingly ill-conditioned, amplifying small numerical errors and leading to unstable assimilation performance. Once the redundant stations are removed, the matrix conditioning improves, and the skill score temporarily recovers; however, the subsequent elimination of dynamically significant offshore stations produces the final sharp decline. Similar effects of correlated observation errors on matrix conditioning and analysis stability have been reported by Goux et al. (2024).

In contrast, the DMD-derived extrema are more spatially sparse and dynamically independent, providing better spatial coverage and numerical stability within the data-assimilation framework. By combining DMD with K-means clustering, we optimized station placement to achieve accuracy comparable to that of a dense observational network while substantially reducing the number of required sensors. The potential locations for offshore stations are represented by the two leading DMD modes, which also capture energy variations influenced by bathymetric effects, as evidenced by the optimized points aligning near the −300 m depth contour (Figure 9c).

Despite the significant reduction in station number, the optimized network reproduces the performance of a dense predefined array, offering a cost-effective strategy for future meteotsunami monitoring. The DMD extrema also capture key wave dynamics in the southern region of Ireland (Figure 9b,c), yielding two additional candidate stations in an area underrepresented by EOF. This agrees with previous findings that DMD outperforms EOF by resolving non-stationary, time-evolving flow structures and providing a more faithful representation of large-scale kinetic-energy variability (e.g., Gugole and Franzke, 2020; Li et al., 2023). Using a straightforward selection strategy, we identified three optimal extrema from the EOF modes, whereas the DMD approach produced five, demonstrating a broader and more flexible range of candidate locations. Overall, DMD offers richer spatial information and greater numerical robustness for optimizing offshore observational networks.

The high costs associated with offshore observational systems require careful station placement to maximize efficiency while minimizing expenses. The financial burden of offshore tsunami detection systems is substantial. For example, a DART station costs approximately $0.5 M per unit, with additional high emergency maintenance costs due to ship-based interventions. Cabled observatories, while supporting high-density monitoring, require an upfront investment of hundreds of millions of dollars, as demonstrated by Japan's $500 M cabled network covering 1000 km with 164 pressure sensors. Similarly, differential GPS buoys cost around $3 M per unit, with additional limitations in deployment flexibility (Bernard and Titov, 2015). Given these expenses, reducing the number of required stations without compromising accuracy is critical. It should be clarified that the proposed offshore monitoring stations are not intended to perform high-performance computations. They function as lightweight data acquisition platforms that collect sea-level or pressure observations and transmit them in near real time. All computationally intensive processes —including numerical simulations, DMD analysis, and data assimilation—are performed centrally at an onshore high-performance computing facility. The present study, therefore, focuses on sensor placement optimization rather than distributed offshore computation.

In determining the optimal placement of observational stations, previous studies have often relied on EOF analysis using multiple hypothetical scenarios based on recorded events and enhanced using a stochastic model (Mulia et al., 2019; Ren et al., 2022). However, in the case of meteotsunamis in Ireland, historical records are scarce or poorly documented, limiting the ability to use past events as a basis for station placement. Consequently, the characteristics of meteotsunamis in this region, as well as their potential future occurrence, remain largely unidentified. This contrasts with the Balearic Islands, where the probability of meteotsunami occurrence has been extensively studied (Šepić et al., 2016).

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

As a result, this study relies solely on a synthetic scenario of the 2022 Ireland event to determine the optimal station locations for meteotsunami detection. While this approach provides a reasonable first-order approximation, future studies could benefit from the development of stochastic scenarios to incorporate a wider range of potential meteotsunami conditions, thereby improving the robustness of station placement strategies. This represents an intriguing direction for future research. Additionally, higher-resolution nested modeling of the Wexford harbor area would help resolve the local dissipation processes that our regional model currently underestimates, potentially improving forecast accuracy in complex coastal environments.

6. Conclusions

This study introduces a data-driven framework for extracting dominant meteotsunami dynamics from sparse observational data, combining dynamic mode decomposition, numerical simulation, and data assimilation. Using the 2022 Ireland meteotsunami as a demonstration case, we show that a minimal, optimally placed network of offshore sensors can effectively capture key wavefield characteristics, even in under-monitored regions. While developed in the context of a specific event, the methodology is broadly applicable and scalable to other meteotsunami-prone coastal areas. We anticipate that the methodology introduced here will encourage the installation of additional offshore stations, ultimately enhancing the detection, forecasting, and understanding of meteotsunamis.

