frontiers in Climate

TYPE Original Research
PUBLISHED 29 May 2026
DOI 10.3389/fclim.2026.1782346

Check for updates

OPEN ACCESS

EDITED BY
Avinash Kumar,
National Centre for Polar and Ocean Research (NCPOR), India

REVIEWED BY
Kristine Skovgaard Madsen,
Danish Meteorological Institute (DMI), Denmark
Michele Bendoni,
National Research Council (CNR), Italy
Kai Yin,
Southeast University, China

*CORRESPONDENCE
Thao Thi Nguyen
✉ thi.nguyen@hereon.de

RECEIVED 06 January 2026
REVISED 27 April 2026
ACCEPTED 06 May 2026
PUBLISHED 29 May 2026

CITATION
Nguyen TT, Staneva J, Ho-Hagemann HTM, Schrum C and Mayer B (2026) Future changes of coastal extremes from the regional wave-ocean coupled model system for the Northern European continental shelf.
Front. Clim. 8:1782346.
doi: 10.3389/fclim.2026.1782346

Future changes of coastal extremes from the regional wave-ocean coupled model system for the Northern European continental shelf

Thao Thi Nguyen[1,2]*, Joanna Staneva[1], Ha T. M. Ho-Hagemann[1], Corinna Schrum[1,3] and Bernhard Mayer[3]

[1]Institute of Coastal Research, Helmholtz-Zentrum Hereon, Geesthacht, Germany, [2]University of Hamburg, Hamburg, Germany, [3]Institute of Oceanography, University of Hamburg, Hamburg, Germany

Extreme storm surges are among the most severe coastal hazards threatening low-lying areas along the Northwest European shelf and the Baltic Sea, while marine heatwaves represent oceanic temperature extremes affecting both open-ocean and coastal environments. These events result from complex interactions between hydrological and meteorological processes and are projected to intensify in several regions under climate change. However, major challenges remain in developing a comprehensive understanding of future coastal extremes and their associated risks. To address this gap, we employed a regional wave-ocean coupled model to dynamically downscale a transient regionalized REMO–MPIOM coupled model under the Representative Concentration Pathway (RCP) 8.5 scenario. Three representative ensemble members from the REMO–MPIOM simulations were selected to investigate the role of internal variability in projected extremes. Model performance was evaluated against tide-gauge observations and satellite-derived sea surface temperatures for the North Sea and Baltic Sea. We then assessed projected changes in extreme coastal events by comparing the future period (2070–2099) with the historical period (1970–1999), including the role of internal variability imposed by the parent model's boundary conditions. The peaks-over-threshold (POT) method was applied to characterize both hazards. The results show that the coupled model reproduces key features of historical coastal extremes across both basins. Projected patterns indicate fewer but more intense storm surges along the North Sea coast, whereas marine heatwaves will occur more frequently but last for shorter periods throughout much of the North Sea. Our findings also emphasize the substantial role of internal variability in shaping future coastal extremes, underscoring the need to account for this uncertainty when interpreting regional climate projections. The insights gained from this study offer valuable guidance for enhancing coastal hazard preparedness, protecting infrastructure, and planning for long-term adaptation in vulnerable maritime regions.

KEYWORDS
internal variability, marine heatwaves, Northwest European continental shelf, regional future changes, storm surges

1 Introduction

Extreme coastal hazards pose significant risks to the Northwest European continental shelf (NWS) and the Baltic Sea (BAS). These regions are exposed to extreme sea levels, storm surges, intense wind waves, and marine heatwaves (MHWs). These events occur individually or as compound events and may cause flooding, erosion, ecological degradation, infrastructure damage, and loss of life (Wolf, 2009). While storm surges directly threaten low-lying and highly exposed coastal regions, including residential areas, transportation infrastructure, ports, industrial facilities, and coastal-protection structures, MHWs primarily affect marine ecosystems, fisheries, and coastal environments. The increasing frequency and intensity of these events under climate change raise concerns about the long-term safety, socio-economic stability, and resilience of coastal communities. Therefore, understanding the mechanisms and future development of these hazards is essential for planning adaptation and risk reduction measures.

Extreme sea level is one of the primary causes of coastal flooding, resulting from the combined effects of storm surges, tides, and, in some cases, waves and river discharge (Woodworth et al., 2007; Vousdoukas et al., 2016; Wahl, 2017; Bernier et al., 2024). Their magnitude is strongly influenced by mean sea level rise (Vitousek et al., 2017; Rasmussen et al., 2018; Vousdoukas et al., 2018b; Frederikse et al., 2020; Kirezci et al., 2020; Tebaldi et al., 2021) as well as by changes in storminess driven by large-scale atmospheric circulation (Morim et al., 2019; Seneviratne et al., 2021). Since these components arise from distinct physical mechanisms and are typically analyzed independently, distinguishing them is essential when assessing extreme sea level events. Notably, several studies have shown that mean sea level rise is the dominant driver of future extreme sea levels, while storm surge changes are secondary and spatially heterogeneous (Arns et al., 2015; Wahl, 2017; Vousdoukas et al., 2018b).

In coastal hazard research, storm surge (SS) is defined as the meteorologically driven increase in sea level above the predicted astronomical tide, caused primarily by wind stress and atmospheric pressure anomalies (Wahl, 2017; Bernier et al., 2024). It does not include tidal contributions, wave setup, or runup. In the present work, we refer to SS as the modeled water level after removing the mean value and tidal signal. In the North Sea, storm surges are influenced by local wind forcing, remote surges propagating from the North Atlantic (Timmerman, 1975; Gönnert and Sossidi, 2011; Bernier et al., 2024), storm characteristics (Arns et al., 2015), and tidal phase (Gönnert and Sossidi, 2011). By contrast, storm surges in the Baltic Sea reflect wind-induced restricted outflow through the Danish Straits (Wiśniewski and Wolski, 2011), as well as local wind conditions and storm characteristics (Leppäranta and Myrberg, 2009; Meier et al., 2022). Moreover, projections of storm surges remain uncertain. Some studies report significant increases in surge heights along the continental coasts (Woth et al., 2006; Vousdoukas et al., 2016), while others emphasize the dominant but spatially varying role of mean sea level rise (Lowe et al., 2009; Arns et al., 2015; Wahl, 2017; Vousdoukas et al., 2018b).

Coasts are at risk of storm surges, which can cause coastal water levels to exceed a critical threshold. This, in turn, can trigger severe coastal flooding, endanger human lives, and cause damage to infrastructure and property in coastal regions. In our region of interest, storm surges are more frequent on the coast in autumn and winter (October to March) and are driven by strong winds and low mean sea-level pressure (MSLP). In contrast, MHWs are more common in summer under weak winds and high MSLP. These extreme temperature events are caused by periods of abnormally high sea-surface temperatures (Hobday et al., 2016) and have intensified substantially worldwide, having significant ecological and socio-economic impacts (Cavole et al., 2016; Hobday et al., 2016; Oliver et al., 2021, 2019, 2018; Smale et al., 2019). MHWs alter the water-column structure, its mixing, and nutrient fluxes (Holbrook et al., 2019; Sen Gupta et al., 2020; Chen et al., 2022), and are projected to increase in frequency with continued warming, as reported in the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2021). The development and persistence of MHW in European shelf regions is influenced by changes in stratification, atmospheric forcing, and ocean circulation.

Although storm surges and MHWs are driven by distinct physical mechanisms, both hazards are closely linked to regional atmospheric variability, large-scale climate forcing, and long-term warming trends. Assessing these extremes consistently within the same regional modeling system is scientifically valuable, yet still uncommon. Our aim is not to imply a direct physical coupling between extreme storm surges and marine heatwaves. Instead, we examine two major coastal hazards projected to intensify in the North Sea, Baltic Sea, and adjacent North Atlantic under climate change (Frölicher et al., 2018; Vousdoukas et al., 2018b). Despite their seasonal differences, these hazards can overlap in time and pose significant risks to coastal communities, impacting infrastructure, ecosystems, and beach safety. Joint assessments provide a more comprehensive view of future coastal risks (Hinkel et al., 2014; Oliver et al., 2019). They illuminate how different types of extremes may overlap, interact, or occur in close succession, as emphasized in recent multi-hazard coastal assessments (IPCC, 2022).

Many studies have examined extreme storm surges and marine heatwaves separately, but far less attention has been given to their potential co-occurrence in the Northern European region. However, the likelihood of combined impacts is increasing, as future projections indicate increases in both the intensity and frequency of MHWs and extreme sea levels (Muis et al., 2020; Seneviratne et al., 2021). These, in turn, are linked to continued population growth and urbanization in coastal zones (Barragán and De Andrés, 2015; Neumann et al., 2015). Zhou and Wang (2024) have highlighted the widespread presence of concurrent extremes. They found that over 87% of global coastlines experienced concurrent extreme sea levels and MHWs, with an average increase of 3.72 concurrent days between 1998 and 2017 relative to 1979–1998. Up to 38 days per year are projected under high-emission scenarios for the period 2025–2049. Such compound events may arise under weakened geopotential height patterns and low MSLP, which can enhance both sea level and warm SST anomalies. Therefore, improving our understanding of these co-occurring extremes is crucial for integrated coastal risk assessments, adaptation strategies, and early warning systems in vulnerable regions.

To realistically capture regional ocean–atmosphere interactions, high-resolution regional models and a dynamical downscaling approach are essential. Schrum et al. (2016) emphasized that regional changes in sea level and temperature are shaped by thermohaline processes, water-mass redistribution, and the combined influence of global boundary forcing and regional atmospheric variability. Fully coupled wave–ocean models, such as the Geesthacht COAst model SysTem (GCOAST: WAM–NEMO), explicitly incorporate wave-induced processes that have been shown to improve simulations of sea surface height, sea surface temperature, and currents in the North Sea and the Baltic Sea (Alari et al., 2016; Staneva et al., 2017; Bonaduce et al., 2020, 2024; Nguyen et al., 2024, 2025). Wind waves play a fundamental role in regulating momentum, heat, and energy exchanges at the air–sea interface, thereby influencing the dynamics of coastal extremes. The inclusion of sea-state-dependent momentum and energy fluxes, Stokes–Coriolis forcing, and wave-induced mixing is key to improving model performance. Collectively, these factors modify upper-ocean stratification and circulation, directly affecting both storm surge behavior and marine heatwave development. In the present study, the wave model is used to provide physically consistent surface forcing within the coupling framework, although the wave model outputs themselves are not analyzed. Previous applications of the WAM–NEMO coupled modeling system have demonstrated its good performance in simulating waves, producing significant wave heights that closely match observations and showing high correlations in both the North Sea (Staneva et al., 2017) and the Baltic Sea (Alari et al., 2016). However, these earlier studies focused on limited subregions and did not comprehensively evaluate storm surges or significant wave heights across the entire domain. To address this limitation, we have conducted a thorough comparison of modeled hindcast simulation and observed significant wave heights, sea surface heights, and storm surges using buoy and tide gauge measurements distributed throughout the North Sea and Baltic Sea. This expanded validation provides direct observational support for the model's performance and strengthens confidence in its suitability for the analyses presented in this study. Consequently, employing a wave-ocean coupled system ensures that key air-sea feedback is consistently represented, providing a robust physical basis for examining extreme storm surges and marine heatwaves within the same modeling system. This integration is particularly relevant because storm surges are primarily driven by wind stress, atmospheric pressure gradients, and coastal geometry, whereas marine heatwaves are governed by air–sea heat fluxes, vertical mixing, and ocean circulation.

Despite these advances, a key scientific challenge remains: the understanding of the role of internal variability (IV). While IV arises from varying lateral open boundary conditions, the variability of storm surges and MHWs arises from IV and differences in atmospheric forcing. IV can lead to substantial differences in modeling regional extremes, even under identical forcing (Hawkins and Sutton, 2009; Deser et al., 2020; Ho-Hagemann et al., 2020; Wiese et al., 2020). To the best of our knowledge, the influence of IV on storm surge hazards and MHWs across the Northwest European shelf has not yet been systematically investigated using a coupled wave–ocean modeling framework.

The scientific novelty of this study is therefore twofold. First, we provide a consistent assessment of extreme storm surges and marine heatwaves using the same high-resolution coupled wave–ocean model, allowing a physically coherent interpretation of projected changes across two major coastal hazards that are often studied separately. Second, we explicitly quantify the influence of internal variability on projected changes, based on three selected members of the REMO–MPIOM regionalized climate ensemble (Mayer et al., 2022). This approach provides new insight into how much of the future changes in coastal extremes may arise from internal variability rather than externally forced trends.

In this work, we use the GCOAST wave–ocean coupled model WAM–NEMO to dynamically downscale REMO–MPIOM results for a historical period (1970–1999) and a future period (2070–2099) under the RCP8.5 scenario. We validate modeled storm surge levels using tide-gauge observations and simulated SST using the Copernicus Marine Environment Monitoring Service (CMEMS) reanalysis. We evaluate historical extreme events and assess projected changes in both storm surge and MHW characteristics. We further estimate the relative contribution of internal variability to these projections.

The paper is structured as follows. Section 2 describes the model configuration, downscaling framework, and methods. Section 3 presents the evaluation and future projections. Section 4 discusses implications and conclusions.

2 Materials and methods

2.1 The numerical modeling framework

2.1.1 The WAM–NEMO coupled model

The downscaling simulation was performed using the GCOAST coupled modeling framework, a regionally wave-ocean system that integrates the Nucleus for European Modeling of the Ocean model NEMO (Madec, 2016), with the spectral wave model WAM (Komen et al., 1994; Warszawski et al., 2014). WAM is a third-generation spectral wave model in which the wave transport equation was solved explicitly without assumptions about the shape of the wave spectrum. A detailed model description can be found in Alari et al. (2016) and Staneva et al. (2017).

The ocean model NEMO has been developed to solve the three-dimensional primitive equations under the hydrostatic and Boussinesq approximations on an Arakawa C horizontal grid and hybrid s-z* vertical levels. The model was designed for the domain covering the Baltic Sea, the Danish Strait, the North Sea and has a large extension in the North Atlantic (Figure 1a), with approximately 3.6 km horizontal resolution and 50 hybrid s-z* vertical levels. The bathymetry was based on the topography of the 1-min global General Bathymetric Chart of the Oceans (GEBCO 2020; http://www.gebco.net).

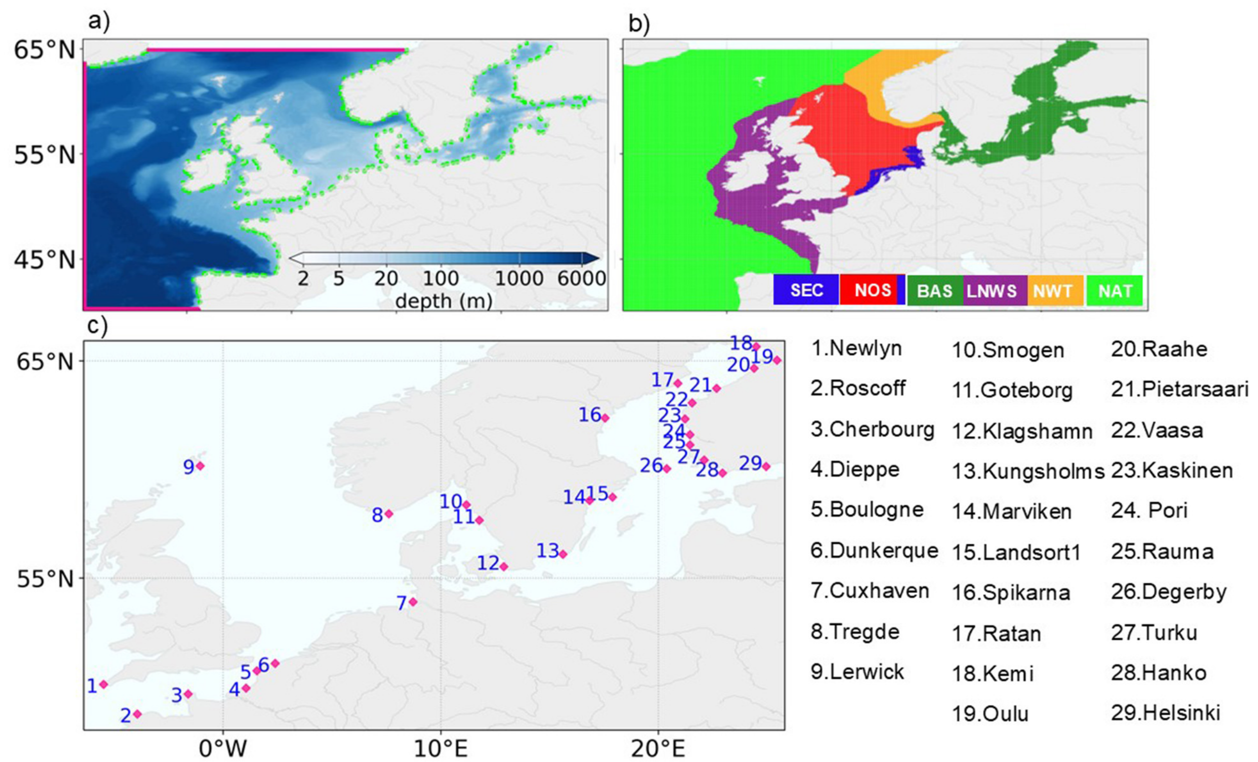

FIGURE 1
(a) Bathymetry and model domain. The open ocean boundaries are shown as pink lines (land is shaded in gray). The river run-off points are indicated by green dots; **(b)** The six sub-domains used in analysis: the Southeastern North Sea coast (SEC: blue), North Sea (NOS: blue+red), the Baltic Sea (BAS: dark green), the Norwegian Trench (NWT: orange), part of the North Atlantic (NAT: light green), and left of the Northwestern European shelf Sea (LNWS: purple); **(c)** Locations of the tide-gauge stations, shown as pink diamonds. See Table 1 for a summary of all abbreviations used.

The wave model WAM was coupled to the NEMO model, taking into account different wave-induced processes (WIPs) including Stokes-Coriolis forcing (Wu et al., 2019b), sea-state dependent momentum and energy fluxes and wave-induced mixing, as documented in Alari et al. (2016) for the Baltic Sea, Staneva et al. (2017) for the North Sea, and Bonaduce et al. (2020) for this study region. Both models run on the same structured computational grid and bathymetry, with a horizontal resolution of approximately 3.6 km. The two-way coupling of WAM and NEMO models was implemented using the OASIS3 Model Coupling Toolkit (OASIS3-MCT, Craig and Banner, 1994; Valcke, 2013), which allows numerical simulation to exchange synchronized information representing different Earth system components (Sterl et al., 2012; Wahle et al., 2017; Varlas et al., 2018). In the WAM–NEMO coupled model, the surface stress, Stokes drift, significant wave height, mean wave period, and wave-induced turbulent energy flux fields are calculated by the wave model WAM and passed to the hydrodynamic model NEMO. WAM receives the sea surface current and water level (i.e., sea surface elevation) from NEMO. The running time steps of WAM and NEMO are 60 s and 90 s, respectively. The coupling time step is 600 s. A schematic of model configuration, and the exchanged variables is shown in Figure 2.

