



Advanced weather typing for downscaling of wave climate and storm surge at a UK nuclear power station

Zehua Zhong¹ · Hachem Kassem¹ · Ivan D. Haigh¹ · Dafni E. Sifnioti² · Ben Gouldby³ · Ye Liu³ · Paula Camus^{1,4}

Received: 23 October 2024 / Accepted: 17 March 2025 / Published online: 28 March 2025
© The Author(s) 2025

Abstract

Evaluating risks from external hazards is crucial for the safety of nuclear power stations throughout their lifecycle. In coastal areas, a key threat arises from the risks of coastal flooding and erosion via a combination of simultaneous processes (e.g., tides, waves, and storm surges) acting on varying spatial and temporal scales. Therefore, an accurate characterisation of local sea state conditions is essential for risk assessment and mitigation. In this paper, we use a weather typing method to downscale local wave climate and storm surge conditions at the Hartlepool nuclear power station. Model validation suggests that the use of 36 weather types can effectively downscale multivariate wave variables (wave height, period, and direction) and storm surge with overall good performance, though the accuracy is limited for wave direction and extreme wave height. Comprehensive sensitivity tests are conducted to investigate key factors influencing the downscaling process, including predictor variable, spatial and temporal definitions, predictor resolution, the number of weather types, and the weighting parameter in semi-supervised classification. For example, we find that the model with sea level pressure and sea level pressure gradient as the predictor has better overall performance in downscaling multivariate predictands than the model using either one individually. These results can facilitate the development of weather typing models to enable efficient and reliable estimations of local predictands in wider applications. This approach links atmospheric conditions to potential coastal threats, which offers a valuable tool for proactive hazard preparedness and risk management in nuclear power and other critical infrastructure sectors.

Keywords Weather types · Wave climate · Storm surge · Sensitivity analysis · Statistical downscaling

1 Introduction

Understanding the driving mechanisms of, and contributing factors to, coastal hazards is fundamental to providing reliable characterisation and predictions of hazard-related parameters, which are essential for robust hazard assessments. Coastal hazards like flooding and erosion often arise from the simultaneous or consecutive occurrence of extreme conditions in multiple processes, including wave conditions, storm surges, astronomical tides, and river discharge (Leonard et al. 2014). This multivariate nature of hazard-related conditions is critical in quantifying risk potential, as events with extreme impacts can also occur without any of the individual processes being extreme on their own. These sea state components exhibit large variabilities and act on a wide range of spatial and temporal scales, which complicates their characterisation and assessments of their local impacts. For instance, coastal wave climates typically constitute local wind waves and swell waves generated by distant winds

Responsible Editor: Antonio Bonaduce

✉ Zehua Zhong
Zehua.Zhong@soton.ac.uk

¹ School of Ocean and Earth Science, National Oceanography Centre, University of Southampton, European Way, Southampton SO14 3ZH, UK

² EDF Research & Development UK Centre, 90 Whitfield Street, London W1T 4EZ, UK

³ HR Wallingford, Howbery Park, Wallingford OX10 8BA, UK

⁴ Geomatics and Ocean Engineering Group, Departamento de Ciencias y Técnicas del Agua y del Medio Ambiente, E.T.S.I.C.C.P., Universidad de Cantabria, 39005 Santander, Spain

several days before their arrival (Hegermiller et al. 2017). Storm surges result from the combined effect of local wind stress and regional barometric pressure anomalies and can occur over several hours to days depending on the weather events. Over interannual time scales, climate oscillation patterns like El Niño Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO) also modulate local wave and water conditions (Odériz et al. 2020; Scott et al. 2021). In addition, climate change can significantly increase exposure to coastal hazards in the long term, with research linking global sea level rise with heightened risks of coastal flooding and erosion hazards (Vousdoukas et al. 2018, 2020).

The characterisation of exposure to coastal hazards is thus a complex problem that requires consideration of the inherent variability of multiple hazard drivers. However, historical records are limited in spatial coverage and often not long enough to have observed the full range of variabilities, especially for extreme conditions (Anderson et al. 2019). Numerous approaches have been proposed to address this issue. For example, stochastic models were designed to generate a large number of synthetic sea state scenarios to allow for more robust and reliable coastal risk analysis (Wahl et al. 2011; Gouldby et al. 2014). Additionally, the ever-increasing computation power has advanced numerical simulations (reanalyses and hindcasts) on regional or even global scales (e.g., Hersbach et al. 2020; Muis et al. 2020), which in turn provide insights into sea state conditions at locations where observational data are absent, as well as the spatial distribution and variations of extreme conditions.

Meanwhile, with advancements in global climate models (GCMs) providing more reliable climate projections for various climate-change scenarios, downscaling techniques have emerged as a promising tool to translate large-scale climate forcings into local sea state responses to facilitate local climate impact studies. These techniques can account for local-scale processes unresolved in GCMs and provide outputs at resolutions suited to coastal applications. Conventional downscaling methods can be broadly classified into dynamical or statistical models. Dynamical downscaling defines approaches to embedding higher-resolution regional climate models (RCMs) into GCMs, using outputs from GCMs as boundary conditions, capable of producing accurate and physically consistent local information (Fowler et al. 2007; Gaur and Simonovic 2019). However, the associated high computational demand has led to interest in applying statistical downscaling (SD) models as an efficient alternative. SD models relate large-scale atmospheric conditions (predictors) to local variables (predictands) through developing empirical statistical relationships. Three categories of SD methods exist: regression-based models, weather-type approaches, and stochastic weather generators (Wilby and Wigley 1997). SD methods can downscale GCM outputs in multiple scenarios with

low computational cost and are thus desirable for estimating uncertainties associated with climate variabilities (Fowler et al. 2007; Gaur and Simonovic 2019).

One SD method based on weather type classification (Camus et al. 2014b), originally proposed to downscale multivariate local wave climates, has been widely applied in coastal settings (e.g., Rueda et al. 2019; Costa et al. 2020; Zhao et al. 2024). A key advantage of this type of approach is the ability to trace back the synoptic conditions (i.e., the atmospheric conditions over a wide area at a given time) responsible for local responses (Rueda et al. 2016). In addition to downscaling local variables, recent studies have demonstrated their potential to provide robust projections of coastal conditions to manage coastal erosion (Toimil et al. 2017) and build operational forecasting tools for coastal flooding (Neal et al. 2018). Other efforts have been made to couple such weather typing methods with climate emulators to account for climate variabilities at multiple temporal scales (e.g., Anderson et al. 2019; Cagigal et al. 2020).

A fundamental requirement for approaches based on weather types and other SD models is that there should be a strong relationship between the predictors and predictands (Schoof 2013; Gaur and Simonovic 2019). Research efforts have been made to identify suitable predictor variables depending on the predictand of interest. For example, variables based on sea level pressure and wind have been considered for downscaling significant wave height (Wang et al. 2010; Casas-Prat et al. 2014) and storm surge (Costa et al. 2020). Besides the choice of predictor variables, factors like the spatial and temporal definitions (i.e., the spatial and temporal extent over which the predictor is defined) also affect the performance of SD models (Camus et al. 2014a; Hegermiller et al. 2017). While these findings are generally transferable to SD model development elsewhere, there could be site-specific factors that need to be accounted for. For instance, the multivariate regression model developed in Cid et al. (2017) for reconstructing global daily surge levels showed reduced accuracy around tropical areas compared to higher latitudes. They attributed this discrepancy to the larger spatial and temporal scales of ocean dynamics at equatorial regions, which their uniform spatial and temporal definitions of predictor could not adequately capture. This highlights the importance of considering the impact of these factors and adjusting the predictor definition to sufficiently account for the processes contributing to the variations of local variables. Despite these insights, only a limited number of studies have investigated the impact of different factors on weather typing and optimised their model configurations accordingly. Most of them considered a few factors and their impacts on significant wave height and storm surge, with little attention paid to other predictands like wave period and direction. Consequently, there remains the need for a

more comprehensive analysis of how different factors affect the downscaling of local multivariate sea state conditions.

Downscaling local wave climates from regional atmospheric conditions has been the focus of extensive research. Advances have been made by recognizing distinct features in the generation of wind waves and swells, and incorporating these differences into predictor definitions (Camus et al. 2014a; Hegermiller et al. 2017). However, most studies have primarily focused on estimating bulk wave parameters (e.g., significant wave height and mean wave period) for the combined wind waves and swells (e.g., Camus et al. 2014b; Rueda et al. 2016; Zhao et al. 2024), while few studies have explicitly evaluated how these SD methods perform for individual wind wave and swell parameters. Characteristics of these two wave components can be crucial to understand nearshore wave processes to support wave impact studies and coastal hazard assessments.

The overall aim of this paper is to apply a weather typing method to downscale local hazard-related conditions at a nuclear power station in the UK. The study has three objectives: (1) develop a set of weather types to characterise the synoptic conditions and their relationships with wave climate and storm surge at the chosen site; (2) validate the model's ability to downscale local multivariate wave climate and storm surge; (3) investigate the impact of various factors on weather typing. Here, we explore the limit of the weather typing model in providing multivariate sea state conditions, including the combined wind waves and swells, the individual wind wave and swell components, and storm surge. Model validation was conducted to ensure model performance and robustness for each predictand. We then identify factors affecting the downscaling process and evaluate their relative importance through sensitivity analysis. The results will provide relevant information for the probabilistic assessment of coastal hazards at a later stage.

This paper is structured as follows. Section 2 describes the data involved and the processes for weather typing, model validation, and sensitivity analysis. Section 3 presents the developed weather types and their associated distributions of local predictands (i.e., the statistical relationships) and evaluates the model performance and its sensitivity to different factors. Section 4 further discusses the impact of these factors, with conclusions given in Section 5.

2 Data and methodology

2.1 Data sources

This study focuses on the EDF Energy UK Hartlepool nuclear power station (Fig. 1), which is located on the north-east coast of England, facing the North Sea. Historically, the North Sea region experienced several intense storm events,

which had caused extensive coastal flooding to many of the adjacent countries (e.g., in 1953 and 2013). Coastal erosion is also a major concern, with 27% of the coastline undergoing erosion in the northeast of England (Masselink et al. 2020). The exposure to these potential risks at Hartlepool needs to be better understood for the safety of daily operations and long-term risk management.

Three types of data are utilised in this study at this site, which are described below (also see Table 1). First, sea level pressure (SLP) is used to represent the regional atmospheric conditions. We used hourly mean SLP fields from the fifth generation of atmospheric reanalysis (ERA5) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF; Hersbach et al. 2020). ERA5 provides detailed estimates of the global atmosphere, land surface and ocean waves from 1940 to present. The spatial resolution of the atmospheric variables is 0.25° . Three spatial domains are considered in this work (see Fig. S1).