Abbreviations

AWS	automatic weather station
DA	data assimilation
DART	deep-ocean assessment and reporting tsunamis
DMD	dynamic mode decomposition
DONET	dense oceanfloor network system for earthquakes and tsunamis
EOF	empirical orthogonal function
NCC	normalized cross correlation coefficient
NLSWE	non-linear shallow water equation
RMSE	root mean square error
S-Net	seafloor observation network for earthquakes and tsunami along the Japan trench

Open peer review. To view the open peer review materials for this article, please visit http://doi.org/10.1017/eds.2026.10036.

Acknowledgments. Figures in this paper were made using GMT software (Wessel and Luis, 2017), which can be downloaded at http://gmt.soest.hawaii.edu/.

Author contribution. AF initiated the research idea, conducted simulations, prepared the manuscript, and produced the figures. ER contributed to the research idea, provided critical comments on the results, and assisted in structuring the manuscript. FD provided critical feedback on the results and contributed to the manuscript's structure. DPZ and TK supplied raw data and contributed to the manuscript's structure. All authors read and approved the final manuscript.

Competing interests. The authors declare that they have no competing interests.

Data availability statement. The raw data underlying this study is not publicly available because it constitutes a key research asset that is still being explored in ongoing work and may form the basis of future publications. Raw data may be shared upon reasonable request, subject to a collaboration agreement. For additional reference, air pressure and tidal datasets used in this study were obtained from the Marine Institute of Ireland (www.marine.ie) and Met Éireann, the Irish Meteorological Service (www.met.ie). The 15-arcsec resolution global bathymetry dataset is publicly available and can be downloaded from GEBCO at https://www.gebco.net/data-products/gridded-bathymetry-data. The source code for the forward modelling, data assimilation, and dynamic mode decomposition to reproduce our experiments can be accessed at https://github.com/ardifauzi/metDA.git.

Ethics statement. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Funding statement. This research was not supported by any external grants/funding.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

References

Anarde K, **Cheng W**, **Tissier M**, **Figlus J and Horrillo J** (2021) Meteotsunamis accompanying tropical cyclone rainbands during hurricane Harvey. *Journal of Geophysical Research: Oceans 126*(1), e2020JC016347. https://doi.org/10.1029/2020JC016347.

Andraud P (2025) Estimation rapide des inondations et courants de tsunami à partir d'algorithmes de Deep Learning. PhD Thesis, Université Paris-Saclay. Available at https://theses.hal.science/tel-05245925v1/file/144141_ANDRAUD_2025_archivage.pdf.

Baba T, **Takahashi N and Kaneda Y** (2014) Near-field tsunami amplification factors in the Kii peninsula, Japan for dense Oceanfloor network for earthquakes and tsunamis (DONET). *Marine Geophysical Research 35*(3), 319–325. https://doi.org/10.1007/s11001-013-9189-1.

Bernard E and Titov V (2015) Evolution of tsunami warning systems and products. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 373*(2053), 20140371. https://doi.org/10.1098/rsta.2014.0371.

Byrne D, **Horsburgh K and Williams J** (2023) Variational data assimilation of sea surface height into a regional storm surge model: Benefits and limitations. *Journal of Operational Oceanography 16*(1), 1–14. https://doi.org/10.1080/1755876X.2021.1884405.

Costa W, **Bryan KR**, **Stephens SA and Coco G** (2023) A regional analysis of tide-surge interactions during extreme water levels in complex coastal systems of Aotearoa New Zealand. *Frontiers in Marine Science 10*. https://doi.org/10.3389/fmars.2023.1170756.

Denamiel C, **Šepić J**, **Huan X**, **Bolzer C and Vilibić I** (2019a) Stochastic surrogate model for Meteotsunami early warning system in the eastern Adriatic Sea. *Journal of Geophysical Research: Oceans 124*(11), 8485–8499. https://doi.org/10.1029/2019JC015574.