2.1.2 Atmospheric forcing and lateral open boundary conditions

In this study, the forcing fields for the GCOAST WAM–NEMO coupled model were obtained from the REMO–MPIOM climate model simulation (Sein et al., 2016), which comprises the regional atmospheric model REMO (Jacob, 2001) and the global ocean model MPIOM (Jungclaus et al., 2013). The REMO–MPIOM outputs were dynamically downscaled to the regional WAM–NEMO coupled model for two 30-year time-slices representing historical condition and RCP8.5 scenario. Wind forcing for the WAM model was taken from REMO 10-m wind fields at a spatial resolution of 0.25°. Atmospheric forcing for the NEMO model comes also from the REMO model results, including 10 m meridional and zonal wind components, air temperature at 2 m height, short wave and longwave radiation, mean sea level pressure, humidity, snow fall, evaporation and precipitation. Initial and open boundary conditions of temperature, salinity, velocities, and sea level for NEMO were taken from the MPIOM model outputs. The global MPIOM ocean model provides dynamic and steric sea-level changes associated with large-scale ocean processes (Jungclaus et al., 2013), which are included in the boundary and initial conditions. However, it does not account for local vertical land motion or other site-specific contributors

TABLE 1 Summary of abbreviations used in the study.

Abbreviation	Full term
SST	sea surface temperature
RCP	Representative concentration pathway
POT	Peaks-over-threshold
NWS	Northwest European continental shelf
LNWS (purple)	Left of the Northwest European continental shelf
NAT (light green)	Part of North Atlantic
SEC (blue)	Southeastern North Sea coast
NOS (red + blue)	North Sea
BAS (dark green)	Baltic Sea
NWT (orange)	Norwegian Trench
UK	The United Kingdom
MHW	Marine heatwave
IPCC	Intergovernmental Panel on Climate Change
IV	Internal climate variability
GCOAST	Geesthacht COAst model SysTem
WAM	The wave model
NEMO	Nucleus for European Modeling of the ocean model
GEBCO	General Bathymetric chart of the oceans
REMO	Regional atmospheric model
MPIOM	Max Plank-Institute Ocean Model
MPI-ESM	Max Plank-Institute Earth System Modeling
CMEMS	The Copernicus Marine Environment Monitoring Service
AR5	Fifth Assessment Report
RMSE	The root-mean-squared errors
CRMSE	The centered root-mean-squared errors
CV	Coefficient of variation
BIAS	Mean bias
NCV	Normalized coefficient of variation
SI	Scatter index
HH	Hanna and Heinold index
SKILL	Skill score
CORR	Correlation coefficient
GEV	Generalized-extreme-value
SSH	Sea surface height
Hs	Significant wave height
FREQ	Frequency
INT	Intensity
DUR	Duration
MAG	Magnitude
ESS	Extreme storm surges
RE/RD	Relative difference
RE_i	Point-by-point relative difference
PROJ	Future period 2070–2099
HIST	Historical period 1970–1999
FDR	False discovery rate
SNR	Signal-to-noise ratio
MSLP	Mean sea level pressure
UHSLC	The University of Hawaii Sea Level Center

to relative sea level rise (Church et al., 2013; Slangen et al., 2014; Wöppelmann and Marcos, 2016). Tidal elevation at the open boundaries was added using harmonic constituents from TPXOv8 database (Egbert and Erofeeva, 2002). The river runoff was prescribed using freshwater discharge fields from the global hydrological discharge model (Hagemann and Gates, 2001), as implemented within the REMO–MPIOM coupled modeling system.

These forcing data from the REMO–MPIOM model offer a significant advantage for high-resolution climate simulations over Central and North Europe, with a 5–12 km oceanic horizontal resolution and an approximately 24 km atmospheric resolution (Mayer et al., 2022). Further details of the REMO–MPIOM climate model system can be found in Sein et al. (2016) and Mathis and Mikolajewicz (2020).

The regionalized REMO–MPIOM climate model simulations were conducted to dynamically downscale a 30-member global climate model ensemble derived from the low-resolution Max Planck-Institute Earth System Model MPI-ESM1.1-LR, which has horizontal resolutions of 1.9° in the atmosphere and 1.5° in the ocean (Giorgetta et al., 2013). The ensemble members, labeled M1 to M30, simulated the period from 1950 to 2099. Each member was initialized with different states taken from the same model system during 1950–1959 (Mayer et al., 2022). The transient simulations span two phases: (1) a historical period from 1950 to 2005, driven by historical atmospheric pCO2; and (2) a scenario period from 2006 to 2099, following the RCP 8.5 scenario. Riahi et al. (2022) noted that RCP 8.5 represents a high-emissions scenario useful for exploring the high-end risks of climate change; however, it should not be considered as a typical "business-as-usual" projection. While it is not the most likely projection under current policy trends, it remains an important criterion for evaluating the worst-case climate impacts. Defined in the IPCC Fifth Assessment Report (AR5), RCP8.5 corresponds to a radiative forcing value of 8.5 W/m^2 in 2100 relative to pre-industrial values (IPCC, 2013; Bowyer et al., 2014).

2.1.3 Experimental design

Three experiments were conducted using the high-resolution regional coupled model WAM–NEMO configuration (Bonaduce et al., 2020; Nguyen et al., 2024, 2025). Each simulation was performed for two 30-year time slices representing past (1970–1999) and future climate conditions (2070–2099). The regional ensemble simulations were forced at the open ocean boundaries and in the atmosphere by three representative ensemble members of REMO–MPIOM climate model ensemble. The three experiments are described as follows:

- EXPhig representing high open boundary sea level variability (forced with REMO–MPIOM member M208).
- EXPmed representing median open boundary sea level variability (forced with REMO–MPIOM member M221).
- EXPlow representing low open boundary sea level variability (forced with REMO–MPIOM member M213).

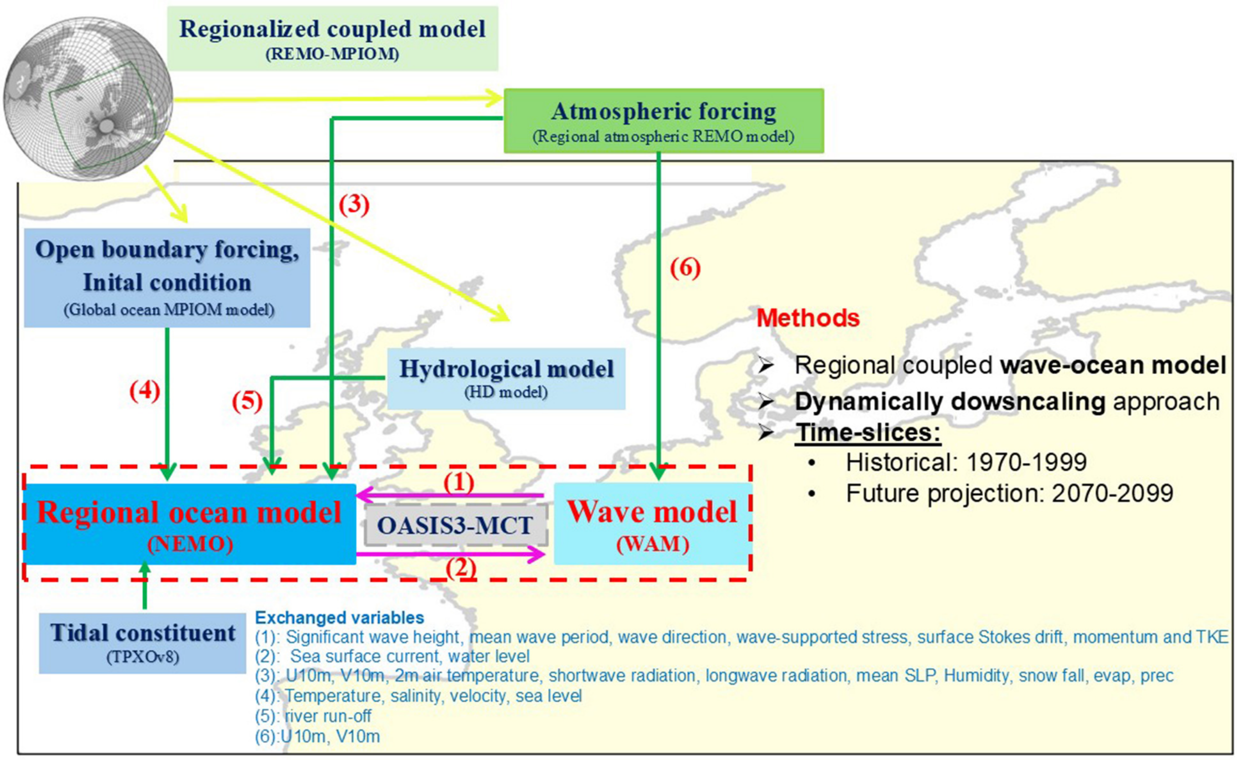

FIGURE 2
Schematic diagram of the regional coupled wave and ocean models. Parameter values exchanged between the wave and ocean models are also represented.

These three REMO-MPIOM ensemble members were selected from the original 30 members of the regionalized transient climate model ensemble based on percentile-based sea level variability at the lateral open boundaries. The ensemble size was reduced from 30 to 3 due to computational constraints associated with the high-resolution coupled simulations. To ensure that the reduced ensemble still represents the spread of internal variability, we applied a percentile-based selection method based on sea level variability at the open boundaries. The selected members represent high, medium, and low variability conditions within the original ensemble. We first computed the spatially averaged sea level variability along the open boundary grid cells and identified the 95th (high), 50th (medium), and 5th (low) percentiles. Each percentile was calculated independently for two 30-year periods mentioned above. For each percentile, we calculated the root mean squared errors (RMSE) between each of the 30 ensemble members and the percentile target. The three members with the smallest RMSE relative to the respective percentile were chosen: M208, M221 and M213, representing high, medium, and low variability, respectively. These members were intentionally selected to represent the upper, median, and lower bounds of internal variability within the full large ensemble. This approach preserves the essential spread of internal variability while keeping the computationally intensive coupled simulation feasible. These procedures follow the approach described by Pein et al. (2023), and the selection of representative members is illustrated in Figure A1.

The model's ability to simulate storm surges and MHWs was assessed through comparison with tide-gauge records, and satellite-derived SST, respectively.

2.2 Analysis method for coastal extremes

2.2.1 Extreme storm surges (ESS)

The simulated storm surges were derived from the hourly sea surface height (SSH) fields by first removing the mean SSH and then filtering out the tidal signal. The resulting storm surges were analyzed for the Northwest European continental shelf and the Baltic Sea. Extreme storm surges were assessed using two well-known approaches: the block-maxima method and the peaks-over-threshold (POT) method.

For the block-maxima approach, yearly maximum storm surges were extracted from fixed-length annual blocks, and a Generalized-extreme-value (GEV) distribution was fitted to the cumulative distribution of these maxima (Coles, 2001). The fitted GEV distribution was used to estimate return periods.

The POT method was applied to identify storm surge hazards for each 30-year period using a quantile-based threshold. For each grid cell and ensemble member, we first computed the annual 95th-percentile storm surge values and then averaged them over the corresponding 30-year period to define the threshold. Extreme events were subsequently identified as surge heights exceeding this threshold. Individual events were separated using a 72-h de-clustering window (Marcos et al., 2011; Serafin and Ruggiero, 2014; Vousdoukas et al., 2016, 2018b).

The selected events are characterized by four indices: the occurrence rate per year (frequency, FREQ [events/year]), the peak surge height (magnitude, MAG [m]), the peak height above threshold (intensity, INT [m]), and the duration of the event (DUR [hours]). The detection of extreme storm surges (ESS) is illustrated in Appendix A6, Figure A7.

The analysis focused on comparing future and historical runs to investigate the projected changes in ESS events. These were assessed by computing the mean ESS differences over 30-year period between the historical (1970–1999) and future (2070–2099) simulations. From the three-member ensemble, the 95th-percentile of storm surge signals, their associated characteristics, and the ensemble mean were calculated. The ensemble means for the two 30-year periods were then used to estimate projected changes relative to the historical simulations. The projected signal was evaluated using the relative differences (RD) between the future and historical periods, defined as follows:

$$RD = 100\% \times \frac{PROJ - HIST}{HIST} \quad (1)$$

PROJ and HIST refer to future and historical runs, respectively. Positive (negative) RD values indicate an increase (a decrease) in future period relative to the past. To assess whether the mean values from the future and historical simulations differ significantly at each grid point for the 95th-percentile of SS, we first applied the student's t-test (Scheer, 1986; Fan et al., 2013) at the 95% confidence level. We then adjusted the resulting p-values using the false discovery rate (FDR) method (Benjamini and Hochberg, 1995; Wilks, 2016) to control false positive rates. The difference is considered statistically significant when the adjusted p-value is $\leq$0.05. Further details can be found in Appendix B in Pham et al. (2026).

2.2.2 Marine heatwaves

Marine Heatwaves (MHWs) were identified following the widely used framework of Hobday et al. (2016), as illustrated in Appendix A6, Figure A8. In this approach, an MHW event is defined as a period of anomalously warm water lasting at least five consecutive days, during which the sea surface temperature (SST) exceeds the 90th-percentile threshold for that day of year. Climatological studies generally use at least a 30-year long baseline. Accordingly, we calculated the daily climatology and their corresponding thresholds for each grid cell using daily SST derived from 6-hourly model outputs for the three experiments over the historical (1970–1999) and the future (2070–2099) periods. When two consecutive MHW events occur with less than 2 days between them, they are merged into a single MHW. Using seasonal variability thresholds and a 30-year baseline enables MHWs to be detected throughout the year rather than solely during the warm seasons. This contrasts with fixed-threshold approaches, which tend to identify events only during the warm season and may underestimate MHW activity in cooler months (Frölicher et al., 2018).

MHW events are characterized by their duration (DUR [days]: the total number of consecutive days measured from the start and end dates of the event), the total MHW days ([days): the total number of MHW days in each year], frequency [FREQ (events/year): the number of extreme events occurring each year], and the maximum intensity (INT [°C]: the highest SST anomalies during each MHW event). These MHW characteristics were explored using the Python Marine Heatwaves module (https://github.com/ecjoliver/marineHeatWaves).

To assess the future changes in MHWs, we calculated the relative difference (RD; see Equation 1) between future and past periods.

2.2.3 Internal variability

The influence of internal variability on the simulation of regional extremes will be investigated in Section 3 using the three selected members mentioned above. Our aim was to estimate the influence of IV transmitted through the open boundary conditions on the simulation of storm surges and MHWs of the regional coupled model. In this study, the regional coupled model WAM–NEMO responds to the variability imposed by the different large-scale atmospheric forcings and oceanic states provided by the selected REMO–MPIOM members. To quantify the role of the IV in shaping the regional extreme events, we evaluated the inter-member spread between the three experiments. Following the methods used in regional climate modeling (Alexandru et al., 2007; Lucas-Picher et al., 2008; Ho-Hagemann et al., 2020; Wiese et al., 2020), we applied the standard deviation across ensemble members as a measure of the IV:

$$\sigma_{\exp} = \frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\left(\overline{x_t} - x_{n,t}\right)^2} \quad (2)$$

with x being frequency, duration and intensity of extreme events from ensemble member n, $\overline{x_t}$ is the ensemble mean at time t, $T = 30$ years, and $N = 3$ is the number of selected members.

This metric reflects how sensitive the regional coupled model is to the variability in the lateral open boundary conditions, thereby providing an estimate of the contribution of IV to regional extremes.

2.2.4 Signal-to-noise ratio (SNR)

We assess the role of internal variability relative to projected changes using the signal-to-noise ratio (Hawkins and Sutton, 2009; Faghih and Brissette, 2023) defined as:

$$SNR = \frac{|PROJ - HIST|}{\sigma_{PROJ-HIST}} \quad (3)$$

where PROJ and HIST refer to the mean values in the future and historical 30-year time slices, respectively.

To quantify the noise component in the signal-to-noise ratio, we define $\sigma_{PROJ-HIST}$ as the standard deviation of the differences between the projected and historical annual values:

$$\sigma_{PROJ-HIST} = SD(X_{PROJ} - X_{HIST}) \quad (4)$$

where X_{PROJ} and X_{HIST} denote the annual values of the interested variables in the future and historical periods, respectively. It is not as simple as the difference between σ_{PROJ} and σ_{HIST}. This metric reflects the variability of extremes across different periods, rather than the difference in standard deviation for each period. Using the standard deviation of the difference series allows us to capture both changes in variability and the influence of internal climate variability on the projected changes. This provides a more robust estimate of the noise against which the projected signal is evaluated.

The SNR quantifies the relative importance of internal climate variability compared with the projected change signal. Values of SNR < 1 indicate IV dominates over the forced signal. As SNR increases, the relative magnitude of IV decreases, and vice versa. SNR > 1 indicates that the forced signal is stronger than IV.

TABLE 2 RMSE between simulations and observation and then averaged over three experiments.

Stations	RMSE (annual maxima)
1. Newlyn	0.150
2. Roscoff	0.247
3. Cherbourg	0.152
4. Dieppe	0.137
5. Boulogne	0.226
6. Dunkerque	0.272
7. Cuxhaven	0.603
8. Tregde	0.186
9. Lerwick	0.150

2.3 Data

2.3.1 Tide gauge data

Hourly sea level observations from tide gauges located in the North Sea and Baltic Sea for the period 1970–1999 were obtained from the Copernicus Marine Environment Monitoring Service (http://marine.copernicus.eu/). The dataset comprises five stations in the North Sea and twenty stations in the Baltic Seas, as indicated in Figure 1c. Additional sea level records were sourced from the University of Hawaii Sea Level Center (UHSLC) database (Caldwell et al., 2001). We downloaded four tide gauge records from https://uhslc.soest.hawaii.edu/data/, specifically for Newlyn, Cuxhaven, Tregde, and Lerwick. After we removed the mean sea level, tidal signals were computed from each time series using UTide (2014), the Python implementation of the Unified Tidal Analysis and Prediction (Codiga, 2011). Then, the surge component of the sea level signal can be taken as the remainder after removing the tidal signal.