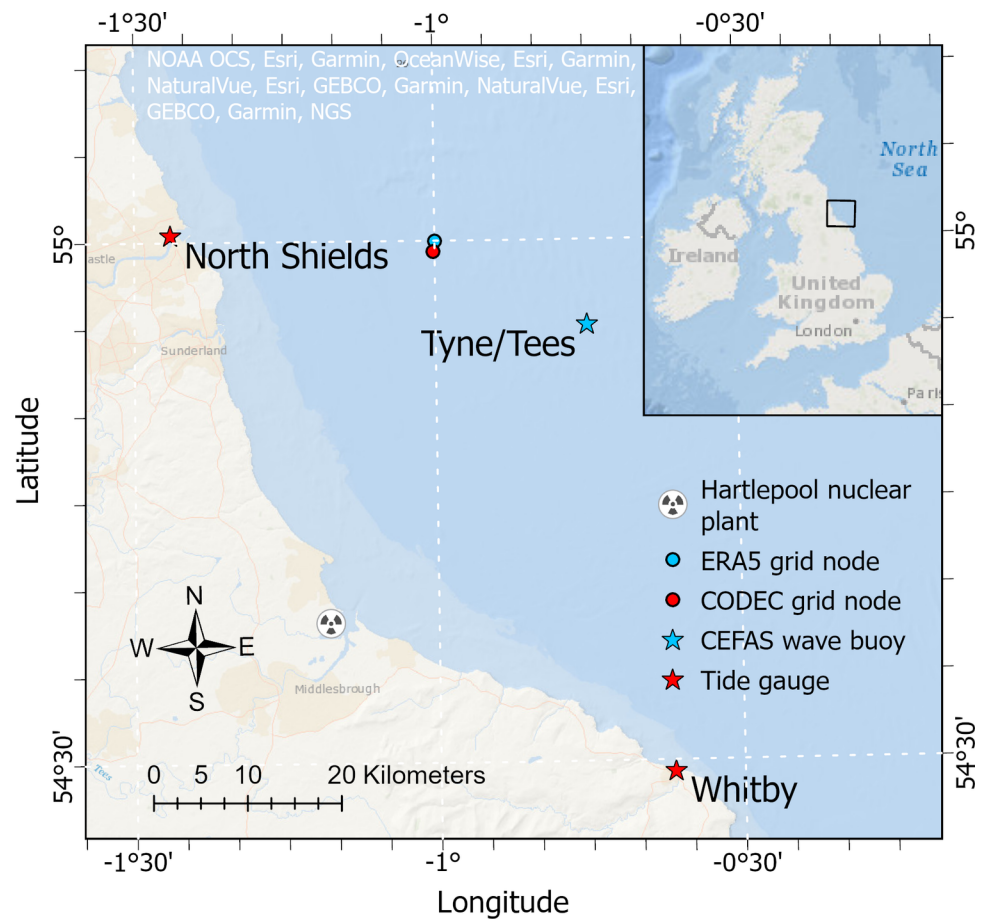
The local wave climate is also obtained from ERA5 at the grid node closest to Hartlepool (55°N , 1°W), approximately 42 km from the site. Wave parameters are generated by an ocean wave model (WAM; The WAMDI Group 1988) embedded in the ECMWF Integrated Forecast System. In addition, wave buoy observations from CEFAS WaveNet site at Tyne/Tees ($54^\circ55'08''\text{N}$, $0^\circ44'55''\text{W}$) are used to validate the ERA5 wave data. The wave buoy has a 30-min resolution while ERA5 provides hourly values.

Last, to derive non-tidal residuals (storm surges), still water levels from two sources are used: (1) modelled still sea levels from the Coastal Dataset for the Evaluation of Climate Impact (CODEC; Muis et al. 2020); and (2) tide gauge observations at North Shields and Whitby from the UK National Tide Gauge Network (obtained from the British Oceanographic Data Centre). Here we use CODEC as the primary data source as it is continuous (i.e., without large gaps in the time series). The grid node closest to the aforementioned ERA5 wave grid node in CODEC is selected for model development. Tide gauge measurements are used to validate and correct bias within the CODEC data. As water levels from CODEC are available only from 1979 to 2018 (40 years in total), the same period is used for SLP and waves for consistency. The links to access these datasets are provided in the Data Availability section.

2.2 Data validation and pre-processing

2.2.1 Atmospheric data

The original high-resolution hourly SLP fields from ERA5 are averaged to produce daily SLP values at a 2° resolution. The spatial averaging needs to account for the change in the area each point represents due to the Earth's curvature. Therefore, each point is weighted by the cosine of its latitude. SLP fields at 0.5° and 1° resolution are also obtained to test the

Fig. 1 Map of the study site and locations of data sources**Table 1** Summary of the datasets used in this work

Data type		Variables	Temporal resolution	Source (Station)	Period considered
Modelled datasets	Atmospheric conditions	Mean SLP	1 h	ERA5	1979–2018 (40 years)
	Wave	Combined wind waves and swells: H_s , T_m , T_p , θ_m	1 h	ERA5	
		Wind waves: H_s , T_m , θ_m			
		Swells: H_s , T_m , θ_m			
Observations	Non-tidal residual	Total water level	1 h	CODEC	2006–2018
	Wave	Combined wind waves and swells: H_s , T_p	30 min	CEFAS WaveNet (Tyne/Tees)	
	Non-tidal residual	Total water level	1h (before 1993), 15 min (1993 onwards)	UK National Tide Gauge Network (North Shields, Whitby)	

H_s significant wave height, T_m mean wave period, T_p peak wave period, θ_m mean wave direction

sensitivity of weather typing to spatial resolutions. In addition to SLP, the squared SLP gradient (SLPG) is also included as the predictor. Mathematically, it is the sum of the squared SLP gradients along both latitude and longitude (Wang and Swail

2006). Thus, SLPG at a grid node can be calculated from SLP values at the surrounding four grid nodes. We derive the daily SLPG fields at 0.5°, 1° and 2° resolutions from their corresponding daily SLP fields with the same spatial resolutions.

2.2.2 Wave data

Wave parameters extracted from ERA5 include significant wave height H_s , mean wave period T_m , peak wave period T_p , and mean wave direction θ_m for combined wind waves and swells. Hereafter, the combined wind waves and swells will be referred to as total waves. In addition, wind wave and swell parameters, including H_s , T_m , and θ_m , are also available from ERA5 and were used for detailed analysis of these two wave components. We observed an anomaly in the wind wave data, characterised by a large concentration of data within a narrow range of T_m from 3.8 to 3.9 s and θ_m from 35° to 36°. Most of these waves have wave heights below 3 mm. This results in an unusually tall and narrow spike in the joint probability distributions. Given the disproportionately high probability density of these data points, we considered them suspicious and excluded them from our analysis.

The CEFAS wave buoy at Tyne/Tees is used to validate the ERA5 wave data. Due to data availability, the validation is only conducted for H_s and T_p (as CEFAS provides peak wave direction instead of mean direction) for the period of 2006–2018. A comparison of the two datasets shows relatively better agreement for H_s (Pearson's correlation $\rho=0.97$, root mean square deviation RMSD=0.35 m) but weaker results for T_p ($\rho=0.62$, RMSD=2.48 s). It is worth noting that ERA5 tends to underestimate extremes in H_s and T_p , a systematic bias caused by factors such as uncertainties in wind forcing and limitations in modelling wave processes (Fanti et al. 2023). Therefore, bias correction of the wave reanalysis is usually considered important to avoid under-predictions for coastal risk assessments where extreme events are of particular focus. However, we use the original wave data from ERA5 based on the following considerations: (1) we are looking at a much wider range of variables than the local wave buoy can provide; and (2) correcting the modelled data independently using the corresponding observations may compromise the physical consistency and dependencies between different wave variables (Bhowmik and Sankarasubramanian 2019). For developing weather types and understanding factors relevant to wave downscaling, using uncorrected reanalysis data is unlikely to introduce significant errors.

2.2.3 Non-tidal residuals

The non-tidal component of sea level is obtained following the approach of Jenkins et al. (2023). First, a linear regression method is applied to still water levels to identify any trend in mean sea level, which is then removed. Next, the astronomical tides are derived from the detrended water levels through a year-by-year harmonic analysis using the MATLAB Unified Tidal Analysis and Prediction Functions (UTide; Codiga 2024). The tidal constituents are selected

using an automated decision tree method. Last, the astronomical component is subtracted from the detrended water levels to leave the non-tidal residuals. Hereafter, the term “storm surge” or “surge” will be used to refer to the non-tidal residual.

To validate the CODEC data, we compare it with records from the two nearby tide gauges. Water levels at grid nodes closest to North Shields and Whitby are extracted from CODEC. Storm surges are calculated for, and compared between, those two grid nodes and two tide gauges. The results indicate that: (1) there is no significant spatial variation in storm surge in this region ($\rho=0.95$, RMSD=0.05 m between observed surges at the two tide gauges); (2) CODEC agrees well with observations ($\rho=0.83$ and RMSD=0.10 m for hourly values, and $\rho=0.89$ and RMSD=0.09 m for daily maximum). It is worth noting that CODEC also tends to slightly underestimate high surge values and overestimate low surge values. To account for this, a bias correction method is applied. We perform a simple linear regression on hourly surges between CODEC (SS_{mod}) and tide gauge observations (SS_{obs}) at North Shields and Whitby (Eq. 1). The resulting gradient k and intercept b are applied to the CODEC to get bias-corrected surges (SS'_{mod}) at the previously selected grid node for model development (Eq. 2). After correction, the linear regression equation between SS'_{mod} and SS_{obs} yielded $k=1$ and $b=0$ (i.e., the regression line aligns perfectly with the 1:1 diagonal line), while ρ remains unchanged but RMSD slightly increases from 0.10 to 0.12 m.

$$SS_{mod} = k \cdot SS_{obs} + b \quad (1)$$

$$SS'_{mod} = (SS_{mod} - b)/k \quad (2)$$

2.3 Statistical downscaling

The SD model applied herein is based on the weather typing method developed by Camus et al. (2014b, 2016) which embraces several data mining techniques. An overview of the method is shown in Fig. 2. The method produces a set of daily synoptic circulation patterns (known as weather types) along with the corresponding empirical distributions of local predictands (i.e., wave and storm surge variables at Hartlepool).

2.3.1 Developing weather types

The predictors are defined by 2° daily mean SLP and SLPG fields with a spatial domain spanning from 26°N to 76°N and from 40°W to 30°E in the North Atlantic Ocean (see Domain 2 in Fig. S1). To account for the latest weather conditions responsible for generating waves and storm surges,

