Department of Civil, Chemical and Environmental Engineering

Curriculum: Risk and Resilience Engineering for the Natural, Industrialized and Built Environment.

Storm surge hazard in the Mediterranean: a consistent hindcast-projections framework

Supervisors:
Prof. Giovanni Besio
Prof. Andrea Lira Loarca

Author:
Filippo Giaroli

PhD Cycle:
XXXVIII

Academic Year
2025-2026

Abstract

Coastal regions worldwide are increasingly exposed to natural hazards, several of which are projected to intensify under global warming, amplifying risks to coastal communities, ecosystems, and critical infrastructure. Storm surges are a key driver of extreme sea levels and, when combined with other contributions to total water level, can trigger damaging coastal flooding. This PhD research develops a basin-wide, high-resolution storm surge modeling framework for the Mediterranean Sea, with the dual aim of producing a long, homogeneous hindcast suitable for climate-scale and extreme-value analyses, and generating an internally consistent ensemble of future projections to investigate potential climate-driven changes in surge hazard.

First, a high spatial and temporal resolution storm surge model is implemented using Delft3D-FLOW over the entire Mediterranean domain. The model produces a hindcast spanning 1979–2023 that is forced by hourly high-resolution atmospheric fields, including 10-m wind and mean sea level pressure, enabling the representation of both synoptic and mesoscale meteorological features controlling surge generation in a semi-enclosed basin. Model performance is evaluated against surge residuals derived from tide gauge observations using standard statistical indicators and seasonally resolved diagnostics. Over the full hindcast period, the model reproduces observed variability with a mean Pearson correlation of approximately 0.6, with systematically higher skill during the storm-active autumn-winter months and reduced skill in summer, when surge magnitudes are small and local-scale processes become comparatively more influential. Model results are benchmarked against existing Mediterranean storm surge datasets from previous studies, showing broadly comparable skill. The representation of peak events relevant for coastal hazard is further examined. This is assessed through biases in extreme values between model results and observations, together with timing offsets and specific event-detection metrics, which demonstrate a satisfactory ability to reproduce the occurrence and key characteristics of surge extremes.

Next, the influence of Atlantic open-boundary conditions and associated barometric forcing is also examined. The analysis demonstrates the crucial role of accounting for the Inverse Barometer Effect in improving the representation of Mediterranean sea level variability, with clear gains relative to configurations neglecting this contribution. The hindcast is then used to characterize basin-scale extremes through spatial diagnostics of maxima and return levels, and a focused case study of the October 2018 Ligurian storm illustrates the model's capability to reproduce the timing, spatial footprint, and intensity of this high-impact surge event.

Then, building on the validated historical reconstruction, the same modeling framework is applied to generate storm surge projections up to the end of the XXI century under a selected climate change scenario. A 17-member ensemble is produced using distinct EURO-CORDEX forcing combinations derived from Global and Regional Climate Models, covering a historical reference period (1970–2005) and a future period (2030–2100). Finally, to enable coherent comparisons with the hindcast and across the ensemble, projections are bias-adjusted using a dedicated methodology and analyzed in terms of changes in surge statistics and extremes, including return levels associated with frequent and rare events.

Ensemble results indicate spatially heterogeneous and generally modest future changes in Mediterranean storm surge extremes, with no basin-wide uniform trend emerging across models. At the basin scale, ensemble-aggregated return level changes typically remain within ± 5 cm for the selected return periods, whereas individual model combinations can yield locally larger increases or decreases, highlighting substantial inter-model spread. Inter-comparison further suggests that the projected spread is more strongly driven by the choice of the driving Global Climate Model than by the Regional Model, with simulations sharing the same global model often exhibiting more similar spatial patterns.

Overall, the datasets and methods developed in this thesis provide a foundation for basin-scale assessments of present and future storm surge hazard in the Mediterranean, supporting both scientific investigations and coastal risk management applications.

Contents

List of figures

List of Tables

List of acronyms

AR ..... Assessment Report

ATM_3 ..... Atmospheric forcing at 3 km resolution

ATM_10 ..... Atmospheric forcing at 10 km resolution

Cd ..... Drag coefficient

CDF ..... Cumulative Distribution Function

CMIP ..... Coupled Model Intercomparison Project

Corr ..... Pearson Correlation Coefficient

ETS ..... Equitable Threat Score

EVA ..... Extreme Value Analysis

EVT ..... Extreme Value Theory

FB ..... Frequency Bias

FAR ..... False Alarm Ratio

GCM ..... Global Climate Model

GEV ..... Generalized Extreme Value (distribution)

HR ..... Hit Rate

IBE ..... Inverse Barometer Effect

IPCC ..... Intergovernmental Panel on Climate Change

NRMSE ..... Normalized Root Mean Square Error

RCM ..... Regional Climate Model

RCP ..... Representative Concentration Pathway

RES_3 ..... Hydrodynamic grid resolution of 3 km

RES_10 ..... Hydrodynamic grid resolution of 10 km

RL ..... Return Level

SLP ..... Sea Level Pressure

SSP ..... Shared Socioeconomic Pathway

TR ..... Return Period

UTC ..... Coordinated Universal Time

1 Introduction

1.1 Background

Coastal regions are among the most densely populated and economically strategic areas worldwide (Oppenheimer et al., 2019; Small and Nicholls, 2003). They are not only characterized by high population densities, but also function as essential hubs for socioeconomic activities, maritime and terrestrial trade networks, key infrastructure systems, and major tourist destinations (Hilmi et al., 2022; Azzopardi et al., 2020). Coastal areas also host a wide range of historical and cultural heritage sites, further enhancing their societal significance (Mejjad et al., 2022).

Given their central role in supporting contemporary economic and social processes, coastal zones require comprehensive protection measures to mitigate environmental hazards (Lira-Loarca et al., 2022). Among these, processes driven by marine storms represent one of the most critical threats to both coastal communi-

ties and ecosystems (Nicholls et al., 2007).

Among the various storm-related hazards, coastal flooding stands out as one of the most severe consequences of intense marine storms, particularly in low-lying coastal areas (Wood et al., 2023). Such events are recognized as some of the most destructive environmental hazards affecting densely populated coastal regions, including those of the Mediterranean basin.

Coastal flooding can inflict extensive damage on built infrastructure, trigger severe erosion processes, and lead to substantial economic losses, degradation of ecosystems, and, albeit rarely, loss of human lives. An increasing number of coastal cities across the Mediterranean already provide clear evidence of the threats associated with coastal flooding. Venice, in Italy, is widely recognized as one of the most exposed urban areas in the region due to the combined effects of sea level rise, which enhances storm-induced marine flooding, and ongoing land subsidence (Vecchio et al., 2024). Similarly, the city of Alexandria in Egypt is experiencing accelerated coastal erosion and the structural deterioration of buildings as a consequence of more frequent marine inundation and saltwater intrusion (Ismail, 2018).

However, this phenomenon is not limited to the Mediterranean region. Major cities worldwide demonstrate how widespread and severe coastal flooding has become. Jakarta, in Indonesia, faces dramatic flooding risks driven by the combined effects of rapid sea level rise and severe land subsidence (Tay et al., 2022), while Miami, in the United States, is considered one of the most vulnerable western cities to storm-driven coastal inundation (Pena et al., 2023). These examples illustrate that coastal flooding represents a global challenge for densely populated coastal zones, threatening cultural heritage, critical infrastructure, local economies, and communities across a wide variety of geographic and environmental settings.

Accurate sea level prediction is therefore essential for the preservation and sustainable management of coastal cities worldwide. While several operational forecasting

systems currently provide short-term predictions on timescales of days, investigating sea level dynamics over longer temporal horizons is equally crucial. Such long-term assessments require the use of climate projections produced by numerical models, which can support the prioritization of adaptation strategies by identifying coastal regions where flooding is expected to become more frequent under future climate change.

1.1.1 Extreme sea levels and storm surges

Coastal flooding is primarily driven by extreme sea level events, which temporarily raise water levels above the normal tidal range. These events result from the combined effects of multiple oceanographic and atmospheric processes, including astronomical tides, storm surges, and wave-induced effects such as wave setup and run-up, which contribute to the total water level above the mean sea level.

Among the drivers of extreme sea levels, meteotsunamis deserve particular mention as a well-documented phenomenon in specific Mediterranean sub-basins. Meteotsunamis are long ocean waves with periods ranging from minutes to hours, generated by rapid atmospheric pressure disturbances, such as those associated with thunderstorms or atmospheric gravity waves, that transfer energy to the water column, creating propagating barotropic waves similar to seismic tsunamis in their coastal amplification but meteorological in origin. These waves undergo resonant amplification in semi-enclosed bays and shallow shelves, leading to sudden sea level oscillations of the order of meters observed in sub-basins like the Adriatic Sea, the Balearic Islands, and the Sicily Channel. Despite their localized nature, meteotsunamis pose significant flooding risks to Mediterranean coasts, with documented impacts including harbor seiches, vessel damage, and infrastructure inundation (Pérez Gómez et al., 2022; Pérez-Gómez et al., 2016; Šepić et al., 2015).

Although similar in coastal flooding impacts, a different atmospheric driver is

the storm surge, caused by strong winds and pressure gradients associated with intense atmospheric systems (Figure 1.1), which plays a central role in amplifying sea level anomalies and triggering inundation, particularly when they coincide with high tides and pronounced wave-induced effects.

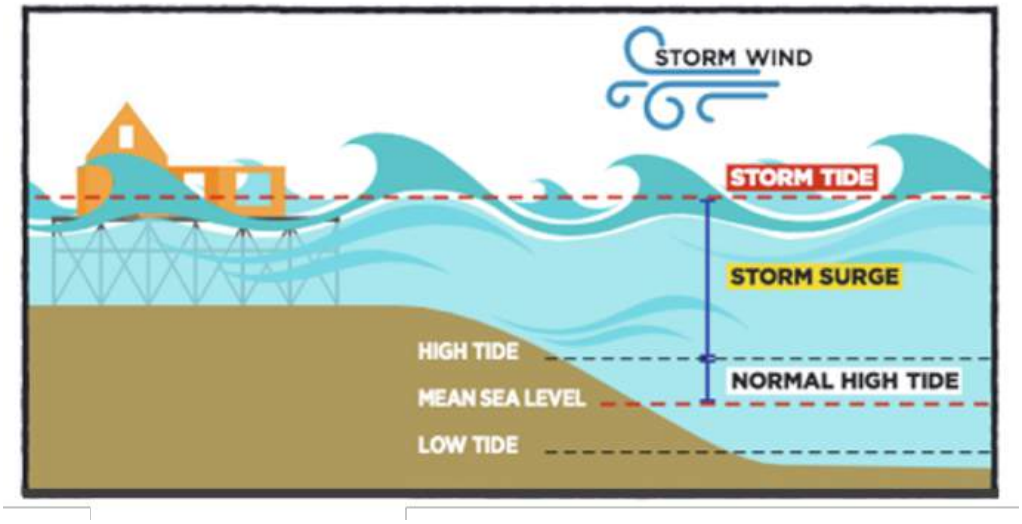

Figure 1.1: Storm surge represents the anomalous sea level setup driven by strong winds and the atmospheric pressure gradient. The surge is superimposed on the normal tidal range and the mean sea level to produce the total storm tide.

This PhD dissertation therefore focuses on the phenomenon of storm surges, considered a key component of extreme sea levels, with the aim of modeling them across the Mediterranean region and assessing their characteristics and impacts.

Unlike long-term, climate-driven sea level changes, such as thermal expansion of seawater or the melting of ice sheets, storm surges develop over short temporal and spatial scales, with substantial variations in intensity occurring over relatively limited distances (Figure 1.2). The magnitude and impacts of extreme sea level events are determined not only by atmospheric and oceanic forcing (e.g., wind intensity, pressure anomalies, and tidal phase) but also by local geographical and morphological characteristics, including coastline geometry and bathymetry. Shallow-water environments and semi-enclosed basins tend to favor surge amplification, whereas

open or steep coastlines generally experience more limited flooding. Such spatial heterogeneity arises from the combined influence of coastal morphology, seabed slope, ocean circulation, and localized atmospheric conditions (Ji et al., 2021; Haigh et al., 2016), which may either enhance or attenuate surge levels. Considering local features is therefore crucial for realistic storm surge simulations allowing an accurate hazard assessment and effective mitigation strategies.

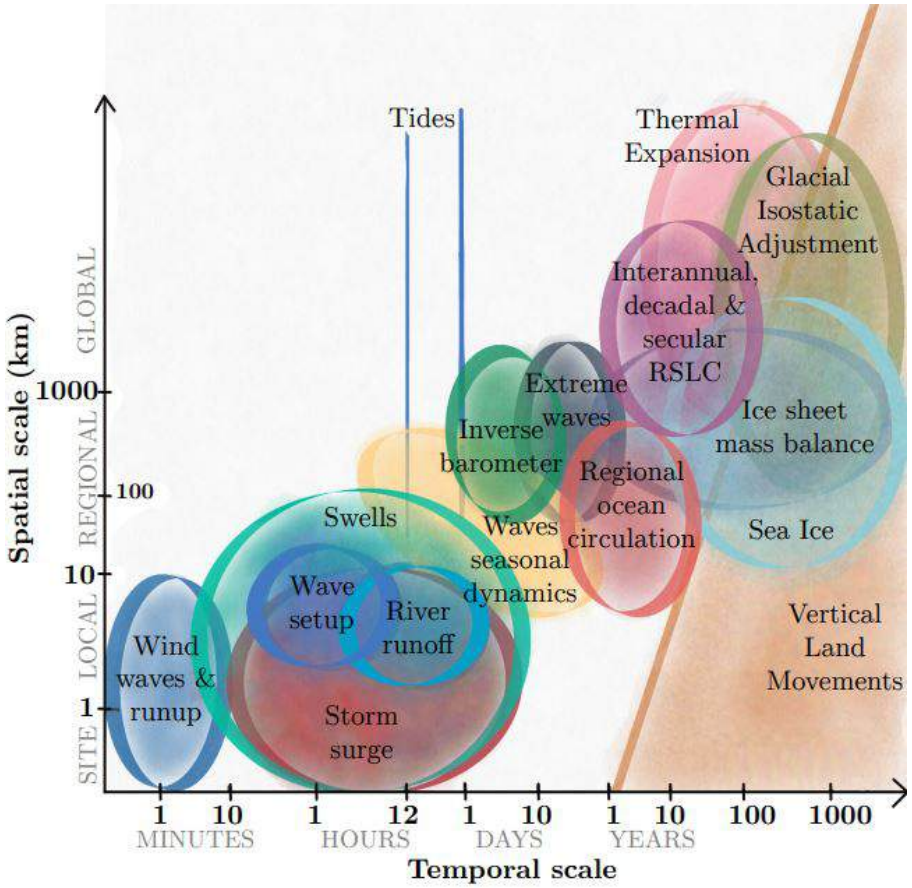

Figure 1.2: Temporal and spatial scales characterizing different natural processes.

1.1.2 Storm surge modeling

Understanding and accurately modeling storm surges is essential for managing coastal risks in regions exposed to extreme sea level events. As these events can

lead to severe flooding, erosion, and damage to infrastructure, many coastal areas must design and implement effective protective measures. To support such efforts, analyses of the frequency, intensity, and probability of extreme sea level events are periodically carried out to identify the most vulnerable zones and quantify their exposure. Various protective strategies have been implemented, ranging from the construction of engineered coastal defenses to the development of predictive models for extreme sea levels and related hazards (Bajo et al., 2007). These analyses are crucial for prioritizing the most vulnerable locations and determining the most effective mitigation measures.

However, a major limitation in characterizing surge events along the Mediterranean coast is the scarcity of observational data, particularly the limited number of tide gauges in the southern and eastern parts of the basin. The sparse spatial coverage and the restricted availability of high-frequency, quality-controlled measurements hinder the accurate representation of storm surge level variability and extremes across the region. This data gap constrains comprehensive risk assessments and limits the development of reliable models essential for effective coastal management and mitigation strategies.

Numerical models play a crucial role in compensating for the lack of observational data, as they can simulate surge levels across an entire region with consistent spatial and temporal coverage. By integrating available measurements and atmospheric forcing, these models help to fill data gaps and improve the understanding and prediction of extreme sea levels and coastal flood risks in areas where tide gauge observations are sparse or unavailable.

Hindcast datasets, which reproduce past events, have a wide range of applications, especially in climate science. They are important tools for analyzing extreme weather phenomena (e.g., hurricanes, extreme rainfall, floods), studying long-term climate and environmental trends, and supporting engineering tasks such as the

design of resilient infrastructure (e.g., coastal defenses, retention basins, hydraulic infrastructures; Smith et al. 2021; Donatini et al. 2015). Because they often extend over several decades and are produced using consistent model configurations, hindcasts provide homogeneous and continuous datasets that are particularly valuable for statistical analyses, return period estimation, and the detection of changes in extreme event frequency or intensity. For these reasons, hindcast datasets are also widely used to estimate the probability of occurrence of rare but high-impact events (Reistad et al., 2011). While hindcast studies help assess current and past risks, anticipating how these hazards will evolve under future climate scenarios is a growing scientific and societal challenge. Particular attention is paid today in identifying if some trends of change related to the global warming exist to understand if extreme events are expected to become more or less frequent in the future in the different coastal zones. Global warming is expected to alter atmospheric circulation, increase the frequency and intensity of extreme events in many regions, and accelerate mean sea level rise, leading to enhanced coastal risk in the coming decades (Bouwer and Bakker, 2019; Oppenheimer et al., 2019). To address this, the scientific community relies on climate projections, which simulate future climate conditions under different scenarios. These projections are periodically revised to reflect the latest scientific evidence, observed trends, and emission trajectories. As a result, the associated forecasts of coastal hazards, including storm surges and extreme sea levels, must also be updated regularly to ensure that coastal planning and infrastructure design remain adequate under changing climatic conditions.

1.2 Research gaps

Although hydrodynamic models for marine variables are widely studied worldwide, high-resolution storm surge models ($\Delta x < 5$ km) typically cover less than 5% of the

Mediterranean basin, such as the Adriatic Sea, or focus on specific coastal stretches (Bajo et al., 2019) or localized events spanning only days to weeks (Ferrarin et al., 2023; Fortelli et al., 2021; Amores et al., 2020; Cavaleri et al., 2020).

As storm surge behavior varies significantly across the Mediterranean, previous studies have typically focused on individual sites of particular scientific or practical interest, most notably the northern Adriatic, including the Venice Lagoon, due to its well-known frequent inundation episodes (Corvaro et al., 2025; Toledo et al., 2024; Yu et al., 1998). Larger-scale studies have been conducted but usually with coarser spatial resolutions of 50-70 km offshore, refining to 4-10 km nearshore (Mentaschi et al., 2023; Fernández-Montblanc et al., 2020). Moreover, accurate detection and modeling of storm surges require high-resolution meteorological data. Coarse datasets smooth out critical local variations in wind patterns and pressure systems, leading to underestimation of surge magnitude, timing, and impacts: daily forcing causes up to 50% underestimation of extreme surges (vs hourly), while coarser spatial resolutions (2° vs 0.25°) add biases up to 0.25 m (Agulles et al., 2024). Comprehensive spatial coverage at sufficient spatio-temporal resolution is therefore essential to capture the complex interactions between atmospheric forces, coastal features, and marine-terrestrial ecosystems that drive storm surge dynamics.

Characterizing coastal sea level variability requires long time series (Toomey et al., 2022), as only data collected over extended periods can distinguish long-term trends from seasonal or inter-annual fluctuations. However, the availability of such long records is often limited to a sparse network of tide gauges, leaving many coastal segments insufficiently monitored.

To develop a comprehensive regional understanding of which coastal zones are most affected by storm surges, it is therefore necessary to rely on basin-scale models rather than localized approaches. At the same time, both the hydrodynamic model and the atmospheric forcing must have sufficiently high resolution to pro-

vide realistic results. Such a configuration would enable the production of spatially continuous storm surge information, allowing for a basin-wide assessment and characterization of storm surge hazards, overcoming the limitations imposed by the uneven distribution of tide gauge observations.

While addressing current knowledge gaps is crucial, anticipating future changes is equally important. Extending the same basin-wide modeling framework to climate projections allows identification of coastal areas that are currently not at risk but may become vulnerable in the coming decades. However, single climate projections are subject to substantial intrinsic uncertainty, as their outcomes depend on the entire modeling chain, from the global or regional climate model used, to the downscaling method, boundary conditions, and forcing datasets applied. For this reason, ensemble approaches are generally preferred over individual projections, as they allow a more comprehensive assessment of the uncertainty associated with climate projections. Multi-model ensembles, commonly used in climate research, can increase confidence in projected changes by sampling the structural uncertainty across different climate models. At the same time, structural differences among models may introduce inconsistencies that can complicate the interpretation of results.

In this context, the approach adopted in this thesis aims to specifically investigate the uncertainty associated with climate forcing. To this end, a large ensemble is generated using the same hydrodynamic model configuration while varying the atmospheric forcing derived from different climate model outputs. This strategy ensures methodological consistency in the hydrodynamic modeling framework and enables a systematic assessment of how differences in climate forcing propagate into storm surge projections.

Finally, future projections must be anchored to a physically consistent reference. A validated hindcast dataset, based on historical observations and consistent model

configurations, provides such a baseline. By comparing future projections to a well-characterized historical reconstruction, changes in storm surge behavior can be robustly quantified. Developing both the hindcast and the ensemble of future projections within the same modeling framework ensures internal consistency and provides a solid basis for a comprehensive assessment of storm surge hazards in the Mediterranean under changing climate conditions.

1.3 Research objectives

This dissertation aims to address the research gaps identified in the previous section and advance the understanding and prediction of storm surge hazards across the Mediterranean basin by developing a comprehensive, high-resolution modeling framework that integrates both past hindcast reanalysis and future projections.

The primary objective of this research is to develop a storm surge model capable of simulating sea level variations with high spatial and temporal resolution. Using this model, the study focuses on the following specific objectives:

i) **Hindcast simulation and validation** – developing a basin-wide storm surge dataset for the Mediterranean, forced by reanalysis atmospheric data and evaluating model performance against tide gauge measurements and analyzing specific case studies of extreme events to ensure that the model reliably captures both typical and high-impact storm surges.

ii) **Future projections using an ensemble of climate model forcing** – generating multiple simulations up to the end of the 21st century using combinations of Global and Regional Climate Models (GCMs–RCMs) to assess variability and robustness while maintaining methodological consistency.

iii) **Trend analysis over future time horizons** – comparing hindcast and

projected datasets to identify changes in storm surge intensity and frequency under a selected climate change scenario.