First, we evaluated the model performance in reproducing ESS based on the comparison of simulated and measured monthly maximum storm surges. This has been done by estimating four metrics: the coefficient of variation (CV), the normalized standard deviation (NSTD), the centered root-mean-square errors (CRMSE), and the mean bias (BIAS or Bias). These metrics are necessary for assessing how well the model reproduces the observed mean state (e.g., CRMSE and BIAS) and variability (CV and NSTD). Here, the observation is considered as a reference. The best match between observation and simulation is indicated when the NSTD is close to 1, and the CV of simulations is close to that of the observations, while BIAS and CRMSE are close to 0. The detailed information can be found in Appendix A2.

Previously, we used monthly maxima because they provide a larger sample size and capture the seasonal variability of storm surges, making them suitable for assessing general model skills (e.g., bias and variability). In contrast, annual maxima represent only the single most extreme event each year and are therefore more appropriate for extreme value analysis (such as return period estimation), but they offer limited statistical robustness for evaluating overall model performance. For this reason, we selected monthly maximum values to quantify model skill, while annual maxima are used to analyze extreme storm surge behavior. An example of Cuxhaven station is included in the study.

In order to evaluate the model's performance in reproducing observed extreme values more accurately, we applied a simple bias correction to the annual maximum storm surges. This step does not aim to generate corrected values for future projections. Instead, it identifies the magnitude and structure of systematic errors in simulations. By comparing the uncorrected and bias-corrected series, we can assess whether the model's deviations from observations are primarily due to systematic biases or limitations in representing the variability and distribution of extremes. This approach provides essential context for interpreting the BIAS and CRMSE metrics and evaluating the model's performance in the present day. As the corrected values are not used in future projections, the analysis does not rely on the stationarity assumption typically associated with bias correction. The corrected values for different stations in the North Sea are shown in Table 2. These values were computed as the root-mean-squared errors (RMSE [m]) between the simulations and observations of the annual maximum sea level and then averaged over the three experiments. For example, the corrected surge levels at Cuxhaven were obtained by adding the corresponding correction values to the annual simulated surge levels.

2.3.2 SST satellite-derived database

We compared the modeled SST from the three historical experiments with satellite-derived Level 4 SST fields from CMEMS for the overlapping period from 1982 to 1999. This CMEMS product, recently published SST dataset by Bashiri et al. (2025), provides high-resolution, optimally interpolated SST fields for the North Sea and Baltic Sea. This dataset offers valuable observational data for evaluating SST and related extreme events in our study region. Incorporating this dataset enhances the assessment of model performance, particularly for extreme SSTs such as MHWs, in areas where reanalysis products are limited or unavailable.

The CMEMS satellite-derived SST dataset spans 42 years (1982–2023) and is based on reprocessed SST observations from the Copernicus Marine Service (product ID: SST_BAL_SST_L4_REP_OBSERVATIONS_010_016)

(Høyer and She, 2007; Høyer and Karagali, 2016; Embury et al., 2024). This multisensory Level4 product is derived from infrared satellite measurements and provides daily SST fields at a high spatial resolution of 0.02° × 0.02° (approximately 2 km). Further details are available in Bashiri et al. (2025). We used this dataset to evaluate the simulated SST across the study region.

The reprocessed SST measured fields were interpolated onto GCOAST grid, which has a spatial resolution of approximately 3.6 km. We evaluated model performance by computing the NSTD, BIAS, RD and NCV of the simulated SST with respect to the satellite-derived SST for the period 1982–1999 (see Appendix A2 for details on the calculation procedures).

3 Results

3.1 Model validation

In this study, the model's performance was first evaluated through a hindcast simulation, using the same configuration that was applied in the main experiments. A dedicated validation against observations is provided in Appendix A3. The evaluation reveals that the model configuration successfully reproduces the key characteristics of significant wave heights (Hs), sea surface heights (SSH), and storm surges. The results in Appendix A3 demonstrate good agreement between the simulations and the observations, with low bias, relative difference (RE), root-mean-square error (RMSE), scatter index (SI), the Hanna and Heinold index (HH), as well as high skill score (SKILL), and correlation coefficient (CORR) (Mentaschi et al., 2013; Willmott et al., 2012). These metrics indicate that the model provides a robust representation of present-day conditions, although some limitations remain in capturing extreme variability, strengthening confidence in its suitability for long-term simulations of the coupled modeling framework.

3.1.1 Storm surges

In this section, the modeled historical storm surges were validated against measured surge levels derived from tide-gauge records for both the North and Baltic Seas. Our objective was to assess how effectively the regional coupled model reproduces the magnitude and variability of historical storm surge extremes. To this end, we first compared the monthly maximum storm surge of simulations with those derived from the tidal gauge stations. The locations of all stations are shown in Figure 1c.

Figure 3 shows the maximum modeled storm surge levels and four evaluation metrics, compared with tide-gauge measurements in the North Sea (left panels) and the Baltic Sea (right panels), for the period 1970–1999. The mean of the monthly maximum surges (MAX) reveals pronounced extreme surge levels along the southeastern North Sea coast, while substantially lower maxima occur throughout the Baltic Sea. The coefficient of variation (CV) ranges from approximately 20% to 45%, increasing toward the northern stations. The CV of simulations generally agrees with observations, but the largest deviation occurs at Cuxhaven. Differences among the experiments in the North Sea are small except at Dieppe and Dunkerque. In the Baltic Sea, however, all experiments systematically underestimate variability at stations, where CV values are much higher, ranging from 40 to 80%, particularly at the northern stations.

The normalized standard deviation (NSTD) ranges from 0.54 to 1.34 across the North Sea stations. Stations 1–5 show moderate overestimation (NSTD > 1), while stations 6–9 drop to around 0.8, indicating that the model underestimates the spread. A similar behavior is found in the Baltic Sea, where NSTD values consistently fall below 1. The centered root-mean-squared errors (CRMSE) range from 20 to 70 cm, with exceptional high values at Cuxhaven and peak errors at the northern Baltic Sea stations (numbers 18–20). These results suggest that the model captures the general shape of storm surge variability reasonably well but performs less effectively in regions with complex coastlines and bathymetry (e.g., the southeastern North Sea) or in areas influenced by sea-ice cover (northern Baltic basin).

Biases fall within approximately −16 to +8 cm. Overall, the evaluation of monthly maximum storm surges indicates that the model reproduces mean surge levels reasonably well, as reflected by the generally low biases across most stations. However, all experiments systematically underestimate variability, as shown by consistently low CV and NSTD values relative to observations. This suggests that the model does not fully capture the magnitude of extreme storm surges, particularly in the Baltic Sea, where surge dynamics are strongly wind-driven and sensitive to basin geometry. Despite this limitation, CRMSE values remain moderate, indicating that the temporal evolution and relative differences are reasonably well represented. Moreover, differences among the three experiments are relatively small, though EXPmed tends to perform slightly better, with slightly lower CRMSE and bias values compared to EXPhigh and EXPlow. In summary, the model captures mean surge conditions well, but challenges remain in representing the full range and intensity of storm surge variability, especially in regions with narrow inlets, complex topography, or strong wind-setup.

At Cuxhaven, the model reproduces the mean monthly maximum surges reasonably well with relatively small bias (8–9 cm) across all experiments. However, it substantially underestimates variability: the CV amounts to 44% in observed data but 22–24% in the simulated data. This means that models capture only about half of CV compared to observation. This is fully consistent with NSTD of a range 0.55–0.6, indicating the modeled STD is 40–45% lower relative to observations. Despite the small bias, the CRMSE of about 68 cm for all runs indicates a large spread to the mean values, pointing to the mismatches in peak or timing of monthly maxima. Differences between experiments are small but systematic. Overall, Cuxhaven shows a realistic mean state of storm surges but is pronounced underrepresentation of variability of monthly maxima. Cuxhaven is located at the Elbe estuary mouth, a highly dynamic transition zone where storm surges are strongly influenced by several physical processes that are not fully resolved in our regional model configuration (Weisse and Plüβ, 2006; Arns et al., 2015). These include the complex estuarine and neighboring Wadden Sea bathymetry, strong tidal amplification, nonlinear tide–surge interactions, and the funnel shaped geometry of the German Bight, all of which can significantly modify surge heights (Weisse and Plüβ, 2006; Horsburgh and Wilson, 2007; Müller et al., 2012; Arns et al., 2015).

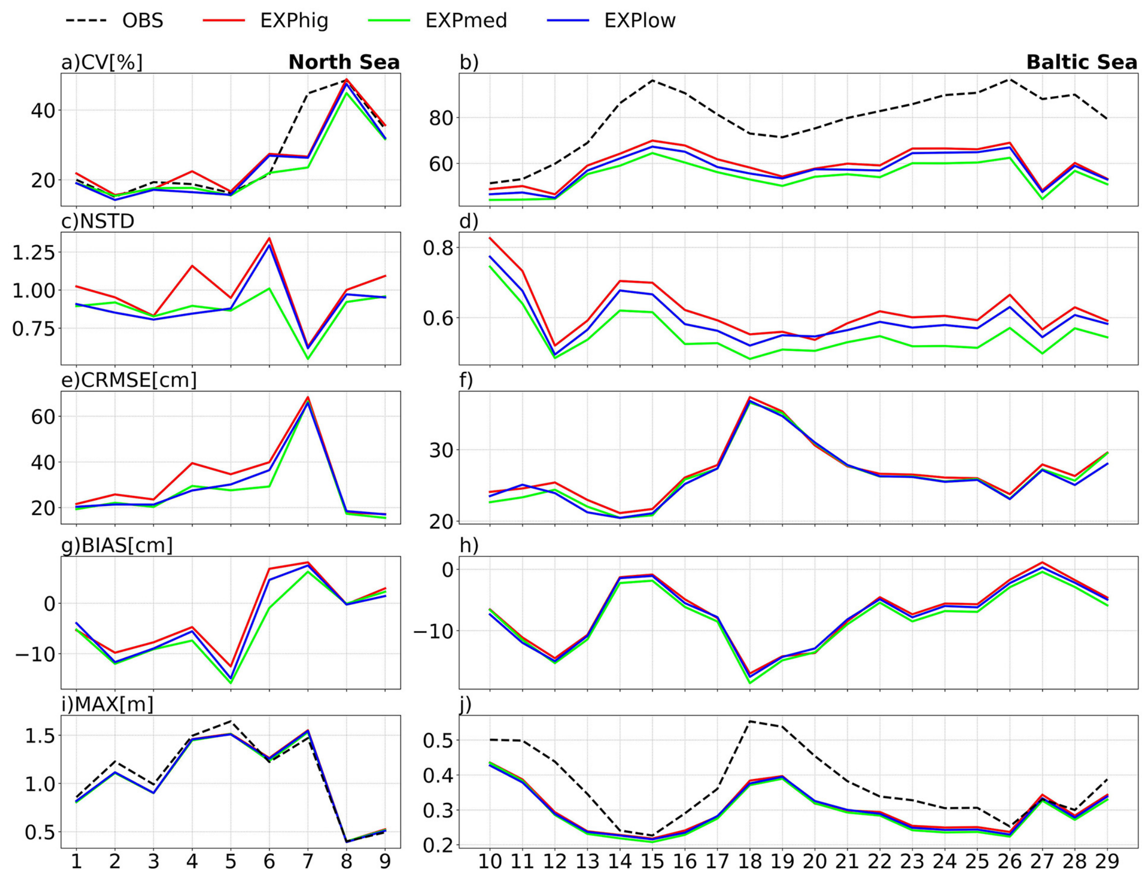

FIGURE 3
Coefficient of variation CV[%] **(a, b)**, normalized standard deviation NSTD **(c, d)**, centered root-mean-squared errors CRMSE [cm] **(e, f)**, mean bias BIAS [cm] **(g, h)**, and the mean of the monthly maximum storm surges MAX [m] **(i, j)**. Results compare simulations with observations at nine stations in the North Sea **(left)** and twenty stations in the Baltic Sea **(right)**. Different colors represent the three experiments: EXPhig (red), EXPmed (green), EXPlow (blue), while observations (OBS) are shown by the black dashed line.

Figure 4 shows the return levels of annual maximum storm surges at the Cuxhaven tide-gauge station for return periods ranging from 1.1 to 30 years. All three experiments (EXPhig, EXPmed, and EXPlow) capture the main statistical behavior of extreme surges. However, the simulations (solid lines) systematically underestimate the observed return levels, with the discrepancy increasing for longer return periods. This is a common feature of regional surge modeling, where unresolved coastal processes and insufficiently resolved nearshore bathymetry tend to suppress extremes (Weisse and Plüβ, 2006; Arns et al., 2015). The discrepancies between the experiments reflect the uncertainty imposed by the selected member forcings, which becomes more pronounced at longer return periods. The results also demonstrate that the monthly maxima and GEV-based methods are highly sensitive to single rare events, particularly when the sample sizes are small. Applying the bias correction (dashed lines) substantially improves the agreement with observed data by shifting the modeled return levels upwards closer to the observed curve. The results are consistent across return periods; EXPhig is always higher than EXPmed. EXPlow exhibits a rapid increase. Its return levels lie lowest with a return period of less than 2.5 years, falling between the return levels of EXPhig and EXPmed for return periods between 2.5 and 6 years. The highest return levels correspond to a return period of more than 6 years. EXPmed generally aligns closest to the observed data after correction. This correction effectively reduces the systematic offset while preserving the shape of the distribution, resulting in more realistic estimates of high-impact, low-frequency surge events. Although some underestimations remain at the longest return periods, the improvement is substantial. The 95% confidence intervals show that the bias-corrected experiments fall within or near the observational uncertainty range for most return periods.

3.1.2 Sea surface temperature

Figure 5 shows four SST skill metrics evaluated against the satellite-derived CMEMS dataset (Bashiri et al., 2025) for the EXPhig, EXPmed and EXPlow experiments, averaged over the entire 30-year historical period. The bias fields (as in panels a–c) reveal minor discrepancies across the domain, including a relatively warm bias along parts of the North Sea coast, a relatively cold bias in the Norwegian Trench, and weak, spatially mixed biases in the Baltic Sea. There are minimal differences among these experiments, indicating that model reproduces the mean SST well and that its performance is not sensitive to the choice of representative forcing

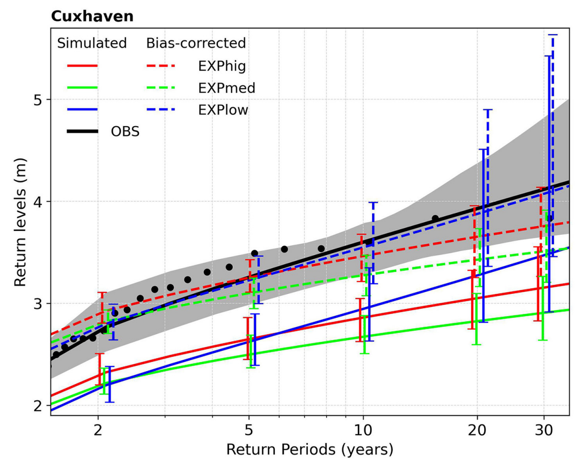

FIGURE 4
Return levels of annual maximum storm surges for return periods ranging from 1.1 to 30 years at the Cuxhaven tide gauge station. The red, green, blue, and black solid lines show the simulated surge return levels for the three experimental and the observational data, while the corresponding dashed lines represent the bias-corrected values. Black dots represent the observed annual maxima. The shaded gray area denotes the 95% confidence interval of the observations, and the colored vertical lines indicate the corresponding 95% confidence intervals for each experiment at several return periods.

member. Relative differences (RD in panels d–f) mainly fluctuate within ±10–20%, with more substantial value observed in the southern North Sea and parts of the Baltic Sea, a phenomenon attributable to the impact of complex coastlines and bathymetry. Both the normalized coefficient of variation (NCV in panels g–i) and normalized standard deviation (NSTD in panels j–l) show that the model reasonably well captures the spatial distribution of SST variability, with values close to 1 across much of the North Sea. However, larger values in the Baltic Sea and Norwegian Trench suggest an overestimation of variability in these more enclosed basins. Crucially, the consistent spatial distribution across all the experiments demonstrates that model performance is mostly not affected by the selected members. Overall, the model skills show a good agreement between simulated and observed SST. The model performs reliably across the region, with the strongest skill in the open North Sea and slightly reduced performance in more complex, semi-enclosed basins, such as the Baltic Sea.

To assess whether the model captures thermal conditions relevant to generating extreme temperature events, we extended the SST evaluation to include SST anomaly variability and 90th-percentile statistics, in addition to CV, NSTD, CRMSE, and BIAS. SST anomalies were computed by removing the monthly climatology, and their standard deviation and 90th-percentile were compared with satellite observations for the period from 1982 to 1999. As summarized in Appendix A5, Figure A6, the observed fields show strong spatial heterogeneity, with the highest values in the German Bight and Baltic Sea. All experiments reproduce key aspects of SST variability and warm anomaly behavior, although they underestimate variability in regions with the strongest observed anomalies. The results support the suitability of the coupled model for representing thermal conditions associated with extreme temperature events.

3.2 Projected changes for coastal extremes

To assess how climate change may influence future coastal hazards across the Northwest European shelf and the Baltic Sea, we compared the future period (2070–2099) with the historical baseline (1970–1999) using the same high-resolution regional model setup. The evaluation focuses on changes in frequency, intensity, and duration of ESS and MHW. The differences reflect the combined effect of changes in atmospheric forcing, ocean conditions, and internal variability.

3.2.1 Extreme storm surges

Historical 95th–percentile storm surge levels, ESS frequency, intensity and duration display a coherent spatial pattern across the northwestern European continental shelf and the Baltic Sea. Figure 6 presents the spatial distribution of the 95th–percentile threshold (a) and the ESS frequency (c), intensity (e), and duration (g) for the historical period and their relative differences (RD; b, d, f, h) to future runs. Only the mean values across the three experiments are shown, as EXPhig, EXPmed, and EXPlow exhibit similar spatial patterns. Areas without hatching indicate the regions where future and historical means differ significantly at the 95% confidence level, based on a student's *t*-test combined with the FDR method (p-value ≤ 0.05).