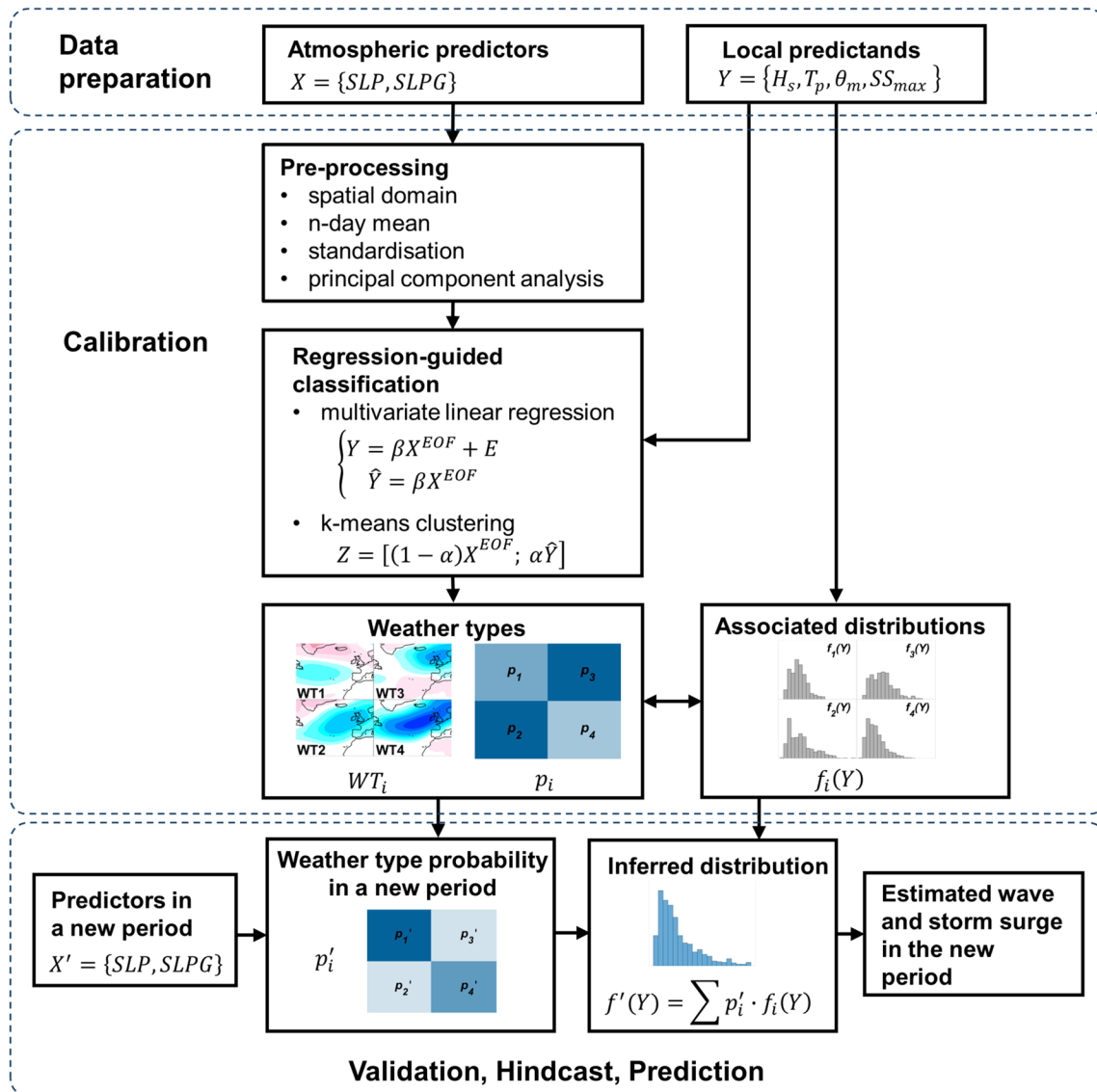


Fig. 2 Flowchart of the weather typing method based on Camus et al. (2014b). Symbols and equations are explained in detail in Sect. 2

the predictors are averaged over 4 days to represent the overall conditions over a given day and its preceding 3 days. The definition of predictors is important in the first place to ensure the validity and robustness of the weather typing model. The sensitivity of the model to different predictor definitions is tested in terms of the choice of predictor variable, the temporal coverage (i.e., the number of days averaged), and the spatial domain. The daily predictors are then standardised to have a mean of 0 and a standard deviation of 1, followed by a principal component analysis (PCA).

PCA is a widely used statistical technique to reduce the dimensionality of complex data and reveal underlying patterns. PCA identifies the directions in which the variance of the data is maximised, which are defined by the eigenvectors (known as Empirical Orthogonal

Functions or EOFs) of the data covariance matrix with the eigenvalue representing the associated explained variance (Jolliffe and Cadima 2016). The projected data onto each vector are referred to as the principal components (PC). The original dataset X can then be reconstructed by

$$X = \sum_{i=1}^n EOF_i \cdot PC_i \quad (3)$$

where n is the total dimension of dataset X . The PCs are sorted in decreasing order of explained variance, with the first PC explaining the most data variance and each subsequent PC explaining progressively less. As such, it is valid to use a reduced number of PCs (i.e., the first d PCs; Eq. 4) and

the corresponding EOFs to represent the original dataset. Typically, PCs explaining 95% of data variance are selected.

$$X^{EOF} = \{PC_1(t), \dots, PC_d(t)\} \quad (4)$$

Next, a regression-guided classification (Camus et al. 2016) is employed to categorize the predictors into a pre-defined number of weather types (WTs). This semi-supervised approach differs from the original method (Camus et al. 2014b) in that it incorporates local predictand information to guide the classification so that weather patterns generating similar local conditions are more likely to be clustered together. The regression-guided classification is achieved by two steps. First, a multivariate linear regression model (Eq. 5) is established between the predictor X^{EOF} and the selected predictand Y which consists of daily mean H_s , T_p and θ_m and daily maximum storm surge SS_{max} (Fig. 2). The fitted coefficient β is then used to make predictions for $\hat{Y}$ (Eq. 6).

$$Y = \beta X^{EOF} + E \quad (5)$$

$$\hat{Y} = \beta X^{EOF} \quad (6)$$

The regression model captures the synoptic-scale influence on the local predictands and filters out irrelevant information contained in observations caused by other processes (Cannon 2012). The predicted $\hat{Y}$ is then concatenated with X^{EOF} to form a new dataset Z (Eq. 7). Here, a weighting parameter α is introduced to adjust the relative importance of predictors and predictands in the clustering process. The classification becomes unsupervised and fully supervised when α is 0 and 1, respectively. We used a value of 0.6 based on previous research (Camus et al. 2016), and conducted sensitivity tests across the range of 0 to 1.

$$Z = [(1 - \alpha)X^{EOF}; \alpha\hat{Y}] \quad (7)$$

The dataset Z is then clustered using a k-means clustering algorithm (KMA), a classification scheme that aims to partition a set of data points into K clusters so that the total within-cluster variance is minimised. KMA iteratively assigns each data point to the nearest cluster centroid and then updates the centroids based on the newly formed clusters (Hastie et al. 2009). While the initial cluster centroids can be randomly selected from the data points, we adopt the maximum dissimilarity algorithm (MDA; Snarey et al. 1997) to ensure that the initially selected centroids are well-spaced and representative of the original dataset, as well as to guarantee a deterministic clustering outcome (Camus et al. 2011, 2014b). The generated clusters (i.e., WTs) are represented by the corresponding cluster centroids projected back to the original high-dimension space, which are organised into a bi-dimensional lattice in a way that WTs with

similar patterns (i.e., lower centroid distances) are placed together for better visualisation (Camus et al. 2014b).

The following step is to define empirical statistical relationships between predictors and predictands. Since each WT is associated with certain dates, hourly wave and surge parameters of those dates are collected to derive the probability distributions $f_i(Y)$ of these parameters associated with each WT. Together with the probability of occurrence p_i for each WT_i , the complete distribution of the local predictands $f(Y)$ during the calibration period can be given as follows:

$$f(Y) = \sum_{i=1}^N p_i \cdot f_i(Y) \quad (8)$$

where N is the number of WTs. The empirical relationships $f_i(Y)$ derived in the calibration period are assumed to be stationary and valid in any period. Hence, the probability distributions of predictands in a new period $f'(Y)$ can be inferred by:

$$f'(Y) = \sum_{i=1}^N p'_i \cdot f_i(Y) \quad (9)$$

where p'_i is the new probability for WT_i in that period.

2.3.2 Model validation

The model's ability to downscale local wave climate and storm surge is evaluated by K-fold cross-validation. Here K means the number of subsets into which the whole dataset is divided and the number of iterations the cross-validation process undergoes. We select $K=5$ to ensure that each training data subset remains representative of the whole dataset while maintaining a manageable computational load for the validation process. First, the entire 40 years of data are split into five subsets. Each subset contains data at five-year intervals (e.g., the first subset contains data from 1979, 1984, ..., and 2014). Each time, four periods are used as calibration datasets to develop a set of weather types, which are then used to estimate local waves and surges for the fifth period (i.e., the validation period). We estimate the monthly mean and/or the monthly and annual n -th percentile of a predictand over the validation period based on the empirical distributions derived in Eq. 9. The subset not included in previous model training can be used as the reference to validate against the model estimations. This process is repeated five times with each subset being used exactly once as the validation dataset. Note that five sets of weather types (i.e., five models) will be developed during the validation and they are different from the weather types derived from the whole period and used for final risk assessment. Results are shown in Sect. 3.2. Three metrics are calculated to evaluate how well the reference values x_i (i.e., values directly

calculated from ERA5 and CODEC) are replicated by the model y_i (i.e., values estimated from the weather typing model). The Pearson's correlation ρ is given by:

$$\rho = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\sigma_x} \right) \left(\frac{y_i - \bar{y}}{\sigma_y} \right) \quad (10)$$

where N is the number of monthly values, $\bar{x}$ is the mean of x_i , and σ is the standard deviation. The bias can be calculated by:

$$\text{BIAS} = \bar{x} - \bar{y} \quad (11)$$

The RMSD is given by:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad (12)$$

Equations 11 and 12 are also applicable to wave directions with some adjustments on calculating the difference and mean. The angular difference in radian is given by the length of the shortest arc from one direction to the other along the unit circle. The mean direction $\bar{\theta}$ is calculated by:

$$\bar{\theta} = \tan^{-1} \left(\frac{\sum_{i=1}^N \sin \theta_i}{\sum_{i=1}^N \cos \theta_i} \right) \quad (13)$$

Note that wave directions are by default in the nautical convention (i.e., measured clockwise from the North or the positive y-axis) in our context and they need to be converted to the Cartesian convention (i.e., measured counterclockwise from the East or the positive x-axis) before applying Eq. 13. Likewise, the resulting mean direction should be transformed back to the nautical convention. For correlation, we use a MATLAB toolbox to calculate circular-circular correlation ρ_{cc} (Berens 2009):

$$\rho_{cc} = \frac{\sum_i \sin(\theta_i - \bar{\theta}) \sin(\varphi_i - \bar{\varphi})}{\sqrt{\sum_i \sin^2(\theta_i - \bar{\theta}) \sum_i \sin^2(\varphi_i - \bar{\varphi})}} \quad (14)$$

where θ_i and φ_i are two directional variables and $\bar{\theta}$ and $\bar{\varphi}$ are the angular means.

2.3.3 Sensitivity analysis

A series of sensitivity analyses was conducted to evaluate the importance of different factors in the downscaling of local wave and surge conditions. Table 2 lists all the factors and values considered. The tests were done by varying one factor at a time while keeping the rest at their baseline values. For each testing factor, several weather typing models with different model settings were

Table 2 Factors and their values tested in the sensitivity analysis. The values in bold are baseline values and are also used in the final model settings. The three spatial domains are shown in Fig. S1

Testing factors	Values
Predictor variable	SLP; SLPG; SLP and SLPG
Temporal coverage (days)	1; 2; 3; 4 ; 5
Spatial domain	Domain1 (50°W–40°E, 20°–80°N) Domain2 (40°W–30°E, 26°–76°N) Domain3 (30°W–20°E, 30°–70°N)
Spatial resolution of predictors (°)	0.5; 1; 2
Number of weather types	16; 25; 36 ; 49; 64
The weighting parameter α in regression-guided classification	0; 0.2; 0.4; 0.6 ; 0.8; 1

developed and validated regarding their abilities to estimate the monthly averages of H_s , T_p , θ_m , and the monthly 95th percentile of storm surge (SS_{95}). Furthermore, the monthly mean of wind wave and swell wave parameters (H_s , T_m , θ_m) were also estimated to investigate how those testing factors impact the model skill in downscaling the two wave components.

The closeness of model estimations to the reference values from ERA5 and CODEC could be assessed by several metrics. Here, we use Taylor diagrams (Taylor 2001) to facilitate the intercomparison of multiple models. Taylor diagrams are designed to visualise three complementary metrics: ρ , RMSD, and σ . These statistics are interrelated with one another in a way that resembles the law of cosines, which allows them to be plotted on a 2D graph simultaneously.