The overarching goal of this research is to provide a robust, high-resolution storm surge modeling framework for the entire Mediterranean basin, capable of supporting both historical analyses and projections under future climate conditions.

1.4 Significance and Contribution

This research provides several important contributions to the understanding and assessment of storm surges in the Mediterranean Sea. The main novelty of the work lies in the combination of long-term simulations and high-resolution atmospheric forcing, which allows for the modeling of storm surges driven by both regional and local low-pressure systems, with particular emphasis on reproducing extreme events.

The scientific significance can be highlighted in three key aspects. First, regarding the hindcast analysis, the integration of high-resolution atmospheric forcing ($\sim$10 km) with a fine computational mesh ($\sim$3 km) over the entire Mediterranean domain provides detailed insights into storm surge dynamics at the regional scale, producing a comprehensive dataset capable of informing basin-scale analyses. Notably, the model has been developed using Delft3D-FLOW (Deltares, 2014), a widely used hydrodynamic modeling software, which, to date, has rarely been applied at the scale of the entire Mediterranean Sea. This represents an important step forward in applying high-resolution hydrodynamic modeling at a basin-wide level.

Second, the development of a large ensemble of future projections offers a comprehensive view of potential changes in coastal risk. The resulting dataset constitutes a valuable resource for both scientific investigations and engineering applications, allowing for robust assessments of the variability and uncertainty of future storm

surge events.

Finally, by coupling storm surge projections with wind-wave data, this research produces an integrated dataset that enables joint analyses of the primary marine drivers of coastal flooding. This is made possible by the availability of an existing dataset of wave projection developed by the MeteOcean Research Group of the University of Genoa, which was generated using the same atmospheric forcings adopted in the present study. This approach facilitates the identification of worst-case scenarios, the assessment of regional vulnerability to compound events, and the prioritization of adaptation and protection strategies in the most at risk areas.

1.5 Dissertation structure

The thesis is organized into seven chapters, each addressing a specific aspect of the research.

Chapter 1 – Introduction: provides the general background and motivation of the study, outlines the main scientific and societal challenges related to coastal flooding and sea level extremes, and clarifies the concept and characteristics of storm surges. This chapter also presents the research objectives, highlights the gaps addressed, and emphasizes the strengths and novelty of the carried out research and the possible application of the obtained output.

Chapter 2 – State of the Art: summarizes the existing literature on storm surge modeling and projected climate change impacts in the Mediterranean region.

Chapter 3 – Material and Methods: describes the high-resolution numerical model developed for storm surge simulations, including the computational grid, model parameters, atmospheric forcing, and boundary conditions. Section 3.2 details the model validation methodology.

Chapter 4 – Results - Hindcast: presents the output of the hindcast simulation,

evaluating historical storm surge events and analyzing spatial and temporal variability across the basin. It includes an extreme event characterization, associating surge magnitudes with return levels. Section 4.3 presents two case studies that occurred in the Ligurian Sea in 1955 and 2018.

Chapter 5 – Results - Projections: presents the output of future storm surge simulations and provides a comparative analysis of past and projected surge levels, examining potential trends induced by climate change.

Chapter 6 - Discussion: it is dedicated to the discussion of the results. It synthesizes the main outcomes of the analyses, interprets their physical meaning, compares them with previous studies, and examines their implications for coastal hazard assessment and future research directions.

Chapter 7 – Conclusion: Summarizes the main findings of the research, discusses its limitations, and outlines potential future developments and applications of the results.

2

State of the Art

2.1 Literature review on storm surges in the Mediterranean

A large body of literature exists on storm surge modeling worldwide, including numerous contributions focusing on the Mediterranean region and employing a variety of numerical approaches. Within the Mediterranean basin, most studies have concentrated on specific sub-regions or local areas where storm surge dynamics are particularly pronounced due to local geomorphological, meteorological, and hydrodynamic characteristics.

One of the most extensively investigated areas is the Adriatic Sea. Owing to its semi-enclosed configuration, its shallow bathymetry, and the long fetch associated with southeasterly Scirocco winds, the basin provides favorable conditions for the

development of intense storm surge events. Consequently, a significant proportion of Mediterranean storm surge research is dedicated to this region (Corvaro et al., 2025; Bajo et al., 2019; Yu et al., 1998), with even more focused studies examining the Venice Lagoon and the city of Venice (Umgiesser et al., 2021; Lionello et al., 2012; Bajo et al., 2007), where storm surges have historically caused severe flooding.

These studies share a common focus on storm surge processes but differ markedly in their methodological approaches, and none of them aim at constructing a long-term reference dataset for subsequent analyses. For instance, Bajo et al. (2019) explores the interaction between storm surges and seiches, standing waves occurring in enclosed or semi-enclosed basins such as lakes, bays, and channels, within the Adriatic Sea. Using an unstructured grid implementation of the SHYFEM shallow-water hydrodynamic model (Umgiesser et al., 2004), the authors demonstrate that many extreme surge events are significantly influenced by pre-existing basin seiches. In contrast, Yu et al. (1998) adopts a more predictive perspective, validating a storm surge forecasting model against a set of historical extreme events in the Adriatic. More recently, Corvaro et al. (2025) investigates the relationship between storm surges and wind waves, analyzing their dependence structure through copula-based statistical models.

Focusing on the Venetian area, Lionello et al. (2012) simulate storm surges using the Hydrostatic Padua Sea Elevation (HYPSE) model over a 30-year hindcast period, restricted to the months from October to May, which are assumed to encompass the majority of extreme surge events. HYPSE is implemented over the entire Mediterranean Sea on a variable-resolution squared grid, with a minimum spacing of $0.05°\times0.04°$ (lon–lat) in the Northern Adriatic and progressively coarsening in both directions up to a maximum of $0.5°\times0.3°$ at the domain boundaries.

The review by Umgiesser et al. (2021) provides an overview of operational models of particular relevance for assessing extreme storm surge events in the context of the

commissioning of the MOSE system in Venice. An example of such an operational framework is the high-resolution SHYFEM-based model developed by Bajo et al. (2007), designed to provide early assessments of potentially hazardous water levels both inside the Venice Lagoon and along adjacent coastal zones.

Other Mediterranean areas considered sensitive from both an ecological and a flood-risk perspective have also been the subject of localized studies. Examples include river deltas, such as the Po Delta in Italy (Favaretto et al., 2025) and the Ebro Delta in Spain (Romero-Martín et al., 2025). The former employs the 2DEF hydrodynamic model (Defina, 2003), based on an unstructured triangular mesh extending offshore down to a bathymetric depth of 20 m. The study simulates specific 12-hour extreme event scenarios to assess their environmental impact on the delta. The latter highlights storm surges as an important driver of flooding in low-lying, sheltered coastal areas, such as deltas and wetlands, in the western Mediterranean, with a focus on the Ebro Delta. However, it does not perform storm surge modeling directly; instead, it relies on data from the Global Ocean Surge (GOS) database (Cid et al., 2014), which provides hourly mean sea level time series from 1950 to 2014 at a horizontal resolution of 1/8° (approximately 14 km). Additional research focusing on the Spanish Mediterranean coast has been carried out by Toledo et al. (2024), who analyses storm surges at locations of the coast where observational measurements are available. Their study highlights that the Mediterranean Sea exhibits greater sensitivity to atmospheric pressure drops than the Atlantic Ocean and emphasizes the importance of event duration and regional geographic factors in governing the correlation between pressure anomalies and surge levels.

Outside these well studied hot-spots, large sections of the Mediterranean coastline remain comparatively under-represented in storm surge research. Exceptions include studies along the Egyptian coast (Torresan et al., 2020) and in the Gulf

of Gabes, in Tunisia (Rizzi et al., 2016), both of which primarily focus on coastal risk assessment through dedicated vulnerability indexes, relying on storm surge estimates from existing datasets rather than numerical modeling. Another exception is the work of Krestenitis et al. (2011), one of the few contributions addressing storm surges in the southern and southeastern sectors of the basin. Their study employs an updated version of the AUT storm surge model (de Vries et al., 1995), implemented over the entire Mediterranean on a $1/10° \times 1/10°$ grid, but with particular emphasis on the Aegean region.

Several other studies have concentrated on reconstructing and modeling specific extreme storm surge events of particular relevance in the Mediterranean, thereby providing valuable insights into the mechanisms driving high impact conditions. Androulidakis et al. (2023) and Ferrarin et al. (2023) both examine Medicane Ianos, which occurred in September 2020 in the Ionian Sea. Ianos was one of the most intense Mediterranean tropical-like cyclones on record, displaying characteristics comparable to a "Category 2" hurricane. The storm generated significant sea level anomalies along its trajectory, producing marked storm surges in the central Ionian Sea and causing substantial impacts along the coasts of the Ionian Islands and western Greece (Causio et al., 2025). In Androulidakis et al. (2023), storm surge simulations are performed using the high resolution HiReSS numerical model on a $0.15°\times0.15°$ grid covering the entire Mediterranean, with a two-month simulation period (August–September 2020) and outputs at 3-hour intervals. On 18^{th} September, the highest surge values of approximately 30 cm are reported near the town of Katakolo, on the western coast of the Peloponnese. Conversely, Ferrarin et al. (2023) employ the SHYFEM model over the full Mediterranean domain using an unstructured grid with spatial resolution ranging from about 12 km in the open sea to roughly 2 km along the cyclone track. Their results confirm similarly elevated surge levels in the Katakolo area, in agreement with the findings of the former

study.

Amores et al. (2020) investigate the effects of Storm Gloria, which occurred in January 2020 in the northwestern Mediterranean, severely impacting the Balearic Islands and the Spanish coastline and causing extensive flooding, particularly in the Ebro Delta. The coupled storm surge and wind wave fields generated by Storm Gloria are simulated using the SCHISM model (Semi-implicit Cross-scale Hydroscience Integrated System Model; Zhang et al. 2016). The model is implemented on an unstructured triangular grid covering the entire western Mediterranean basin, with a horizontal resolution ranging from approximately 15 km in the open ocean to 1–2 km along most coastlines, and reaching 30 m in the Ebro Delta. The simulation spans nine days, from 17$^{\text{th}}$ to 26$^{\text{th}}$ January 2020, and reproduces storm surges exceeding 50 cm along the Spanish mainland between 20$^{\text{th}}$ and 23$^{\text{rd}}$ January.

Cavaleri et al. (2020) once again draw attention to the Venice area, describing the extreme event that occurred in November 2019 and resulted in severe flooding across the city. During this episode, the storm surge component reached approximately 70 cm. However, the most critical factor was its coincidence with the astronomical high tide. This concurrence underscores the fact that storm surge and tidal elevation are distinct contributions which must be combined when assessing coastal hazard.

Complementary to these predominantly local studies, the literature also includes numerous contributions focusing on storm surge simulations for the whole European and Mediterranean regions. Several authors have developed hindcast datasets with varying spatial resolutions and atmospheric forcings, providing broader regional perspectives on sea level variability and the dynamics of extreme surge events (Agulles et al., 2024; Mentaschi et al., 2023; Toomey et al., 2022; Fernández-Montblanc et al., 2020; Cid et al., 2014).

Most of these studies, and particularly those by Agulles et al. (2024), Mentaschi et al. (2023), Toomey et al. (2022), and Fernández-Montblanc et al. (2020), employ

the SCHISM circulation model on unstructured grids to simulate storm surges across the Mediterranean, using different model configurations and atmospheric forcing resolutions.

Agulles et al. (2024) implement the model on a grid with a resolution ranging from approximately 25 km in the open ocean to about 5 km along the coast, also extending over the northeastern Atlantic Ocean. The study aims to assess the influence of the spatial and temporal resolution of atmospheric inputs on storm surge simulations. To this end, the model is forced with atmospheric fields of different temporal (hourly and daily) and spatial (from 0.25° to 2°) resolutions from the ERA5 reanalysis dataset (Hersbach et al., 2020), spanning 1940–2021. The results show that simulations forced with daily atmospheric data substantially underestimate the magnitude of extreme coastal sea level events by up to 50% compared to hourly-forced simulations, although low resolution simulations still capture the general temporal variability of storm surges. The correlation with tide gauge observations across the Mediterranean is approximately 0.6 for the simulation using the highest spatio-temporal resolution.

A comparable geographical domain is investigated by Fernández-Montblanc et al. (2020), who cover the entire European coastline. Their work is implemented on an unstructured grid with spatial resolution ranging from 70 km offshore to 10 km near the coast. The model is forced by the ERA-Interim reanalysis (Dee et al., 2011) from the European Centre for Medium-Range Weather Forecasts (ECMWF), with a horizontal resolution of 0.75°×0.75° and 6-hourly temporal sampling. The resulting hindcast dataset spans 40 years, from 1979 to 2018, providing outputs at 3-hour intervals. The validation shows correlation values ranging between 0.4 and 0.8.

The contribution of Mentaschi et al. (2023) is not limited to the Mediterranean and European region but extends to the global coastline. In this study, storm surges

are simulated globally, coupled with wind waves, on an unstructured grid with a resolution ranging from approximately 50 km offshore to 2–4 km nearshore. The simulations span the period 1950–2022 and are forced by hourly ERA5 reanalysis atmospheric fields at 0.25° spatial resolution. Comparison with tide gauge observations indicates a global correlation of approximately 0.55. The same atmospheric forcing is also used in the work of Toomey et al. (2022), who model storm surges over almost the same period (1950–2021) for the entire Mediterranean Sea at very high spatial resolution. The unstructured grid ranges from roughly 20 km in the open ocean down to about 200 m along the coast, providing storm surge values at hourly temporal sampling. This hindcast achieves a mean correlation of 0.64 when validated against observations from 93 tide gauge stations.

While SCHISM is a widely used hydrodynamic model, other Mediterranean-scale studies have been developed using different numerical modeling tools. For instance, Cid et al. (2014) developed an hindcast dataset named GOS using the Regional Ocean Model System (ROMS; Shchepetkin and McWilliams 2005). The model domain encloses Southern Europe, including the Mediterranean Sea and the Atlantic coast, with a horizontal resolution of 1/8° (about 14 km). It is composed by two sub-datasets: the first spanning 62 years (1948-2009), while the second 21 years (1989–2009). They are driven with two different regional atmospheric forcings: SeaWind I (30 km of horizontal resolution) and SeaWind II (15 km of horizontal resolution). Both are the result of a dynamical down-scaling from global atmospheric reanalysis: NCEP global reanalysis (Toth et al., 1997) and ERA-Interim global reanalysis, respectively. Also in this study, the correlation between modeled and observed surge levels is consistent with the values reported in the previous works.

From this brief review of existing studies carried out at various spatial scales within the Mediterranean region, it becomes evident that the literature addresses multiple aspects of the storm surge phenomenon, with particular emphasis on its

modeling through a wide variety of numerical tools.However, within the literature reviewed here, Delft3D appears to be rarely used as the core modelling system for dedicated Mediterranean storm-surge hindcast and reanalysis developments. Moreover, the atmospheric forcings adopted in previous studies typically rely on datasets with coarser spatial resolution, 0.25° being the finest, compared to the higher resolution products that will be presented in the context of this research. This is particularly relevant for the reconstruction of past conditions and for the generation of the hindcast dataset that will serve as a reference to assess future variations driven by climate change.

2.2 Climate change scenarios

Climate change today represents one of the major environmental, social, economic, and politic challenges at the global level. It refers to significant and long-term variations in the Earth's climate system, observable over decadal or centennial timescales, resulting from both natural factors and anthropogenic activities. Among the latter, the rise in concentrations of greenhouse gases such as carbon dioxide (CO_2) and methane (CH_4) alters the planet's energy balance, causing an increase in global mean temperature, changes in precipitation patterns, sea level rise, diffuse drought and desertification, and a variation in the frequency and the intensity of extreme climatic events. These changes have significant impacts on both natural ecosystems and human communities, requiring a multidisciplinary approach to understand their dynamics and develop effective mitigation and adaptation strategies.

In this context, the Intergovernmental Panel on Climate Change (IPCC), established in 1988 by the United Nations Environment Program (UNEP) and the World Meteorological Organization (WMO), plays a central role in synthesizing and assessing scientific knowledge on climate. The IPCC does not carry original research

on but analyzes and integrates international scientific results, producing reports intended to inform global climate policies. The main tool through which the IPCC disseminates its assessments is the so-called Assessment Report (AR), periodic documents summarizing the state of climate science, observed and projected impacts, and potential mitigation and adaptation strategies.

The Assessment Reports are developed within the framework of the Coupled Model Intercomparison Project (CMIP), an international initiative that coordinates climate-model simulations across research institutes worldwide. CMIP provides standardized simulations of past, present, and future climate under different greenhouse gas and aerosol emission scenarios, enabling systematic comparison and evaluation of model performance and projections.

Since the first Assessment Report in 1990, six assessment cycles have been published, the most recent being the Sixth Assessment Report (AR6) released in 2021–2022. Each report represents a critical update of scientific knowledge, with increasing attention to observational evidence, climate modeling, and future scenario projections. Notably AR5, developed in the context of the CMIP5 and published in 2013-2014, introduced the Representative Concentration Pathways (RCPs) (on Climate Change , IPCC), scenarios based on future greenhouse gas concentration levels in atmosphere, designed to explore how different mitigation policies could influence global climate change (Figure 2.1).

The four main RCPs (RCP2.6, RCP4.5, RCP6.0, and RCP8.5) describe low, intermediate, medium–high, and high greenhouse gas emission pathways, providing projections of temperature, precipitation, and sea level rise through the end of the 21^{st} century. These scenarios are named after their associated radiative forcing levels in 2100 relative to preindustrial conditions, corresponding to +2.6, +4.5, +6.0, and +8.5 W/m^2, respectively.

AR6, developed in the context of CMIP6, represents a significant evolution in

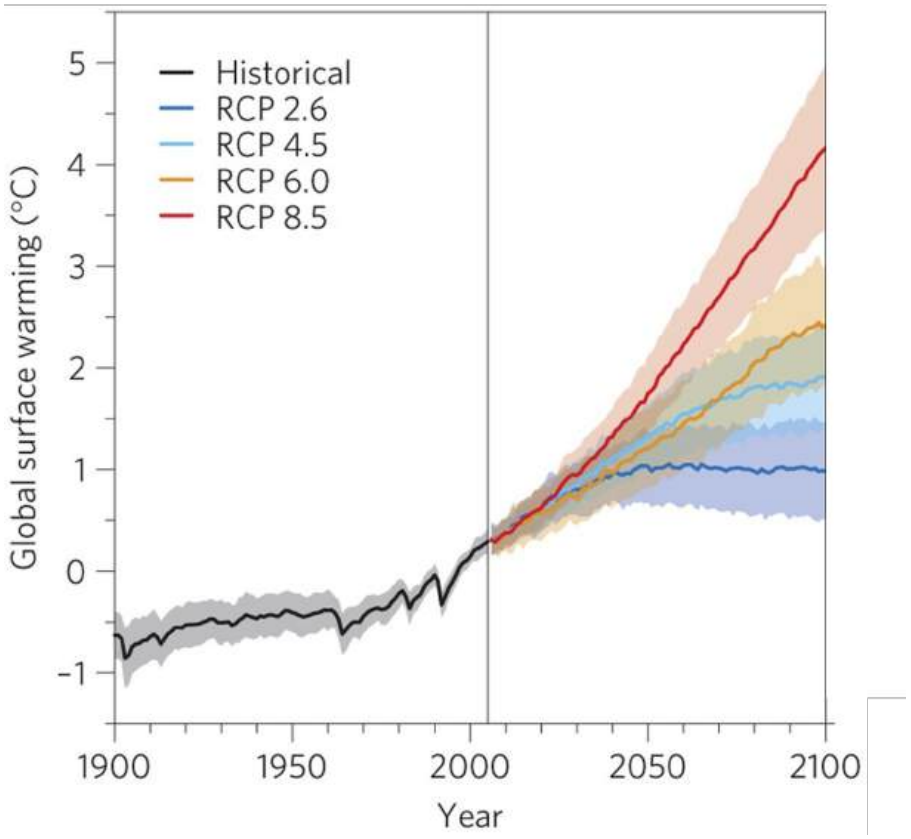

Figure 2.1: Projected global mean surface temperature change under CMIP5 RCPs scenarios (AR5).

climate scenario modeling through the introduction of Shared Socioeconomic Pathways (SSPs) (Intergovernmental Panel on Climate Change (IPCC), 2022). SSPs (Figure 2.2) combine emission trajectories with detailed assumptions about global socioeconomic development, integrating variables such as economic growth, demographic change, urbanization, technological innovation, land-use patterns, and environmental policies (Tebaldi et al., 2021).

The combination of SSPs and RCPs allows for the creation of more realistic and diversified scenarios, reflecting not only potential levels of greenhouse gas emissions but also the influence of human decisions on climate dynamics. This integrated approach enhances the ability to estimate climate risks and to plan mitigation and adaptation strategies at local, national, and international scales.