As shown in the left column of Figure 6, the greatest values occur along the North Sea and much of the UK coastal areas, with their maxima in the English Channel. This confirms that these areas are particularly prone to extreme storm surge activity, as found in previous studies (Vousdoukas et al., 2016). In contrast, ESS in BAS is generally less frequent and less intense, although their durations are comparable to those in NOS. NOS is directly exposed to North Atlantic storm tracks and therefore experiences more frequent and often more severe surges, sometimes multiple times within a single season. BAS is semi-enclosed and therefore not exposed to open-ocean storms. Long-term analyses confirm a lower surge frequency (Wolski and Wiśniewski, 2021). Nevertheless, persistent wind patterns can push water into narrow sub-basins, such as the Gulf of Finland or the western Baltic Sea, producing exceptional high surges (Groll et al., 2025).

The 95th-percentile threshold (Figure 6b) shows that RDs are generally small, fluctuating within a range of ±6.3%. However, noticeable increases in RDs appear in the offshore North Atlantic, Norwegian Sea and along the NWT, while negative values occur along the southern boundary, the Celtic Sea and the English Channel. Projected changes in ESS characteristics vary regionally substantially depending on the metric used. Overall, RDs increase in intensity by up to 20% in regions above 50°N, while decreases remain below 10%. The strongest increases occur in the southeastern North Sea, the northern UK and the southern Gulf of Bothnia, even though the projected changes in 95th-percentile are not statistically significant.

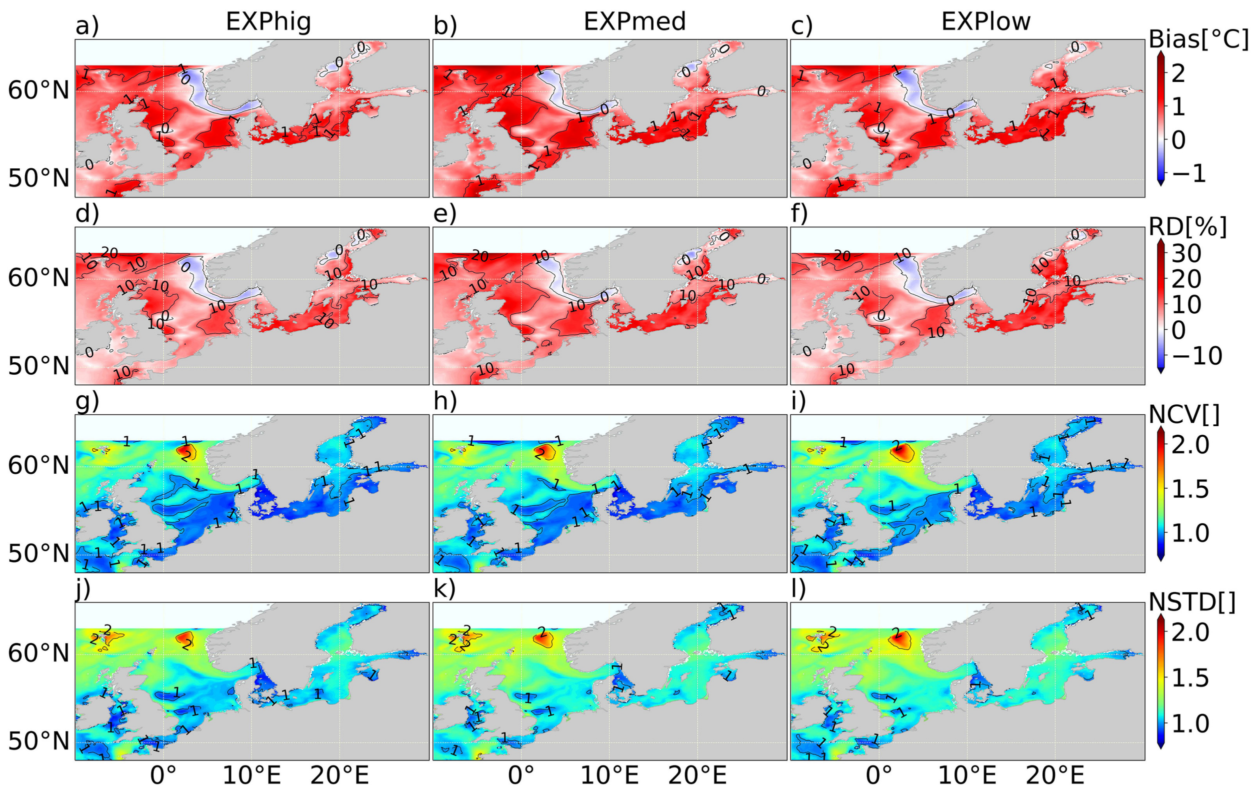

FIGURE 5
Spatial distribution of SST skill metrics for the period 1982–1999 for EXPhig (panels in first column); EXPmed (panels in second column) and EXPlow (panels in third column): mean differences [Bias: **(a–c)**], relative differences [RD: **(d–f)**], normalized coefficient of variation [NCV: **(g–i)**] and normalized standard deviation [NSTD: **(j–l)**].

A closer examination of regional patterns reveals that ESS projections for the Irish Sea indicate less frequent but more intense and longer-lasting events. This behavior is consistent with projected changes in storm-track dynamics over the North Atlantic (Chaigneau et al., 2024), increased persistence of high-energy wave and surge conditions, and the basin's semi-enclosed geometry, which enhances event duration. A similar pattern emerges in the NWS, where the highest surge levels coincide with the most frequent ESS events in the historical period. Future projections indicate a reduction in ESS frequency but an increase in surge intensity, particularly along the German Bight coast. This region is strongly influenced by westerly wind regimes (Gaslikova et al., 2013) and pronounced interannual variability (Dangendorf et al., 2014), although earlier studies reported no significant long-term change in storm surge (Lowe et al., 2009). Comparable behavior is evident along the northern UK coast. In contrast, there are some small areas in the southwest UK and the southeast Baltic Sea exhibiting declining 95th-percentile surge levels but increasing surge intensity. This reflects the fact that different parts of the extreme distribution respond to different physical drivers. ESS events using the 95th-percentile represent moderate extremes. These events can decline when typical storm tracks, prevailing wind directions, or seasonal storm characteristics become less favorable for generating frequent surges (e.g., Woollings et al., 2012). Conversely, the most intense and rare events depend on the combined local effects of wind, atmospheric pressure, and coastline geometry, and these can intensify even when moderate events weaken (Arns et al., 2013; Vousdoukas et al., 2018a).

Figure 7 presents the temporal evolution of ESS frequency, magnitude, and duration, averaged for six sub-areas defined in Figure 1b, over the 30-year period from 1970 to 1999. The time series reveals substantial interannual variability across all three metrics, with some regions exhibiting more pronounced fluctuations than others. All regions show weaker relationships among these metrics, underscoring notable spatial heterogeneity in storm surge dynamics. Overall, ESS activity exhibits marked temporal and regional variability, with the largest surge heights and highest occurrence rates of extreme events concentrated along the German Bight coast. Historical storm surges in the Baltic Sea are generally less frequent and less intense but tend to persist for longer durations.

The spatial distribution of internal variability (Figure 8, left column) and signal-to-noise ratios (SNR; right column) for extreme storm surge characteristics reveals a clear regional inverse relationship between the Northwest European continental shelf (NWS) and the Baltic Sea. The NWS exhibits substantially higher IV in frequency and intensity. This high variability is most pronounced in the Irish Sea, southern and central North Sea, including German Bight and adjacent coastal regions, where exposure to North Atlantic storm tracks drives strong year-to-year fluctuations. In contrast, the Baltic Sea exhibits exceptionally high IV in ESS duration and comparatively high IV along the northern

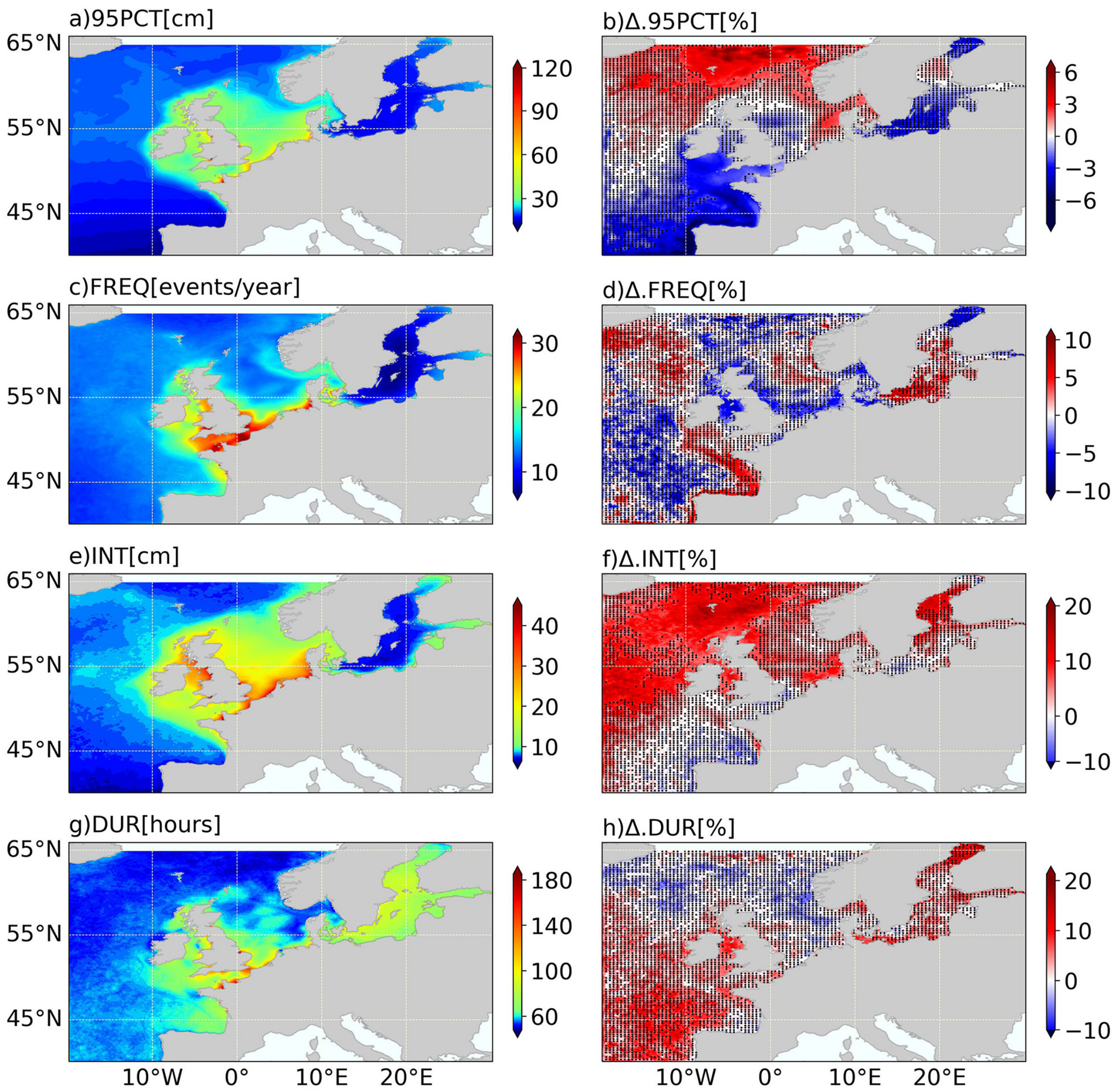

FIGURE 6
Spatial distribution of **(a)** the 95th−percentile, **(c)** frequency FREQ [events/year], **(e)** intensity INT [cm], and **(g)** duration DUR [hours] of ESS during the historical period. Panels **(b, d, f, h)** show the corresponding relative differences between the future projection (2070–2099) and the historical period (1970–1999). The hatched area in panels **(b, d, f, h)** indicates regions where the differences are statistically insignificant ($p > 0.05$).

Gulf of Bothnia and the eastern coastal Baltic Sea. These patterns indicate that internal climate processes exert strong control on ESS behavior, especially in closed basins and near coastal areas. The corresponding SNR quantifies the magnitude of projected changes relative to IV. The projected changes are relatively insignificant in comparison to the IV, ranging from 0 to 1, indicating that the IV exerts a dominant influence in shaping ESS characteristics. The largest SNR values occur for intensity in the offshore outer Northern NOS, indicating a comparable impact of forced signal and IV in these regions. ESS frequency shows SNR values up to 0.4 in the Baltic Sea and the southwestern UK. These findings demonstrate that projected changes in ESS characteristics are generally smaller than the magnitude of internal variability itself, highlighting the dominant role of IV for the whole domain, making projected changes less certain.

3.2.2 Extreme marine heatwaves

In this section, projected changes in MHWs are assessed by comparing future MHW characteristics with those of the historical period. The analysis focuses on the key MHW metrics: frequency, intensity, duration, and annual total MHW days across the North Sea and Baltic Sea. This provides a comprehensive view of how MHW behavior may change under future climate conditions. Figure 9 illustrates the historical distribution of these characteristics and their relative differences to the future.

Historical MHWs occurred relatively rarely in the southern North Sea, particularly the Southern Bight and along eastern English coast, where only about 1–1.5 events per year were recorded. Despite their low frequency, these events tended to persist for long periods, often exceeding 16 days. Such prolonged

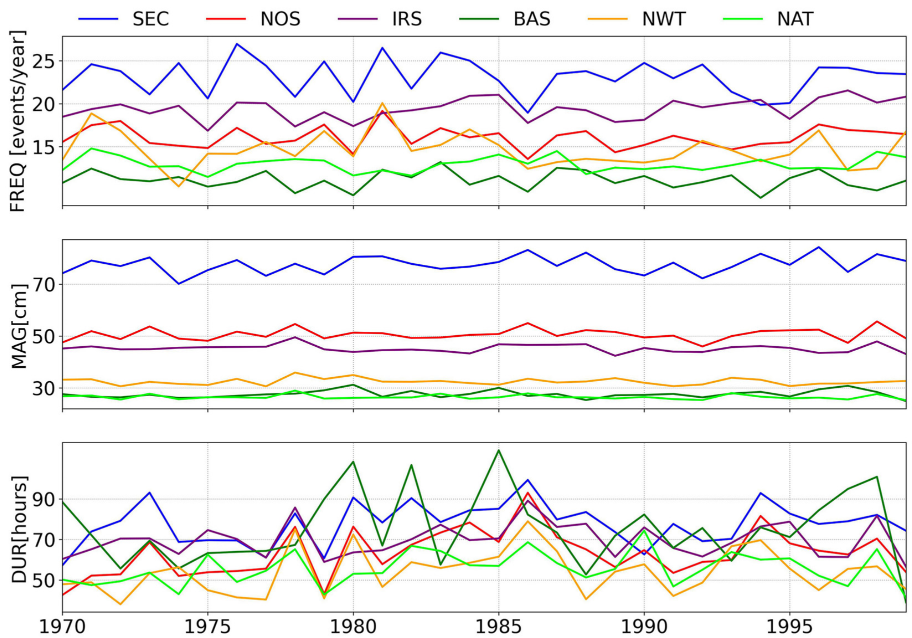

FIGURE 7
Frequency FREQ [events/year], mean magnitude MAG [cm] and mean duration DUR [hours] of storm surge events for each year, spatially averaged over the subregions defined in Figure 1b. Results are shown for the 30-year period from 1970 to 1999. Different colors represent individual subregions.

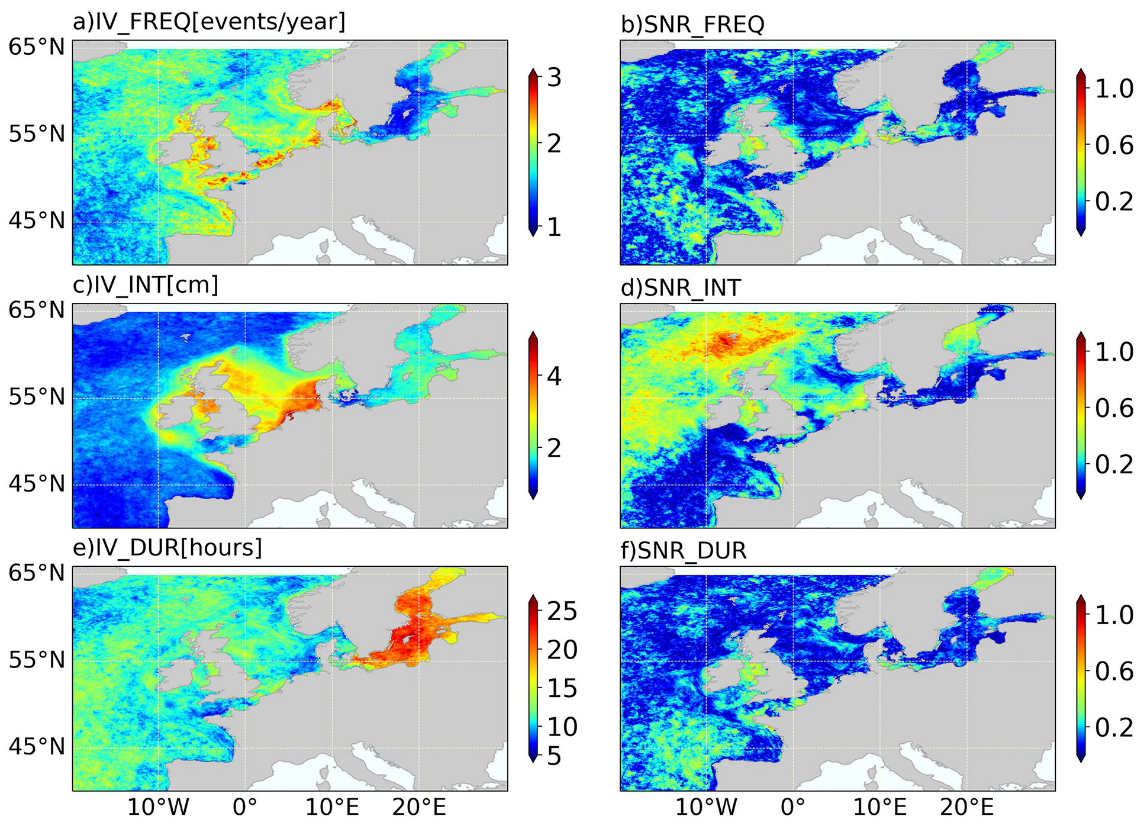

FIGURE 8
Spatial distribution of historical internal variability [IV; panels **(a, c, e)**] and signal-to-noise ratio [SNR; panels **(b, c, f)**] of extreme storm surges, shown for **(a, b)** frequency; **(c, d)** intensity; and **(e, f)** duration.