However, the original Taylor diagram is constructed for scalar quantities, such as wave height and period, and does not apply to circular quantities like wave direction. To address this issue, a Vector Field Evaluation (VFE) diagram was proposed (Xu et al. 2016) which is essentially a generalized form of the Taylor diagram for evaluating vector quantities. Likewise, three interrelated statistical metrics are visualized to compare vector fields: vector similarity coefficient, root mean square length (RMSL), and root mean square vector difference (RMSVD). Detailed descriptions of these three statistics can be found in Xu et al. (2016). Note that wave direction in our context is not a vector as it does not have magnitude. Nonetheless, we can still use the VFE diagrams for evaluating wave directions by assuming the vector magnitude to be one. As a result, the RMSL for all models will be equal to one since this parameter measures the mean of vector magnitude. The results will be discussed in Sect. 3.3.

3 Results

3.1 Weather types and their associated distributions

The first objective of the study is to develop a set of weather types to characterise the synoptic conditions and their relationships with local wave climates and storm surge at Hartlepool. A set of 36 WTs was generated as the representative synoptic circulation patterns over the British Isles and the surrounding Atlantic and European region. These WTs are displayed in a bi-dimensional lattice in Fig. 3 where WTs exhibiting similar spatial SLP patterns are placed in adjacent locations in the lattice for better visualization. The seasonal probability of occurrence for each WT is shown in Fig. 4. WTs at the top of the lattice are generally characterised by a low-pressure system in the north and a high-pressure system in the south (e.g., WT1, 7, 25, and 31), though the location, size and intensity of these pressure systems vary across different WTs. WTs with such strong pressure gradients are most likely to occur in winter months. On the other hand, WTs with weak pressure gradients (e.g., WT16, 22, and 28) are found in the middle of the lattice and have higher probabilities of occurrence during summer. Notice that weather conditions in summer can be described by much fewer

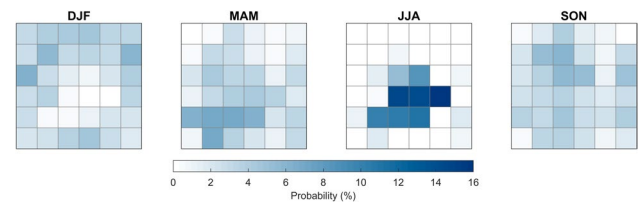
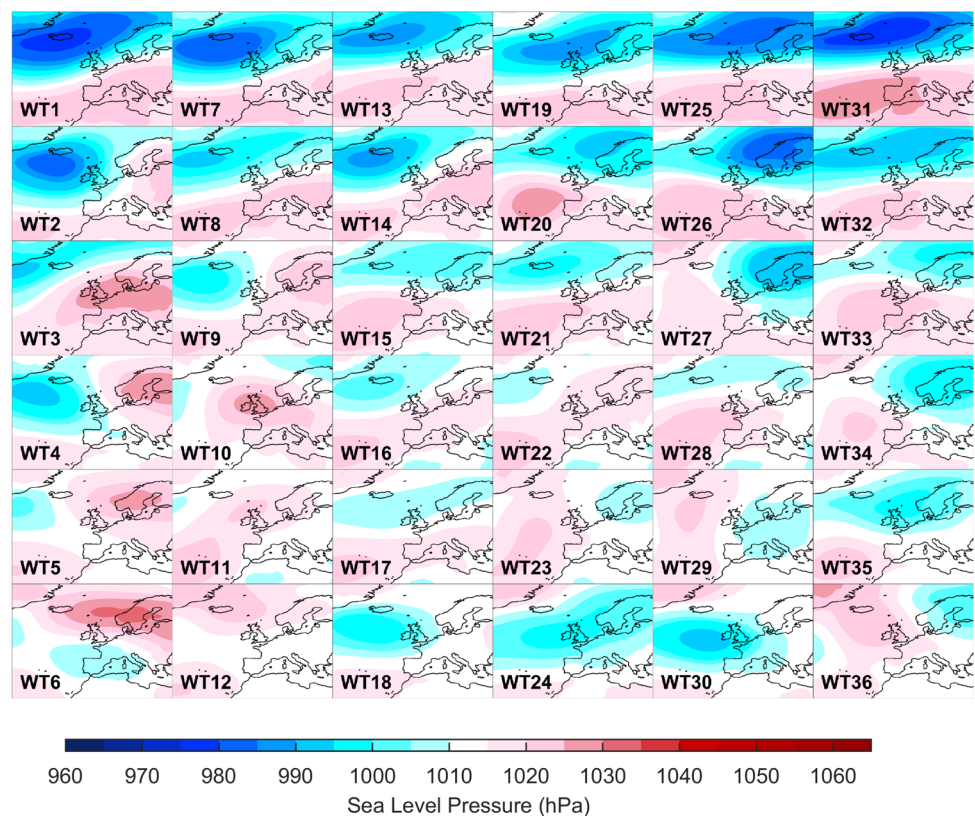


Fig. 4 Seasonal probability of occurrence of each WT. Each cell in the box corresponds to the WT of the same location in Fig. 3

WTs than those in other seasons (Fig. 4), which is also observed in Camus et al. (2014b) and Zhao et al. (2024). This indicates that the summer synoptic circulation patterns exhibit less variations.

Each individual WT is associated with unique distributions of sea state conditions. The distribution of hourly H_s for each WT is given in Fig. 5. Some distributions are characterised by a narrow and high peak below 2 m, indicating mild wave climate. Widespread distributions with a long tail are also found for certain WTs, suggesting the potential of extreme wave conditions. Together with other sea state parameters, these empirical distributions allow the investigation of the relationships between regional synoptic patterns and the local wave climates and surge response at Hartlepool. We illustrate this by comparing two WTs with contrasting synoptic patterns (Fig. 6). WT10 exhibits

Fig. 3 36 WTs represented by SLP. WTs with similar patterns are placed in neighbouring locations



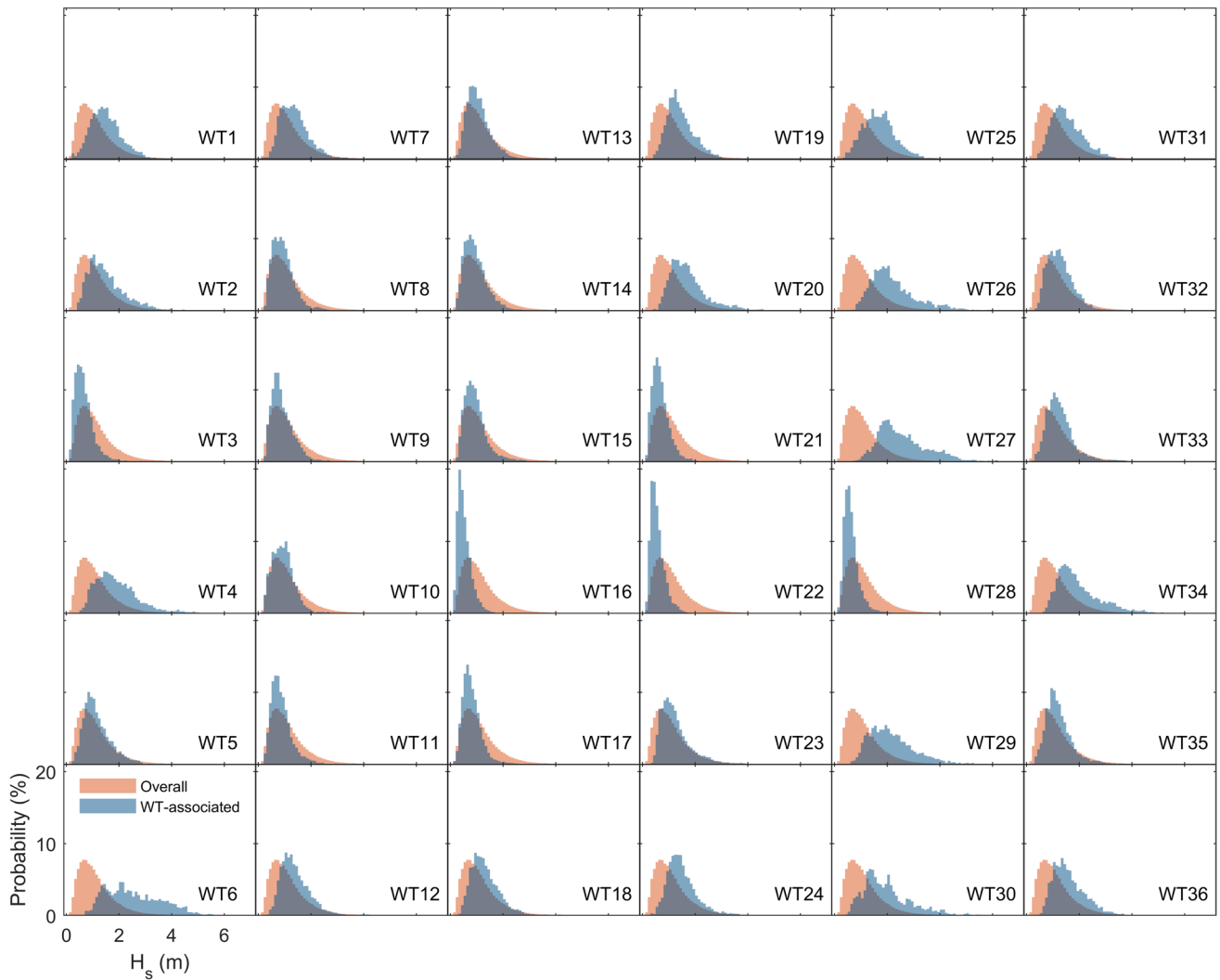


Fig. 5 The distribution of hourly H_s at Hartlepool for each WT (blue), along with the overall distribution of all 40 years of wave data (orange) for comparison

a high-pressure centre over the British Isles which results in waves mainly from the north with H_s generally below 2 m (0.94 m on average). In the H_s - T_p joint distribution, two peaks can be observed which mainly differ in T_p , with one roughly at 6 s and the other at 10 s, representing wind waves and swells respectively. This highlights the significance of both wave components under this synoptic condition. On the other hand, WT31 features a stronger-than-average pressure difference between the Icelandic Low and the Azores High. Under this condition, waves are mainly from the west with higher H_s (1.55 m on average) and a defined peak in T_p at 5s, which suggests the presence of strong wind wave activity. These two WTs also lead to very distinct surge behaviours. The surge distribution of WT10 has a more peaked shape ($\sigma=0.14$) which sits more on the negative side (84% of the surge is negative), whereas the distribution of WT31 is more widespread ($\sigma=0.33$) and more on the positive side (80%

of the surge is positive). The contrast in wave and surge behaviours demonstrates how local sea state parameters are affected by different regional synoptic conditions.

3.2 Validation of the weather typing model

The second objective of the study is to validate the weather typing model's ability to downscale local multivariate wave climate and storm surge. The model represented in Fig. 3 is designed to balance its overall performance on different sea state parameters. In Fig. 7, the model's ability to downscale ten parameters is evaluated, including the monthly means of H_s , T_p , T_m , θ_m , the monthly 95th and 99th percentiles of H_s for total waves (H_{s95} , H_{s99}) and storm surges (SS_{95} , SS_{99}), and the annual H_{s99} and SS_{99} .