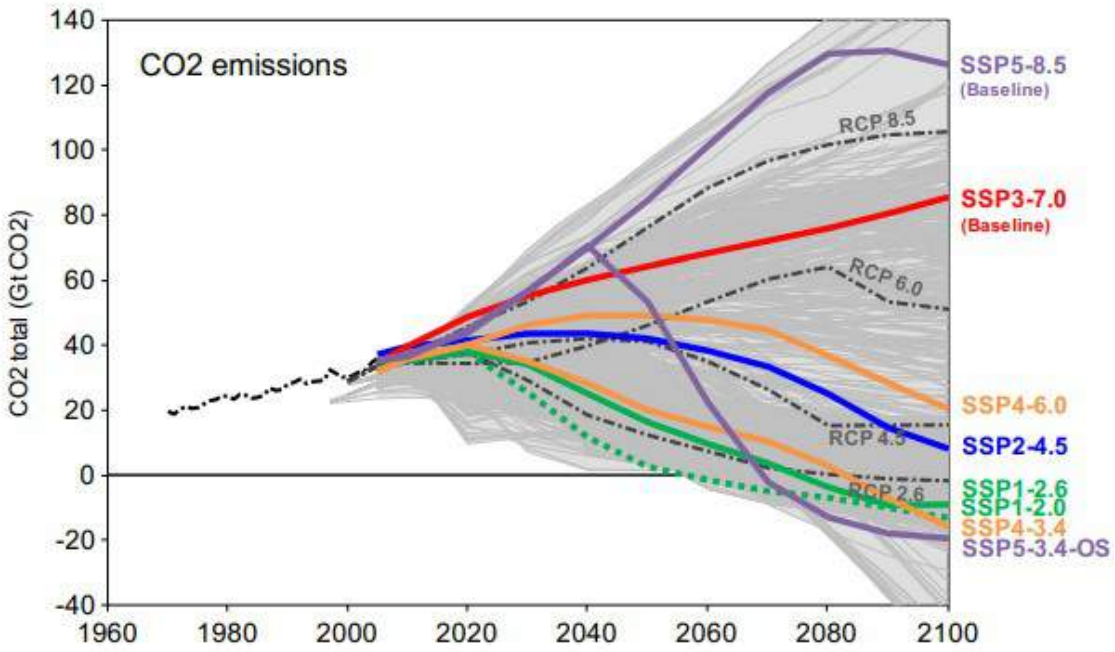

Figure 2.2: Greenhouse gas emission trajectories under different combined RCPs-SSPs (AR6) up to 2100.

Projections based on RCPs and SSPs show significant differences depending on the chosen pathway. High-emission scenarios (RCP8.5, SSP5-8.5) predict global temperature increases exceeding 4°C by the end of the century, along with intensified extreme events such as heatwaves, droughts, and heavy precipitation. In contrast, low-emission scenarios and ambitious mitigation policies (RCP2.6, SSP1-1.9) suggest that warming could be limited to 1.5–2°C, resulting in correspondingly smaller impacts on environmental systems.

According to IPCC regional assessments, by 2050 the mean sea level in the Mediterranean basin is likely to rise by approximately 0.15–0.33 meters relative to the period 1995–2014, depending on the scenario. By 2100, under low-emission and strong mitigation pathways (e.g., SSP1-1.9, RCP2.6), the projected rise is about 0.3–0.6 meters, while under high-emission scenarios (e.g., SSP5-8.5, RCP8.5), it

may reach 0.6–1.1 meters, driven by thermal expansion, glacier melt, and ice-sheet loss.

In this context, the storm surge component that is modeled in this thesis must be added to the projected mean sea level rise, together with tidal and wave-induced components, to estimate the total potential coastal water levels during extreme events.

2.2.1 Choice of the climate change scenario

Although the latest climate change scenarios are provided by CMIP6 and the IPCC Sixth Assessment Report (AR6), storm surge modeling in the Mediterranean requires high resolution atmospheric forcing to accurately capture local and coastal dynamics (Agulles et al., 2024). Currently, such high resolution atmospheric forcing datasets derived from AR6 scenarios are not yet fully accessible for regional hydrodynamic modeling.

For this reason, the projections used in the present work are based on CMIP5 and the Fifth Assessment Report (AR5), specifically referring to the RCP8.5 scenario. At the time of AR5's publication, RCP8.5 was considered the "business-as-usual" pathway, representing the most likely trajectory of greenhouse gas emissions in the absence of effective mitigation policies. This scenario therefore provides an upper bound for potential future changes in extreme sea level events and is particularly useful for assessing the maximum hazard to Mediterranean coastal zones under extreme climate conditions.

Sea level rise and storm surge projections rely on long-term simulations of the atmospheric system using Global Climate Models (GCMs). GCM outputs provide large-scale boundary conditions, but their coarse spatial resolution (typically 100–250 km) is insufficient to resolve the complex meteorological and hydrodynamic processes in semi-enclosed basins such as the Mediterranean Sea. To overcome this

limitation, down-scaled datasets from Regional Climate Models (RCMs), such as those from the EURO-CORDEX initiative (Jacob et al., 2014, 2020), are employed. These datasets achieve horizontal resolutions of approximately 10-15 km, allowing for a more accurate representation of atmospheric forcing at regional and coastal scales.

In this context, Mediterranean storm surge projections are generated by forcing hydrodynamic models with atmospheric fields from EURO-CORDEX simulations based on AR5 RCP8.5. Despite being derived from an older scenario framework, these high resolution datasets remain the most suitable for capturing local storm surge dynamics. The aim of the present study is to exploit these datasets to produce storm surge simulations extending to the end of the XXI century, thereby enabling the assessment of future changes in extreme sea level events and the identification of Mediterranean regions most at risk from coastal flooding.

2.3 Existing surge projections

Since the investigation of the future evolution of environmental variables is widely recognized as a key scientific challenge, several previous studies have focused on simulating storm surges under future climate conditions across a range of temporal horizons and spatial scales. These studies differ in terms of the climate change scenarios adopted, as well as in the spatial and temporal resolution of the atmospheric forcing derived from Global or Regional Climate Models (GCMs or RCMs), depending on the geographical extent and objectives of each investigation.

At the European scale, Vousdoukas et al. (2016) forced the hydrodynamic model Delft3D-FLOW with surface wind and atmospheric pressure fields from an eight-member GCMs ensemble to evaluate changes in storm surge levels along the European coastline throughout the 21st century under the RCP4.5 and RCP8.5 scenarios.

The periods 2010–2040 and 2070–2100 were considered as short- and long-term future reference windows. The results indicate minimal changes or slight decreases in projected storm surge levels along most European coasts. Model outputs were produced at 3-hour intervals and at approximately 25 km spacing along the coastline. Extreme surge levels derived from the ensemble exhibit similar spatial patterns across scenarios and return periods. In the Mediterranean Sea, only minor changes are projected, with the largest increases confined to the western basin, while in the central Mediterranean the frequency of the present-day 100-year event is projected to decrease both by mid-century and towards the end of the century. This research was later extended by Vousdoukas et al. (2017, 2018) to the analysis of extreme sea levels in selected global hot spots. These studies employed a six-member GCMs ensemble with characteristics similar to the previous work, but additionally coupled with wave datasets and mean sea level rise projections. The results consistently demonstrate that mean sea level rise represents the dominant driver of future changes in extreme sea levels, whereas storm surge and wave climate are not expected to undergo substantial long-term variations.

More recently, attention has shifted towards basin-scale assessments specifically targeting the Mediterranean Sea using higher forcing resolution. Makris et al. (2023) investigated the impacts of projected climate change on episodic sea level elevation events along Mediterranean coasts using three RCMs from the Med-CORDEX initiative (Ruti et al., 2016). Simulations were conducted under the RCP4.5 and RCP8.5 emission scenarios and used to force a barotropic storm-surge model (MeCSS) for the period 2071–2100. The authors project an overall attenuation of storminess by the end of the 21st century, while identifying pronounced regional differences in storm surge maxima. Extreme surge values are estimated to range between 0.35 and 0.50 m, with the highest magnitudes occurring along parts of the northern Mediterranean coastline, such as the Venice Lagoon, the Gulf of Lions, the north-

ern Adriatic and the Aegean Seas, as well as in the Gulf of Gabes in the southern basin. Differences between RCP4.5 and RCP8.5 mainly affect the magnitude of extremes rather than their spatial distribution.

Complementary to these regional scale investigations, other studies have focused on more localized coastal settings. For instance, Lin-Ye et al. (2020) analysed storm surge dynamics along the northeastern Spanish coast, specifically the Catalan coastline, over the period 1950–2100 under the high emission RCP8.5 scenario. The study relies on atmospheric forcing from a single RCM and employs the HYPSE hydrodynamic model to generate 3-hourly storm surge time series. The results suggest that, while average storm surge characteristics may remain relatively stable under climate change, the likelihood of more variable and extreme surge events may increase.

Although most storm-surge projection studies to date are based on AR5 Representative Concentration Pathways (RCPs), mainly due to the limited availability of high resolution RCM outputs derived from AR6 scenarios, recent efforts have begun to explore projections based on the newer SSPs framework using GCM-scale atmospheric inputs.

In this context, Irazoqui Apecechea et al. (2025) developed storm-surge projections for the entire Mediterranean region based on a 17-member CMIP6 GCMs ensemble under the SSP5-8.5 scenario, covering the period 2080–2099. Their results indicate negligible changes in 10-year return level extreme sea states by mid-century, while robust changes (defined as agreement in the sign of change across at least 13 out of 17 models) emerge by the end of the century in several regions. For the Mediterranean Sea, a robust negative change of approximately 7% is projected.

Similarly, Muis et al. (2023) produced global storm surge projections up to 2050 using the Global Tide and Surge Model (Muis et al., 2016), forced by a five-member climate model ensemble from the CMIP6 HighResMIP experiment at a spatial res-

olution of approximately 25–50 km, under the SSP5-8.5 scenario. Their findings suggest a general tendency towards reduced storm surge magnitudes in the Mediterranean region.

The results of these previous studies provide an important benchmark for the present work and will be used for comparative evaluation. Nevertheless, it is evident that storm surge projections for the Mediterranean based on large, high resolution ensembles of dynamically down-scaled atmospheric forcing remain scarce. Addressing this gap represents one of the key innovative aspects of the present study.

3

Material and Methods

3.1 Delft3D-FLOW storm surge modeling

Delft3D is an integrated modeling system widely adopted for simulating hydrodynamic, sediment transport, and morphodynamic processes in coastal, estuarine, and fluvial environments. Developed by Deltares, the system comprises multiple modules, each dedicated to specific physical processes. In this study, the Delft3D-FLOW module (Deltares, 2014) is employed to simulate storm surge dynamics across the Mediterranean basin.

Delft3D-FLOW solves the unsteady, depth-averaged (2D) or fully three dimensional Navier-Stokes equations for incompressible fluids under the shallow-water and Boussinesq assumptions. The governing system consists of the horizontal momentum equations and the continuity equation. These equations describe free-surface flow and incorporate the effects of tides, wind forcing, atmospheric pressure gra-

dients, and density variations due to non-uniform temperature and salinity fields. The partial differential equations are discretized using a finite-volume method on a structured, rectangular or curvilinear grid, together with appropriate initial and boundary conditions. This numerical approach enables an accurate representation of the hydrodynamic processes relevant to storm surge dynamics, particularly in complex coastal environments where wind stress and pressure anomalies play a dominant role.

The FLOW module is designed to model flow phenomena for which horizontal spatial and temporal scales are much larger than the vertical scales, as is typically the case in shallow seas, coastal zones, estuaries, lagoons, rivers, and lakes. A two-dimensional mode, based on a single computational layer and corresponding to the depth-averaged formulation, is appropriate when the water column is vertically homogeneous. Typical applications of the 2D depth-averaged equations include tidal propagation, storm surges, tsunamis, harbour oscillations, and transport of pollutants in vertically well-mixed flow regimes.

3.1.1 Domain configuration and numerical set-up

The storm surge model is configured over a domain encompassing the entire Mediterranean Sea and a portion of the adjacent Atlantic Ocean, extending slightly west of the Strait of Gibraltar. This configuration enables the model to capture wind-driven sea level gradients between the Atlantic Ocean and the Mediterranean Sea through the Strait (Menemenlis et al., 2007), which is essential for a realistic representation of basin-scale surge dynamics.

A structured computational grid with uniform horizontal resolution is employed, defined in WGS84 geographic coordinates (EPSG:4326). Two spatial resolutions, 10 km and 3 km, were initially tested to assess the sensitivity of the model to grid refinement. Based on comparative analyses, the 3 km grid, comprising more than

308,000 nodes, is selected for both the hindcast simulations and the future projections. This resolution improves the representation of hydrodynamic variability, particularly in complex coastal regions, while maintaining a reasonable compromise between accuracy and computational cost.

The model is depth-averaged (two-dimensional) and uses a fixed computational time step of 30 seconds. Bathymetric data are obtained from the ETOPO2022 dataset (MacFerrin et al., 2025), which has a native spatial resolution of 15 arc-seconds (∼450 m). Elevations in ETOPO2022 are referenced to the EGM2008 geoid (EPSG:3855), meaning that negative values represent bathymetric depths below the geoid, which approximates mean sea level. In this study, the geoid-based vertical reference is consistently adopted as the model vertical datum. The bathymetric data are interpolated onto the computational grid using bilinear interpolation, ensuring a smooth and spatially consistent representation.

Storm surges are recognized as the dominant drivers of extreme sea levels in the Mediterranean Sea, while tide–surge interactions are generally negligible (Marcos et al., 2009). For this reason, tidal forcing is not included in the simulations, allowing the modeling effort to focus exclusively on the meteorologically-driven component of sea level variability.

Delft3D-FLOW allows atmospheric forcing to be specified either as spatially uniform but time-varying fields or as fully space- and time-dependent inputs, depending on the scale and objectives of the application. In this study, wind and atmospheric pressure are provided as spatially and temporally varying datasets defined on a regular meteorological grid and subsequently interpolated onto the computational grid. This approach ensures that the spatial heterogeneity of atmospheric forcing, essential for accurately representing storm surge dynamics, is preserved in the model.

Wind stress at the free surface is calculated using a bulk aerodynamic formulation,

in which the surface shear stress depends on the 10-meters wind speed, air density, and a drag coefficient that varies with wind intensity:

$$\boldsymbol{\tau}_w = \rho_a \, C_D \, |\mathbf{U}_{10}| \, \mathbf{U}_{10},$$

where $\boldsymbol{\tau}_w$ is the wind stress vector, ρ_a is air density, $\mathbf{U}_{10}$ is the 10 m wind velocity vector, and C_D is the dimensionless drag coefficient.

While Delft3D-FLOW provides a default drag coefficient parameterization, it also allows the use of alternative formulations. In the present work, the parameterization proposed by Shankar and Behera (2021) is adopted, as it offers a more realistic description of the drag coefficient across a wide range of wind speeds and performs well under high wind conditions typical of extreme events. The formulation is expressed as:

$$C_d = \left(0.42 + 3.86\tilde{U} - 2.53\tilde{U}^2 + 0.4\tilde{U}^3\right) \cdot 10^{-3}, \quad \text{with} \quad \tilde{U} = \frac{U_{10}}{U_{\text{ref}}} \tag{3.1}$$

where U_{10} is the 10-meters wind speed, and U_{ref} is a reference wind speed experimentally set to 37.5 m/s. This formulation avoids the unrealistic increase of C_D at very high wind speeds that affects other bulk formulas, preventing the overestimation of wind stress during intense storms. Furthermore, this formulation captures nonlinear variations in wind stress more effectively, which is crucial for simulating storm surge events with high fidelity.

3.1.2 Boundary conditions and barometric forcing

At the open ocean boundary, sea level variability is prescribed via the Inverse Barometer Effect (IBE), which relates atmospheric pressure anomalies to changes in sea surface height under the assumption of hydrostatic equilibrium. This approach is especially relevant in the open ocean, where barotropic sea level responses

to pressure forcing are dominant (Ponte, 2006). The IBE is expressed mathematically as:

$$\Delta\eta = -\frac{1}{\rho g}(P - \overline{P}_{ocean}) \tag{3.2}$$

where:

- $\Delta\eta$ is the sea level change (m) relative to mean sea level;
- P is the local atmospheric pressure (Pa);
- $\overline{P}_{ocean}$ is the time-dependent global mean sea level pressure averaged over the oceans;
- ρ is the seawater density;
- g is the gravitational acceleration.

The inclusion of IBE forcing at the Atlantic boundary allows the associated sea level variability to affect not only the Strait of Gibraltar region, but also the broader Mediterranean basin.

While the developed hindcast dataset includes the IBE contribution, for comparison purposes, an earlier simulation was also carried out neglecting this effect. This dual configuration allows for the assessment of the influence of IBE and the main related findings will be discussed and presented in Section 4.1.1.

3.1.3 Forcing data

Since storm surges are fast-evolving and spatially localized phenomena, high resolution atmospheric input, both in space and time, is essential for realistic modeling. Global-scale atmospheric datasets often lack the resolution required for coastal applications, particularly in semi-enclosed basins like the Mediterranean. Therefore,

high resolution atmospheric forcing from climate models specific for the studied region are adopted for both the hindcast dataset and the future projections (Figure 3.1), relying on the highest resolution datasets available, particularly for the climate projections.

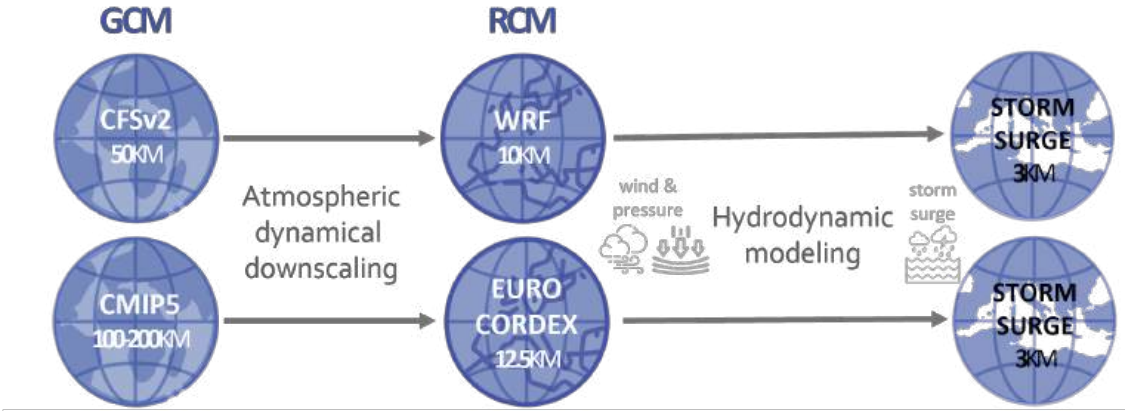

Figure 3.1: Schematics of the modeling chains for the hindcast storm surge dataset (top) and future projections (bottom), illustrating the progressive increase in spatial resolution from the GCMs providing the atmospheric forcing to the regional storm surge simulations.

As previously outlined, storm surge dynamics are governed mainly by strong surface winds and atmospheric pressure gradients. Accordingly, the Delft3D-FLOW model is forced with spatially and temporally varying 10-meter wind fields and mean sea level pressure to reproduce storm surge levels across the Mediterranean basin.

The hindcast simulations are driven by hourly atmospheric input fields with a horizontal resolution of about 10 km derived from simulations performed with the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), specifically tailored for the Mediterranean basin in the studied period (1979–2023) by the MeteOcean Research Group at DICCA, at the University of Genoa (Ferrari et al., 2020; Cassola et al., 2015). The WRF model is dynamically down-scaled from the Climate Forecast System Version 2 (CFSv2), a global climate model with a native resolution of approximately 50 km (Saha et al., 2014). The use of regional modeling enables the accurate reproduction of localized wind systems and mesoscale pressure

patterns, which are essential for simulating storm surge at regional and sub-regional scales.

3.1.4 Storm surge projections

Future storm surge projections are developed under the high-emission RCP8.5 climate change scenario, introduced in the Fifth Assessment Report (AR5) of the IPCC. This scenario represents the emission-based pejorative trajectory and is widely used in impact studies to explore upper-bound changes in future climate extremes.

The projections applied in the context of this research are derived from combinations of Global and Regional Climate Models (GCMs and RCMs, respectively), specifically those belonging to the EURO-CORDEX initiative (Jacob et al., 2014, 2020). EURO-CORDEX provides dynamically down-scaled atmospheric fields over Europe and the Mediterranean region at a spatial resolution of approximately 10-15 km. The increase in resolution, from the 100-150 km typical of GCMs to the much finer RCMs grid, allows a substantially improved representation of regional dynamics relevant for storm surge generation (Figure 3.2). Each EURO-CORDEX simulation provides atmospheric fields every 6 hours up to the year 2100. Differences among projections arise from the driving GCM and the RCM used for down-scaling.

Since the future is inherently uncertain, a multi-model ensemble approach is necessary to assess potential variations. Individual projections, influenced by the specific model chain used, may vary significantly. To account for this, future surge levels projections are developed using the 17 distinct GCM-RCM combinations illustrated in Figure 3.3.

Eight different GCMs and four RCMs are considered in this work and each tick corresponds to a specific GCM-RCM simulation. Ticks on the same row represent projections sharing the same driving GCM but differing in the RCM used for down-

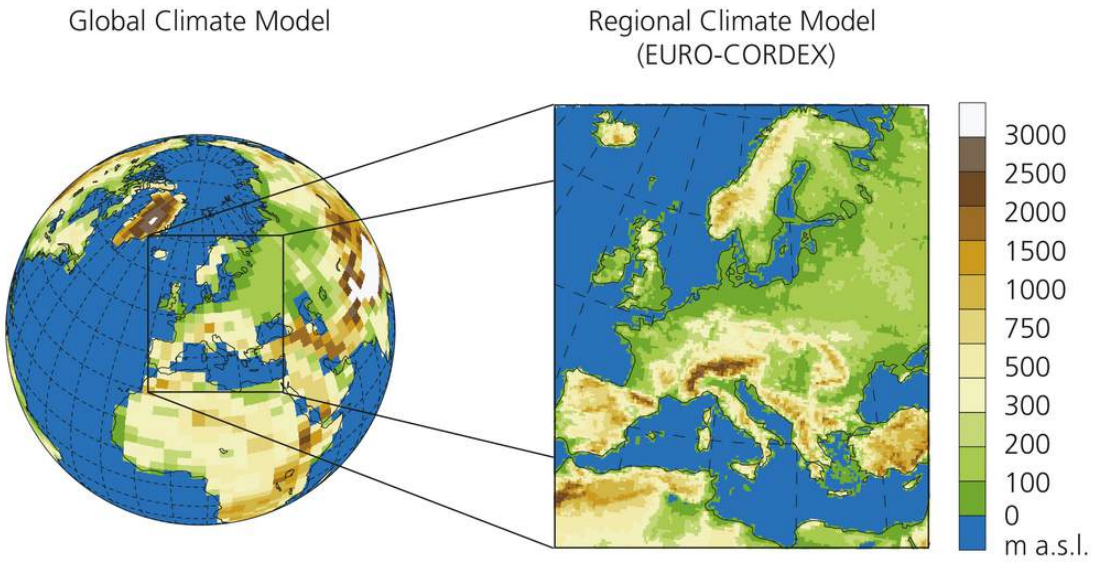

Figure 3.2: Dynamic down-scaling framework adopted by EURO-CORDEX: a Global Climate Model (about 100-150 km) provides the boundary conditions to a Regional Climate Model (about 10-15 km), enabling higher-resolution simulations of atmospheric processes over Europe and the Mediterranean basin.

scaling, while ticks in the same column represent simulations using the same RCM but driven by different GCMs.