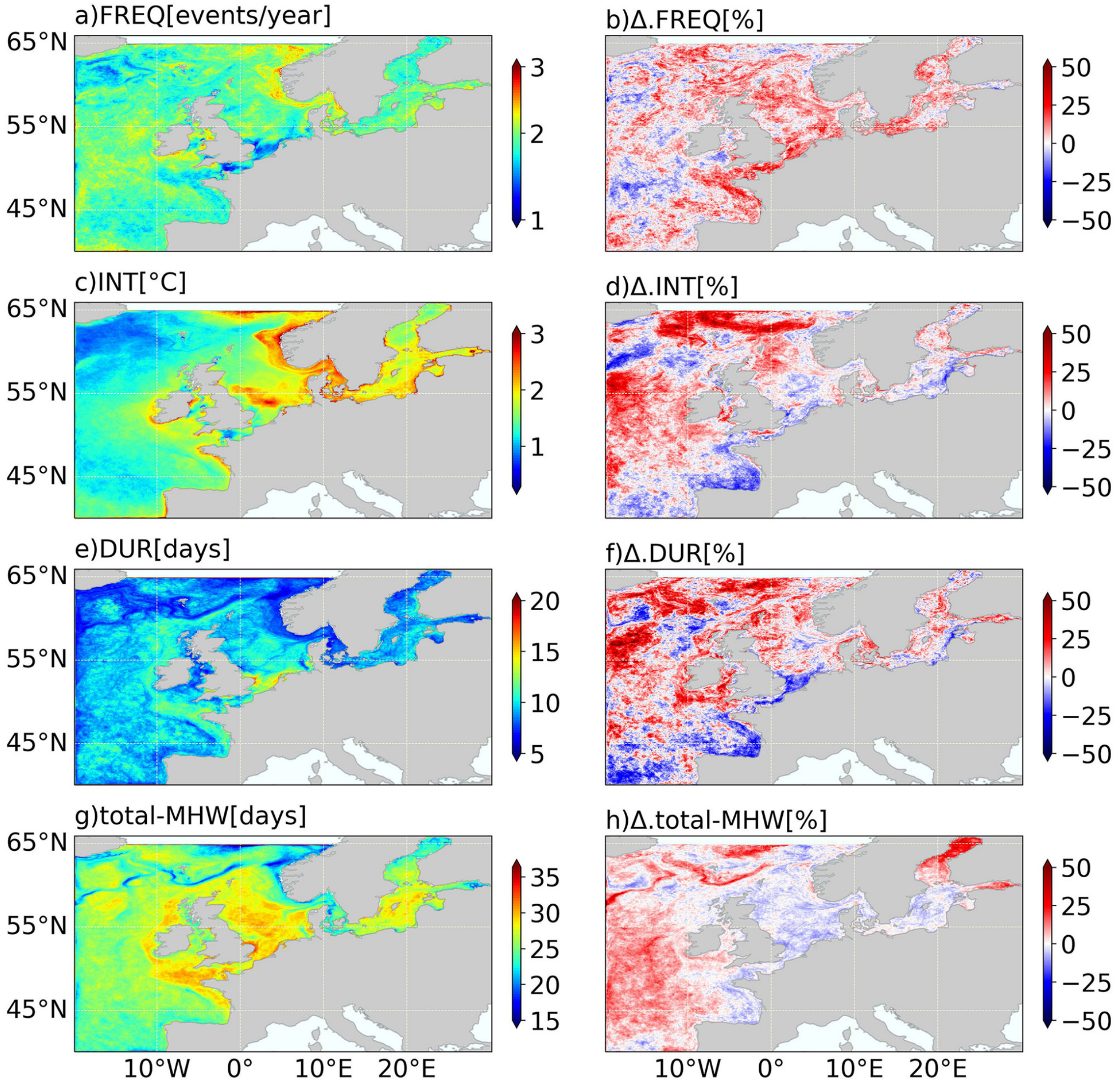

FIGURE 9
Projected changes in MHW characteristics: **(a, b)** frequency, **(c, d)** total MHW days, **(e, f)** duration, and **(g, h)** intensity. Left-hand panels show mean values across the three experiments averaged over 1970–1999, while right-hand panels show relative differences between the future period (2070–2099) and the historical period (1970–1999).

durations are consistent with the physical characteristics of these shallow and seasonally stratified regions, where warm anomalies are sustained by strong local heat retention and weak vertical mixing (Frölicher et al., 2018). In contrast, the NWT experienced more frequent but shorter MHW events, reflecting the influence of deeper waters, stronger currents, and more efficient heat exchange with the open ocean (Williams-Kerslake et al., 2025). By the end of the 21 century, the southern North Sea is projected to experience more frequent but shorter and less intense MHW, resulting in only modest changes in annual total MHW days (Frölicher et al., 2018). In the Baltic Sea, historical MHWs generally occurred more often, but projected changes in annual total MHW days vary substantially between sub-basins, with increases in the north and decrease in the south. These results highlight the strong influence of regional circulation, stratification patterns and basin-scale heat dynamics on future MHW responses (Travkin et al., 2024; Bashiri et al., 2025).

Figure 10 presents the evolution of MHW characteristics for both the historical (1970–1999) and the future periods (2070–2099). The values represent the spatial mean over the subregions defined in Figure 1b. During the historical period, MHW activity generally remains low, but all regions exhibit pronounced interannual variability in frequency, duration, intensity, and total MHW days. Notable peaks occur around 1980, 1991, and 1997, affecting most regions, but with varying magnitudes. Total MHW days (panel g) show a similar pattern, indicating years when events were both more frequent and longer lasting. MHW duration varies substantially between regions (panel e), with some subregions experiencing particularly longer-lasting events, consistent with the above-mentioned peaks. In contrast, MHW intensity (panel c) shows a gradual upward trend over the three decades, suggesting a climate warming trend. Regional differences are evident: the NAT region generally exhibits lower MHW intensities, whereas

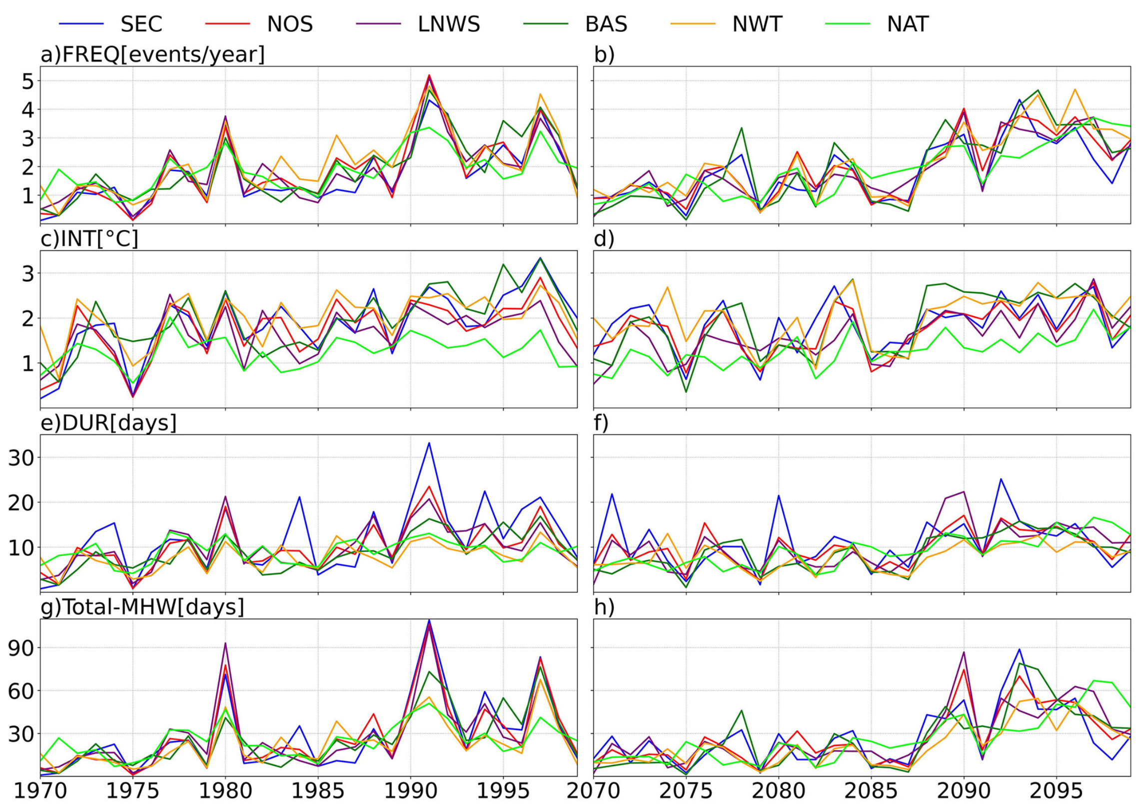

FIGURE 10
Evolution of marine heatwave (MHW) characteristics (yearly means over the three experiments) for the past and the future periods: **(a, b)** frequency, **(c, d)** intensity, **(e, f)** duration, and **(g, h)** total MHW days. Values show spatial averages across the subregions shown in Figure 1b.

the Northwest European continental shelf (including NOS, SEC, LNWS, and NWT) shows more pronounced fluctuations in all metrics except intensity. In the future period, the time series reveals a temporary reduction in MHW frequency and total heatwave days before values rise again later in the century. The results indicate that a future MHW activity for the first half of the scenario projection will be lower in the last part of the historical period. A clear upward shift across all metrics consistent with climate warming trends. However, projected changes vary regionally, affecting different metrics. The frequency and intensity of MHWs increase more strongly in the southern domain, while northern areas show a pronounced increase in total MHW days toward the end of the century. This reflects the transition from a cooler, more ice-affected environment in the past to a predominantly open-ocean environment in the future.

The internal variability (IV) and the signal-to-noise ratios (SNR) of the four MHW metrics are presented in Figure 11 as averages over a 30-year period (1970–1999). The Baltic Sea exhibits consistently high IV across all metrics, indicating substantial internal variability in this region. The strongest IV associated with total-MHW days and duration appears in the English Channel, while the NWT exhibits comparatively low IV for these two metrics. Overall, the spatial patterns of the ensemble-mean MHW characteristics and their SNR are consistent across the three ensemble simulations, reflecting dominant influences of IV over the projected changes in MHW characteristics.

4 Discussions

4.1 Interpretation of projected changes in extreme storm surges

Projected changes in extreme storm surges show pronounced spatial variability across the Northwest European continental shelf and the Baltic Sea. Along several parts of the North Sea coast, including the German Bight and sections of the UK coastline, our simulations indicate a decrease in the frequency of extreme storm surge events but an increase in their intensity. This pattern is physically plausible, reflecting changes in the distribution and characteristics of meteorological forcing rather than uniform shifts in storm occurrence. This behavior is also consistent with projected changes in storm track behavior and storm intensity under future climate scenarios (Schade et al., 2025). Although mean sea level rise does not substantially modify storm surge dynamics, it elevates the local mean sea level, thereby increasing the total water level associated with the most intense surge events. This trend has been documented across European coastlines, where sea level rise is identified as the dominant driver of increases in extreme sea levels (Calafat et al., 2022).

Extreme storm surges in this region are typically driven with persistent and intense wind forcing, strong pressure gradients, and specific storm track configurations that favor long fetch conditions (Timmerman, 1975; Woodworth et al., 2007; Gönnert and Sossidi,

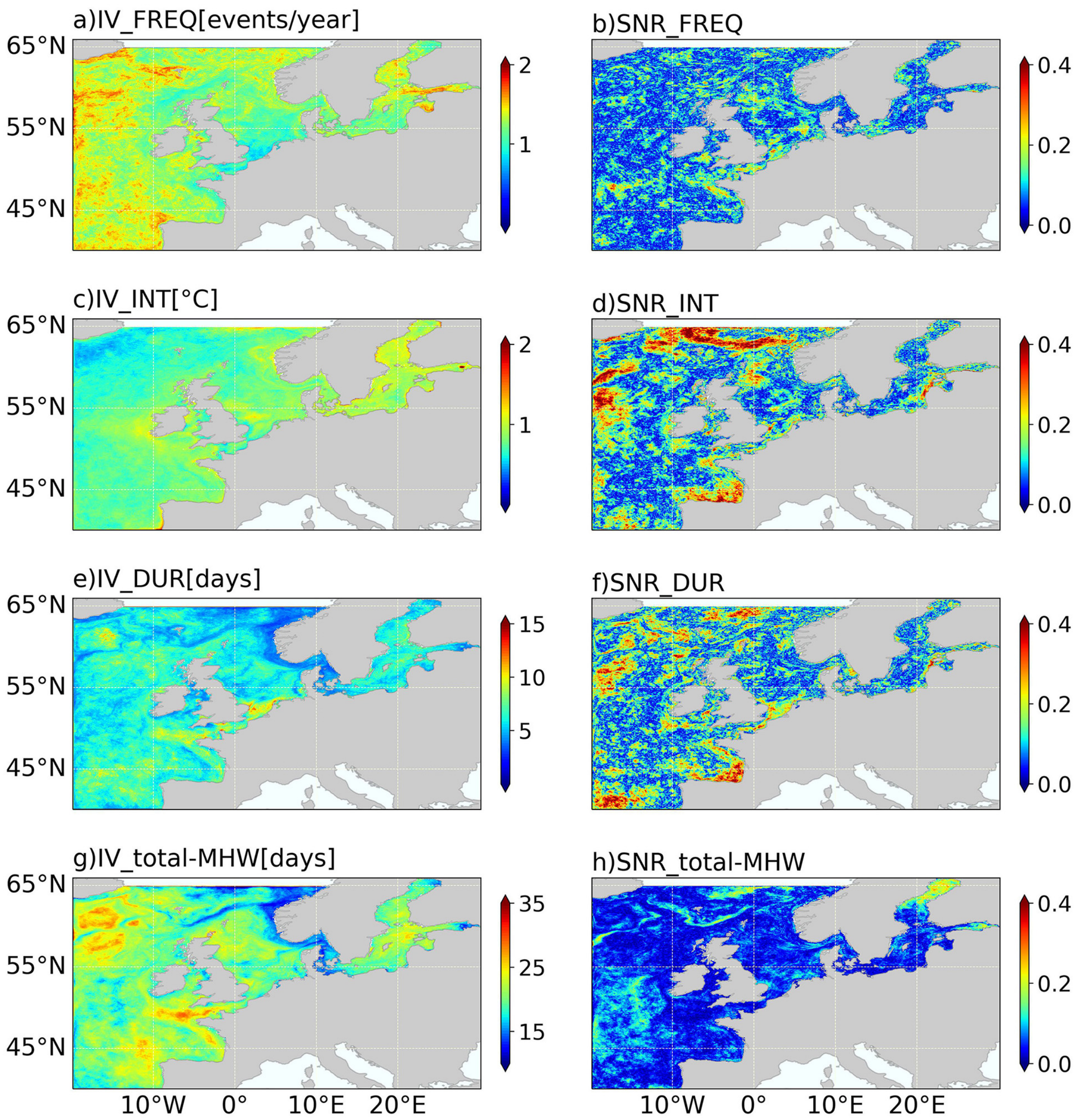

FIGURE 11
Spatial distribution of historical internal variability [IV; panels **(a, c, e, g)**] and signal-to-noise ratio [SNR; panels **(b, d, f, h)**] for MHWs: **(a, b)** frequency; **(c, d)** intensity; **(e, f)** duration, and **(g, h)** total MHW days.

2011). Even modest changes in the upper tail of wind speed or pressure gradient distributions can disproportionately affect the highest surge levels, while exerting a weaker influence on the total number of exceedance events. Future climate projections indicate a poleward and eastward shift of North Atlantic storm tracks (Woollings et al., 2012; Zappa et al., 2013; Harvey et al., 2020; IPCC, 2021). As a result, fewer low-pressure cyclones follow the classical southwest–northeast trajectory that typically drives strong southwesterly winds into the North Sea, leading to a reduction in the frequency of storm surges (Pinto et al., 2007, 2009). At the same time, extratropical cyclones are projected to intensify, with lower central pressure and stronger wind speed (Ulbrich et al., 2009; Zappa et al., 2013; Priestley and Catto, 2022). These changes enhance the upper tail of the wind speed and pressure gradient distributions, increasing the likelihood of producing the most extreme surges even if the total number of events declines. Moreover, nonlinear interactions between wind forcing, shallow-water dynamics, coastal bathymetry, and tide–surge coupling can amplify peak surge heights without a corresponding increase in event frequency (Arns et al., 2015; Vousdoukas et al., 2016). Similar behavior has been identified in earlier regional climate projections, where changes in storm characteristics and internal variability were found to modify the intensity of rare events more strongly than their occurrence rate (Woth et al., 2006; Gaslikova et al., 2013; Dangendorf et al., 2014). Collectively, these findings suggest that future storm surge risks along parts of the North Sea coast may increasingly be characterized by fewer but more severe events, particularly in regions such as the German Bight, where coastal geometry and shallow-water effects enhance sensitivity to extreme wind forcing.

Storm surges in the Baltic Sea are strongly influenced by the geometry of the semi-enclosed basins, local wind forcing, and weak tidal effects. The Baltic Sea is not affected by the storm surges entering from the North Atlantic, because it is only connected to the North Sea through the narrow Danish Straits. Furthermore, extreme sea levels in the region primarily result from local wind setup, storm direction, and the spatial distribution of wind fields. These factors have a significant impact on the magnitude and spatial pattern of surge events (Leppäranta and Myrberg, 2009; Madsen et al., 2019; Meier et al., 2022). These studies have demonstrated that strong, persistent westerly or southwesterly winds can push large volumes of water into the Baltic Sea. This results in pronounced sea level gradients and produces extreme surges, particularly in the Gulf of Finland, the Gulf of Riga, and the Bothnian Sea. In addition, basin-scale oscillations and seiches can further amplify sea-level responses to local storms. These oscillation modes, which are caused by the Baltic Sea's geometry and its limited exchange with the North Sea, can resonate with storm forcing to enhance surge levels under specific meteorological conditions (Wolski et al., 2014; Meier et al., 2022). Projected changes in large-scale atmospheric circulation, including shifts in the North Atlantic Oscillation (NAO), will continue to affect extreme storm surges in the Baltic Sea (Meier et al., 2004; Rutgersson et al., 2022). Studies by Hünicke and Zorita (2008); Meier et al. (2022) have suggested that intense wind events may become more episodic. However, climate models do not show a consistent increase in storm frequency. This could lead to higher risks of extreme storm surge events, even without a systematic increase in overall storminess, meaning that the Baltic Sea may experience intense storm surge events, driven by short-term atmospheric variability rather than long-term increases in storm activity. Future extreme sea levels in this region are profoundly affected by the mean sea level rise, specifically during the storm events (Kapsi et al., 2023; Dieterich and Radtke, 2024; Groll et al., 2025). Extreme sea levels will increase substantially, and extreme storm surges are projected to become more frequent, even if the magnitude of storm surges may not increase significantly in the future.