For the four parameters representing mean wave conditions (Fig. 7a-d), the model estimations agree well with the

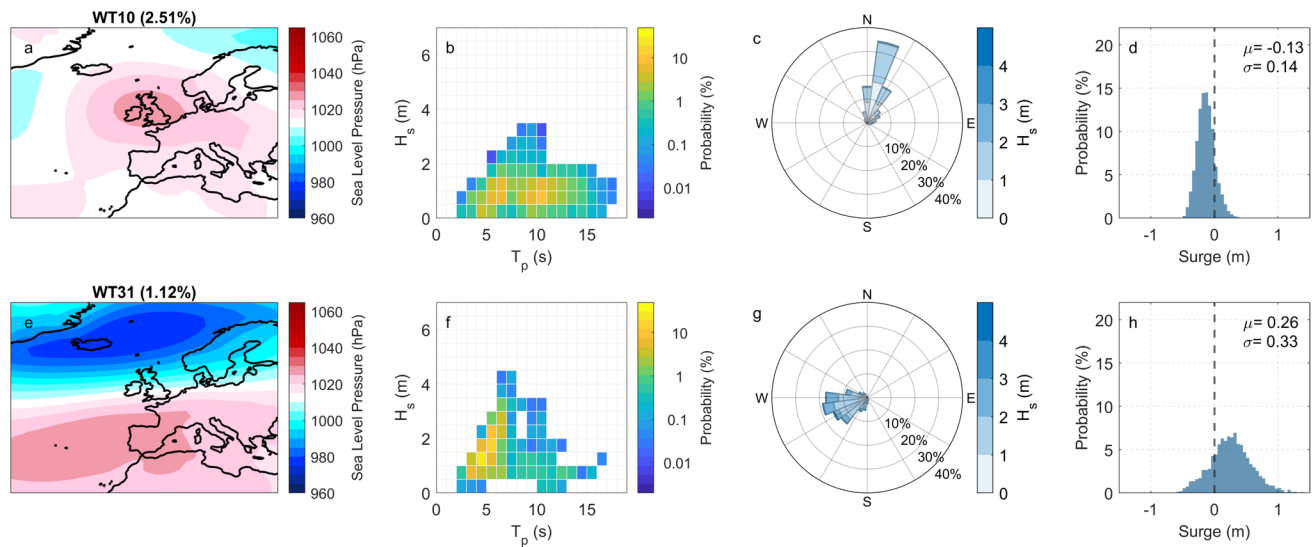


Fig. 6 Two WTs and their associated distributions of local predictands at Hartlepool. **a** The SLP pattern of WT10, with the probability of occurrence indicated. **b** The joint distribution of hourly H_s and T_p associated with WT10. **c** The joint distribution of hourly H_s

and θ_m associated with WT10, represented by a wave rose. **d** The distribution of hourly storm surge associated with WT10, with the mean μ and standard deviation σ displayed in metres. Similarly, panels (e) to (h) are for WT31

reference values from ERA5 and CODEC with relatively high ρ (0.87–0.94), low RMSD, and low BIAS. The wave direction appears to be the least accurate, with several points deviating from the 1:1 line and the RMSD reaching nearly 30°, but the overall performance remains good ($\rho_{cc}=0.87$, BIAS=0.63°). Besides the mean conditions, the model's ability to capture more extreme conditions is crucial in applications like coastal hazard assessments. We selected the 95th and 99th percentiles at monthly and annual scales to represent various extreme scenarios. The results suggest good model performance in representing extreme storm surges, with relatively low RMSD (within ~0.1 m). However, the performance for extreme wave events is less satisfactory compared to storm surge, as indicated by the broader scattering of data points, lower ρ , and larger RMSD (Fig. 7f and i). In general, the validation across a wide range of sea state parameters demonstrates the weather typing model's overall ability to downscale local conditions of total waves and storm surges, with reduced accuracy in wave direction and extreme wave height.

Although the development of the weather typing model is guided by total wave conditions, it can also be applied to downscale wind wave and swell variables. Similar validation was conducted regarding the model's skill in estimating the monthly means of H_s , T_m , and θ_m for wind wave and swell (Fig. 8). For both wave components, the estimations in H_s and T_m are almost as accurate as those for the total wave. For θ_m , the model performs much better for swells ($\rho=0.88$, RMSD=14.20°) than for wind waves ($\rho=0.72$, RMSD=47.08°). Note that wind wave θ_m is widely scattered

and covers all possible directions whereas swells mainly come from the northeast. As a result, the higher variation of wind wave direction is harder for the model to reproduce. Overall, the model can downscale the individual wind wave and swell components with reasonably good accuracy.

3.3 Sensitivity analysis

The third objective of the study is to investigate the impact of different factors on weather typing. The performance of SD methods largely depends on the quality of predictor definition, which includes the choice of predictor variables and how they are defined spatially and temporally. For weather-type approaches, the number of WTs and other method-specific factors can also affect downscaling skills. In this section, several aspects of the model configuration are investigated in detail as to their impacts on downscaling different sea state parameters. We use Taylor diagrams and VFE diagrams to visualise the differences in model skills. In these diagrams, models with higher ρ or vector similarity coefficients, lower RMSD or RMSVD, and standard deviations or RMSL closer to the reference values are considered more accurate, which can be indicated by their proximity to the reference points on the diagram.

3.3.1 Downscaling total wave and storm surge

First, we focus on the downscaling of total wave and storm surge conditions. In terms of predictor variables (Fig. 9), using the combination of SLP and SLPG as the predictor

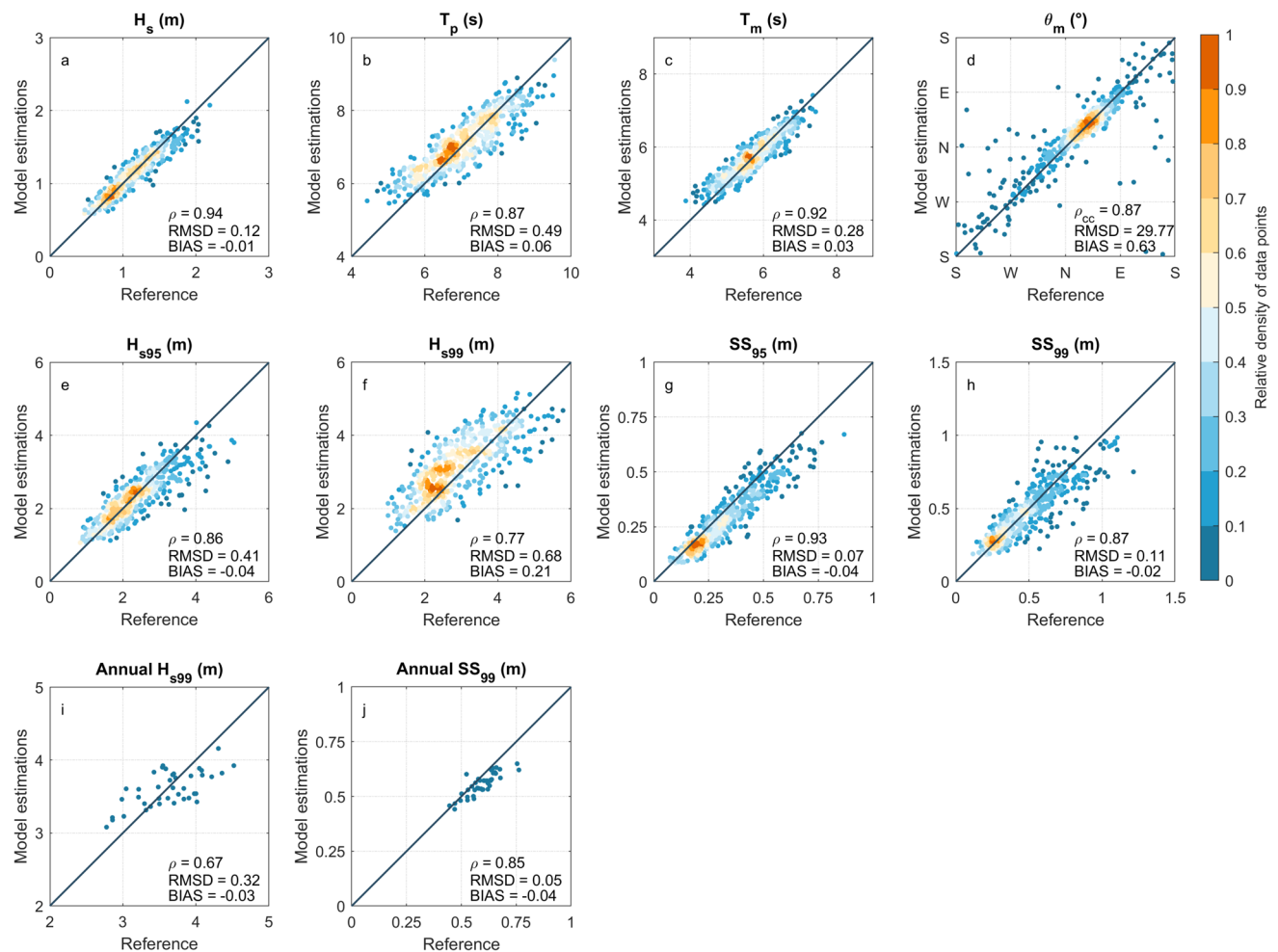


Fig. 7 Density scatterplots of monthly (panels (a) to (h)) and annual (panels (i) and (j)) sea state parameters in the period of 1979–2018 between model estimations and the corresponding reference values calculated from ERA5 and CODEC. The diagonal line in black is the 1:1 line. The axes in panel (d) for θ_m are adjusted so that most data

points are in the center of the figure. Three metrics (ρ , RMSD, and BIAS) are displayed to indicate model performance. Circular statistics are used for θ_m . The colour bar only applies to monthly parameters. See Sect. 2.3.2 for detailed explanations on model validation

yields better model performance compared to using either one individually, indicating that both variables are related to the local wave and surge conditions. For temporal coverage (Fig. S2), models with 2 to 5 days generally exhibit comparable accuracy at H_s , θ_m , and SS_{95} . However, the estimations of T_p improve significantly as more preceding days are included to define the predictor, suggesting that T_p is sensitive to the amount of recent synoptic conditions captured by the daily WTs. Similar results are observed for spatial domain (Fig. S3). The model with a bigger domain performs marginally better in T_p than that with a smaller domain. For the other three parameters, there is no significant difference among models with different values.

Regarding the spatial resolution of the atmospheric data used to define predictors, the model does not appear to be sensitive to this factor, as only a minor improvement in θ_m is found when using finer-resolution data (Fig. S4).

This indicates the use of low-resolution data in weather typing can also produce robust results while maintaining a low computational cost. In contrast, the estimations of all four predictands improve to varying degrees with a larger number of WTs, but this improvement becomes less pronounced as the number of WTs increases (Fig. S5). Finally, the model performance turns out to be highly sensitive to the value of α (Fig. S6), which represents the relative weight of local environmental information against atmospheric predictors in weather typing. The impact of α is not monotonic. For example, H_s is better reproduced as α increases from 0 to 0.6; for $\alpha > 0.6$, the model does not yield better results with a higher α , even when the model becomes fully supervised ($\alpha=1$). This pattern is also observed for the remaining predictands, with different threshold values ranging between 0.4 and 0.6. Waves seem to be more sensitive to α than surges.