Atmospheric projections produced using the regional models CCLM4-8-17 and COSMO are provided by the Climate Limited-area Modeling Community (CLM-com), an international community of researchers who develop and apply regional climate models. Simulations performed with the RCA4 RCM are provided by the Swedish Meteorological and Hydrological Institute (SMHI), whereas those using the HIRHAM5 RCM are developed by the Danish Meteorological Institute (DMI).

A detailed description of each model chain employed in this study is provided in Appendix A, while the differences among the 17 projections will be analyzed and discussed in Chapter 5.

Each projection is examined either individually and jointly to build a large ensemble, enabling a statistically robust assessment of future storm surge conditions.

This ensemble allows for robust comparisons with the hindcast dataset and the

GCM / RCM	CCLM4-8-17	RCA4	HIRHAM5	COSMO
CCCma-CanESM2	✓			
MIROC-MIROC5	✓			
CNRM-CERFACS-CNRM-CM5		✓	✓	
IPSL-IPSL-CM5A-MR		✓	✓	
MOHC-HadGEM2-ES		✓	✓	✓
MPI-M-MPI-ESM-LR		✓	✓	
NCC-NorESM1-M		✓	✓	✓
ICHEC-EC-EARTH		✓	✓	✓

Figure 3.3: Combinations of Global and Regional Climate Models used to develop future storm surge projections. Rows correspond to the driving GCMs, columns to the RCMs, and ticks identify the available GCM-RCM combinations.

computation of ensemble statistics, providing a comprehensive view of possible future trends.

Each projection includes a historical baseline (1970–2005) and a projection period (2006–2100) (Figure 3.4). However, since the running of the simulations requires a certain amount of time, for the future part of each projection, just the period from 2030 to 2100 has been run, in order to ensure coverage of the mid-century (around 2050) and end-century periods, which are the two target time windows specifically analyzed in this work.

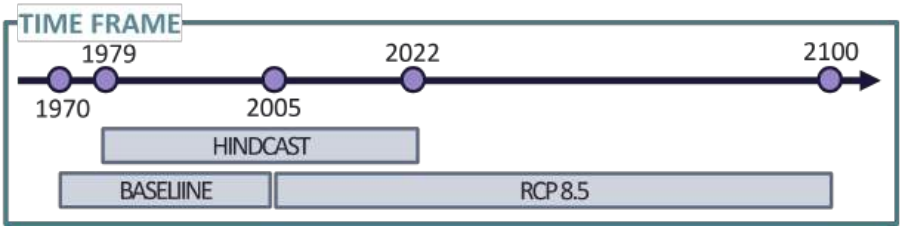

Figure 3.4: Temporal coverage of the hindcast dataset and the historical and future portions of the EURO-CORDEX projections used in this study.

3.2 Model validation

3.2.1 Observed data - the GESLA project

High-quality sea level observations are essential for validating storm surge models and for understanding coastal sea level variability. Among the most comprehensive global datasets available, the Global Extreme Sea Level Analysis (GESLA) project provides a collection of open-source high-frequency (hourly or sub-hourly) tide gauge records from stations worldwide (Haigh et al., 2023). Its main objective is to assemble as many reliable sea level records as possible into a consistent, quality-controlled format, facilitating research on tides, storm surges, and extreme sea levels at global and regional scales.

Since its inception in 2009, GESLA has undergone several major updates. The most recent version, GESLA-4, released in July 2025, contains over 127,000 station-years of data from 6,474 coastal records contributed by more than 40 international and national providers. To ensure consistency, all records are converted to a common format: sea levels are expressed in meters, timestamps are standardized to Coordinated Universal Time (UTC), and provider-specific quality-control (QC) flags are mapped to GESLA's unified QC system. Each data point is therefore associated with a reliability score, allowing users to identify or exclude potentially erroneous or unreliable measurements. Each GESLA file also includes metadata such as station

name, geographic coordinates, provider institution, temporal coverage, and measurement frequency (typically hourly, though higher frequency records exist for several sites). These standardized and quality-controlled observations make GESLA an invaluable dataset for the validation and calibration of storm surge models, enabling comparisons between simulated and observed extreme sea levels and supporting robust coastal hazard assessments.

3.2.2 Selection of tide gauges

The storm surge model output is validated against observational sea level time series extracted from the GESLA dataset. The available Mediterranean tide gauge records are not homogeneous either spatially or temporally: most stations are located in the northwestern Mediterranean, whereas the eastern basin is sparsely sampled. In addition, the observational records vary significantly in length, with several eastern stations providing data only for limited time windows.

Among the available stations, a screening was performed to remove stations located in areas not adequately resolved by the model domain: this is the case for gauges located on small islands that are omitted from the computational grid. Furthermore, only stations with at least two consecutive years of observations were retained for the validation analysis. Figure 3.5 shows the distribution of the approximately seventy tide gauges that are ultimately retained in this study for the validation process, with a higher concentration of stations in the northwestern Mediterranean, from the Strait of Gibraltar to the Italian Peninsula.

For each of the selected stations, the corresponding time series was cleaned by retaining only data points flagged as reliable within the GESLA quality-control system and discarding measurements marked as invalid or doubtful.

While model outputs at grid cells near each tide gauge are available at 5-minute resolution, the analysis conducted in this study is based on hourly values for consis-

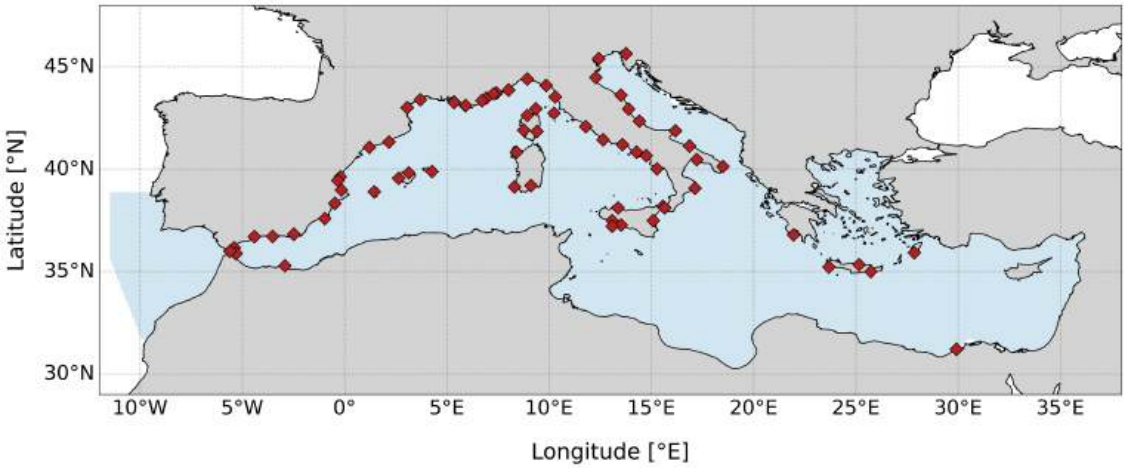

Figure 3.5: Mediterranean storm surge model domain (light blue area) and tide gauge stations (red diamonds) used for model validation. Information of the temporal coverage of the tide gauges is included in Table 1 in Appendix B.

tency with the observations. Nevertheless, the higher-frequency simulations remain valuable for potential local-scale or specialized applications.

3.2.3 Tidal Analysis and Residual Computation

The storm surge model developed in this study simulates only the meteorological contribution to sea level, as described in Section 3.1.1, and therefore does not include the astronomical tidal component. Tide gauge observations, however, record the total sea level, which can be expressed as the sum of two distinct contributions:

$$\eta_{\text{tot}}(t) = \eta_{\text{tide}}(t) + \eta_{\text{res}}(t), \tag{3.3}$$

where η_{tide} represents the deterministic astronomical tide, and η_{res} corresponds to the meteorologically driven residual, i.e., the storm surge. To enable a consistent comparison between modeled storm surge and observations, the tidal signal must be accurately estimated and removed from the tide gauge sea level time series.

Raw tide gauge data may contain missing values, irregular sampling, or spurious

measurements. After merging the available records, the time series were therefore uniformly resampled to an hourly frequency using linear interpolation. This step ensures compatibility with the tidal analysis tools and avoids aliasing effects in the subsequent spectral analysis.

Astronomical tides were reconstructed using four complementary approaches, adopted with different objectives. Three methods were employed to validate and cross-check the tidal signal at selected reference stations, while a fourth method was used for the systematic and operational processing of the full tide gauge network. These approaches together provide a consistent and robust estimate of the tidal component.

- TPXO-derived tidal predictions: a first estimate of η_{tide} was obtained from the global TPXO model (Egbert and Erofeeva, 2002), which supplies harmonic tidal constituents on a global ocean grid. By interpolating these components at the coordinates of a tide gauge, a physically consistent tidal prediction can be obtained. While the model resolution limits the accuracy near complex coastlines, TPXO provides a robust external large-scale tidal reference, being independent of local observations.

- Harmonic analysis: a second reconstruction was obtained by performing a classical harmonic analysis of the observed sea level using the `pytide` Python library (Le Provost et al., 2023). According to standard tidal theory (Pugh et al., 2014), the astronomical tide is represented as the sum of K sinusoidal constituents:

$$\eta_{\text{tide}}(t) = \sum_{k=1}^{K} A_k \cos(\omega_k t + \phi_k)\,, \tag{3.4}$$

  where A_k is the amplitude, ω_k the angular frequency, and ϕ_k the phase of the k-th constituent.

`pytide` automatically computes the astronomical arguments and nodal factors, the harmonic constants (A_k, ϕ_k) via least-squares fitting and a continuous tidal prediction $\eta_{\text{tide}}(t)$ for arbitrary times. Comparison against TPXO showed high correlation for the major constituents (M2, S2, K1, O1), confirming the physical realism and reliability of the harmonic decomposition.

- Spectral validation using Fast Fourier Transform: to further validate the dominant tidal frequencies, a Fast Fourier Transform (FFT) was applied to the observed signal. Let $\hat{\eta}(f)$ be the Fourier transform of $\eta(t)$; the amplitude spectrum

$$A(f) = \frac{2}{N} \left| \hat{\eta}(f) \right| \tag{3.5}$$

was analysed to identify peaks at the expected tidal frequencies (M2, S2, N2, K1, O1, etc.). Each spectral peak was matched to its corresponding astronomical constituent. A synthetic tidal signal,

$$\eta_{\text{tide}}^{\text{FFT}}(t) = \sum_j A_j \cos(2\pi f_j t + \phi_j) \,, \tag{3.6}$$

was then reconstructed using only the dominant constituents identified from the spectrum. This provides an independent validation of the harmonic model.

The three validation approaches (TPXO, harmonic analysis, and FFT-based reconstruction) exhibited strong mutual agreement, particularly for the major tidal constituents. This consistency confirms that the observed sea level signal follows classical tidal dynamics and is not affected by instrumental artefacts within the tidal frequency band. Among these methods, the harmonic prediction obtained with `pytide` showed the best agreement with the observed dynamics and was therefore selected as a reference for method comparison.

The meteorological residual is defined as

$$\eta_{\mathrm{res}}(t) = \eta_{\mathrm{obs}}(t) - \eta_{\mathrm{tide}}(t). \tag{3.7}$$

Before subtraction, the observed sea level was demeaned,

$$\eta'_{\mathrm{obs}}(t) = \eta_{\mathrm{obs}}(t) - \overline{\eta}_{\mathrm{obs}}, \tag{3.8}$$

to avoid projecting long-term or non-tidal mean shifts into the residual. A final quality check comparing `pytide`-based residuals with those obtained using TPXO showed correlations above 0.9, confirming the robustness of the extraction.

While the above procedures were applied to selected stations for validation purposes, the large-scale and multi-year processing of the entire tide gauge network requires a method specifically designed to handle uneven sampling, data gaps, and varying record lengths. For this reason, the UTide toolbox (Codiga, 2025) was adopted for the operational reconstruction and removal of the tidal signal.

UTide implements a refined least-squares harmonic analysis that is robust to short gaps and irregular sampling and is therefore well suited for long-term observational datasets. For each tide gauge station, UTide was applied on a yearly basis to estimate and remove the astronomical tide. The annual (SA) and semiannual (SSA) constituents were excluded from the tidal reconstruction to avoid incorporating large-scale seasonal variability, which may mix with steric and meteorological contributions (Pugh et al., 2014).

A final quality assessment comparing UTide-based residuals with those obtained from TPXO and `pytide` reconstructions yielded correlations exceeding 0.9, confirming that the choice of UTide does not introduce methodological bias but instead ensures robustness and consistency in the residual extraction.

The resulting residual time series therefore represent the high frequency, meteo-

rologically driven sea level variability and can be directly compared with the storm surge model output to assess model performance.

3.2.4 Statistical indexes

The performance of the Delft3D-FLOW model against the available observations is evaluated using three standard complementary statistical indicators: the Pearson correlation coefficient (Corr), the mean bias (Bias), and the Normalized Root Mean Square Error (NRMSE). These metrics quantify different aspects of model skill, namely the ability to reproduce temporal variability, the presence of systematic deviations, and the overall magnitude of the model–data mismatch. Together, they provide a robust and widely adopted framework for validating hydrodynamic models, particularly in coastal and oceanographic applications where both timing and intensity of sea level fluctuations are important.

The considered skill metrics are defined as:

$$Corr = \frac{\sum_{i=1}^{n}(M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(M_i - \overline{M})^2}\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}}, \tag{3.9}$$

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(M_i - O_i), \tag{3.10}$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2}}{\max(M, O) - \min(M, O)}, \tag{3.11}$$

where M denotes the modeled surge values obtained from Delft3D-FLOW and O represents the observed surge levels at the tide gauge stations, with $\overline{M}$ and $\overline{O}$ indicating their respective mean values.

The Pearson correlation coefficient measures the degree to which the model reproduces the observed temporal fluctuations. Its values range from -1 (perfect

anti-correlation) to +1 (perfect correlation), while 0 indicates the absence of a linear relationship. In the context of storm surge validation, higher correlations imply that the model successfully captures the timing and relative magnitude of meteorologically driven sea level variations.

The mean bias quantifies systematic overestimation or underestimation by the model. A positive bias indicates that the model tends to over-predict surge levels, whereas a negative bias reflects a systematic under-prediction. An ideal model would exhibit a bias tending to zero, meaning that no persistent offset exists between simulations and observations.

The NRMSE assesses the overall discrepancy between modeled and observed values, normalized by their range. This normalization allows comparisons across stations with different surge variability and ensures that the metric is dimensionless. Lower NRMSE values indicate better agreement, with zero representing a perfect match. Because it emphasizes larger deviations more strongly than the bias, the NRMSE is particularly informative for evaluating model performance during high-energy events and extreme sea level conditions.

In Section 4.1, the performance of the model, evaluated through these three statistical indices, is presented and discussed.

3.2.5 Peak events detection and skill metrics

In addition to the described continuous statistics, the performance of the model was further evaluated through a dedicated analysis of extreme surge events. These extreme events represent the conditions associated with the greatest coastal hazard and are, in general, the most critical to accurately reproduce for applications such as coastal risk assessment, early-warning systems, and long-term impact studies.

This complementary evaluation focuses on the model's ability to reproduce the timing and magnitude of the most relevant surge episodes. The methodology is

based on a Peak Over Threshold (POT) approach (Coles et al., 2001), combined with a de-clustering procedure and a time-window event matching algorithm, following practices commonly adopted in hydrodynamic and meteorological verification studies (e.g., Sterl and Caires, 2005).

Extremes from both observed and modeled surge levels are extracted using a percentile-based threshold, specifically the 98^{th} percentile of the observed series of each gauge. Denoting the observed surge level by $X(t)$, the threshold is defined as

$$u = \text{quantile}_{98}(X), \tag{3.12}$$

and all peaks satisfying $X(t) > u$ were initially retained.

Because consecutive exceedances may belong to the same physical event, a de-clustering procedure was applied: peaks separated by less than 24 h were considered part of a single event, and only the maximum value within each cluster was retained. Additionally, an event had to exceed the threshold for a minimum duration of 6 h, consistent with typical storm surge timescales in the Mediterranean region, to avoid erroneous outliers. The same procedure was applied to the modeled time series $M(t)$.

Each observed event was then paired with at most one modeled event using a ± 24 h temporal matching window. For each matched pair, two continuous error metrics were computed:

$$\Delta\eta = \eta_{\text{peak}_{\text{mod}}}(t_{\text{peak}_{\text{mod}}}) - \eta_{\text{peak}_{\text{obs}}}(t_{\text{peak}_{\text{obs}}}), \tag{3.13}$$

$$\Delta t = t_{\text{peak}_{\text{mod}}} - t_{\text{peak}_{\text{obs}}}, \tag{3.14}$$

which quantify, respectively, the peak height error and the timing error.

Model skill in reproducing the occurrence of extreme events was further evalu-

ated using standard dichotomous forecast metrics (Wilks, 2011). Each time step is classified as event (YES) or non-event (NO) in both observations and model output, leading to the contingency table reported in Table 3.1.

	Observed YES	**Observed NO**
Modeled YES	Hits (H)	False Alarms (F)
Modeled NO	Misses (M)	Correct Negatives (C)

Table 3.1: Contingency table used for the verification of extreme-event occurrence.

The elements of the contingency table are defined as follows:

- H — number of hits (correctly detected events)
- M — number of misses (observed but not detected events; Type II error, false negative; Wilks, 2011)
- F — number of false alarms (modeled events non actually observed, Type I error, false positive; Wilks, 2011)
- C — number of correct negatives (correctly detected non-events)
- $N = H + M + F + C$ — total number of time samples

Using the contingency counts, the following skill scores are computed to evaluate the model performance in reproducing extreme events:

- Hit Rate (HR)

$$\mathrm{HR} = \frac{H}{H + M}. \tag{3.15}$$

  HR measures the fraction of observed events that are correctly modeled. Values range from 0 (no events detected) to 1 (all events detected). Therefore, higher HR indicates better model detection capability.

- False Alarm Ratio (FAR)

$$\text{FAR} = \frac{F}{H + F}. \tag{3.16}$$

FAR quantifies the proportion of modeled events that did not actually occur. Values range from 0 (no false alarms) to 1 (all modeled extremes were false). Lower FAR indicates better model reliability.

- Frequency Bias (FB) — tendency to over- or under-estimate the peaks occurrence:

$$\text{FB} = \frac{H + F}{H + M}. \tag{3.17}$$

FB compares the total number of modeled events to the total number of observed events. A value of 1 indicates an unbiased model; values > 1 indicate an over-estimate of the peaks occurrence, while values < 1 indicate an under-estimate of the extreme events frequency.

- Equitable Threat Score (ETS) — evaluates the ability of the model to correctly detect extreme events while explicitly accounting for hits that may occur purely by random chance:

$$H_{\text{rand}} = \frac{(H + F)(H + M)}{N}, \tag{3.18}$$

$$\text{ETS} = \frac{H - H_{\text{rand}}}{H + F + M - H_{\text{rand}}}, \tag{3.19}$$

where H_{rand} is the number of hits expected by random chance. ETS ranges from $-\frac{1}{3}$ to 1:

  - ETS $= 1 \rightarrow$ perfect peaks detection
  - ETS $= 0 \rightarrow$ no skill beyond random chance
  - ETS $< 0 \rightarrow$ worse than random

The application of these metrics to the model validation and the obtained performance is presented in Section 4.1.2.

3.3 Bias correction and Extreme Value Analysis

Extreme Value Theory (EVT) provides a rigorous statistical framework for estimating the probability of rare and high-impact events. However, its applicability to geophysical hazards such as storm surges is often limited by the scarcity of available data. Extreme events occur infrequently, and observational records typically span only a few decades, resulting in small samples of annual maxima and, consequently, large statistical uncertainties.

One way to mitigate this limitation is to increase the sample length through hindcast simulations, which extend the observational period by reconstructing past conditions using numerical models. Even so, most studies evaluating historical or future changes have relied on relatively short records, generally not exceeding 30-40 years, thereby constraining the robustness of return level estimates.

An alternative approach involves pooling extremes from multiple climate simulations to form a single, larger sample. This ensemble-based methodology substantially increases sample size and reduces statistical uncertainty. However, different models typically exhibit systematic biases relative to each other and to the reference dataset. As emphasized in Toomey et al. (2025), pooling uncorrected extremes can lead to misleading return level estimates. Thus, a consistent and spatially coherent bias correction is required before combining simulations.

The approach adopted for bias correction follows the methodology used for wave climate extremes in Toomey et al. (2025) and adapts it to storm surge maxima. It performs a bias correction just on the annual maxima values of the projections

(historical and future periods).

First, for each dataset (hindcast, historical simulations, and future projections) annual maxima are obtained by selecting the highest surge value within each year:

$$M(y) = \max_{t \in \text{year } y} \{\eta(t)\}, \tag{3.20}$$

where $\eta(t)$ denotes the storm surge at time t.

This yields, at each grid point, three annual-maximum series:

$$\{H_y\}, \qquad \{X_y\}, \qquad \{F_y\},$$

corresponding respectively to the hindcast (45 values, 1979-2023), the historical climate simulation (36 values, 1970-2005), and the future projections (71 values, 2030-2100).

3.3.1 Bias correction: quantile-mapping approach

To ensure comparability across simulations and with the hindcast reference, a quantile–mapping (QM) procedure is applied independently at each grid point. This method adjusts the empirical CDF of each projection so that it matches the empirical CDF of the hindcast dataset. The procedure is carried out by assuming the CDFs, indicated with F, follow either:

- a Gumbel distribution:

$$F(x) = \exp\left[-\exp\left(-\frac{x-\mu}{\beta}\right)\right],$$

- a Generalized Extreme Value (GEV) distribution:

$$F(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \qquad 1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0.$$

The two distributions differ in flexibility: the Gumbel distribution is a special case of the GEV family with a fixed tail shape ($\xi = 0$), and is therefore fully described by a location parameter μ and a scale parameter β. In contrast, the GEV distribution introduces an additional shape parameter ξ, which controls the behavior of the upper tail. This added flexibility makes the GEV distribution more suitable for representing a wider range of extreme value behaviors.