It is important to emphasize that our analysis focuses on storm surge residuals derived from sea surface height after removal of mean sea level and tidal signals. Consequently, the projected changes describe the meteorologically driven component of extreme sea levels rather than total extreme water levels. From a coastal flooding perspective, changes in total extreme sea level depend on the combined effects of storm surge residuals, tides, wave-related contributions, and mean sea level rise. Even where the frequency of storm surge residual exceedances decreases, the frequency of total water level exceedances is expected to increase under rising mean sea level (Vitousek et al., 2017; Wahl, 2017; Vousdoukas et al., 2018b; Frederikse et al., 2020; Tebaldi et al., 2021). The projected changes in storm surge residuals should therefore be interpreted as one component of future coastal flood hazards.

The GEV-based return level estimates of annual maximum storm surges for the historical period (1970–1999) and the future period (2070–2099) across the three experiments are shown in Appendix A4, Figure A5. The projected return level curve (dashed lines) in EXPmed lies above the historical curve (solid line), indicating a general increase in extreme storm surges across most return periods. This upward shift becomes more pronounced at longer return periods, reflecting a stronger intensification of rare, high-impact events under future climate conditions. The EXPhigh and EXPlow experiments exhibit more complex behavior. EXPhig shows a slight decrease in return levels for the return period from 2.2 to 20 years, while exhibiting increases for both shorter and longer return periods. In EXPlow, shorter return periods show a clear increase in future return levels; however, the return levels associated with longer return periods are noticeably lower than the historical estimates. This pattern implies that under EXPlow forcing, moderate extremes may become more frequent, while the most severe events may not intensify, or may even weaken, relative to the historical baseline. The 95% confidence intervals, shown as colored vertical lines, widen with increasing return period due to the rarity of extreme events, yet the separation between historical and future estimates remains evident in EXPmed. The results demonstrate that projected changes in extreme storm surges depend strongly on the choice of atmospheric forcing. EXPmed indicates substantial increases across the full range of return periods, while EXPhigh and EXPlow reveal more nuanced responses, highlighting the influence of internal variability on future surge projections. Despite these uncertainties, the general increase in return levels across several experiments suggests that high-impact, low-frequency surge events may become more likely in the future. These findings underscore the importance of considering both mean sea level rise and changes in the upper-tail behavior of storm surge distributions when assessing future coastal risk, as relying solely on mean changes would underestimate the potential severity of extreme events.

4.2 Projected changes in extreme marine heatwaves

The simulations indicate a substantial increase in the frequency of extreme marine heatwaves across large parts of the study region toward the end of the 21 century. This behavior is consistent with global and regional assessments that report more frequent marine heatwaves under continued ocean warming (Hobday et al., 2016; Frölicher et al., 2018; Oliver et al., 2018, 2019, 2021; IPCC, 2021). The spatial patterns show regional contrasts, with some areas experiencing more frequent but shorter events, while others show increases in both frequency and cumulative duration.

Such contrasts are consistent with the physical characteristics of shelf seas. Changes in stratification, vertical mixing, and circulation pathways can modulate the persistence of sea surface temperature anomalies, particularly in regions with strong seasonal stratification or limited ventilation, such as parts of the North Sea and the Baltic Sea (Holbrook et al., 2019; Sen Gupta et al., 2020; Chen et al., 2022; Chen and Staneva, 2024). The projected increase in the total number of marine heatwave days in several subregions implies longer cumulative exposure of marine ecosystems and coastal sectors to thermal stress, with potential consequences for biogeochemical cycling, oxygen conditions, and ecosystem functioning (Cavole et al., 2016; Smale et al., 2019).

A notable feature is the temporary reduction in MHW activity at the beginning of the future period compared with the end of the historical period. This behavior is not caused by colder atmospheric or ocean boundary conditions, nor by the model restart. The wave–ocean coupled model was forced using two 30-year time slices extracted from the transient climate simulation. This means that both the historical and future periods begin from fully dynamic, initially consistent ocean–atmosphere states. As a result, the future simulation does not start from a colder state. Instead, the temporary decline in MHW activity in the early future reflects the dominant influence of IV within the coupled system rather than a forcing or initialization, particularly multi-year fluctuations in upper-ocean heat content, stratification, and circulation superimposed on the warming trends. Moreover, in this study, the percentile used to detect MHWs is calculated based on their periods. However, using this approach may filter out the potential changes in storm tracks, wind regimes, cloud cover, and sea-ice evolution, thereby limiting the model's ability to capture emergent future drivers of MHWs. Therefore, the temporary decline in early future MHW activity should be interpreted as a manifestation of internal ocean variability superimposed on the overall warming trends, rather than as a robust signal of reduced future MHW risks.

4.3 Role of internal variability in projected coastal extremes

Internal variability emerges as a dominant source of uncertainty in the projected changes of both extreme storm surges and extreme marine heatwaves. The inter-member spread among the simulations indicates that internal variability can be comparable to, or even exceed, the mean projected changes in several regions. This finding is consistent with previous studies showing that internal variability can strongly modulate regional climate signals and delay or obscure the emergence of externally forced trends (Alexandru et al., 2007; Lucas-Picher et al., 2008; Hawkins and Sutton, 2009; Deser et al., 2012).

For storm surges, the importance of internal variability aligns with evidence of pronounced decadal and multi-decadal variability in North Sea storminess and surge conditions (Gaslikova et al., 2013; Dangendorf et al., 2014; Schrum et al., 2016). For marine heatwaves, internal variability reflects the sensitivity of surface temperature extremes to fluctuations in atmospheric forcing, ocean circulation, and background stratification (Holbrook et al., 2019; Sen Gupta et al., 2020; Oliver et al., 2021). These results imply that projections based on single realizations may not adequately represent the range of plausible future outcomes at regional scales, particularly for extremes that depend on the tail of the distribution.

4.4 Co-occurrence of extreme storm surges and marine heatwaves

Improving our understanding of co-occurring extremes is crucial for advancing integrated coastal risk assessments, as well as for developing effective adaptation strategies and early warning systems in vulnerable regions. While this study does not provide full compound events analysis, we demonstrate the potential for temporal overlap between ESS and MHWs using examples identified in the simulations. An illustrative case study of the Lerwick station in the northern North Sea shows how simultaneous extremes can occur locally. To assess whether this behavior is spatially consistent, we additionally examined modeled time series from several stations across the Northwest European continental shelf, including Roscoff and Newlyn. These sites span a range of hydrographic settings and exposure conditions, enabling us to evaluate whether co-occurrence patterns are unique to a single location or emerge more broadly across the region.

Using the methodological framework described earlier, years in which both ESS and MHWs occurred concurrently were identified across the 30-year period from 1970 to 1999. Figure 12 illustrates representative examples of such overlaps in different years and locations (see Table 3). These events can occur in different seasons. During these periods, elevated water levels driven by prevailing west-southwesterly winds coincide with winter MHWs, which are influenced by warm air advection and the mixing of deeper, warmer waters. These overlaps are noteworthy because many marine species, including cold-water corals and kelp forests, are highly sensitive to extreme temperatures. The combination of storm-driven flooding and thermal stress could endanger ecological and socio-economic systems.

Although Figure 12 only shows representative years for visualization, all statistical assessments use the full 30-year dataset. This approach ensures that the results reflect long-term variability rather than short-term anomalies. Across the entire dataset, co-occurring events are relatively rare. A quantitative assessment indicates that co-occurrence events represent only a small fraction of the total record, confirming that these compound extremes are relatively rare but potentially impactful. These events can be explained using the same conditions that generate strong storm surges, such as prolonged wind forcing and persistent pressure gradients, which may contribute to the development or persistence of warm ocean anomalies. However, the strength of this relationship varies both regionally and interannually. This indicates that co-occurrence is not governed by a single mechanism. Instead, the interplay between atmospheric circulation, ocean stratification, and seasonal conditions is important. While this analysis provides first insights into potential compound extremes, the limited number of co-occurrence events indicates the necessity of further investigation. Broader spatial coverage and longer records are required.

While ESS and MHWs often peak in different seasons, the simulations suggest that they can overlap, particularly during transitional periods. Severe sea surface temperatures associated with long-term warming can persist into autumn and winter, while storm surge events continue to occur under strong wind forcing and low mean sea level pressure (Wahle et al., 2017). Even limited temporal overlaps can intensify impacts, such as storm-driven flooding and infrastructure stress coincide with thermal stress on marine ecosystems.

Recent global analyses indicate that concurrent extreme sea levels and marine heatwaves are already widespread and are projected to increase under high-emission scenarios

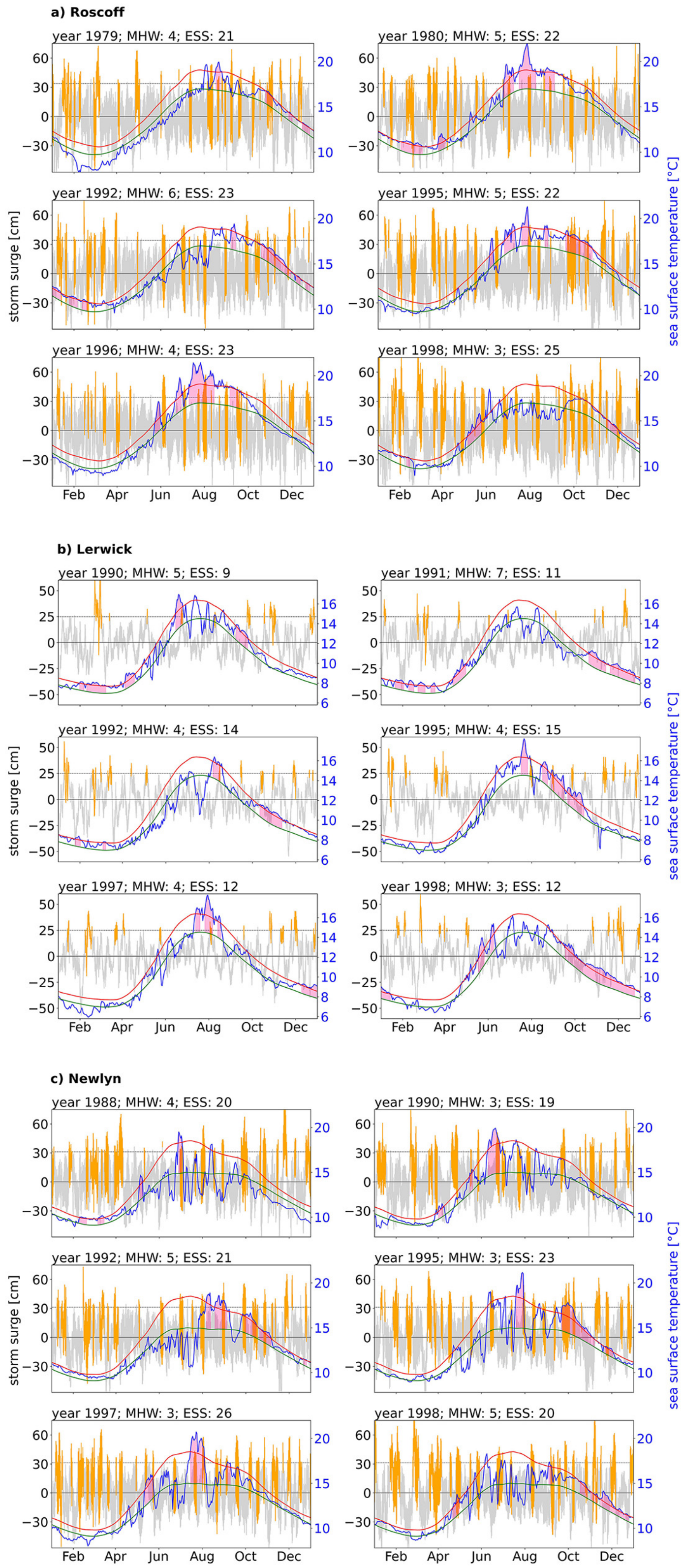

FIGURE 12
Time series of hourly storm surge and daily temperature at several stations in Northwest European shelf, **(a)** Roscoff, **(b)** Lerwick, and **(c)** Newlyn, extracted from the EXPhig experiment. Orange curves indicate extreme surge events exceeding the 95th–percentile threshold, and the shaded pink highlights periods identified as MHW.

TABLE 3 Years with ESS and MHW co-occurrences at various stations across the Northwest European shelf.

st	73	74	76	77	79	80	82	83	86	88	90	91	92	94	95	96	97	98	99
1		x	x		x	x				x	x		x		x		x	x	
2					x	x							x		x	x		x	
3						x				x			x		x			x	x
4						x								x	x	x		x	x
5				x		x			x			x		x	x		x		x
6	x			x		x						x	x		x	x		x	
7				x								x	x	x	x		x	x	x
8			x	x		x	x				x		x	x	x	x		x	
9								x			x	x	x			x	x	x	x

The first column lists the station numbers (corresponding to the locations shown in Figure 1b), and the subsequent columns indicate the years in which ESS-MHW co-occurrences were detected (e.g., 77 = 1977; 99 = 1999).

(Zhou and Wang, 2024). The simulated examples of co-occurrence in the Northwest European shelf suggest that such compound events may represent an additional and growing source of coastal risk, particularly for sensitive ecosystems and densely populated coastal regions (Glynn and D'Croz, 1990; Wernberg et al., 2016). A systematic quantification of compound-event probabilities and joint drivers would further strengthen integrated coastal risk assessments.

Finally, while this study focuses on ESS–MHW co-occurrence, we acknowledge that this is only one of several possible compound hazards affecting coastal regions. Other combinations, such as storm surges with heavy rainfall or MHWs with oxygen depletion, may produce more severe or widespread consequences. Our aim here is not to compare the relative severity of different compound events, but to demonstrate that ESS and MHW overlaps do occur within the coupled modeling framework and therefore represent an additional hazard worth considering. A comprehensive assessment of interactions among hydrological extremes, biogeochemical stressors, and physical hazards would require dedicated analyses and remains an important direction for future research.

4.5 Limitations and outlook

In the present work, the wave model is used to provide physically consistent surface forcing within the coupled wave–ocean modeling framework, but wave outputs are not analyzed or interpreted. Nevertheless, the coupled model provides a robust physical basis for examining extreme storm surges and extreme marine heatwaves within the same modeling system. This integration is particularly relevant because storm surges are primarily driven by wind stress, atmospheric pressure gradients, and coastal geometry, whereas marine heatwaves are governed by air–sea heat fluxes, vertical mixing, and ocean circulation. Previous studies have intensively focused on wave-induced processes affecting ocean dynamics, while neglecting their influence on the atmosphere, particularly how surface waves modify sea surface roughness and affect the ocean–atmosphere boundary layers (e.g., Wahle et al., 2017; Wiese et al., 2020). Wind-generated surface waves exert a crucial impact on both oceanic and atmospheric components of Earth system models. Fully coupled wave–ocean–atmosphere modeling frameworks (Warner et al., 2010; Wu et al., 2019a; Grayek et al., 2023) enable these interactions to be represented consistently across multiple hazards avoiding the inconsistencies that may arise when separate modeling systems are used for different extremes.

A major limitation of this study is that it is based on the same model, the same forcing scenario, and only three ensemble members, which considerably increases uncertainty represented for both marine heatwaves and extreme storm surges. When only one emissions scenario is used, the analysis is limited to a single greenhouse gas trajectory and therefore does not capture the broader range of plausible future forcing conditions, even though the selected scenario profoundly affects the projected marine ecosystem response and climate-driven hazards (Payne et al., 2016). Furthermore, the use of a single model statistically constrains the robustness of the results, as different models vary in their representation of key processes such as atmospheric circulation, storm track behavior, and coastal dynamics. These differences can lead to substantial inter-member spread in projections of marine extremes (Denamiel and Vilibić, 2023; Muis et al., 2023). In addition, the use of only three ensemble members provides a partial sampling of internal variability. It represents a limitation as natural fluctuations in the atmosphere–ocean system significantly modulates both MHW occurrence and storm surge intensity on interannual to decadal time scales. Consequently, some of the projected changes may reflect ensemble-specific noise rather than the forced climate signal. Future work would benefit from incorporating multiple emissions scenarios, a multi-model ensemble, and larger ensemble sizes to fully characterize scenario uncertainty, model uncertainty, and internal variability. This would provide a more comprehensive and robust assessment of future extremes.

Further limitations should be considered when interpreting the results. The horizontal resolution of the regional ocean model, while high for climate applications, may still be insufficient to fully resolve estuarine dynamics and shallow nearshore processes that influence local surge levels. In addition, the modeling chain does not include a fully coupled atmosphere–wave feedback.

Previous studies have shown that wave–atmosphere coupling can modify near-surface winds and improve the representation of extreme wind and wave conditions, thereby affecting storm surge forcing and associated uncertainty (Wahle et al., 2017; Wiese et al., 2020). Our present results therefore likely represent a conservative estimate of uncertainty regarding the wind-driven extremes.

The annual mean of the 95th-percentile is a statistically robust and internally consistent metric for comparing high storm surge conditions across ensemble members and time periods. It captures the upper tail of the storm surge distribution and provides sufficient data points to ensure stable statistics. This is essential for quantifying IV and ensemble spread (Alexandru et al., 2007; Lucas-Picher et al., 2008). However, this metric does not fully capture the most extreme events in the tail of distribution. Higher percentile thresholds (e.g., the 99th percentile) would better represent the most severe surge events but become statistically unreliable for small sample sizes (Coles, 2001; Katz et al., 2002). Moreover, using only three representative ensemble members introduces uncertainty in estimating internal variability, particularly for extreme events. Our selection was constrained by computational resources and may therefore lead to either an over- or underestimation of internal variability. These limitations should be considered when interpreting the results, particularly in the context of coastal risk assessments.

Next, translating storm surge residual changes into actionable flood-risk information requires integration with mean sea level rise and additional contributors to total water level, including wave setup and river discharge, which are known to play an important role in low-lying estuarine environments (Wolf, 2009; Wahl, 2017; Vousdoukas et al., 2018b).