Denamiel C, **Šepić J**, **Ivanković D and Vilibić I** (2019b) The Adriatic Sea and coast modelling suite: Evaluation of the meteotsunami forecast component. *Ocean Modelling 135*, 71–93. https://doi.org/10.1016/j.ocemod.2019.02.003.

Denamiel C, **Šepić J and Vilibić I** (2018) Impact of geomorphological changes to harbor resonance during Meteotsunamis: The vela Luka Bay test case. *Pure and Applied Geophysics 175*(11), 3839–3859. https://doi.org/10.1007/s00024-018-1862-5.

Di Giacomo D, **Storchak DA**, **Safronova N**, **Ozgo P**, **Harris J**, **Verney R and Bondar I** (2014) A new ISC service: The bibliography of seismic events. *Seismological Research Letters 85*(2), 354–360. https://doi.org/10.1785/0220130143.

Fauzi A and Mizutani N (2020a) Machine learning algorithms for real-time tsunami inundation forecasting: A case study in Nankai region. *Pure and Applied Geophysics 177*, 1437–1450. https://doi.org/10.1007/s00024-019-02364-4.

Fauzi A and Mizutani N (2020b) Potential of deep predictive coding networks for spatiotemporal tsunami wavefield prediction. *Geoscience Letters 7*(1), 20. https://doi.org/10.1186/s40562-020-00169-1.

Gee LS and Leith WS (2011) *The Global Seismographic Network*. U.S. Geological Survey. https://doi.org/10.3133/fs20113021.

Goux O, **Gürol S**, **Weaver AT**, **Diouane Y and Guillet O** (2024) Impact of correlated observation errors on the conditioning of variational data assimilation problems. *Numerical Linear Algebra with Applications 31*(1). https://doi.org/10.1002/nla.2529.

Greenspan M (1956) Propagation of sound in five monatomic gases. *The Journal of the Acoustical Society of America 28*(4), 644–648. https://doi.org/10.1121/1.1908432.

Gugole F and Franzke CLE (2020) Spatial covariance Modeling for stochastic subgrid-scale parameterizations using dynamic mode decomposition. *Journal of Advances in Modeling Earth Systems 12*(8). https://doi.org/10.1029/2020MS002115.

Gusiakov VK (2021) Meteotsunamis at global scale: Problems of event identification, parameterization and cataloguing. *Natural Hazards 106*(2), 1105–1123. https://doi.org/10.1007/s11069-020-04230-2.

Gusman AR and Tanioka Y (2013) W phase inversion and tsunami inundation Modeling for tsunami early warning: Case study for the 2011 Tohoku event. *Pure and Applied Geophysics 171*(7), 1409–1422. https://doi.org/10.1007/s00024-013-0680-z.

Heidarzadeh M, **Wang Y**, **Satake K and Mulia IE** (2019) Potential deployment of offshore bottom pressure gauges and adoption of data assimilation for tsunami warning system in the western Mediterranean Sea. *Geoscience Letters 6*(19). https://doi.org/10.1186/s40562-019-0149-8.

Hibiya T and Kajiura K (1982) Origin of the Abiki phenomenon (a kind of seiche) in Nagasaki Bay. *Journal of the Oceanographical Society of Japan 38*(3), 172–182. https://doi.org/10.1007/BF02110288.

Imamura F, **Yalçiner AC and Ozyurt G** (2006) Tsunami modelling manual. In UNESCO IOC International training course on Tsunami Numerical Modelling. https://www.tsunami.irides.tohoku.ac.jp/media/files/_u/project/manual-ver3_1.pdf (accessed October 3, 2024).

Kalnay E (2002) *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press. https://doi.org/10.1017/cbo9780511802270.

Kaneda Y, **Kawaguchi K**, **Araki E**, **Matsumoto H**, **Nakamura T**, **Kamiya S**, **Ariyoshi K**, **Hori T**, **Baba T and Takahashi N** (2015) Development and application of an advanced ocean floor network system for megathrust earthquakes and tsunamis. In *Seafloor Observatories: A New Vision of the Earth from the Abyss*. Springer. https://doi.org/10.1007/978-3-642-11374-1_25.