Under this framework, the bias correction applied to the annual maxima is defined as:

$$X^* = F_{Hind}^{-1}(F_{Hist}(X))\,. \tag{3.21}$$

where X denotes the set of annual maxima to be corrected (either historical, X_{hist}, or future, X_{future}), and X^* is the corresponding bias-adjusted series.

This corresponds to assuming bias stationarity, meaning that the systematic bias of the historical period remains valid and constant in the future. Although this assumption is standard in climate impact studies, it is a acknowledged limitation. Another limitation of this methodology lies in the range of applicability of the X values. If some values exceed the hindcast maximum or fall below the hindcast minimum, the quantile-mapping relationship cannot be directly applied. Therefore, an additive extrapolation is introduced to preserve the scale of extremes.

The empirical historical CDF values are constrained to fall within the range of the empirical hindcast CDF. Values exceeding the maximum hindcast CDF are set equal to that maximum, while those below the minimum are replaced with the corresponding minimum value.

For future values lying outside the historical range, the correction is:

$$X^*_{\text{future}} = \begin{cases} X_{\text{future}} + \Delta_{\max}, & \text{if } X_{\text{future}} > X_{hist_{max}}, \\ X_{\text{future}} + \Delta_{\min}, & \text{if } X_{\text{future}} < X_{hist_{min}}. \end{cases} \tag{3.22}$$

being

$$\Delta_{\max} = X^*_{hist_{max}} - X_{hist_{max}}, \qquad \Delta_{\min} = X^*_{hist_{min}} - X_{hist_{min}}.$$

This prevents artificial compression of the tail while ensuring continuity with the corrected historical range.

As already pointed out, the bias correction is performed twice for each grid point, assuming respectively a Gumbel and a GEV distribution to represent the series of annual maxima. This dual approach provides a measure of structural uncertainty and allows assessing the sensitivity of return levels to the chosen extreme value model.

This bias correction procedure provides several key advantages for the subsequent analysis of extreme storm surge levels. First, it ensures full consistency between the historical portions of the climate projections and the hindcast reference dataset. Without this adjustment, each model chain would retain its own systematic errors, leading to discrepancies in both the magnitude and statistical distribution of storm surge extremes. By aligning the distribution of historical maxima from each projection with that of the hindcast, all datasets are placed on a common statistical basis.

Second, the correction guarantees direct comparability among the 17 GCM-RCM combinations used in this study. Given that each model pair is subject to different sources of bias, related to the driving Global Climate Model, the Regional Climate Model, or their coupling, unadjusted projections would exhibit inconsistencies that could mask or distort the climate change signal. The bias correction thus homog-

enizes the datasets, allowing differences in future projections to be interpreted as resulting from climate dynamics rather than from systematic model errors.

Third, the procedure prepares the datasets for a pooled Extreme Value Analysis (EVA). Pooling extremes from multiple climate simulations into a single distribution is advantageous because it greatly enlarges the available sample of extreme events. However, such pooling is only meaningful if all simulations share the same statistical characteristics in the historical period, which is precisely what the bias correction achieves.

After correction, datasets can be analyzed individually to assess model-specific responses or combined into a single ensemble-based dataset. The pooled approach reduces sampling uncertainty and stabilizes the estimation of high return levels, improving the reliability of projections for rare events, such as 100-year or 200-year storm surges. Overall, bias correction is essential for producing coherent, statistically consistent, and physically meaningful projections of future storm surge hazards across the Mediterranean basin.

3.3.2 Computation of Return Levels

Return levels (RLs) are a fundamental metric for quantifying the magnitude of rare storm surge events and for assessing changes in coastal hazard under present and future climate conditions. In this study, RLs are estimated both for the hindcast dataset, to characterize present day storm surge hazard, and for each climate projection, following the application of the bias correction procedure described in Section 3.3.1, in order to assess future hazard levels. The analysis is performed independently at each ocean grid point using a classical Extreme Value Analysis (EVA) based on the distribution of annual maxima.

For each GCM-RCM projection, the bias corrected annual maximum series is extracted for three periods of equal length (27 values, one annual maximum per

year), chosen to represent past and future conditions with horizons around 2050 and 2100:

- *historical* (1979-2005),
- *mid-century* (2034-2060),
- *end-century* (2074-2100).

As with the bias-correction procedure, return levels are evaluated assuming two alternative distributions for the annual maxima: Gumbel and Generalized Extreme Value (GEV). Distribution parameters (μ, β) for Gumbel or (μ, σ, ξ) for GEV are estimated by maximum likelihood. For a given return period T (in years), the associated return level η_T is the value expected to be exceeded on average once every T years. For the fitted distributions, return levels are computed as:

$$\eta_T = \begin{cases} \mu - \beta \, \ln\left[-\ln\left(1 - \frac{1}{T}\right)\right], & \text{Gumbel,} \\ \mu + \frac{\sigma}{\xi}\left\{\left[-\ln\left(1 - \frac{1}{T}\right)\right]^{-\xi} - 1\right\}, & \text{GEV.} \end{cases}$$

Return levels are computed for a comprehensive set of return periods, ranging from 2 to 200 years, allowing assessment of both frequent and rare events. The computation yields three spatially distributed return level fields, one for each period (historical, mid-century, end-century). Analyses are performed first for each projection individually and then for the ensemble of all 17 models pooled into a single coherent framework. In the pooled approach, the 27 annual maxima from each projection are aggregated across all 17 models, resulting in a considerably larger sample and more robust return level estimates.

The resulting return level fields provide the basis for assessing changes in storm surge hazard from present to future climate, comparing model-specific sensitivities,

quantifying uncertainty across the ensemble of 17 GCM-RCM chains, and identifying hot-spots of extreme sea level amplification. Combined with the bias corrected maxima, this EVA framework provides a statistically robust and spatially consistent approach for evaluating future storm surge extremes across the Mediterranean basin.

3.3.3 Assessment of robustness of projected changes

In the context of climate change impact studies, robustness refers to the degree to which projected changes are consistent and reliable across different model realizations, methodological choices, and sources of uncertainty. Rather than aiming at deterministic predictions, robustness assessment seeks to identify signals that emerge coherently despite the inherent uncertainties of climate projections.

In this study, robustness is addressed through the multi-model ensemble approach, comprising the 17 developed climate projections, which allows the separation of systematic signals from model-specific variability. As already pointed out, the use of a consistent numerical framework for both hindcast and future simulations ensures that projected changes in storm surge characteristics are not influenced by differences in model configuration.

The robustness of projected changes is evaluated following the guidelines of the IPCC AR6. This approach combines information on inter-model agreement and the emergence of the signal from internal variability. Specifically, projected changes are considered robust when at least 80% of the ensemble members agree on the sign of change and at least 66% of the models simulate a change exceeding a model-specific threshold representative of internal variability. This threshold, denoted as γ, represents the minimum magnitude of change in the annual maxima that can be distinguished from natural interannual variability. Mathematically, γ is expressed as:

$$\gamma = \sqrt{\frac{2}{N}} \cdot z \cdot \delta_{1\text{yr}} \tag{3.23}$$

where $\delta_{1\text{yr}}$ is the interannual standard deviation of the annual maxima, N is the number of years in the considered period (27 per period), and z is the quantile corresponding to the chosen confidence level (1.645 for the 80% confidence, consistent with IPCC AR6 guidance). γ thus provides an estimate of the expected amplitude of the natural variability, allowing for a clear distinction between natural variability and model-projected changes. Changes exceeding this threshold are considered to emerge from the forced climate signal rather than from internal variability. Based on these criteria, projected changes are classified into three categories:

- Robust change: when it is both consistent across the ensemble and sufficiently large to emerge from natural variability. This condition is satisfied when at least 80% of the ensemble members agree on the sign of change (increase or decrease) and when at least 66% of the models simulate a change exceeding the threshold γ.

- Conflicting change: when its magnitude is large enough to exceed natural variability, but the ensemble does not show a clear agreement on the direction of change. In this case, fewer than 80% of the models agree on the sign of change, while at least 66% of the ensemble members simulate a change exceeding γ.

- No robust change: when its magnitude is small relative to natural variability. This occurs when fewer than 66% of the models simulate a change exceeding γ, regardless of the level of agreement on the sign of change.

This classification enables the identification of regions where the climate change signal is both strong and consistent across the ensemble, as well as areas where uncertainty remains substantial either due to weak signals or lack of model agreement.

4

Results - Hindcast

This chapter presents the results obtained from storm surge hindcast simulations performed with the Delft3D-FLOW hydrodynamic model. The resulting dataset spans the period from 1st January 1979 to 31st December 2023, providing hourly estimates at 3 km resolution across the Mediterranean basin. It is validated against available tide gauge observations, considering both the full time series and, in particular, extreme events, which represent the primary source of coastal hazard. Model performance is assessed using the statistical metrics introduced in Section 3.2.4, with special attention given to the model's ability to reproduce surge peaks.

Furthermore, the model outputs are compared with storm surge levels produced by other authors (Mentaschi et al., 2023; Toomey et al., 2022), who kindly provided their results for this analysis.

Subsequently, an Extreme Value Analysis (EVA) is carried out to estimate return levels and return periods of storm surge extremes, providing a quantitative

characterization of present-day coastal hazard across the Mediterranean Sea.

Finally, the added value of a long-term, high-resolution hindcast dataset is illustrated through the analysis of two selected historical storm surge events. These case studies allow for a detailed investigation of the spatial and temporal evolution of past extremes, enhancing the understanding of the underlying processes and their coastal impacts.

4.1 Model performance

The performance of the storm surge model is assessed through a direct comparison between the hindcast dataset and observed surge levels at selected tide gauge stations across the Mediterranean Sea. Both qualitative and quantitative evaluations are carried out to assess the model's ability to reproduce the temporal evolution and magnitude of storm surge events under a wide range of meteorological conditions.

Figure 4.1 provides an initial qualitative validation of the model by comparing observed and simulated storm surge time series at six tide gauges located in different Mediterranean sub-basins. These locations, listed from east to west, are Barcelona (Spain), Marseille (France), Genova (Italy), Ajaccio (Corse, France), Venezia (Italy), and Hrakleio (Crete, Greece). They are selected as representative of different parts of the Mediterranean and will be used as reference sites in the following analysis. It is important to highlight that their selection is random, meaning that they do not represent optimal or worst-case conditions, but rather provide information on the spatial variability of the results.

The figure shows a representative two-year period, from January 1st, 2019, to December 31st, 2020, during which both moderate and extreme surge events occur. This comparison enables a visual assessment of the agreement between modeled and observed signals in terms of phase, variability, and intensity. On the right,

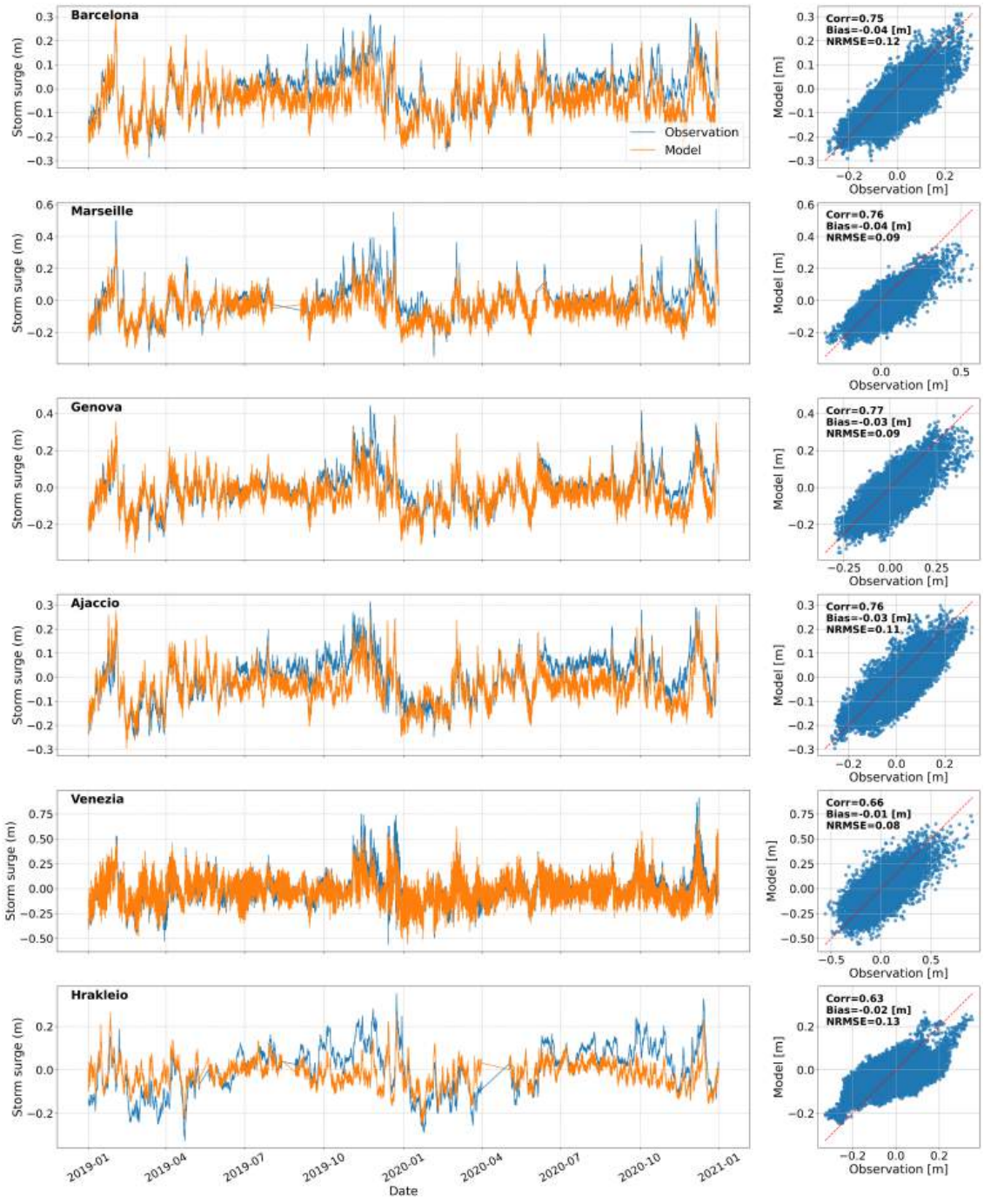

Figure 4.1: Time series of observed (blue) and modeled (orange) storm surge levels (left) and corresponding scatter plots (right) for selected stations during 2019–2020. Main statistical indexes are shown in each scatter plot.

the corresponding scatter plots provide additional visual information on the quality of the results. They represent observed values on the x-axis and modeled values on the y-axis. The more accurate the reproduction, the more the cloud of points aligns along the red line, which represents the ideal 1:1 condition. Each scatter plot also reports the values of Pearson correlation coefficient, bias and NRMSE, which provide a concise quantitative measure of model skill at each station over the same period. Overall, the model provides a realistic representation of the observed storm surge variability at all analyzed stations, although some local underestimations occur over limited time intervals. The correlation coefficient between observed and modeled surge levels is generally between 0.6 and 0.8, in line with previous studies (Androulidakis et al., 2023; Mentaschi et al., 2023; Toomey et al., 2022), indicating satisfactory overall model performance. The bias is predominantly negative and of the order of a few centimeters, meaning that the model tends to underestimate the observations, although this difference is not significant. The error is around 10 cm, again in line with previous studies.

However, the agreement between modeled and observed surge levels is not spatially uniform across the Mediterranean Sea. As shown in Figure 4.2, the statistical performance exhibits marked spatial variability among the validation stations, highlighting regions where model skill is higher and others where larger discrepancies persist. The highest correlation values are observed in the northwestern Mediterranean, from the Ligurian coast in Italy to the northeastern coast of Spain, including the Balearic Islands and Corsica, where values over the full 45-year simulation reach or exceed 0.7. More pronounced discrepancies are found in other areas, such as along the Adriatic coast and in the Alboran Sea. The bias exhibits a regional pattern and is predominantly negative, indicating that the model tends to underestimate the observed phenomenon. The bias assumes larger absolute values near the Strait of Gibraltar, likely due to the highly complex flow dynamics between the Atlantic and

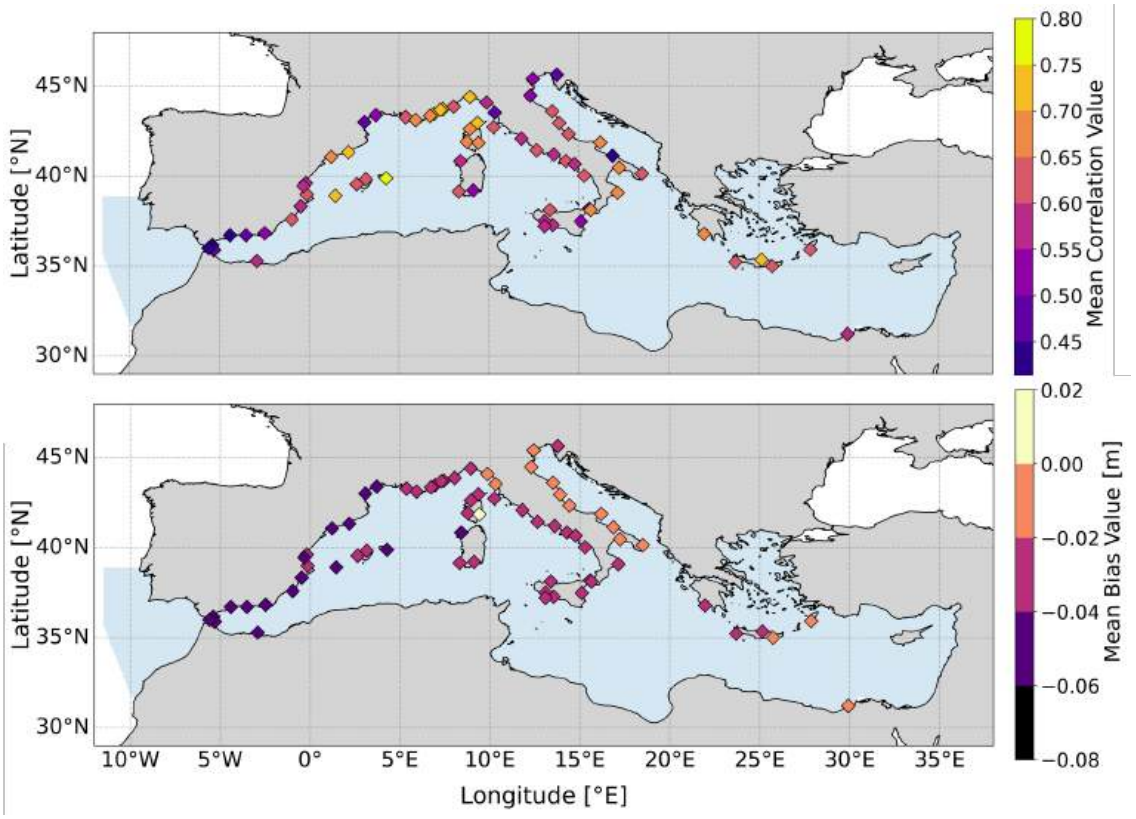

Figure 4.2: Spatial distribution of correlation (top) and bias (bottom) computed at the tide gauge stations over the 45-year period.

the Mediterranean. Consequently, surge levels near the strait tend to exhibit larger discrepancies relative to observations, which gradually decrease with distance from the strait, where bias values approach the ideal value of zero.

While the previous figure highlights the geographical heterogeneity of model skill across the Mediterranean, Figure 4.3 focuses instead on how model performance varies throughout the year by examining the seasonal variability of the same statistical indicators. The figure presents monthly box plots of the key statistical metrics, aggregating results from all analyzed tide gauge stations over the entire 45-year simulation period. As previously noted, the availability of observations varies over time, meaning that some stations provide information for a larger portion of the hindcast period, while others cover only specific time windows, with a greater availability of

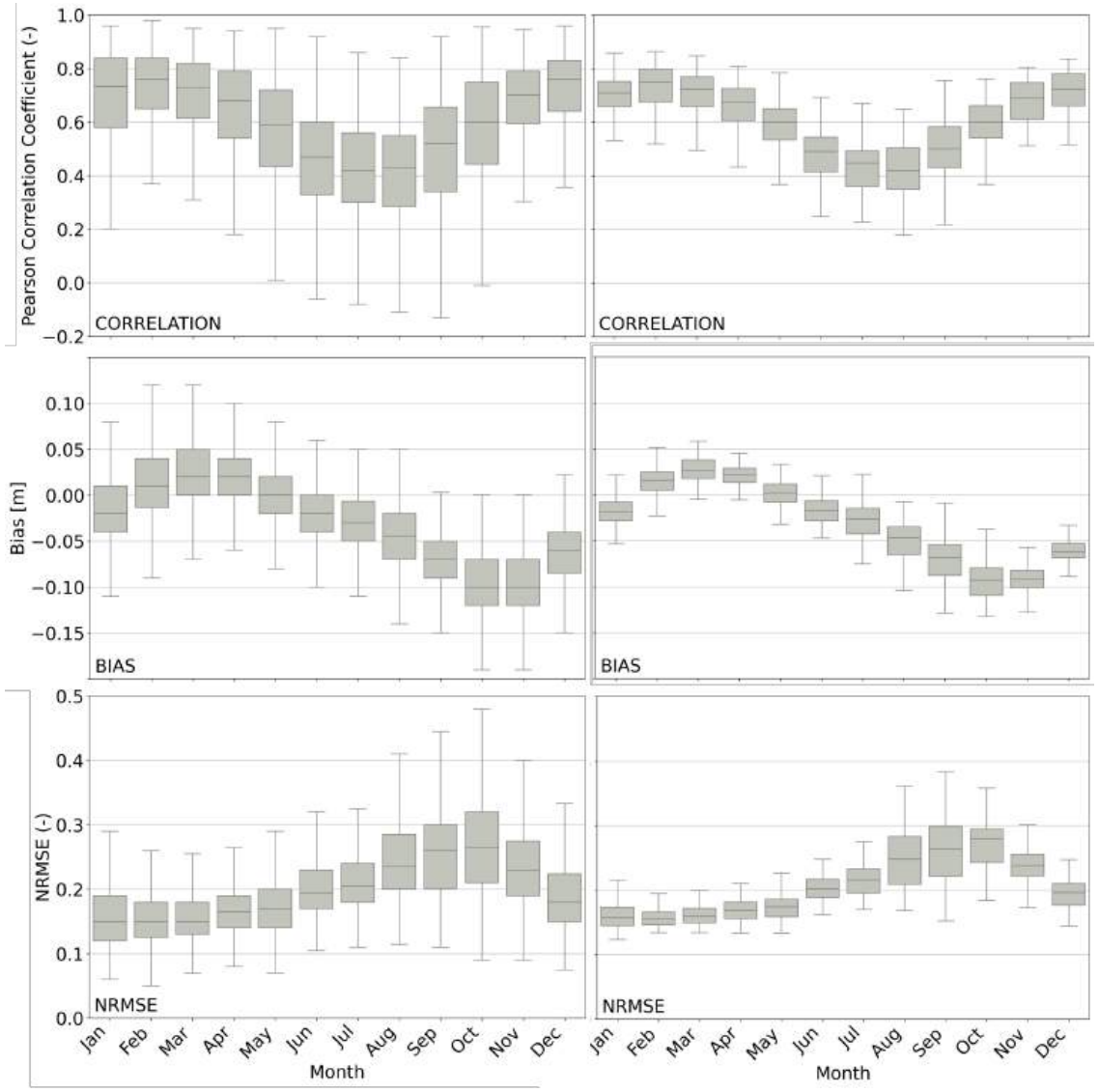

Figure 4.3: Monthly distribution of correlation, bias, and NRMSE derived from model–observation comparisons at tide gauge stations over the period 1979–2023. In the left panels, box plots are constructed using all monthly values from each station and each year, representing the combined interannual and inter-station variability for each month. In the right panels, box plots are based on monthly values averaged over the years at each station, representing inter-station variability only.

data in the last 20-25 years.