A key limitation of the method is its reduced reliability in regions and time periods affected by sea ice, such as the Bothnian Sea and the Gulf of Finland in the Baltic Sea (Eriksson et al., 2025). While ESS can still occur when the sea is partially or fully ice-covered. The ice strongly modifies surge dynamics by damping waves, reducing wind stress transfer, and isolating the ocean surface from atmospheric forcing. Consequently, both the simulation of water levels and the detection of ESS and MHWs become unreliable during ice-covered periods (Leppäranta, 2011; Merkouriadi and Leppäranta, 2014; Soomere and Pindsoo, 2016). For this reason, the method should be applied with consideration of ice-affected seasons, and additional treatments are required to prevent the detection of false extreme events. This limitation is particularly relevant in the Baltic Sea, where sea ice cover influences winter conditions and modulates storm surge dynamics (Soomere and Pindsoo, 2016). Projected decreases in ice extent and duration (Höglund et al., 2017), together with the continued occurrence of partially ice-covered periods (Meier et al., 2004), will continue to influence the accuracy of the detection method under future climate conditions. As long-term warming alters the timing and characteristics of ice formation and melting, careful consideration of ice–ocean interactions will remain crucial. Incorporating ice–ocean interactions more explicitly, for example, through coupled ice–ocean modeling or by applying alternative metrics during ice-covered periods, would strengthen the robustness of the analysis in these regions.

5 Conclusion

This study assessed future changes in two major coastal hazards, extreme storm surges and extreme marine heatwaves, over the Northwest European shelf and the Baltic Sea using the high-resolution GCOAST wave–ocean coupled model (WAM–NEMO) driven by dynamically downscaled REMO–MPIOM boundary conditions under the RCP8.5 scenario. Three representative members were selected to quantify the role of internal variability in shaping projected changes from a historical period (1970–1999) to a future period (2070–2099).

The historical evaluation shows that the coupled model reproduces key features of storm surge residual variability compared with the observed tide-gauges data and captures the main spatial patterns of SST when compared with satellite-derived products provided by CMEMS, supporting its use for analyzing regional extremes. The projections indicate spatially heterogeneous responses of coastal extremes. For storm surge residual extremes, results suggest a tendency toward reduced event frequency but increased event intensity along parts of the North Sea coast, including the German Bight, while other regions exhibit different combinations of changes. For MHWs, the simulations show an increase in frequency across much of the domain. There are regional differences in persistence and cumulative heat exposure, including areas where both frequency and the number of annual MHW days increase toward the end of the century.

A central outcome of this study is that internal variability substantially modulates projected changes in both ESS and MHW characteristics. In several regions, the inter-member spread, and the IV-related changes are larger than the projected mean change, indicating that regional hazard evolution cannot be described reliably from a single realization. This emphasizes the need for ensemble-based approaches and for careful communication of uncertainty in regional projections of extremes.

The storm surge results presented here describe changes in surge residual extremes (meteorological surge component). For coastal flood risk assessment, these changes must be interpreted together with mean sea level rise, tides, and wave-related contributions to the total water level. Future work should therefore integrate mean sea level rise and additional coastal processes, expand ensemble sampling, and develop compound-event metrics to better support coastal risk management and adaptation planning in the Northern European shelf seas.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

TTN: Methodology, Conceptualization, Validation, Writing – original draft, Investigation, Software, Visualization, Formal analysis, Writing – review & editing. JS: Resources, Funding acquisition, Supervision, Project administration, Methodology, Writing – review & editing, Conceptualization. HTMH-H: Writing – review & editing. CS: Supervision, Writing – review & editing. BM: Writing – review & editing, Data curation.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the Hydrodynamics and Data assimilation department, the Helmholtz-Zentrum Hereon (Hereon). TTN acknowledges the Horizon Europe COAST-SCAPES project. JS acknowledges Horizon-Europe Project SeaCLIM and FOCCUS. The study is a contribution to the CoastalFutures-II project that was funded by the Federal Ministry of Research, Technology and Space (BMFTR) under the funding code 03F0980A. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Executive Agencies (EAs). Neither the European Union nor the EAs can be held responsible for them.

Acknowledgments

This study was supported by the Hydrodynamics and Data assimilation department, the Helmholtz-Zentrum Hereon (Hereon), grant funded by the CLICCS-A6 project: the earth system variability and predictability in changing climate. Further, the production of the results presented in this study used computational resources of the LEVANTE cluster system of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID gg0028. We would like to express our gratitude to Dr. Sebastian Grayek for his invaluable support in preparing the model setup. We thank Dr. Antonio Bonaduce for fruitful discussions and Dr. Pham Thanh Nam for his crucial assistance with the wave model validation.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fclim.2026.1782346/full#supplementary-material

References

Alari, V., Staneva, J., Breivik, Ø., Bidlot, J.-R., Mogensen, K., and Janssen, P. (2016). Surface wave effects on water temperature in the Baltic Sea: simulations with the coupled NEMO-WAM model. *Ocean Dyn.* 66, 917–930. doi: 10.1007/s10236-016-0963-x

Alexandru, A., De Elia, R., and Laprise, R. (2007). Internal Variability in Regional Climate Downscaling at the Seasonal Scale. *Mon. Weather. Rev.* 135, 3221–3238. doi: 10.1175/MWR3456.1

Arns, A., Wahl, T., Dangendorf, S., and Jensen, J. (2015). The impact of sea level rise on storm surge water levels in the northern part of the German Bight. *Coast. Eng.* 96, 118–131. doi: 10.1016/j.coastaleng.2014.12.002

Arns, A., Wahl, T., Haigh, I. D., Jensen, J., and Pattiaratchi, C. (2013). Estimating extreme water level probabilities: a comparison of the direct methods and recommendations for best practise. *Coast. Eng.* 81, 51–66. doi: 10.1016/j.coastaleng.2013.07.003

Barragán, J. M., and De Andrés, M. (2015). Analysis and trends of the world's coastal cities and agglomerations. *Ocean Coast. Manage.* 114, 11–20. doi: 10.1016/j.ocecoaman.2015.06.004

Bashiri, B., Barzandeh, A., Männik, A., Uiboupin, R., and Raudsepp, U. (2025). Marine heatwave event maps in the Baltic sea (1982–2023): a gridded dataset from satellite-derived L4 SST. *Sci. Data* 12:2003. doi: 10.1038/s41597-025-06251-7

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bernier, N. B., Hemer, M., Mori, N., Appendini, C. M., Breivik, O., De Camargo, R., et al. (2024). Storm surges and extreme sea levels: review, establishment of model intercomparison and coordination of surge climate projection efforts (SurgeMIP). *Weather Clim. Extrem.* 45:100689. doi: 10.1016/j.wace.2024.100689

Bonaduce, A., Pham, N. T., Staneva, J., Grayek, S., Raj, R. P., and Breivik, Ø. (2024). Sea state contributions to thermosteric sea-level in high-resolution ocean-wave coupled simulations. *Ocean Dyn.* 74, 743–761. doi: 10.1007/s10236-024-01632-9

Bonaduce, A., Staneva, J., Grayek, S., Bidlot, J.-R., and Breivik, Ø. (2020). Sea-state contributions to sea-level variability in the European Seas. *Ocean Dyn.* 70, 1547–1569. doi: 10.1007/s10236-020-01404-1

Bowyer, P., Schaller, M., Bender, S., and Jacob, D. (2014). "Adaptation as climate risk management: methods and approaches," in *Handbook of Climate Change Adaptation*, ed. W. Leal Filho (Berlin/Heidelberg: Springer Berlin Heidelberg), 1–18. doi: 10.1007/978-3-642-40455-9_28-1

Calafat, F. M., Wahl, T., Tadesse, M. G., and Sparrow, S. N. (2022). Trends in Europe storm surge extremes match the rate of sea-level rise. *Nature* 603, 841–845. doi: 10.1038/s41586-022-04426-5

Caldwell, P. C., Merrifield, M. A., and Thompson, P. R. (2001). *Sea Level Measured by Tide Gauges from Global Oceans as Part of the Joint Archive for Sea Level (JASL) Since 1846*, National Centers for Environmental Information.

Cavole, L., A.M. Demko, Diner, R., Giddings, A., Koester, I., Pagniello, C., et al. (2016). Biological impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers, and the future. *Oceanog* 29, 273–285. doi: 10.5670/oceanog.2016.32

Chaigneau, A. A., Melet, A., Voldoire, A., Irazoqui Apecechea, M., Reffray, G., Law-Chune, S., et al. (2024). Dynamic projections of extreme sea levels for western Europe based on ocean and wind-wave modelling. *Nat. Hazards Earth Syst. Sci.* 24, 4031–4048. doi: 10.5194/nhess-24-4031-2024

Chen, W., and Staneva, J. (2024). Characteristics and trends of marine heatwaves in the northwest European Shelf and the impacts on density stratification. *State Planet* 4-osr8, 1–11. doi: 10.5194/sp-4-osr8-7-2024

Chen, W., Staneva, J., Grayek, S., Schulz-Stellenfleth, J., and Greinert, J. (2022). The role of heat wave events in the occurrence and persistence of thermal stratification in the southern North Sea. *Nat. Hazards Earth Syst. Sci.* 22, 1683–1698. doi: 10.5194/nhess-22-1683-2022

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., et al. (2013). Sea-level rise by 2100. *Science* 342, 1445–1445. doi: 10.1126/science.342.6165.1445-a

Codiga, D. (2011). *Unified Tidal Analysis and Prediction using the UTide Matlab Functions*. Available online at: https://www.semanticscholar.org/paper/Unified-Tidal-Analysis-and-Prediction-Using-the-Codiga/0be34c47564f7b94eb4d09bb7824da4afd41d22b (Accessed February 26, 2025).

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. London: Springer London. doi: 10.1007/978-1-4471-3675-0

Craig, P. D., and Banner, M. L. (1994). Modeling wave-enhanced turbulence in the ocean surface layer. *J. Phys. Oceanogr.* 24, 2546–2559. doi: 10.1007/1520-0485(1994)024<2546:MWETIT>2.0.CO;2

Dangendorf, S., Müller-Navarra, S., Jensen, J., Schenk, F., Wahl, T., and Weisse, R. (2014). North Sea storminess from a novel storm surge record since AD 1843*. *J. Clim.* 27, 3582–3595. doi: 10.1175/JCLI-D-13-00427.1

Denamiel, C., and Vilibić, I. (2023). Next generation Atmosphere-Ocean Climate Modelling for Storm Surge Hazard Projections. *EGUsphere* doi: 10.5194/egusphere-2023-913

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S. (2012). Communication of the role of natural variability in future North American climate. *Nat. Clim. Chang.* 2, 775–779. doi: 10.1038/nclimate1562

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Chang.* 10, 277–286. doi: 10.1038/s41558-020-0731-2

Dieterich, C., and Radtke, H. (2024). Higher quantiles of sea levels rise faster in Baltic Sea Climate projections. *Clim. Dyn.* 62, 3709–3719. doi: 10.1007/s00382-023-07094-x

Egbert, G. D., and Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Ocean. Technol.* 19, 183–204. doi: 10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2

Embury, O., Merchant, C. J., Good, S. A., Rayner, N. A., Høyer, J. L., Atkinson, C., et al. (2024). Satellite-based time-series of sea-surface temperature since 1980 for climate applications. *Sci. Data* 11:326. doi: 10.1038/s41597-024-03147-w

Eriksson, P. B., Vainio, J., Tollman, N., Jokiniemi, A., Arola, A., Mäkynen, M., et al. (2025). The Finnish Ice Service, its sea-ice monitoring of the Baltic Sea and operational concept. *Front. Mar. Sci.* 12:1561461. doi: 10.3389/fmars.2025.1561461

Faghih, M., and Brissette, F. (2023). The role of internal climate variability on future streamflow projections. *J. Hydrol. (Amst.)* 625:130101. doi: 10.1016/j.jhydrol.2023.130101

Fan, Y., Held, I. M., Lin, S.-J., and Wang, X. L. (2013). Ocean warming effect on surface gravity wave climate change for the end of the twenty-first century. *J. Clim.* 26, 6046–6066. doi: 10.1175/JCLI-D-12-00410.1

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V., et al. (2020). The causes of sea-level rise since 1900. *Nature* 584, 393–397. doi: 10.1038/s41586-020-2591-3

Frölicher, T. L., Fischer, E. M., and Gruber, N. (2018). Marine heatwaves under global warming. *Nature* 560, 360–364. doi: 10.1038/s41586-018-0383-9

Gaslikova, L., Grabemann, I., and Groll, N. (2013). Changes in North Sea storm surge conditions for four transient future climate realizations. *Nat. Hazards* 66, 1501–1518. doi: 10.1007/s11069-012-0279-1

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., et al. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *J. Adv. Model. Earth Syst.* 5, 572–597. doi: 10.1002/jame.20038

Glynn, P. W., and D'Croz, L. (1990). Experimental evidence for high temperature stress as the cause of El Niño-coincident coral mortality. *Coral Reefs* 8, 181–191. doi: 10.1007/BF00265009

Gönnert, G., and Sossidi, K. (2011). "A new approach to calculate extreme storm surges: analysing the interaction of storm surge components," in *Coastal Processes 2011* (Naples, Italy), 139–150. doi: 10.2495/CP110121

Grayek, S., Wiese, A., Ho-Hagemann, H. T. M., and Staneva, J. (2023). Added value of including waves into a coupled atmosphere–ocean model system within the North Sea area. *Front. Mar. Sci.* 10:1104027. doi: 10.3389/fmars.2023.1104027

Groll, N., Gaslikova, L., and Weisse, R. (2025). Recent Baltic Sea storm surge events from a climate perspective. *Nat. Hazards Earth Syst. Sci.* 25, 2137–2154. doi: 10.5194/nhess-25-2137-2025

Hagemann, S., and Gates, L. D. (2001). Validation of the hydrological cycle of ECMWF and NCEP reanalyses using the MPI hydrological discharge model. *J. Geophys. Res.* 106, 1503–1510. doi: 10.1029/2000JD900568

Harvey, B. J., Cook, P., Shaffrey, L. C., and Schiemann, R. (2020). The response of the northern hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models. *JGR Atmos.* 125:e2020JD032701. doi: 10.1029/2020JD032701

Hawkins, E., and Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.* 90, 1095–1108. doi: 10.1175/2009BAMS2607.1

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J., et al. (2014). Coastal flood damage and adaptation costs under 21st century sea-level rise. *Proc. Natl. Acad. Sci. U. S. A.* 111, 3292–3297. doi: 10.1073/pnas.1222469111

Hobday, A. J., Alexander, L. V., Perkins, S. E., Smale, D. A., Straub, S. C., Oliver, E. C. J., et al. (2016). A hierarchical approach to defining marine heatwaves. *Prog. Oceanogr.* 141, 227–238. doi: 10.1016/j.pocean.2015.12.014

Höglund, A., Pemberton, P., Hordoir, R., and Schimanke, S. (2017). Ice conditions for maritime traffic in the Baltic Sea in future climate. *Boreal Environ. Res.* 22, 245–265. doi: 10.60910/3ysb-99rn

Ho-Hagemann, H. T. M., Hagemann, S., Grayek, S., Petrik, R., Rockel, B., Staneva, J., et al. (2020). Internal model variability of the regional coupled system model GCOAST-AHOI. *Atmosphere* 11:227. doi: 10.3390/atmos11030227

Holbrook, N. J., Scannell, H. A., Sen Gupta, A., Benthuysen, J. A., Feng, M., Oliver, E. C. J., et al. (2019). A global assessment of marine heatwaves and their drivers. *Nat. Commun.* 10:2624. doi: 10.1038/s41467-019-10206-z

Horsburgh, K. J., and Wilson, C. (2007). Tide-surge interaction and its role in the distribution of surge residuals in the North Sea. *J. Geophys. Res.* 112:2006JC004033. doi: 10.1029/2006JC004033

Høyer, J. L., and Karagali, I. (2016). Sea surface temperature climate data record for the North Sea and Baltic Sea. *J. Clim.* 29, 2529–2541. doi: 10.1175/JCLI-D-15-0663.1

Høyer, J. L., and She, J. (2007). Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *J. Mar. Syst.* 65, 176–189. doi: 10.1016/j.jmarsys.2005.03.008

Hünicke, B., and Zorita, E. (2008). Trends in the amplitude of Baltic Sea level annual cycle. *Tellus A Dyn. Meteorol. Oceanogr.* 60:154. doi: 10.1111/j.1600-0870.2007.00277.x

IPCC (2021). "Climate change 2021: the physical science basis," in *Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge/New York, NY: Cambridge University Press), 2391.

IPCC (ed.) (2022). "Extremes, abrupt changes and managing risks," in *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press), 589–656. doi: 10.1017/9781009157964.008

IPCC (2013). "Climate change 2013: the physical science basis," in *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. T. F. Stocker, T. F. D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (Cambridge/New York, NY: Cambridge University Press), 1535. Available online at: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_all_final.pdf (Accessed April 12, 2025).

Jacob, D. (2001). A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorol. Atmos. Phys.* 77, 61–73. doi: 10.1007/s007030170017

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.* 5, 422–446. doi: 10.1002/jame.20023

Kapsi, I., Kall, T., and Liibusk, A. (2023). Sea level rise and future projections in the Baltic Sea. *JMSE* 11:1514. doi: 10.3390/jmse11081514

Katz, R. W., Parlange, M. B., and Naveau, P. (2002). Statistics of extremes in hydrology. *Adv. Water Resour.* 25, 1287–1304. doi: 10.1016/S0309-1708(02)00056-8

Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D., et al. (2020). Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. *Sci. Rep.* 10:11629. doi: 10.1038/s41598-020-67736-6

Komen, G. J., Hasselmann, K., Cavaleri, L., Donelan, M., Janssen, P. A. E. M., and Hasselmann, S. (eds.) (1994). *"Frontmatter," in Dynamics and Modelling of Ocean Waves* (Cambridge: Cambridge University Press), i–vi. Available online at: https://www.cambridge.org/core/books/dynamics-and-modelling-of-ocean-waves/frontmatter/0F906002AED7EE306B179E75C608856A (Accessed May 9, 2024).

Leppäranta, M. (2011). *The Drift of Sea Ice*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-04683-4

Leppäranta, M., and Myrberg, K. (2009). *Physical Oceanography of the Baltic Sea*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-79703-6

Lowe, J., Howard, T., Pardaens, A., Tinker, J., Holt, J., Wakelin, S., et al. (2009). *Climate Projections Science Report: Marine and Coastal Projections*, Tech. rep., Met Office Hadley Centre. http://ukclimateprojections.defra.gov.uk/22570 (Accessed May 16, 2013).

Lucas-Picher, P., Caya, D., De Elía, R., and Laprise, R. (2008). Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain. *Clim. Dyn.* 31, 927–940. doi: 10.1007/s00382-008-0384-8

Madec, G. (2016). *NEMO reference manual 3_6_STABLE: NEMO ocean engine*. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France.