Kim J, **Choi B-J and Omira R** (2022) On the Greenspan resurgence of meteotsunamis in the Yellow Sea—Insights from the newly discovered 11–12 June 2009 event. *Natural Hazards 114*(2), 1323–1340. https://doi.org/10.1007/s11069-022-05427-3.

Kim J and Omira R (2021) The 6–7 July 2010 meteotsunami along the coast of Portugal: Insights from data analysis and numerical modelling. *Natural Hazards 106*(2), 1397–1419. https://doi.org/10.1007/s11069-020-04335-8.

Kubota T, **Saito T**, **Chikasada NY and Sandanbata O** (2021) Meteotsunami observed by the Deep-Ocean seafloor pressure gauge network off northeastern Japan. *Geophysical Research Letters 48*(21). https://doi.org/10.1029/2021GL094255.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

Lewis C, **Smyth T**, **Williams D**, **Neumann J and Cloke H** (2023) Meteotsunami in the United Kingdom: The hidden hazard. *Natural Hazards and Earth System Sciences 23*(7), 2531–2546. https://doi.org/10.5194/nhess-23-2531-2023.

Li L, **Mémin E and Tissot G** (2023) Stochastic parameterization with dynamic mode decomposition. In Chapron B, Crisan D, Holm D, Mémin E and Radomska A (eds.), *Stochastic Transport in Upper Ocean Dynamics. STUOD 2021. Mathematics of Planet Earth*, Vol. 10. Cham: Springer, pp. 179–193. https://doi.org/10.1007/978-3-031-18988-3_11.

Lloyd S (1982) Least squares quantization in PCM. *IEEE Transactions on Information Theory 28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489.

Maeda T, **Obara K**, **Shinohara M**, **Kanazawa T and Uehira K** (2015) Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters 42*, 7923–7932. https://doi.org/10.1002/2015GL065588.

Mansar A, **Turner MR**, **Bridges TJ and Dias F** (2025) Effect of real-world perturbations on wave breaking due to a sharp-crested Superharmonic instability. *Water Waves 7*(1), 159–198. https://doi.org/10.1007/s42286-025-00116-7.

Monserrat S, **Vilibić I and Rabinovich AB** (2006) Meteotsunamis: Atmospherically induced destructive ocean waves in the tsunami frequency band. *Natural Hazards and Earth System Sciences 6*(6), 1035–1051. https://doi.org/10.5194/nhess-6-1035-2006.

Mulia IE, **Gusman AR and Satake K** (2017) Optimal Design for Placements of tsunami observing systems to accurately characterize the inducing earthquake. *Geophysical Research Letters 44*(24). https://doi.org/10.1002/2017GL075791.

Mulia IE, **Gusman AR**, **Williamson AL and Satake K** (2019) An optimized Array configuration of tsunami observation network off southern Java, Indonesia. *Journal of Geophysical Research: Solid Earth 124*(9), 9622–9637. https://doi.org/10.1029/2019JB017600.

Obara K (2007) Hi-net: High sensitivity seismograph network, Japan. In *Methods and Applications of Signal Processing in Seismic Network Operations*. Berlin Heidelberg: Springer, pp. 79–88. https://doi.org/10.1007/BFb0117698.

Proudman J (1929) The effects on the sea of changes in atmospheric pressure. *Geophysical Journal International 2*, 197–209. https://doi.org/10.1111/j.1365-246X.1929.tb05408.x.

Rabinovich AB (2020) Twenty-seven years of Progress in the science of meteorological tsunamis following the 1992 Daytona Beach event. *Pure and Applied Geophysics 177*(3), 1193–1230. https://doi.org/10.1007/s00024-019-02349-3.

Ren Z, **Wang Y**, **Wang P**, **Zhao X**, **Hu G and Li L** (2022) Optimal deployment of seafloor observation network for tsunami data assimilation in the South China Sea. *Ocean Engineering 243*, 110309. https://doi.org/10.1016/j.oceaneng.2021.110309.

Renzi E, **Bergin C**, **Kokina T**, **Pelaez-Zapata DS**, **Giles D and Dias F** (2023) Meteotsunamis and other anomalous "tidal surge" events in Western Europe in summer 2022. *Physics of Fluids 35*(4). https://doi.org/10.1063/5.0139220.