For each tide gauge, correlation values are first computed on a monthly basis for every year. Then the box plots are built in two different ways: the plots in the left column are constructed using all the values collected for each specific months from

all stations, thus representing the combined interannual and inter-station variability for each month; the plots in the right column are constructed using monthly values averaged over the full 45-year period, yielding one representative value per station and month. These station-averaged values therefore summarize the inter-station variability of model performance for each month. More specifically, all plots in the first column are constructed using a number of values equal to the number of validation station multiplied by the number of available monthly observation, while those in the second columns are based only on the number of stations. As expected, the boxes on the left are wider since they are composed of a larger range number and a wider range of data.

To facilitate interpretation, each box plot summarizes the data distribution using the quartiles of the distribution. The box extends from the first to the third quartile (interquartile range, IQR) and contains 50% of the data, with the median indicated by the horizontal line inside the box. The whiskers extend to the most extreme values within 1.5 times the IQR. A wider box indicates a larger variability in the data, whereas a narrower box reflects a more concentrated distribution around the median.

Correlation values exhibit marked seasonal variability, reaching higher values during the autumn and winter months, when storm surges are more intense. In these seasons the median correlation ranges between 0.7 and 0.8, indicating a good representation of the phenomenon. In contrast, significantly lower values are observed in summer, indicating that other processes not captured by the model become more relevant. The model's bias remains relatively small, ranging from about -20 and +15 cm when all values are considered and from -12 cm to +5 cm when only inter-station variability is considered. The predominance of negative values indicates a slight overall underestimation of storm surge levels, a behavior that is consistent with previous modeling studies in the Mediterranean region (Campos-Caba et al.,

2024; Fernández-Montblanc et al., 2020). A seasonal trend is evident for both the bias and the NRMSE, which increases during summer months. This common pattern reflects that storm surges, which are well captured by the model, are the dominant factor in sea level variability during the autumn and winter months; conversely, other processes not included in the simulations (e.g. baroclinic sea level changes) are more prominent in the summer. Importantly, it is during extreme events that accurate modeling is most critical, as these events pose the greatest risk to coastal areas. Reliable simulation of such occurrences is therefore essential for coastal protection and risk mitigation as will be discussed in Section 4.1.2.

Considering again at the six location used as reference, the values of the statistical indications are shown in Figure 4.4, which provides both the mean and the variability (through the standard deviation) of the indicators across the year for the 45-year period. In addition to the previous figure, this representation provides complementary information on model performance, as it allows characterization of the accuracy with which each station is reproduced. Correlation values are generally higher in the first and last months of the year, indicating a greater ability of the model to reproduce storm surge in these seasons. It is important to recognize that this analysis does not account for the number of available measurements, so Hrakleio appears to have the highest correlation, even though observations at this station are available only for some recent years. Overall, it is evident that summer months generally exhibit lower average correlation and higher variability over the years. The bias is often positive in the first months of the years and negative in the last part. Errors are again larger in summer, with higher standard deviations during this period.

To further investigate model performance, attention is focused on the stations exhibiting the highest and lowest correlation values, examining their seasonal trends and assessing whether common features emerge. Figure 4.5 shows the statistical

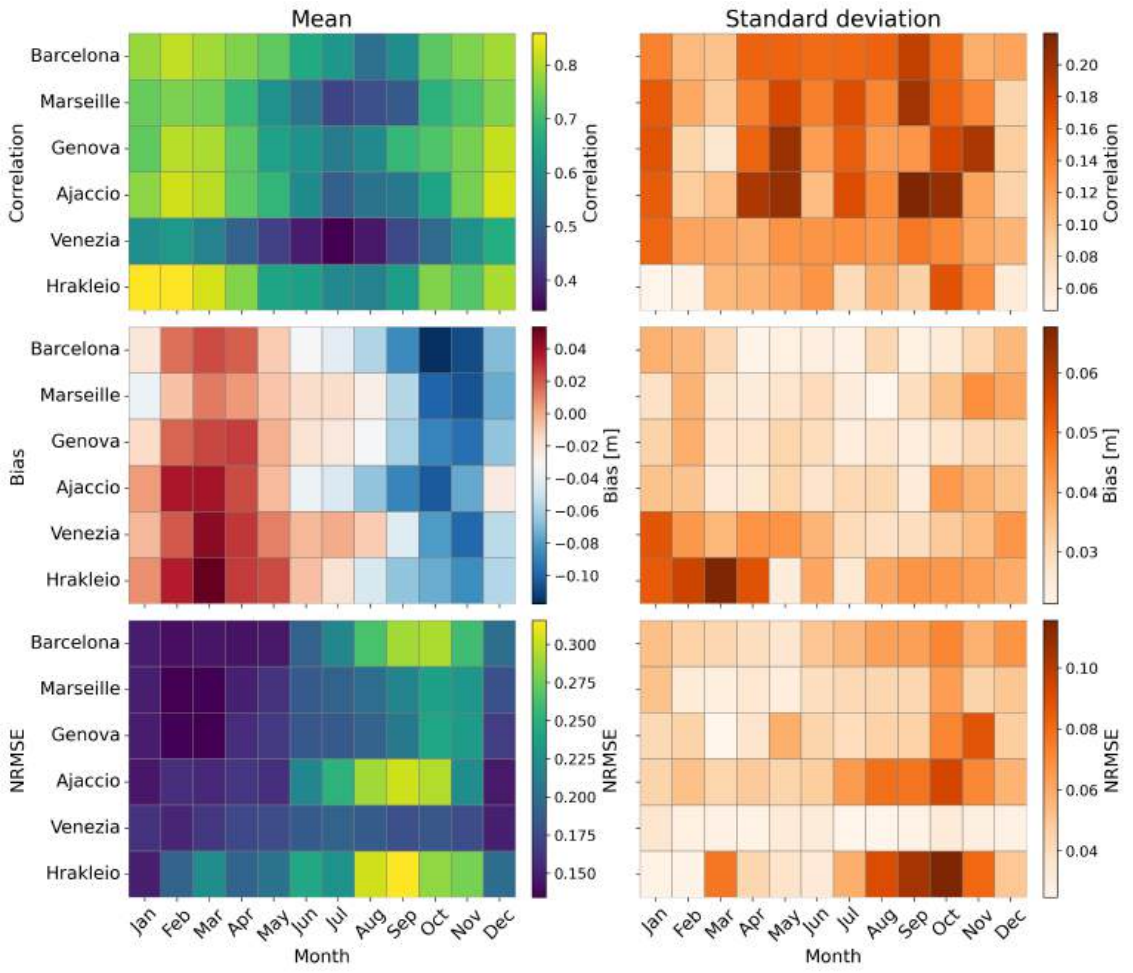

Figure 4.4: Heatmaps of correlation, bias, and NRMSE at the selected tide gauge stations across the different seasons. Both the mean values (left) and the standard deviation (right) are shown.

metrics for the three best reproduced locations (Mahon and Ibiza in Spain, and Hrakleio in Greece) as well as the three least accurately reproduced sites (Algeciras, Malaga, and Tarifa in Spain).

For the best-performing locations, correlation coefficients generally range from 0.5 to above 0.9, while biases remain below ±0.1 m during the autumn and winter months, although they reach higher values in summer. Conversely, the worst-performing stations exhibit considerably lower skill, with correlation coefficients that are in some cases even negative, and show larger discrepancies in terms of

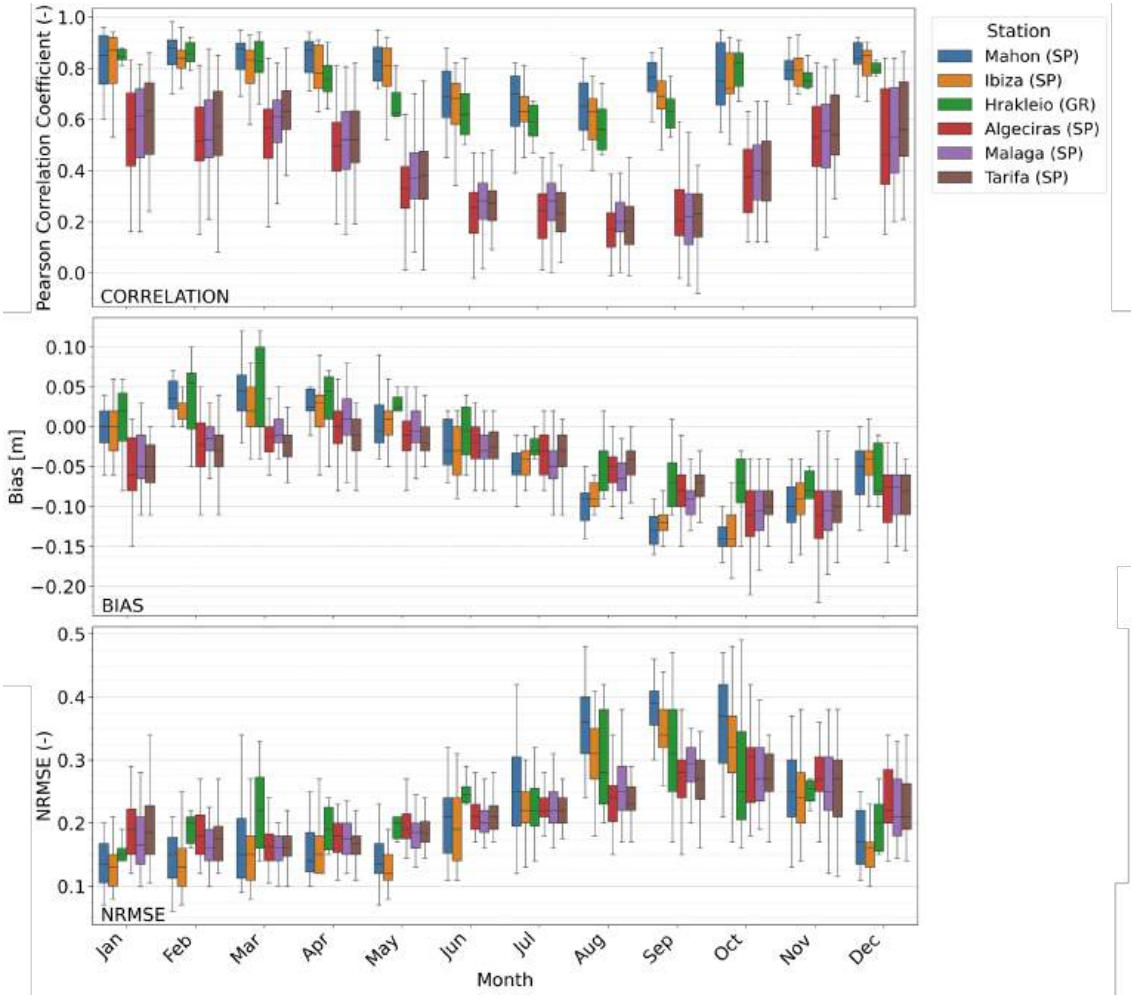

Figure 4.5: Monthly distribution of correlation, bias, and NRMSE derived from model–observation comparisons at the three best- and worst-performing tide gauge stations in terms of correlation over the period 1979–2023. Boxes represent temporal variability.

bias as well. A different behavior is observed for the NRMSE: stations with high correlation values also display the highest NRMSE values, especially in summer. This highlights how these metrics are complementary and provide different information: the NRMSE quantifies the error in terms of amplitude, while the Pearson correlation coefficient characterizes the temporal evolution, i.e. the shape of the signal.

Overall, the differences between the two groups are more evident in terms of

correlation: the best-performing sites consistently achieve high correlations, while the poorer-performing ones display greater variability across years, as highlighted by the wider spread of the box plots, and generally settle at lower values. These consideration are also confirmed by Figure 4.6, in which the statistical metrics are represented both in terms of mean values and standard deviations.

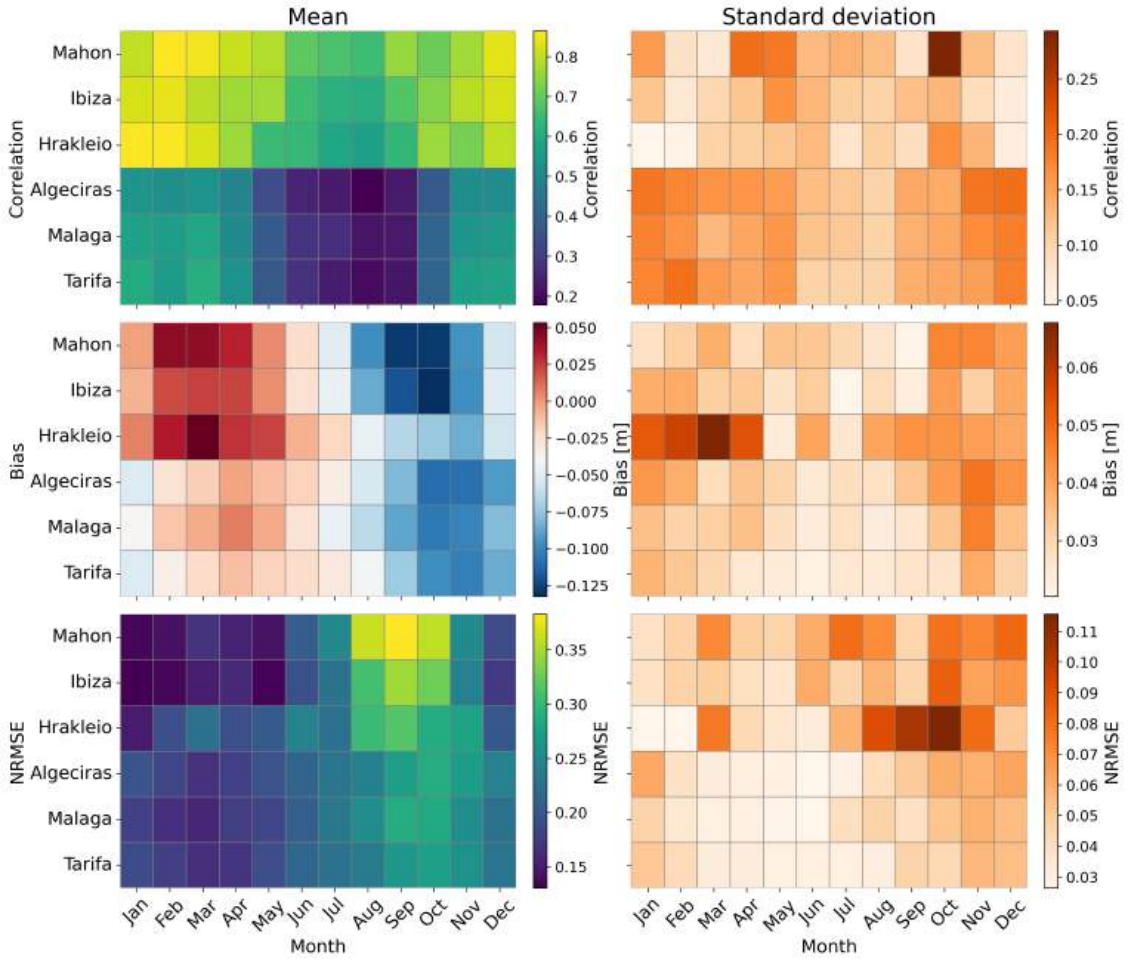

Figure 4.6: Heatmaps of the correlation, bias, and NRMSE at the tide gauges with the highest and lowest correlation values across the different seasons. Both mean values (left) and standard deviation (right) are shown.

Here the higher mean correlation values obtained at Mahon, Ibiza, and Hrakleio compared to the other three stations are even more evident for all months of the

years. Furthermore, the bias also tends to be larger in absolute value, as already highlighted above. The heatmaps therefore appear to be a useful tool to identify the characteristics of each station in terms of monthly statistical metrics, allowing the identification of areas and seasons in which the model perform better and others in which additional physical processes interact and are not properly captured by the model.

4.1.1 The role of the Inverse Barometer Effect

While the final simulation performed over the 45-year period includes the Inverse Barometer Effect (IBE) at the ocean open boundary, a comparison is carried out with an earlier version of the simulation, identical in all respects except for the omission of the IBE as a boundary condition. Both datasets are validated against the same set of tide gauges to assess the improvements resulting from the inclusion of this effect. The inclusion of the IBE has implications beyond the open ocean and the coastal regions near the Strait of Gibraltar, substantially enhancing model performance throughout the entire Mediterranean basin.

Figure 4.7 presents the time series of observed and modeled storm surge levels at the reference tide gauge stations, covering the period from 1st January to 31st March 2019. Both model outputs, including the IBE effect (green) and neglecting it (orange), are shown in order to qualitatively assess the model's ability to reproduce observations. The time series indicate that the hindcast dataset including the IBE effect provides a more accurate representation of observed values, particularly under high-pressure conditions, associated with negative surge levels. This improvement is evident across all analyzed locations, suggesting that the IBE effect influences not only areas near the Strait of Gibraltar but also more distant regions. The scatter plots further support this finding: for all stations, the cloud of points corresponding to the simulation including IBE (in green) is consistently closer to the 1:1 red

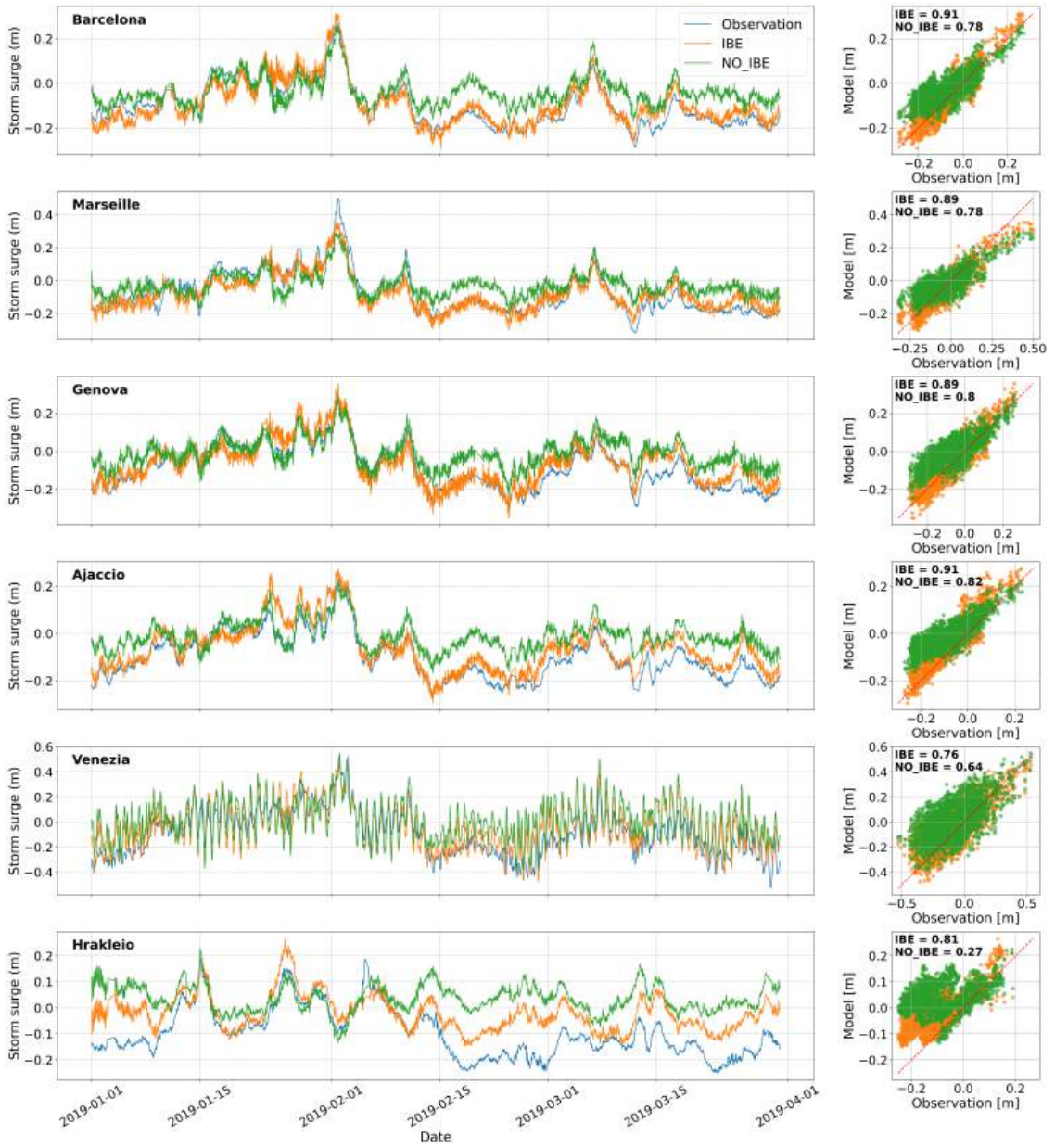

Figure 4.7: Time series of observed (blue) and modeled (orange: including IBE; green: neglecting IBE) storm surge levels (left) and corresponding scatter plots (right) at selected stations during January-March 2019. Pearson correlation coefficient for both simulations are shown in each scatter plot.

reference line than the one corresponding to the simulation without IBE (in orange), and the reported correlation values indicate systematically higher agreement with observations when the IBE effect is included.