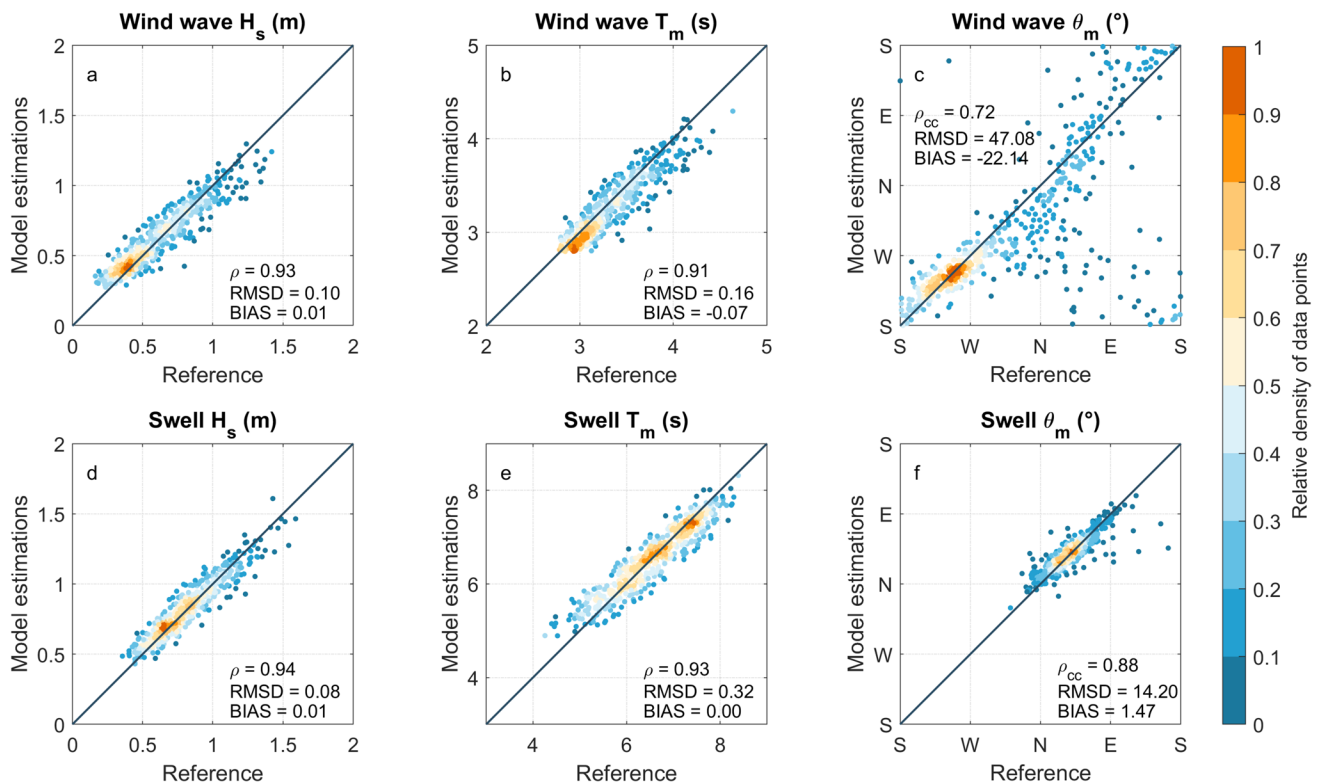


Fig. 8 The same as Fig. 7, but for monthly averages of wind wave and swell parameters

3.3.2 Downscaling wind wave and swell

The model sensitivity in downscaling wind wave and swell is also evaluated to reveal how the impact of each testing factor differs between these two wave components. For predictor variables, the model developed with SLP and SLPG exhibits the best overall accuracy (Fig. 10). Considering the cases with a single variable, using SLPG alone performs better than using SLP in downscaling wind wave H_s and T_m , with results nearly as accurate as those using both predictor variables. Conversely, for wind wave θ_m , SLP is more beneficial than SLPG. Unlike wind waves, the difference between SLP and SLPG is not that pronounced in swell. In this case, using both variables is important to achieve better model results.

In terms of temporal coverage (Fig. S7), the optimal value varies across different predictands, but wind waves generally favour shorter days, while swell parameters tend to benefit from longer days. Specifically, wind wave parameters achieve the best accuracy with a 2-day coverage, and their accuracy gradually decreases with more days included. Swells are better estimated when the predictors are averaged over 3 to 5 days. Among all predictands, the swell T_m seems to be the most sensitive. This predictand also exhibits some sensitivity to the chosen spatial domain, as indicated in Fig. S8. A larger domain leads to improved estimations for swell T_m , but the remaining predictands do not appear to be

affected by the spatial extent. Lastly, the impacts of predictor resolution, the number of WTs, and the value of α on wind wave and swell are similar to those on the total wave (i.e., without distinguishing wind from swell; Fig. S9–11).

4 Discussion

In this paper, we applied a weather typing method to downscale local hazard-related conditions, including wave climates and storm surges, at the Hartlepool nuclear power station. The weather types and their associated distributions offer valuable insights into the linkage between regional synoptic circulations and the local response of wave climate and storm surge. For example, the SLP pattern of WT31 (Fig. 6e) resembles the standard positive phase of NAO, which is characterised by strengthened westerlies. Such wind conditions result in waves mostly coming from the west (Fig. 6g) with larger H_s than normal conditions (Fig. 5). These waves are mostly young wind seas because the fetch (e.g., the area of ocean over which the wind blows) west of our selected grid node is limited (Fig. 1). On the other hand, the regional atmospheric pressure at Hartlepool in WT31 is relatively low which leads to more positive surges (Fig. 6h) due to the inverse barometer effect (Pugh and Woodworth 2014). Additionally, the higher frequency of extreme storms

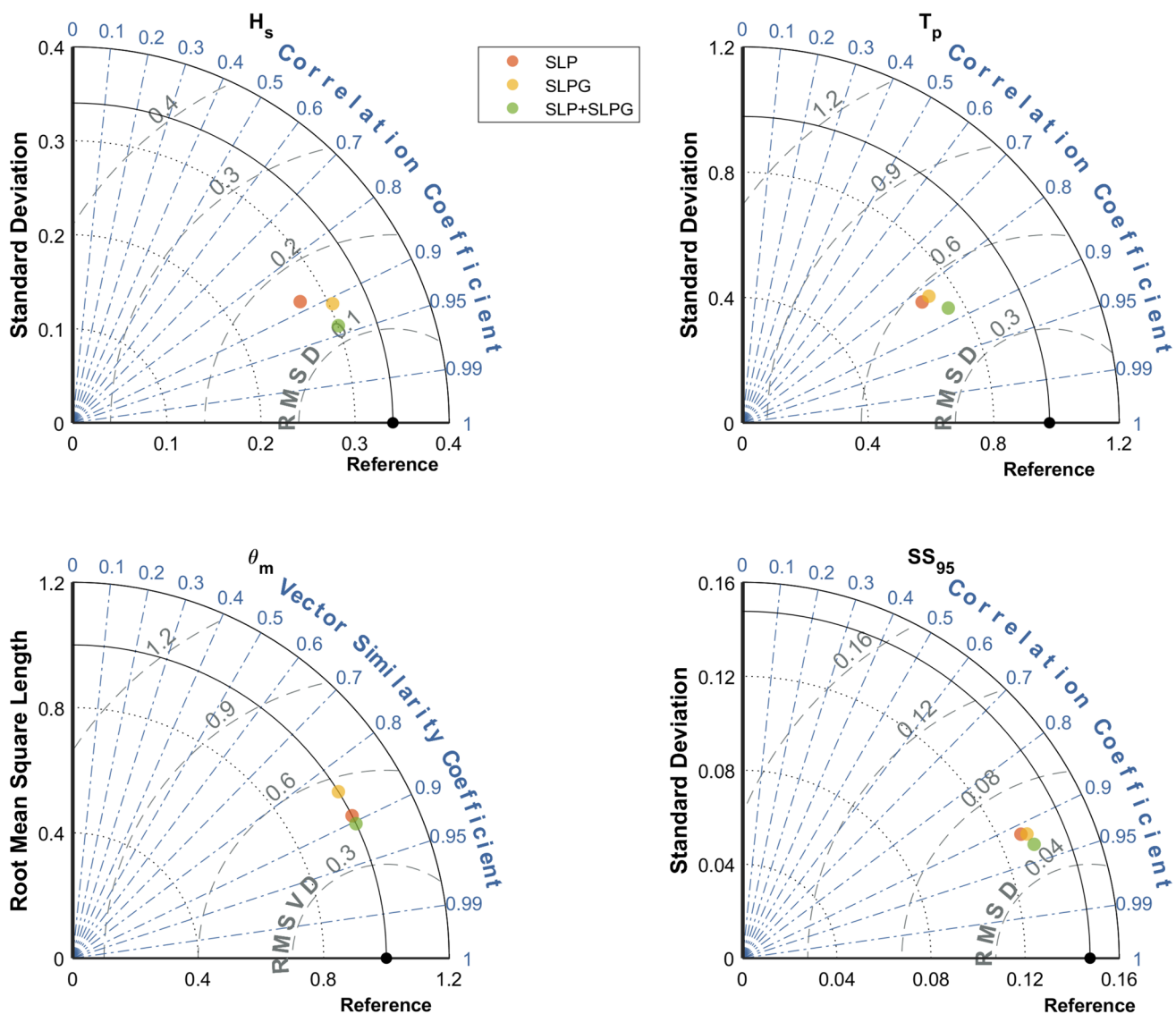


Fig. 9 Diagrams for comparing model performance in estimating total wave and storm surge conditions with different predictor variables. Taylor diagrams are used for scalar quantities whereas the VFE diagrams are used for circular quantities. Each coloured dot represents a model with a different configuration. The black reference points are calculated based on ERA5 or CODEC. The correlation

coefficient and vector similarity coefficient are indicated by the azimuthal angle (blue dash-dotted contours); the RMSD and RMSVD are proportional to the distance between each coloured point and the reference point (grey dashed contours); the standard deviation and the RMSL are proportional to the radial distance from the origin (black dotted contours)

in positive NAO phases (Pinto et al. 2009) likely contributes to the increased probability of large surges.

In addition to providing intuitive understanding of the local impact of regional synoptic systems, the statistical relationships represented by the WTs and their corresponding empirical distributions are the foundation for predicting local sea state conditions in practical applications. In this work, we validated the model's ability to estimate multiple sea state parameters, including both the mean and extreme conditions. While the model performs well in predicting mean wave conditions and extreme storm surges, its

accuracy in estimating extreme wave height remains limited. This limitation could stem from several factors: (1) the use of mean wave height rather than extreme conditions (such as daily maxima) in the regression-guided classification; (2) the model's dependence on historical extreme events associated with each WT, which is constrained by dataset length; and (3) the consideration of all hourly values in developing the empirical distributions, which may not suit extreme event modelling. A potential way to improve performance on extremes is to focus on daily maximum predictands rather than all hourly values of each WT, fit probability distribution