RPS (2021) *Wexford and Burrow Coastal Wave And Water Level Modelling Study 2020.* RPS Group Ltd., Dublin, Ireland. https://s3-eu-west-1.amazonaws.com/docs.floodinfo.opw/floodinfo_docs/CWWS/IBE1729_CWWS_Rp01_Wexford_F02.pdf (accessed February 2, 2025).

Satake K (1989) Inversion of tsunami waveforms for the estimation of heterogeneous fault motion of large submarine earthquakes: The 1968 Tokachi-oki and 1983 Japan Sea earthquakes. *Journal of Geophysical Research 94*(B5), 5627–5636. https://doi.org/10.1029/JB094iB05p05627.

Šepić J, **Rabinovich AB and Sytov VN** (2018a) Odessa tsunami of 27 June 2014: Observations and numerical modelling. *Pure and Applied Geophysics 175*(4), 1545–1572. https://doi.org/10.1007/s00024-017-1729-1.

Šepić J, **Vilibić I and Monserrat S** (2016) Quantifying the probability of meteotsunami occurrence from synoptic atmospheric patterns. *Geophysical Research Letters 43*(19). https://doi.org/10.1002/2016GL070754.

Šepić J, **Vilibić I**, **Rabinovich A and Tinti S** (2018b) Meteotsunami ("Marrobbio") of 25–26 June 2014 on the southwestern coast of Sicily, Italy. *Pure and Applied Geophysics 175*(4), 1573–1593. https://doi.org/10.1007/s00024-018-1827-8.

Thompson J, **Renzi E**, **Sibley A and Tappin DR** (2020) UK meteotsunamis: A revision and update on events and their frequency. *Weather 75*(9), 281–287. https://doi.org/10.1002/wea.3741.

Titov VV, **Gonzalez FI**, **Bernard EN**, **Eble MC**, **Mofjeld HO**, **Newman JC and Venturato AJ** (2005) Real-time tsunami forecasting: Challenges and solutions. *Natural Hazards 35*(1), 35–41. https://doi.org/10.1007/s11069-004-2403-3.

Torrence C and Compo GP (1998) A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society 79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079%3c0061:APGTWA%3e2.0.CO;2.

Tu JH, **Rowley CW**, **Luchtenburg DM**, **Brunton SL and Nathan Kutz J** (2014) On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics 1*(2), 391–421. https://doi.org/10.3934/jcd.2014.1.391.

Vilibić I, **Monserrat S**, **Rabinovich A and Mihanović H** (2008) Numerical modelling of the destructive Meteotsunami of 15 June, 2006 on the coast of the Balearic Islands. *Pure and Applied Geophysics 165*(11–12), 2169–2195. https://doi.org/10.1007/s00024-008-0426-5.

Wang Y, **Satake K**, **Maeda T and Gusman AR** (2018) Data assimilation with dispersive tsunami model: A test for the Nankai trough. *Earth, Planets and Space 70*(131). https://doi.org/10.1186/s40623-018-0905-6.

Wessel P and Luis JF (2017) The GMT/MATLAB toolbox. *Geochemistry, Geophysics, Geosystems 18*(2), 811–823. https://doi.org/10.1002/2016GC006723.

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press

Yamamoto N, **Hirata K**, **Aoi S**, **Suzuki W**, **Nakamura H and Kunugi T** (2016) Rapid estimation of tsunami source centroid location using a dense offshore observation network. *Geophysical Research Letters 43*(9), 4263–4269. https://doi.org/10.1002/2016GL068169.

Yang Y, **Dunham EM**, **Barnier G and Almquist M** (2019) Tsunami wavefield reconstruction and forecasting using the ensemble Kalman filter. *Geophysical Research Letters 46*(2), 853–860. https://doi.org/10.1029/2018GL080644.

Cite this article: Fauzi A, Renzi E, Dias F, Pelaez-Zapata DS and Kokina T (2026). Data-driven discovery of meteotsunami patterns from sparse observations. *Environmental Data Science*, 5: e9. doi:10.1017/eds.2026.10036

https://doi.org/10.1017/eds.2026.10036 Published online by Cambridge University Press