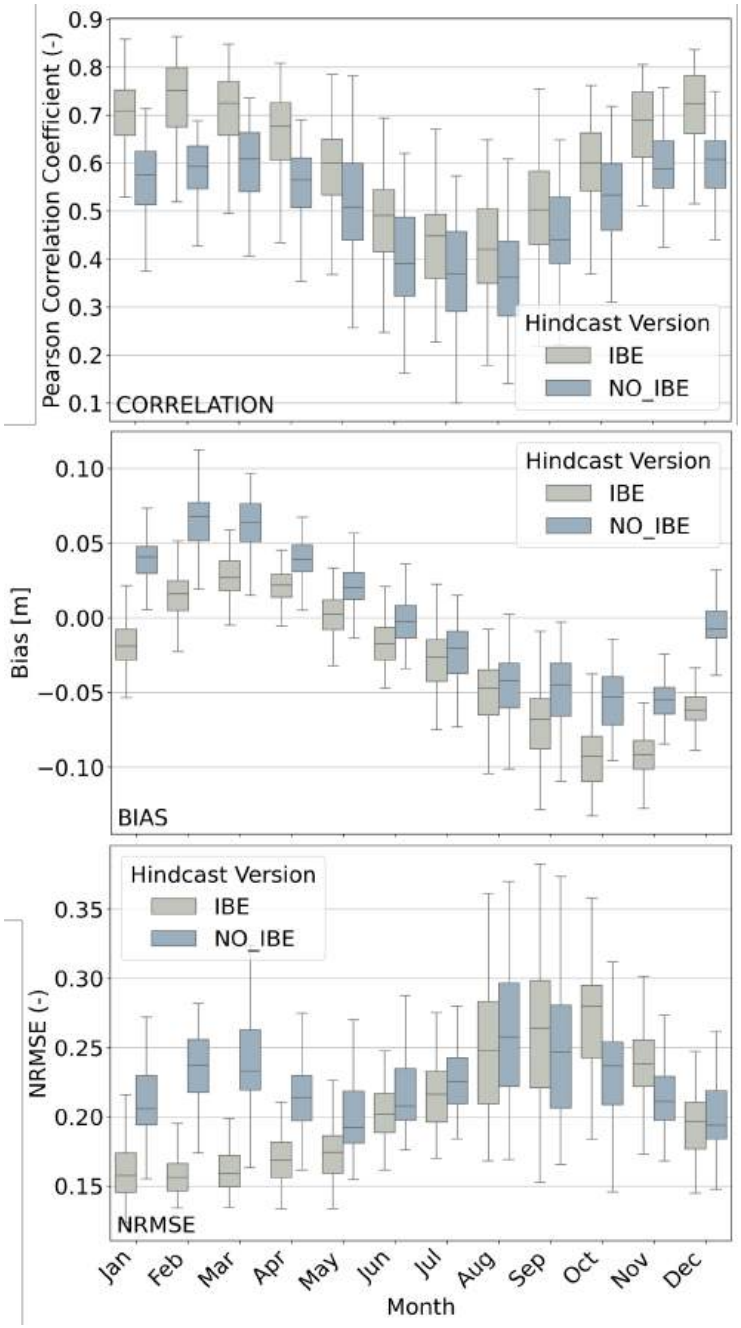

Figure 4.8: Monthly distribution of correlation, bias, and NRMSE derived from model–observation comparisons at tide gauge stations over the period 1979–2023. Both simulations considering (sage green) and neglecting (steel blue) the IBE contribution are shown. Boxes represent inter-station variability for each month.

Figure 4.8 compares the results of the models with and without the IBE contribution using the same box-plot approach described in the previous section, i.e. extending the analysis to all the validation station and over all the available month of observation. Overall, the model run including the Inverse Barometer Effect (IBE) demonstrates improved performance throughout the year: correlation values are consistently higher across all months, while NRMSE values are lower compared to the run without IBE, particularly during winter and spring. The bias distribution is slightly shifted toward negative values; however, the magnitude of this shift remains relatively small and does not indicate a systematic, large-scale underestimation.

4.1.2 Simulation of the extremes

Having assessed the general model performance across the full range of surge variability, a more specific analysis focuses on the extreme surge levels to evaluate the model's ability to capture irregular but severe events.

The model performance in terms of extreme surge values is assessed through the bias, whose values are compared across the full dataset, peak events, and annual maxima for each tide gauge station. As in the case for the previously presented analysis, larger bias values indicate a reduced model capability to accurately reproduce extreme events. Positive bias values denote an overestimation of extremes by the model with respect to tide gauge observations, whereas negative values indicate an underestimation.

Figure 4.9 displays the bias evaluated at all validation stations under three different configurations. The top panel shows the bias computed over the entire 45-year time series (corresponding to the bias analysis previously shown in the bottom panel of Figure 4.2); the middle panel refers to peak events; and the bottom panel is related to annual maxima. Peak events are identified from the observed surge time series at each tide gauge station as exceedances above the station-specific 98^{th}

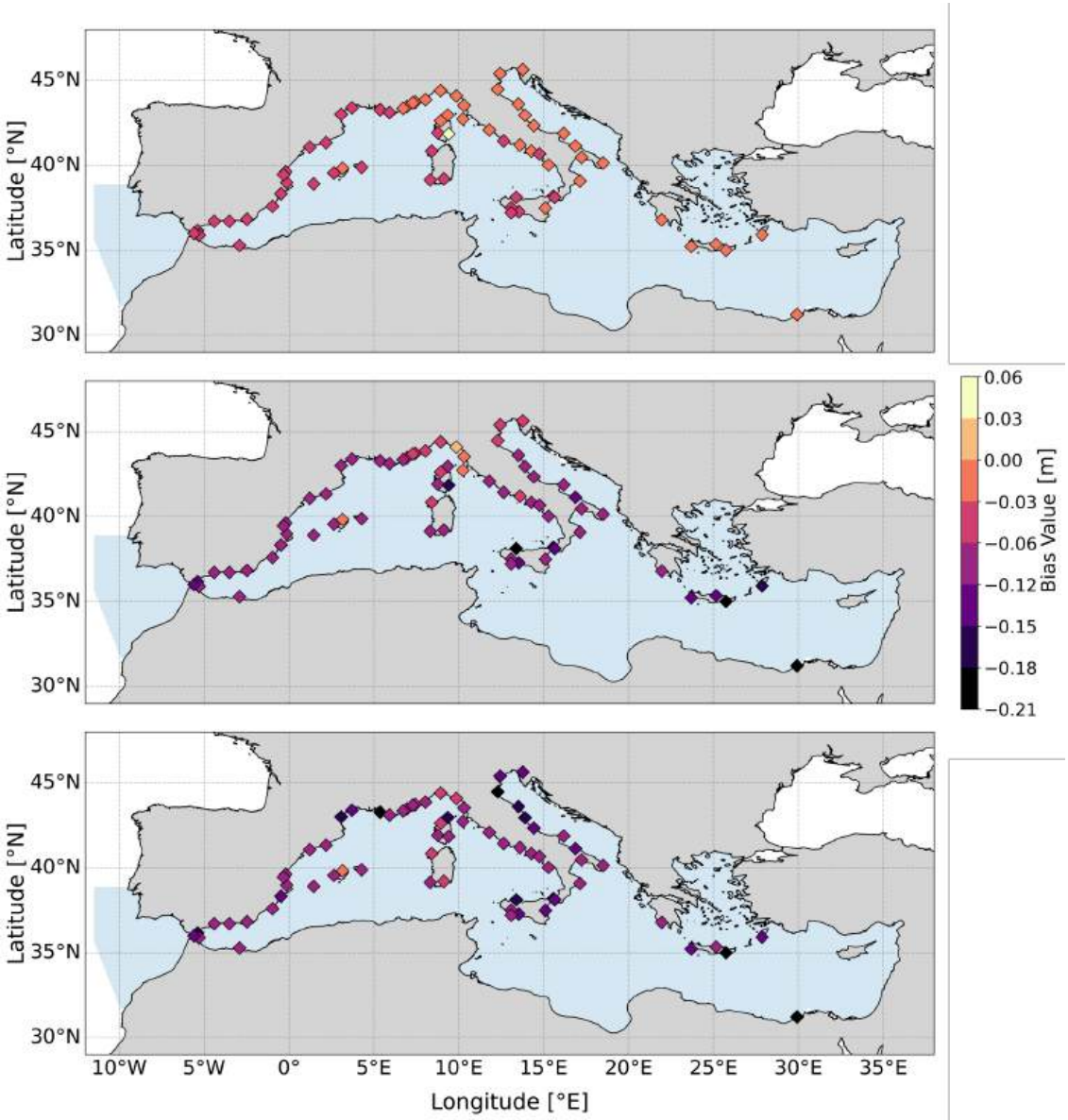

Figure 4.9: Bias values at the tide gauge stations over the 45-year period. Top: full time series; middle: peaks; bottom: annual maxima.

percentile (typically around 0.2 m), persisting for at least six consecutive hours, in order to exclude spurious or erroneous measurements. For each observed peak event, the corresponding (time-concurrent) model surge values are then extracted and used to compute the bias. Annual maxima are defined from observations as the

largest surge value recorded in each year. To account for possible timing offsets between the simulated and observed peaks, the model annual-maximum counterpart is taken as the maximum model surge within a ±12-hour window centred on the time of the observed annual maximum.

While the bias computed over the full time series exhibits a regional pattern, as already discussed for Figure 4.2, with predominantly negative values in the range of 0 to −6 cm, its magnitude increases when considering peak events, reaching values of up to −15 cm. Even larger discrepancies are observed for the annual maxima. This increase arises from two main factors. First, extreme surges are inherently harder to accurately predict, as they depend on very specific combinations of forcing conditions, such as strong winds from particular directions, that in occasions the model cannot fully resolve. Second, there are fewer extreme events compared to moderate ones, meaning that individual modeling errors have a larger influence on the average bias when focusing only on extremes. Figure 4.9 provides a qualitative spatial assessment across the Mediterranean basin of the model limitations in fully capturing extreme surge events.

To better quantify these differences, Figure 4.10 presents the same bias values shown in Figure 4.9, but focusing on the six reference tide gauge stations defined previously and using a bar representation that facilitates comparison. For each station, the first bar refers to the bias evaluated over the full period, the second to peak events, and the third to annual maxima.

While some stations do not show substantial changes in bias behavior, such as Genova, where bias values remain within the range of −0.02 to −0.05 m, indicating a satisfactory model performance also for extreme events, other locations exhibit larger discrepancies. At the Venezia tide gauge, the bias associated with peak events remains close to zero but increases when annual maxima are considered, possibly due to specific extreme events that are not adequately reproduced by the model.

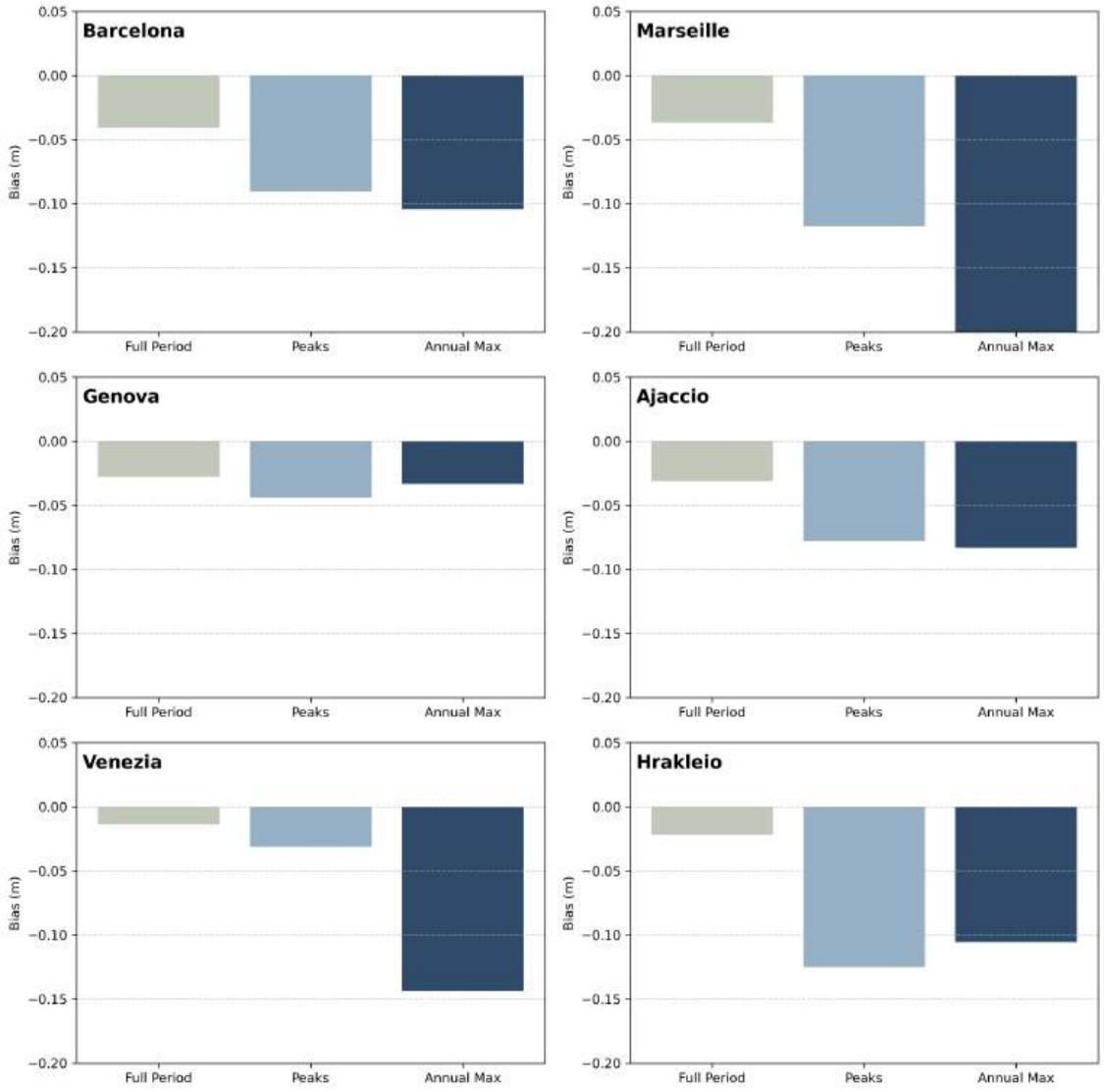

Figure 4.10: Bias values for full time series (first bar), peaks (second bar), and annual maxima (third bar) at six reference tide gauges.

Other stations, such as Marseille and Hrakleio, show consistently negative bias values around -0.1 m for extreme events, indicating a reduced capability of the model to capture intense surge conditions in these areas.

Beyond evaluating the model performance in terms of bias, the timing accuracy of the modeled surge peaks is also analyzed in order to investigate the presence of

systematic temporal offsets. For each observed peak event and for each station, a 12-hour time window before and after the observed peak time is considered, and the temporal shift ΔT between the modeled and observed peaks is identified.

Figure 4.11 shows violin plots of the temporal shift ΔT for peak events and annual maxima at the six reference stations. Violin plots combine the features of box plots and kernel density estimates, providing information on both the distribution and the probability density of the data. The width of each violin is proportional to the density of occurrences at a given ΔT value, with wider sections indicating more frequent temporal offsets. The central markers indicate the median and mean time shifts, while the overall extent of each violin is limited to the interval ± 12 hours. These plots allow the identification of systematic advances or delays, as well as the variability and symmetry of the temporal mismatch between simulated and observed extreme surge events.

The stations of Genova, Ajaccio, and Venezia show distributions concentrated around zero (or within ± 5 hours), indicating a good model performance in reproducing the average timing of peak events. Other stations, such as Marseille and Hrakleio, exhibit more uniform distributions, indicating more variable offsets. In general, many stations display distributions closer to zero for annual maxima than for individual peaks. This is because annual maxima represent the most dominant extreme events, which are better captured by the model, whereas more frequent peak events can be influenced by transient local phenomena, resulting in greater variability in the temporal shift.

A final way in which the model skill in reproducing peak events is evaluated concerns its ability to correctly detect extreme events, rather than to accurately reproduce their exact magnitude. For this purpose, the indices introduced in Section 3.2.5 are used, namely: the model ability to detect events above a given threshold (Hit Rate, HR), the tendency to generate extreme events that are not actually

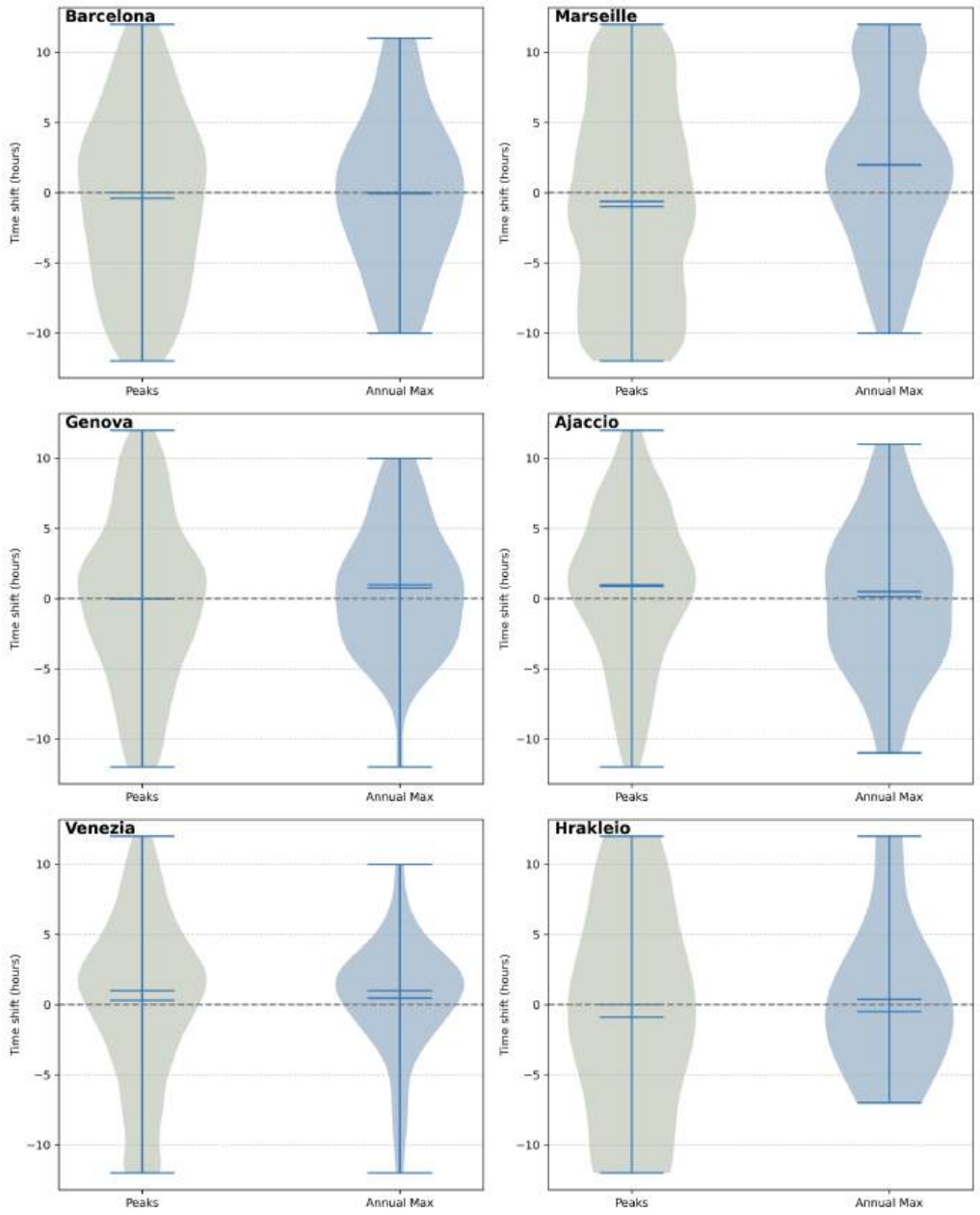

Figure 4.11: Violin plots of peak timing offsets (Δt) and annual maxima offsets for the six reference stations. Each violin shows the full distribution density: wider areas indicate higher data frequency; the central box shows quartiles with median line; whiskers extend to extremes.

observed (False Alarm Ratio, FAR), the tendency to over- or underestimate the overall number of peaks (Frequency Bias, FB), and the capability of the model to

detect extreme events beyond what would be expected by chance (Equitable Threat Score, ETS). As previously discussed, these indices are based on a contingency table and count correctly detected and missed extreme events within a time window of ±24 hours around the observed peak time.

The values obtained from this analysis are summarized by the box plots in Figure 4.12. For each station, single values of HR, FAR, FB, and ETS are computed over the 45-year period. The box plots therefore represent the distribution of these scores across all stations, highlighting both the typical behavior and the interannual variability of the model performance.