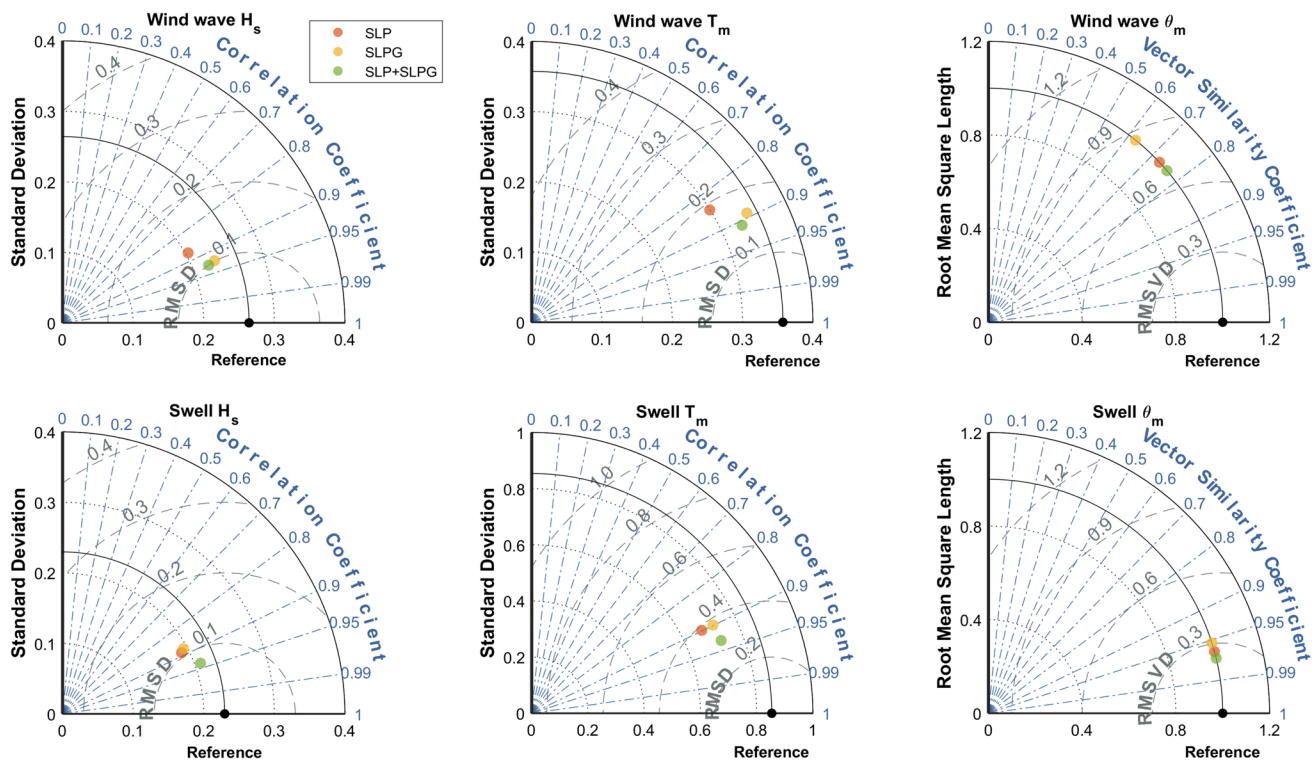


Fig. 10 Diagrams for comparing model performance in estimating wind wave and swell variables with different predictor variables. For detailed explanations, please refer to Fig. 9

functions to model their potential of extreme, and incorporate these fitted distributions into model simulations instead of relying on empirical ones. This approach has been successfully implemented by Rueda et al. (2016) and Lucio et al. (2020) in their stochastic weather emulators to better simulate extreme events, and will be considered in our future work when the focus shifts to simulating the impact of extreme events. In this study, the objective is to evaluate the overall capability of the method. Our validation confirms these statistical relationships as reliable indicators of sea state conditions associated with specific WTs, which enables proactive risk management. For instance, when weather forecasts predict WTs linked to elevated storm surge levels or intensified wave activity, nuclear power stations can implement timely preparedness measures to avoid or mitigate potential damages.

In the sensitivity analysis, we evaluated the importance of different factors to the downscaling of local wave climates and storm surges. First, the selection of predictor variables is fundamental, as they should capture the primary driving mechanisms influencing the local predictands. Previous work has used regional SLP and SLPG to downscale local wave climates (Wang et al. 2012; Casas-Prat et al. 2014; Camus et al. 2014b) as well as storm surges (Rueda et al. 2016; Costa et al. 2020). Our results demonstrate the added value of using both predictors in wave and surge

downscaling, compared to using either one individually. SLPG is shown to be more relevant than SLP for estimating wind wave H_s and T_m . This is reasonable as H_s can be estimated from wind speed in a fully developed sea state (Janssen et al. 2002), while SLPG is closely related to the geostrophic winds and is often used as a proxy for surface wind (Casas-Prat et al. 2014). On the other hand, SLPG is relatively less useful for wind wave direction θ_m , possibly because it relates to the magnitude of wind speed and lacks directional information. The explanatory power of wind has been demonstrated by Wang et al. (2010), who suggested the use of the wind-based predictor alone is sufficient to represent the relationship between atmospheric conditions and H_s . However, SLP fields are less biased than wind in climate models (Wang et al. 2010), which leads to the common use of SLP-based predictors in SD models (Wang et al. 2012; Camus et al. 2014b; Neal et al. 2016). Variations in the definition of SLP-based predictors were found in the literature. For example, Costa et al. (2020) used daily minimum SLP and daily maximum SLPG to downscale daily maximum storm surges. This definition is perhaps more suitable for extreme conditions than the commonly used mean values, although they did not present a comparison to that effect. Moreover, Wang et al. (2010) suggested the use of predictor anomalies instead of predictor values to reduce the effects of model climate bias. This adjustment is possibly equivalent

to standardising the predictor, as presented in our work. The effect of these variations in predictor definition is not within the scope of this research. Future work may consider these aspects by doing similar sensitivity analysis or fitting regression models to improve the predictor definition.

The spatial and temporal scales used to define predictors are also crucial factors to consider (Camus et al. 2014a). Waves observed at the coast usually comprise wind waves and swell waves. Swell waves can be generated from a far-away location and the weather conditions responsible for developing those waves may have occurred several days prior. Similarly, storms that cause local surges can originate from a distant place, travel for several days before reaching the site, and last for some time. Consequently, synoptic conditions over the primary region of wave or storm generation and during the average travel time are often required to downscale local sea state conditions. In our case, most waves reaching Hartlepool originate from the North Sea, which is included in all three domains tested in the sensitivity analysis. This explains the marginal differences among their wave estimations, as a larger domain does not provide much more relevant information in this case. For a given spatial domain, the waves generated at different locations do not arrive at the coast at the same time. When we include more days of atmospheric conditions into the predictor definition, the waves generated further away from our site (and thus have longer travel time) are more likely to be captured. These waves are mostly swells, which explains the improvement in swell parameters with longer temporal coverage. The increased accuracy in T_p is also likely due to a better representation of swell. Wind waves, on the other hand, are more linked to short-term atmospheric conditions. When more days are included, the detailed information contained in those more recent days is smoothed out, leading to decreased model performance. Consequently, the optimal spatial domain and temporal coverage of predictors depends on the predictand of interest. To achieve the most accuracy for both wave components, one may consider using two or more predictors with different spatial and temporal definitions (Camus et al. 2014a; Hegermiller et al. 2017). In areas experiencing multi-modal wave climate (e.g., in large ocean basins), such consideration can be crucial to explicitly account for waves generated at multiple discrete regions.

Our results suggest that spatial resolution is not a critical factor in weather typing. Using lower-resolution data can produce outcomes that are as accurate and robust as those from higher-resolution data, which agrees with Hegermiller et al. (2017). This may indicate that processes occurring at finer spatial scales (e.g., convective-scale features) are not particularly relevant to local sea state conditions, or that the weather typing method is unable to preserve such detailed information as the clustering tends to smooth these out, only leaving synoptic-scale patterns in the final WTs (which is

especially true when the number of WT is small). In either situation, using higher-resolution data would introduce redundant information that increases computational costs but does not entail better model performance. Cid et al. (2017) found lower model skills in downscaling surges at semi-enclosed areas and suggested to use higher-resolution predictors. However, our findings confirm that the benefits of increasing the resolution can be minimal. In fact, the effect of local bathymetry can be significant in these shallow basins which is often unrelated to the synoptic conditions. In such cases, incorporating the impact of local features into the model (e.g., by doing supervised classification in weather typing) can be more effective.

Decisions on the number of WTs are quite subjective, with previous studies choosing a wide range of values (Camus et al. 2014a; Neal et al. 2016; Rueda et al. 2019). The choice represents a compromise between minimising climate variability within each group and ensuring distinct characteristics across different clusters. Our results suggest that model improves as the number of WTs increases, possibly because more details in the synoptic conditions are preserved and represented, but the improvement is less significant for larger numbers of WTs, which is also reported in Ireland et al. (2024). In our analysis, the model with 64 WTs has the best accuracy. However, it should be highlighted that having more WTs means each one is associated with fewer days. As a result, the empirical distributions will be developed from a smaller amount of data, making them less representative of the typical conditions caused by certain weather patterns, and perhaps more prone to overfitting. In realistic applications, the number of WTs needs be manageable for efficient decision-making. For example, the weather regime forecasting tool applied in the UK Met Office is based on a set of 30 WTs (Neal et al. 2016). When combined with weather forecast outputs from ensemble prediction models, the forecasting tool can reduce the number of scenarios to a few weather types for further detailed analysis. Therefore, we choose to use 36 WTs which can achieve comparable model performance to 64 WTs but with fewer WTs and lower computational demand.

Last, the impact of α is quite significant in weather typing, which is understandable as local sea state conditions are not only a response to the atmospheric conditions but also influenced by local bathymetry and coastal morphology. By incorporating the impact of these local processes, which are independent of the synoptic conditions, into the clustering process, the model is greatly improved. The improvement reaches its limit when α exceeds 0.4–0.6, depending on the specific predictands. Similar results were also obtained in Camus et al. (2016). Although we only include total waves and surges in the regression-guided classification, the benefit is not limited to those predictands and observed for wind wave and swell as well.

These findings can support the broader application of weather typing models in other locations. For works concerning wave climate and/or storm surge, SLP-based predictors are sufficient to downscale a wide range of variables collectively in a single model. When applied to an enclosed sea similar to the North Sea, the spatial domain does not need to be extended far beyond the main ocean basin. If T_p is the predictand of interest or swells are the dominant wave component, using a longer temporal coverage would be beneficial. The use of low-resolution data which still preserves the main synoptic-scale patterns would not significantly compromise model performance. The inclusion of local conditions to guide the classification of weather types is generally suggested, as it improves the overall model accuracy.

5 Conclusions

In this paper, we applied a weather typing approach to downscale local, coastal hazard-related conditions at the Hartlepool nuclear power station. A set of 36 WTs was generated to characterise the variabilities of synoptic circulation patterns. Each WT is associated with empirical distributions of multivariate predictands which represent the statistical relationships. These relationships demonstrate how changes in synoptic patterns can result in distinct wave and surge responses. The validation proves the model's capability to downscale multivariate sea state conditions, including the combined wind waves and swells, storm surges, and the individual wind wave and swell components, with varying levels of accuracy. Wave direction, particularly wind wave direction, is less accurately estimated due to its high variability, making it more challenging to reproduce. We also demonstrate the model's ability to downscale extreme conditions of wave height and storm surge, which is particularly relevant for hazard assessments at nuclear power stations. While the model performs well for extreme storm surge events, its accuracy in estimating extreme wave heights is comparatively lower.

A series of sensitivity analyses was undertaken to assess the significance of various parameters in downscaling local wave and surge conditions. We find that the choice of predictor variables is generally important for all predictands. The combination of SLP and SLPG as predictors outperforms using either one individually. The impact of temporal coverage can also be quite significant to certain wave parameters, such as T_p and swell T_m . Model performance is not very sensitive to spatial domain, as long as the atmospheric conditions over the main region of wave and storm generation are included. Similarly, the impact of the spatial resolution of predictors is negligible. Last, increasing the number of WTs and the value of α can improve the model performance to a certain extent.

While our findings stem from a case study in Hartlepool, their relevance extends beyond this specific area. Our research offers insights that can inform the development of weather typing models for broader applications, particularly in defining predictors and optimizing model configurations. The empirical distributions are the foundation for evaluating potential flooding and erosion hazards. They can help identify WTs with higher probabilities of local extreme wave climates and storm surge conditions so that the nuclear station can implement preparedness measures once such WTs are anticipated in the short term. Furthermore, these empirical distributions can be used as boundary conditions in dynamical downscaling to simulate the hazards resulted from the corresponding WTs. This will be the focus of our future work, where we aim to provide probabilistic assessments of coastal hazard exposure at nuclear power stations, contributing to more effective risk management and mitigation strategies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10236-025-01682-7>.