Madsen, K. S., Høyer, J. L., Suursaar, Ü., She, J., and Knudsen, P. (2019). Sea level trends and variability of the baltic sea from 2D statistical reconstruction and altimetry. *Front. Earth Sci.* 7:243. doi: 10.3389/feart.2019.00243

Marcos, M., Jordà, G., Gomis, D., and Pérez, B. (2011). Changes in storm surges in southern Europe from a regional model under climate change scenarios. *Glob. Planet. Change* 77, 116–128. doi: 10.1016/j.gloplacha.2011.04.002

Mathis, M., and Mikolajewicz, U. (2020). The impact of meltwater discharge from the Greenland ice sheet on the Atlantic nutrient supply to the northwest European shelf. *Ocean Sci.* 16, 167–193. doi: 10.5194/os-16-167-2020

Mayer, B., Mathis, M., Mikolajewicz, U., and Pohlmann, T. (2022). RCP8.5-projected changes in German Bight storm surge characteristics from regionalized ensemble simulations for the end of the twenty-first century. *Front. Clim.* 4:992119. doi: 10.3389/fclim.2022.992119

Meier, H., Broman, B., and Kjellström, E. (2004). Simulated sea level in past and future climates of the Baltic Sea. *Clim. Res.* 27, 59–75. doi: 10.3354/cr027059

Meier, H. E. M., Kniebusch, M., Dieterich, C., Gröger, M., Zorita, E., Elmgren, R., et al. (2022). Climate change in the Baltic Sea region: a summary. *Earth Syst. Dynam.* 13, 457–593. doi: 10.5194/esd-13-457-2022

Mentaschi, L., Besio, G., Cassola, F., and Mazzino, A. (2013). Problems in RMSE-based wave model validations. *Ocean Model. (Oxf.)* 72, 53–58. doi: 10.1016/j.ocemod.2013.08.003

Merkouriadi, I., and Leppäranta, M. (2014). Long-term analysis of hydrography and sea-ice data in Tvärminne, Gulf of Finland, Baltic Sea. *Clim. Change* 124, 849–859. doi: 10.1007/s10584-014-1130-3

Morim, J., Hemer, M., Wang, X. L., Cartwright, N., Trenham, C., Semedo, A., et al. (2019). Robustness and uncertainties in global multivariate wind-wave climate projections. *Nat. Clim. Chang.* 9, 711–718. doi: 10.1038/s41558-019-0542-5

Muis, S., Aerts, J. C. J. H., Á. Antolínez, J. A., Dullaart, J. C., Duong, T. M., Erikson, L., et al. (2023). Global projections of storm surges using high-resolution CMIP6 climate models. *Earth's Future* 11:e2023EF003479. doi: 10.1029/2023EF003479

Muis, S., Apecechea, M. I., Dullaart, J., De Lima Rego, J., Madsen, K. S., Su, J., et al. (2020). A high-resolution global dataset of extreme sea levels, tides, and storm surges, including future projections. *Front. Mar. Sci.* 7:263. doi: 10.3389/fmars.2020.00263

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., and Von Storch, J. -S. (2012). Global M_2 internal tide and its seasonal variability from high resolution ocean circulation and tide modeling. *Geophys. Res. Lett.* 39:2012GL053320. doi: 10.1029/2012GL053320

Neumann, B., Vafeidis, A. T., Zimmermann, J., and Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding - a global assessment. *PLoS One* 10:e0118571. doi: 10.1371/journal.pone.0118571

Nguyen, T. T., Staneva, J., Bonaduce, A., Jacob, B., and Pein, J. (2025). Seamless integration of the land–ocean continuum: the complex interplay of wave-induced processes and estuarine influences. *Ocean Dyn.* 75:58. doi: 10.1007/s10236-025-01705-3

Nguyen, T. T., Staneva, J., Grayek, S., Bonaduce, A., Hagemann, S., Pham, N. T., et al. (2024). Impacts of extreme river discharge on coastal dynamics and environment: insights from high-resolution modeling in the German Bight. *Reg. Stud. Mar. Sci.* 73:103476. doi: 10.1016/j.rsma.2024.103476

Oliver, E. C. J., Benthuysen, J. A., Darmaraki, S., Donat, M. G., Hobday, A. J., Holbrook, N. J., et al. (2021). Marine heatwaves. *Annu. Rev. Mar. Sci.* 13, 313–342. doi: 10.1146/annurev-marine-032720-095144

Oliver, E. C. J., Burrows, M. T., Donat, M. G., Sen Gupta, A., Alexander, L. V., Perkins-Kirkpatrick, S. E., et al. (2019). Projected marine heatwaves in the 21st century and the potential for ecological impact. *Front. Mar. Sci.* 6:734. doi: 10.3389/fmars.2019.00734

Oliver, E. C. J., Donat, M. G., Burrows, M. T., Moore, P. J., Smale, D. A., Alexander, L. V., et al. (2018). Longer and more frequent marine heatwaves over the past century. *Nat. Commun.* 9:1324. doi: 10.1038/s41467-018-03732-9

Payne, M. R., Barange, M., Cheung, W. W. L., MacKenzie, B. R., Batchelder, H. P., Cormon, X., et al. (2016). Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES J. Mar. Sci.* 73, 1272–1282. doi: 10.1093/icesjms/fsv231

Pein, J., Staneva, J., Mayer, B., Palmer, M. D., and Schrum, C. (2023). A framework for estuarine future sea-level scenarios: response of the industrialised Elbe estuary to projected mean sea level rise and internal variability. *Front. Mar. Sci.* 10:1102485. doi: 10.3389/fmars.2023.1102485

Pham, N. T., Gramcianinov, B. C., Ricker, M., and Staneva, J. (2026). Wave climate projections in the Western Black Sea using regional EUROCORDEX wind forcing data. *Ocean Eng.* 349:124126. doi: 10.1016/j.oceaneng.2025.124126

Pinto, J. G., Ulbrich, U., Leckebusch, G. C., Spangehl, T., Reyers, M., and Zacharias, S. (2007). Changes in storm track and cyclone activity in three SRES ensemble experiments with the ECHAM5/MPI-OM1 GCM. *Clim. Dyn.* 29, 195–210. doi: 10.1007/s00382-007-0230-4

Pinto, J. G., Zacharias, S., Fink, A. H., Leckebusch, G. C., and Ulbrich, U. (2009). Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Clim. Dyn.* 32, 711–737. doi: 10.1007/s00382-008-0396-4

Priestley, M. D. K., and Catto, J. L. (2022). Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure. *Weather Clim. Dynam.* 3, 337–360. doi: 10.5194/wcd-3-337-2022

Rasmussen, D. J., Bittermann, K., Buchanan, M. K., Kulp, S., Strauss, B. H., Kopp, R. E., et al. (2018). Extreme sea level implications of 1.5 °C, 2.0 °C, and 2.5 °C temperature stabilization targets in the 21st and 22nd centuries. *Environ. Res. Lett.* 13:034040. doi: 10.1088/1748-9326/aaac87

Riahi, K., Schaeffer, R., Calvin, K., Guivarch, C., Hasegawa, T., Jiang, K., et al. (2022). "Mitigation pathways compatible with long-term goals. In IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," in *Climate Change 2022 - Mitigation of Climate Change* (Cambridge/New York, NY: Cambridge University Press), 295–408. doi: 10.1017/9781009157926.005

Rutgersson, A., Kjellström, E., Haapala, J., Stendel, M., Danilovich, I., Drews, M., et al. (2022). Natural hazards and extreme events in the Baltic Sea region. *Earth Syst. Dyn.* 13, 251–301. doi: 10.5194/esd-13-251-2022

Schade, N. H., Jensen, C., and Kruschke, T. (2025). Large scale atmospheric conditions favoring storm surges in the North and Baltic Seas and possible future changes. *Front. Environ. Sci.* 13:1601836. doi: 10.3389/fenvs.2025.1601836

Scheer, B. T. (1986). The significance of differences between means. An empirical study. *Comp. Biochem. Physiol. A Comp. Physiol.* 83, 405–408. doi: 10.1016/0300-9629(86)90123-4

Schrum, C., Lowe, J., Meier, H. E. M., Grabemann, I., Holt, J., Mathis, M., et al. (2016). "Projected change—North Sea," in *North Sea Region Climate Change Assessment*, eds. M. Quante, and F. Colijn (Cham: Springer International Publishing), 175–217. doi: 10.1007/978-3-319-39745-0_6

Sein, D. V., Danilov, S., Biastoch, A., Durgadoo, J. V., Sidorenko, D., Harig, S., et al. (2016). Designing variable ocean model resolution based on the observed ocean variability. *J. Adv. Model. Earth Syst.* 8, 904–916. doi: 10.1002/2016MS000650

Sen Gupta, A., Thomsen, M., Benthuysen, J. A., Hobday, A. J., Oliver, E., Alexander, L. V., et al. (2020). Drivers and impacts of the most extreme marine heatwave events. *Sci. Rep.* 10:19359. doi: 10.1038/s41598-020-75445-3

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. D., et al. (2021). "Weather and climate extreme events in a changing climate. In climate change 2021: the physical science basis," in *Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge/New York, NY: Cambridge University Press), 1513–1766.

Serafin, K. A., and Ruggiero, P. (2014). Simulating extreme total water levels using a time-dependent, extreme value approach. *JGR Oceans* 119, 6305–6329. doi: 10.1002/2014JC010093

Slangen, A. B. A., Carson, M., Katsman, C. A., Van De Wal, R. S. W., Köhl, A., Vermeersen, L. L. A., et al. (2014). Projecting twenty-first century regional sea-level changes. *Clim. Change* 124, 317–332. doi: 10.1007/s10584-014-1080-9

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M., Harvey, B. P., Straub, S. C., et al. (2019). Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nat. Clim. Chang.* 9, 306–312. doi: 10.1038/s41558-019-0412-1

Soomere, T., and Pindsoo, K. (2016). Spatial variability in the trends in extreme storm surges and weekly-scale high water levels in the eastern Baltic Sea. *Cont. Shelf Res.* 115, 53–64. doi: 10.1016/j.csr.2015.12.016

Staneva, J., Alari, V., Breivik, Ø., Bidlot, J.-R., and Mogensen, K. (2017). Effects of wave-induced forcing on a circulation model of the North Sea. *Ocean Dyn.* 67, 81–101. doi: 10.1007/s10236-016-1009-0

Sterl, A., Bintanja, R., Brodeau, L., Gleeson, E., Koenigk, T., Schmith, T., et al. (2012). A look at the ocean in the EC-Earth climate model. *Clim. Dyn.* 39, 2631–2657. doi: 10.1007/s00382-011-1239-2

Tebaldi, C., Ranasinghe, R., Vousdoukas, M., Rasmussen, D. J., Vega-Westhoff, B., Kirezci, E., et al. (2021). Extreme sea levels at different global warming levels. *Nat. Clim. Chang.* 11, 746–751. doi: 10.1038/s41558-021-01127-1

Timmerman, H. (1975). On the importance of atmospheric pressure gradients for the generation of external surges in the North Sea. *Dtsch. Hydrogr. Z.* 28, 62–71. doi: 10.1007/BF02232250

Travkin, V. S., Tikhonova, N. A., and Zakharchuk, E. A. (2024). Characteristics of marine heatwaves of the Baltic Sea for 1993–2022 and their driving factors. *Pure Appl. Geophys.* 181, 2373–2387. doi: 10.1007/s00024-024-03504-1

Ulbrich, U., Leckebusch, G. C., and Pinto, J. G. (2009). Extra-tropical cyclones in the present and future climate: a review. *Theor. Appl. Climatol.* 96, 117–131. doi: 10.1007/s00704-008-0083-8

UTide (2014). Available online at: https://github.com/wesleybowman/UTide, https://github.com/wesleybowman/UTide (Accessed February 26, 2025).

Valcke, S. (2013). The OASIS3 coupler: a European climate modelling community software. *Geosci. Model Dev.* 6, 373–388. doi: 10.5194/gmd-6-373-2013

Varlas, G., Katsafados, P., Papadopoulos, A., and Korres, G. (2018). Implementation of a two-way coupled atmosphere-ocean wave modeling system for assessing air-sea interaction over the Mediterranean Sea. *Atmos. Res.* 208, 201–217. doi: 10.1016/j.atmosres.2017.08.019

Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., and Storlazzi, C. D. (2017). Doubling of coastal flooding frequency within decades due to sea-level rise. *Sci. Rep.* 7:1399. doi: 10.1038/s41598-017-01362-7

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Bianchi, A., Dottori, F., and Feyen, L. (2018a). Climatic and socioeconomic controls of future coastal flood risk in Europe. *Nat. Clim. Change* 8, 776–780. doi: 10.1038/s41558-018-0260-4

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., et al. (2018b). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nat. Commun.* 9:2360. doi: 10.1038/s41467-018-04692-w

Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., and Feyen, L. (2016). Projections of extreme storm surge levels along Europe. *Clim. Dyn.* 47, 3171–3190. doi: 10.1007/s00382-016-3019-5

Wahl, T. (2017). Sea-level rise and storm surges, relationship status: complicated! *Environ. Res. Lett.* 12:111001. doi: 10.1088/1748-9326/aa8eba

Wahle, K., Staneva, J., Koch, W., Fenoglio-Marc, L., Ho-Hagemann, H. T. M., and Stanev, E. V. (2017). An atmosphereandndash;wave regional coupled model: improving predictions of wave heights and surface winds in the southern North Sea. *Ocean Sci.* 13, 289–301. doi: 10.5194/os-13-289-2017

Warner, J. C., Armstrong, B., He, R., and Zambon, J. B. (2010). Development of a coupled ocean–atmosphere–wave–sediment transport (COAWST) modeling system. *Ocean Model. (Oxf.)* 35, 230–244. doi: 10.1016/j.ocemod.2010.07.010

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J. (2014). The inter-sectoral impact model intercomparison project (ISI–MIP): project framework. *Proc. Nat. Acad. Sci.* 111, 3228–3232. doi: 10.1073/pnas.1312330110

Weisse, R., and Plüβ, A. (2006). Storm-related sea level variations along the North Sea coast as simulated by a high-resolution model 1958–2002. *Ocean Dyn.* 56, 16–25. doi: 10.1007/s10236-005-0037-y

Wernberg, T., Bennett, S., Babcock, R. C., De Bettignies, T., Cure, K., Depczynski, M., et al. (2016). Climate-driven regime shift of a temperate marine ecosystem. *Science* 353, 169–172. doi: 10.1126/science.aad8745

Wiese, A., Staneva, J., Ho-Hagemann, H. T. M., Grayek, S., Koch, W., and Schrum, C. (2020). Internal model variability of ensemble simulations with a regional coupled wave-atmosphere model GCOAST. *Front. Mar. Sci.* 7:596843. doi: 10.3389/fmars.2020.596843

Wilks, D. S. (2016). "The stippling shows statistically significant grid points": how research results are routinely overstated and overinterpreted, and what to do about it. *Bull. Am. Meteorol. Soc.* 97, 2263–2273. doi: 10.1175/BAMS-D-15-00267.1

Williams-Kerslake, M., Langehaug, H. R., Skogseth, R., Nilsen, F., Samuelsen, A., Gonzalez, S., et al. (2025). *Characterising Marine Heatwaves in the Svalbard Archipelago and Surrounding Seas*. *EGUsphere* doi: 10.5194/egusphere-egu25-2165

Willmott, C. J., Robeson, S. M., and Matsuura, K. (2012). A refined index of model performance. *Intl. J. Climatol.* 32, 2088–2094. doi: 10.1002/joc.2419

Wiśniewski, B., and Wolski, T. (2011). Physical aspects of extreme storm surges and falls on the Polish coast. *Oceanologia* 53, 373–390. doi: 10.5697/oc.53-1-TI.373

Wolf, J. (2009). Coastal flooding: impacts of coupled wave–surge–tide models. *Nat. Hazards* 49, 241–260. doi: 10.1007/s11069-008-9316-5

Wolski, T., and Wiśniewski, B. (2021). Characteristics and long-term variability of occurrences of storm surges in the Baltic Sea. *Atmosphere* 12:1679. doi: 10.3390/atmos12121679

Wolski, T., Wiśniewski, B., Giza, A., Kowalewska-Kalkowska, H., Boman, H., Grabbi-Kaiv, S., et al. (2014). Extreme sea levels at selected stations on the Baltic Sea coast. *Oceanologia* 56, 259–290. doi: 10.5697/oc.56-2.259

Woodworth, P. L., Flather, R. A., Williams, J. A., Wakelin, S. L., and Jevrejeva, S. (2007). The dependence of UK extreme sea levels and storm surges on the North Atlantic Oscillation. *Cont. Shelf Res.* 27, 935–946. doi: 10.1016/j.csr.2006.12.007

Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M., and Brayshaw, D. J. (2012). Response of the North Atlantic storm track to climate change shaped by ocean–atmosphere coupling. *Nat. Geosci.* 5, 313–317. doi: 10.1038/ngeo1438

Wöppelmann, G., and Marcos, M. (2016). Vertical land motion as a key to understanding sea level change and variability. *Rev. Geophys.* 54, 64–92. doi: 10.1002/2015RG000502

Woth, K., Weisse, R., and von Storch, H. (2006). Climate change and North Sea storm surge extremes: an ensemble study of storm surge extremes expected in a changed climate projected by four different regional climate models. *Ocean Dyn.* 56, 3–15. doi: 10.1007/s10236-005-0024-3

Wu, L., Breivik, Ø., and Rutgersson, A. (2019a). Ocean-wave-atmosphere interaction processes in a fully coupled modeling system. *J. Adv. Model. Earth Syst.* 11, 3852–3874. doi: 10.1029/2019MS001761

Wu, L., Staneva, J., Breivik, Ø., Rutgersson, A., Nurser, A. J. G., Clementi, E., et al. (2019b). Wave effects on coastal upwelling and water level. *Ocean Model. (Oxf.)* 140:101405. doi: 10.1016/j.ocemod.2019.101405

Zappa, G., Shaffrey, L. C., Hodges, K. I., Sansom, P. G., and Stephenson, D. B. (2013). A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the cmip5 climate models*. *J. Clim.* 26, 5846–5862. doi: 10.1175/JCLI-D-12-00573.1

Zhou, M., and Wang, S. (2024). The risk of concurrent heatwaves and extreme sea levels along the global coastline is increasing. *Commun. Earth Environ.* 5:144. doi: 10.1038/s43247-024-01274-1