Acknowledgements We would like to thank the European Centre for Medium-Range Weather Forecasts (ECMWF), British Oceanographic Data Center (BODC), and Center for Environment Fisheries and Aquaculture Science (CEFAS) for providing the atmospheric, storm surge, and wave data used in this research. We would also like to thank Luke J. Jenkins for kindly providing the code and assistance to calculate the non-tidal component of sea level.

Author contributions Z.Z.: Conceptualization, Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing—original draft, Writing—review & editing. H.K.: Conceptualization, Funding acquisition, Supervision, Writing—review & editing. I.D.H.: Supervision, Writing—review & editing. D.E.S.: Funding acquisition, Supervision, Writing—review & editing. B.G.: Supervision, Writing—review & editing. Y.L.: Supervision, Writing—review & editing. P.C.: Methodology, Software, Writing—review & editing.

Funding Zehua Zhong is funded through the INSPIRE Doctoral Training Partnership by the Natural Environment Research Council (NERC) (NE/S007210/1; project reference: 2740403), and co-sponsored by EDF Energy.

Data availability The SLP and wave climate data from ERA5 reanalysis (DOI: <https://doi.org/10.24381/cds.adbb2d47>), and the still water level data from CODEC (DOI: <https://doi.org/10.24381/cds.a6d42d60>) are publicly available at the Climate Data Store of the Copernicus Climate Change Service at <https://cds.climate.copernicus.eu/>. Tide gauge observations from the UK National Tide Gauge Network are available at https://www.bodc.ac.uk/data/hosted_data_systems/sea_level/uk_tide_gauge_network/. The CEFAS WaveNet buoy data can be downloaded at <https://www.cefass.co.uk/data-and-publications/wavenet/> (DOI: <https://doi.org/10.14466/CefasDataHub.11>).

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson D, Rueda A, Cagigal L et al (2019) Time-varying emulator for short and long-term analysis of coastal flood hazard potential. *J Geophys Res Oceans* 124:9209–9234. <https://doi.org/10.1029/2019JC015312>
- Berens P (2009) CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software* 31:1–21
- Bhowmik RD, Sankarasubramanian A (2019) Limitations of univariate linear bias correction in yielding cross-correlation between monthly precipitation and temperature. *Int J Climatol* 39:4479–4496. <https://doi.org/10.1002/joc.6086>
- Cagigal L, Rueda A, Anderson D et al (2020) A multivariate, stochastic, climate-based wave emulator for shoreline change modelling. *Ocean Model* 154:101695. <https://doi.org/10.1016/j.ocemod.2020.101695>
- Camus P, Méndez FJ, Medina R, Cofiño AS (2011) Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coast Eng* 58:453–462. <https://doi.org/10.1016/j.coastaleng.2011.02.003>
- Camus P, Méndez FJ, Losada IJ et al (2014a) A method for finding the optimal predictor indices for local wave climate conditions. *Ocean Dyn* 64:1025–1038. <https://doi.org/10.1007/s10236-014-0737-2>
- Camus P, Menéndez M, Méndez FJ et al (2014b) A weather-type statistical downscaling framework for ocean wave climate. *J Geophys Res Oceans* 119:7389–7405. <https://doi.org/10.1002/2014JC010141>
- Camus P, Rueda A, Méndez FJ, Losada IJ (2016) An atmospheric-to-marine synoptic classification for statistical downscaling marine climate. *Ocean Dyn* 66:1589–1601. <https://doi.org/10.1007/s10236-016-1004-5>
- Cannon AJ (2012) Regression-guided clustering: a semisupervised method for circulation-to-environment synoptic classification. *J Appl Meteorol Climatol* 51:185–190. <https://doi.org/10.1175/JAMC-D-11-0155.1>
- Casas-Prat M, Wang XL, Sierra JP (2014) A physical-based statistical method for modeling ocean wave heights. *Ocean Model* 73:59–75. <https://doi.org/10.1016/j.ocemod.2013.10.008>
- Cid A, Camus P, Castanedo S et al (2017) Global reconstructed daily surge levels from the 20th Century reanalysis (1871–2010). *Glob Planet Change* 148:9–21. <https://doi.org/10.1016/j.gloplacha.2016.11.006>
- Codiga D (2024) UTide unified tidal analysis and prediction functions (<https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions>), MATLAB Central File Exchange. Retrieved September 17, 2024
- Costa W, Idier D, Rohmer J et al (2020) Statistical prediction of extreme storm Surges Based on a Fully Supervised Weather-Type Downscaling Model. *J Mar Sci Eng* 8:1028. <https://doi.org/10.3390/jmse8121028>
- Fanti V, Ferreira Ó, Kümmerer V, Loureiro C (2023) Improved estimates of extreme wave conditions in coastal areas from calibrated global reanalyses. *Commun Earth Environ* 4:1–11. <https://doi.org/10.1038/s43247-023-00819-0>
- Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int J Climatol* 27:1547–1578. <https://doi.org/10.1002/joc.1556>
- Gaur A, Simonovic SP (2019) Chapter 4 - Introduction to Physical Scaling: A Model Aimed to Bridge the Gap Between Statistical and Dynamic Downscaling Approaches. In: Teegavarapu R (ed) *Trends and Changes in Hydroclimatic Variables*. Elsevier, pp 199–273
- Gouldby B, Méndez FJ, Guanche Y et al (2014) A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis. *Coast Eng* 88:15–26. <https://doi.org/10.1016/j.coastaleng.2014.01.012>
- Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York
- Hegermiller CA, Antolinez JAA, Rueda A et al (2017) A Multimodal Wave Spectrum-Based Approach for Statistical Downscaling of Local Wave Climate. *J Phys Oceanogr* 47:375–386. <https://doi.org/10.1175/JPO-D-16-0191.1>
- Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. *Q J R Meteorol Soc* 146:1999–2049. <https://doi.org/10.1002/qj.3803>
- Ireland LG, Robbins J, Neal R et al (2024) Generating weather pattern definitions over South Africa suitable for future use in impact-orientated medium-range forecasting. *Intl J Climatol*. <https://doi.org/10.1002/joc.8396>
- Janssen P, Doyle JD, Bidlot J et al (2002) Impact and feedback of ocean waves on the atmosphere. *Adv Fluid Mech* 33:155–198
- Jenkins LJ, Haigh ID, Camus P et al (2023) The temporal clustering of storm surge, wave height, and high sea level exceedances around the UK coastline. *Nat Hazards* 115:1761–1797. <https://doi.org/10.1007/s11069-022-05617-z>
- Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci* 374:20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Leonard M, Westra S, Phatak A et al (2014) A compound event framework for understanding extreme impacts. *Wires Clim Change* 5:113–128. <https://doi.org/10.1002/wcc.252>
- Lucio D, Tomás A, Lara JL et al (2020) Stochastic modeling of long-term wave climate based on weather patterns for coastal structures applications. *Coast Eng* 161:103771. <https://doi.org/10.1016/j.coastaleng.2020.103771>
- Masselink G, Russell P, Rennie A et al (2020) Impacts of climate change on coastal geomorphology and coastal erosion relevant to the coastal and marine environment around the UK. *MCCIP Science Review* 2020 32 pages. https://doi.org/10.14465/2020_ARC08.CGM
- Muis S, Apecechea MI, Dullaart J et al (2020) A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections. *Front Marine Science* 7:263. <https://doi.org/10.3389/fmars.2020.00263>
- Neal R, Fereday D, Crocker R, Comer RE (2016) A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorol Appl* 23:389–400. <https://doi.org/10.1002/met.1563>
- Neal R, Dankers R, Saulter A et al (2018) Use of probabilistic medium- to long-range weather-pattern forecasts for identifying periods with an increased likelihood of coastal flooding around the UK. *Meteorol Appl* 25:534–547. <https://doi.org/10.1002/met.1719>

- Odériz I, Silva R, Mortlock TR, Mori N (2020) El Niño–Southern Oscillation Impacts on Global Wave Climate and Potential Coastal Hazards. *J Geophys Res Oceans* 125. <https://doi.org/10.1029/2020JC016464>
- Pinto JG, Zacharias S, Fink AH et al (2009) Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Clim Dyn* 32:711–737. <https://doi.org/10.1007/s00382-008-0396-4>
- Pugh D, Woodworth P (eds) (2014) Storm surges, meteotsunamis and other meteorological effects on sea level. In: *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes*. Cambridge University Press, Cambridge, pp 155–188
- Rueda A, Camus P, Tomás A et al (2016) A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Model* 104:242–251. <https://doi.org/10.1016/j.ocemod.2016.06.008>
- Rueda A, Cagigal L, Antolínez JAA et al (2019) Marine climate variability based on weather patterns for a complicated island setting: The New Zealand case. *Int J Climatol* 39:1777–1786. <https://doi.org/10.1002/joc.5912>
- Schoof JT (2013) Statistical Downscaling in Climatology. *Geography Compass* 7:249–265. <https://doi.org/10.1111/gec3.12036>
- Scott T, McCarroll RJ, Masselink G et al (2021) Role of Atmospheric Indices in Describing Inshore Directional Wave Climate in the United Kingdom and Ireland. *Earth's Future* 9:e2020EF001625. <https://doi.org/10.1029/2020EF001625>
- Snarey M, Terrett NK, Willett P, Wilton DJ (1997) Comparison of algorithms for dissimilarity-based compound selection. *J Mol Graph Model* 15:372–385. [https://doi.org/10.1016/S1093-3263\(98\)00008-4](https://doi.org/10.1016/S1093-3263(98)00008-4)
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres* 106:7183–7192. <https://doi.org/10.1029/2000JD900719>
- The WAMDI Group (1988) The WAM Model—A Third Generation Ocean Wave Prediction Model
- Toimil A, Losada JJ, Camus P, Díaz-Simal P (2017) Managing coastal erosion under climate change at the regional scale. *Coast Eng* 128:106–122. <https://doi.org/10.1016/j.coastaleng.2017.08.004>
- Vousdoukas MI, Mentaschi L, Voukouvalas E et al (2018) Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nat Commun* 9:2360. <https://doi.org/10.1038/s41467-018-04692-w>
- Vousdoukas MI, Ranasinghe R, Mentaschi L et al (2020) Sandy coastlines under threat of erosion. *Nat Clim Chang* 10:260–263. <https://doi.org/10.1038/s41558-020-0697-0>
- Wahl T, Muddersbach C, Jensen J (2011) Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas: a stochastic storm surge model. *Nat Hazard* 11:2925–2939. <https://doi.org/10.5194/nhess-11-2925-2011>
- Wang XL, Swail VR (2006) Climate change signal and uncertainty in projections of ocean wave heights. *Clim Dyn* 26:109–126. <https://doi.org/10.1007/s00382-005-0080-x>
- Wang XL, Swail VR, Cox A (2010) Dynamical versus statistical downscaling methods for ocean wave heights. *Int J Climatol* 30:317–332. <https://doi.org/10.1002/joc.1899>
- Wang XL, Feng Y, Swail VR (2012) North Atlantic wave height trends as reconstructed from the 20th century reanalysis. *Geophys Res Lett* 39. <https://doi.org/10.1029/2012GL053381>
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography: Earth and Environment* 21:530–548. <https://doi.org/10.1177/030913339702100403>
- Xu Z, Hou Z, Han Y, Guo W (2016) A diagram for evaluating multiple aspects of model performance in simulating vector fields. *Geoscientific Model Development* 9:4365–4380. <https://doi.org/10.5194/gmd-9-4365-2016>
- Zhao G, Li D, Camus P et al (2024) Weather-type statistical downscaling for ocean wave climate in the Chinese marginal seas. *Ocean Model* 187:102297. <https://doi.org/10.1016/j.ocemod.2023.102297>