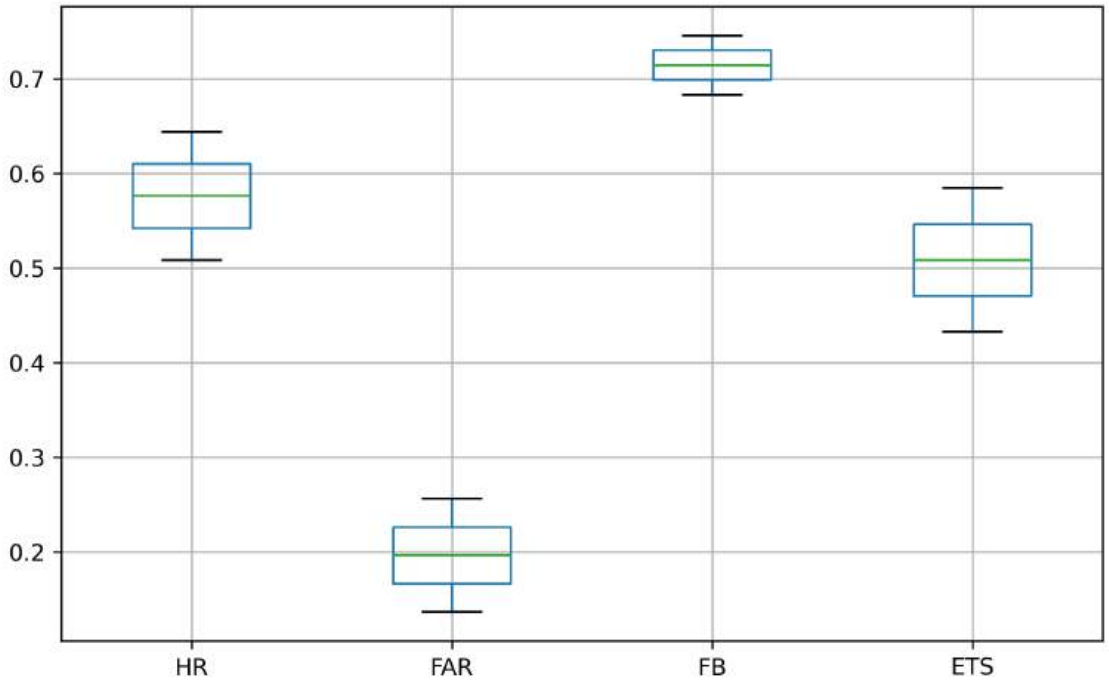

Figure 4.12: Distribution of peak detection skill scores (HR, FAR, FB, ETS) across all tide gauge stations and years.

Overall, the model shows a good ability to detect extreme events. Hit Rate (HR) values generally lie between about 0.5 and 0.6, with medians around 0.55-0.60, indicating that more than half of the observed peaks are correctly identified

within the selected time window. In contrast, the False Alarm Ratio (FAR) remains comparatively low, with median values around 0.20, suggesting that most simulated extreme events correspond to observed ones and that the number of false alarms is low.

The Frequency Bias (FB) exhibits mean values close to 0.70, indicating a slight tendency of the model to underestimate the total number of extreme events (FB $<$ 1). The Equitable Threat Score (ETS) reaches mean values of about 0.50 pointing to an overall good capability of the model to reproduce peak events beyond what would be expected from random agreement. Taken together, these indices confirm that the model is able to, not only represent storm surge levels reasonably well, but also to reliably capture the occurrence of extreme events that are most relevant for coastal risk assessment.

4.1.3 Comparison with previous studies

While the validation presented so far provides an assessment of model performance against tide gauge observations, it is also important to place these results within the context of existing studies.

To this end, the hindcast developed in this thesis is compared with three previous storm surge datasets produced by Toomey et al. (2022), Mentaschi et al. (2023), and Muis et al. (2022). The first two are based on the SCHISM hydrodynamic model, although with different spatial and temporal resolutions, as described in Section 2.1. The latter dataset is derived from the Global Tide and Surge Model (GTSM), implemented using the Delft3D Flexible Mesh suite, and provides global coverage with a spatial resolution ranging from approximately 1° in the open ocean to 0.1° in coastal areas for the period 1950–2024, forced by ERA5 reanalysis data.

All datasets provide storm surge simulations over the entire Mediterranean Sea. However, for consistency, the comparison is restricted to ten tide gauge locations

for which data from all datasets are made available: Barcelona, Marseille, Genova, Venezia, Ancona, Trieste, Valencia, Nice, Malaga, and Napoli.

Figure 4.13 shows time series and scatter plots for a subset of the selected tide gauge stations. All models reproduce the main temporal variability and most surge peaks, with modelled and observed values clustering close to the 1:1 line. Statistical metrics indicate high correlations (generally above 0.8) and relatively small biases, confirming that all datasets provide a realistic representation of storm surge variability at these sites.

A more systematic assessment is provided in Figure 4.14, which presents the monthly distribution of correlation, bias, and NRMSE across the ten common validation stations for the period 1979-2023. For each model and month, box plots summarize both inter-station and interannual variability, being built with all the single monthly values for each station.

All datasets exhibit broadly consistent seasonal patterns, although differences in performance are evident. The datasets by Toomey, Mentaschi, and GTSM generally show higher median correlation values throughout the year, while the hindcast developed in this study presents slightly lower correlations, particularly during spring and summer. During autumn and winter, however, differences among the datasets are reduced, and the present hindcast approaches comparable skill levels.

Bias values remain close to zero for all models, with a slight tendency toward underestimation, more evident for the present hindcast during winter. NRMSE shows a clear seasonal cycle, increasing during summer for all datasets, with the hindcast presented here generally exhibiting moderately higher median values.

To further investigate spatial and seasonal differences among the four datasets, Figures 4.15 and 4.16 present heatmaps of the mean values and standard deviations of the same statistical metrics for all tide gauge stations.

The mean heatmaps (Figure 4.15) show that all datasets capture the main spatial

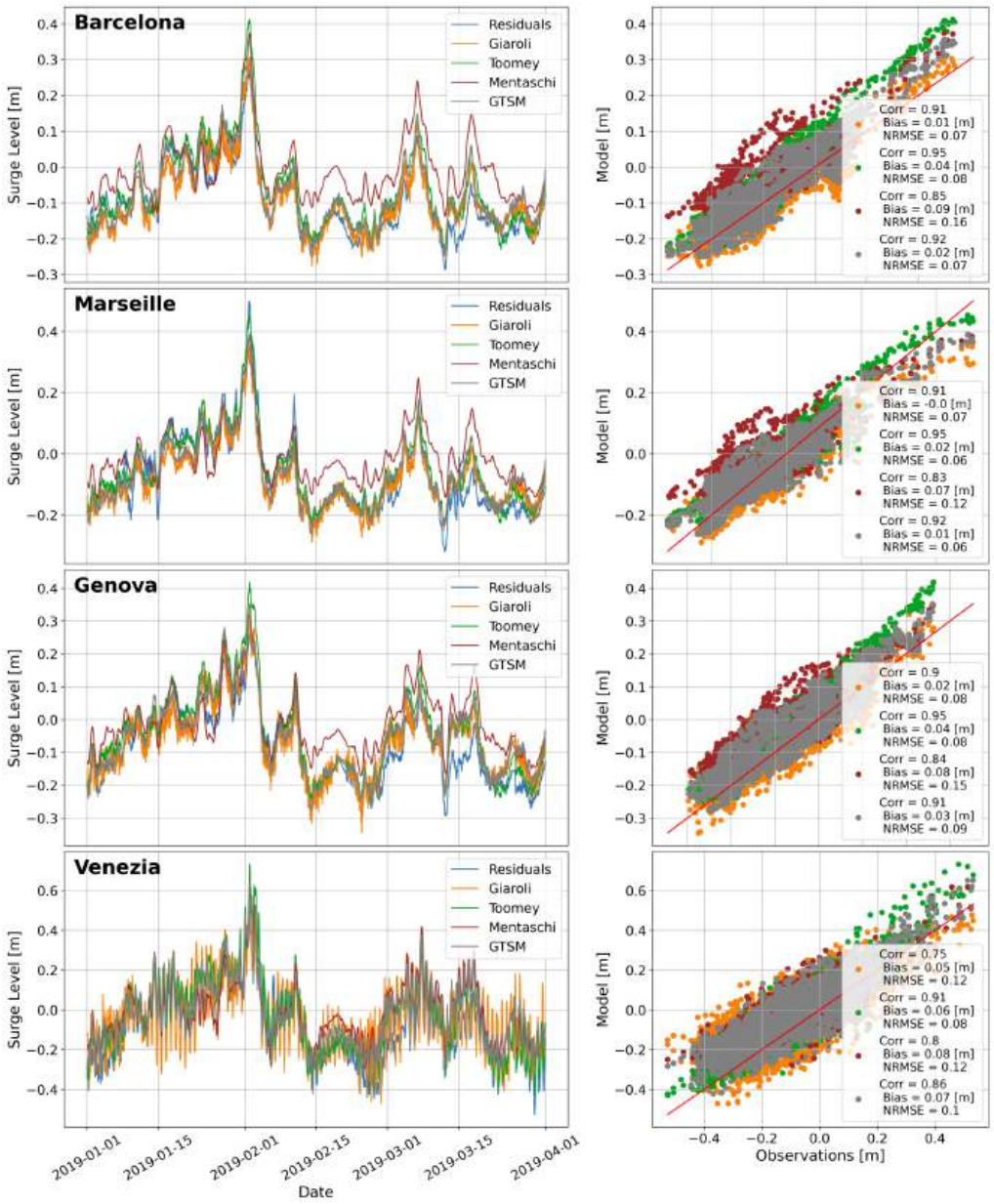

Figure 4.13: Comparison between observed surge residuals (blue) and modeled surge levels from this study (Giaroli, orange), Toomey (green), Mentaschi (brown), and the Global Tide and Surge Model (GTSM; grey) during January-March 2019 (left panels). Corresponding scatter plots (right panels) show model values against observations, together with correlation, bias, and NRMSE for each dataset.

patterns, with higher correlations and lower errors at several northwestern Mediterranean stations and reduced performance at sites characterized by more complex

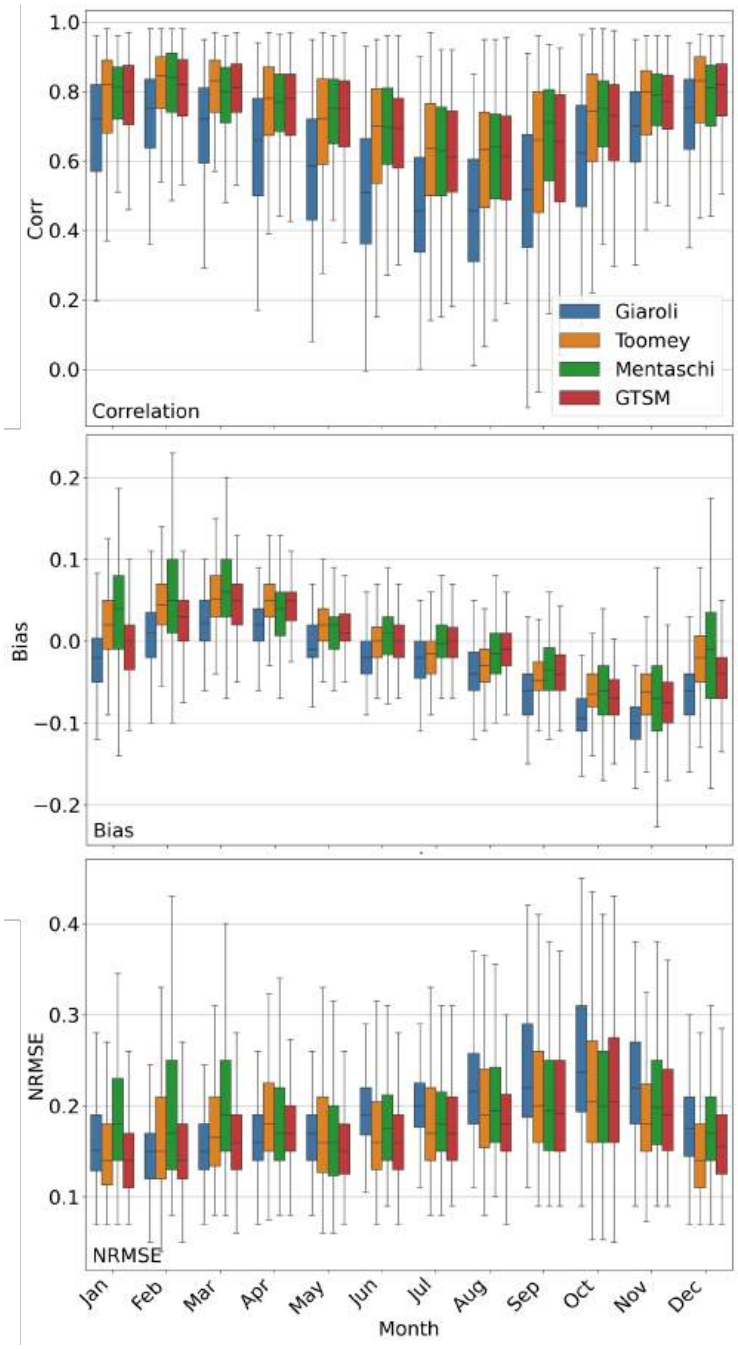

Figure 4.14: Monthly distribution of correlation, bias, and NRMSE derived from model–observation comparisons at tide gauge stations over 1979–2023. Box plots show inter-station variability for each month for the hindcast developed in this study (Giaroli, blue), Toomey (orange), Mentaschi (green), and GTSM (grey).

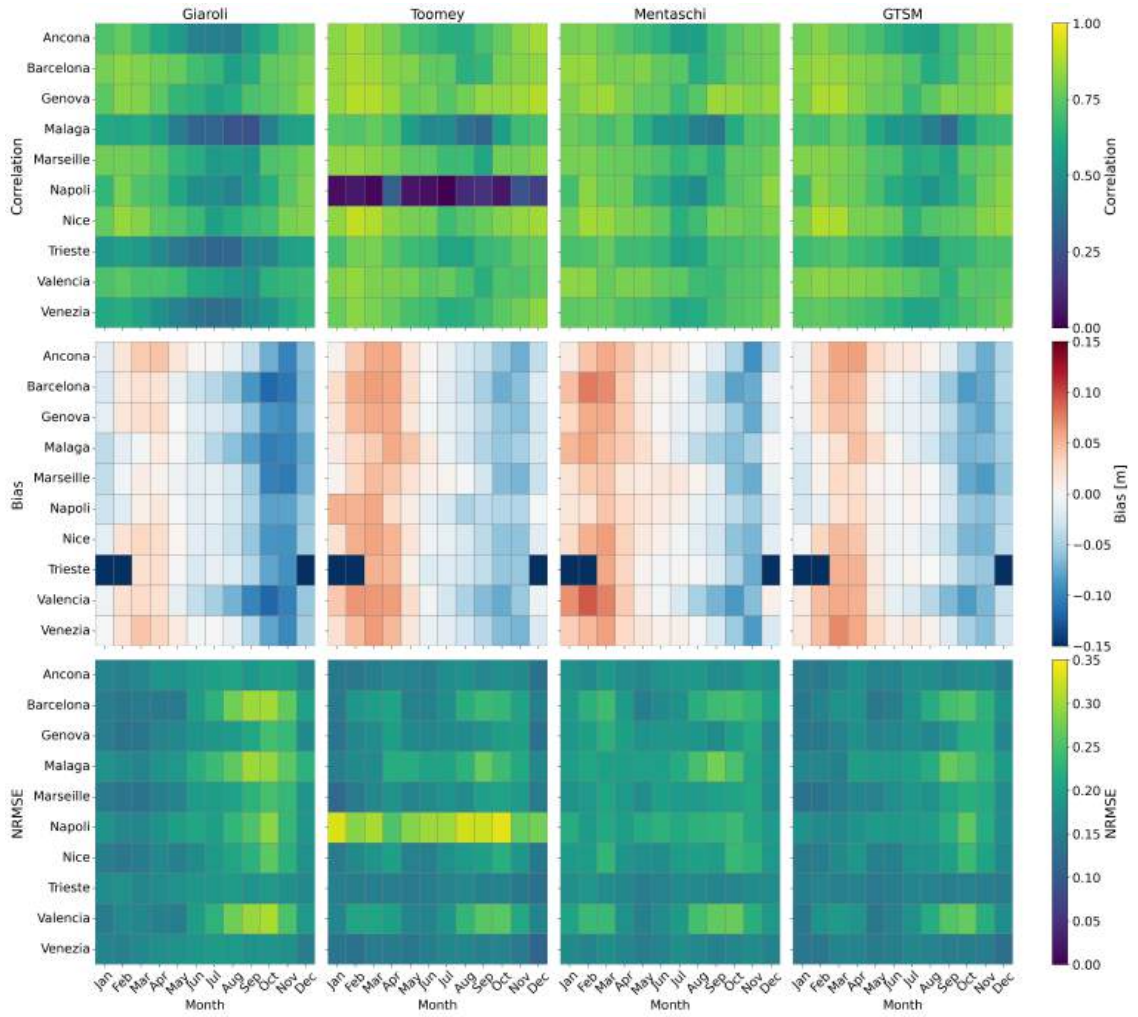

Figure 4.15: Heatmaps of mean monthly correlation, bias, and NRMSE at tide gauge stations for the hindcast developed in this thesis (first column), and for the datasets by Toomey et al. (2022) (second column), Mentaschi et al. (2023) (third column), and Muis et al. (2022) (fourth column). Rows correspond to stations and columns to calendar months.

local dynamics. In many cases, the performance of the present hindcast is comparable to that of the other datasets, and in some instances it shows slightly higher correlation or lower NRMSE. The standard deviation fields (Figure 4.16) indicate similar levels of interannual variability across the datasets, suggesting that differences are primarily related to local and seasonal factors rather than systematic long-term biases.

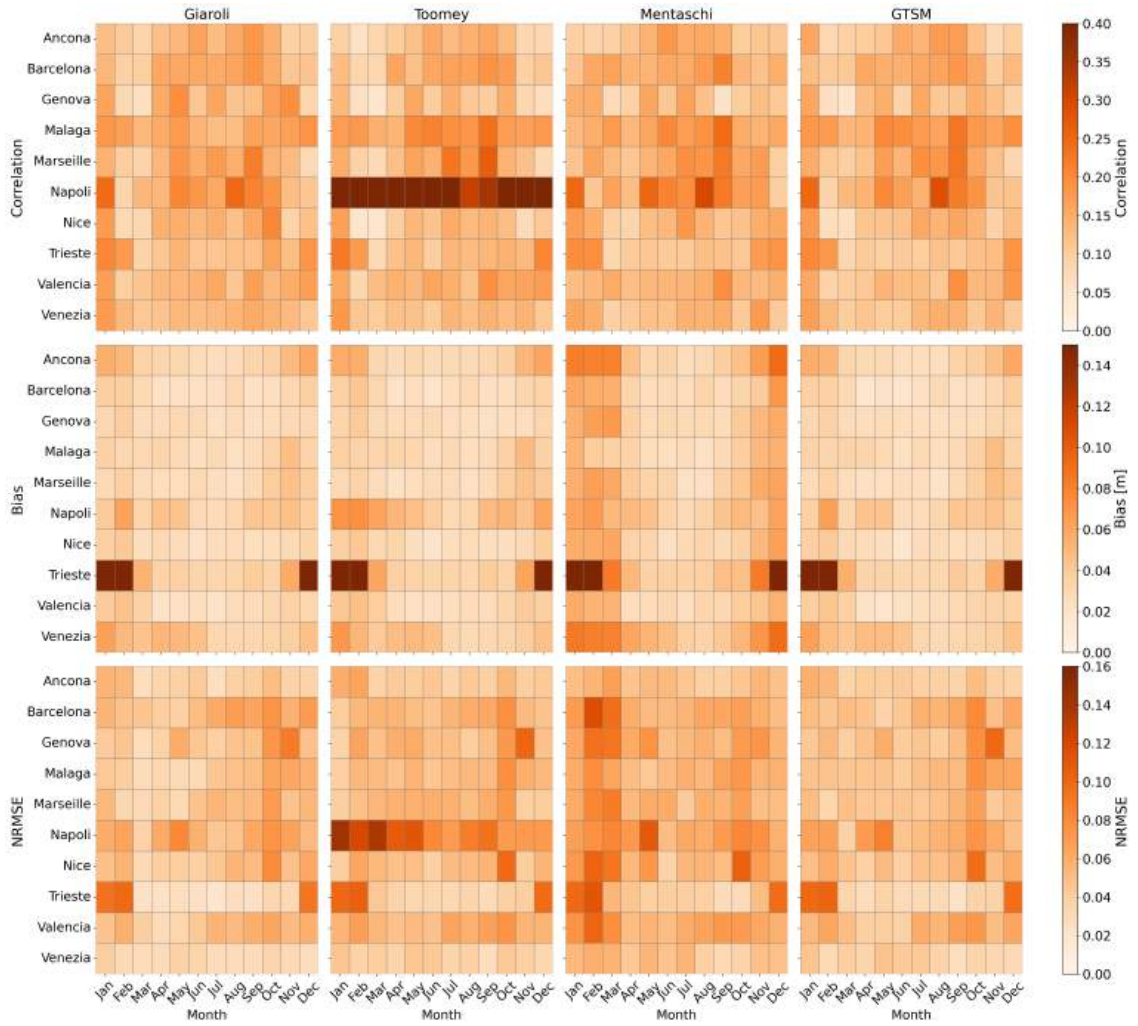

Figure 4.16: Same as Figure 4.15, but showing the standard deviation of monthly correlation, bias, and NRMSE, thus highlighting interannual variability for each station and month.

In some locations, the hindcast developed in this study exhibits lower performance compared to previous datasets. These differences can be attributed to several factors related to model configuration and forcing. In particular, discrepancies may arise from differences in spatial resolution and coastal representation, including bathymetry, which strongly influence nearshore surge dynamics. Additional factors include differences in atmospheric forcing datasets, as well as in the parameterization of wind stress and the treatment of open boundary conditions and large-scale

pressure effects.

Overall, these differences reflect variations in modelling choices rather than a systematic deficiency of the model, and highlight the sensitivity of storm surge simulations to both atmospheric forcing and hydrodynamic configuration.

Despite these differences, the results indicate that the hindcast developed in this work performs comparably to existing Mediterranean and global storm surge datasets at the selected validation sites. The resulting 45-year, consistently validated dataset can therefore be reliably used for statistical analyses, such as the estimation of return levels and return periods of storm surge extremes, and will serve as a robust basis for the future projections presented in the following chapters.

4.2 Extreme events characterization

After assessing the overall performance of the hindcast and comparing it with previous Mediterranean storm surge datasets, the analysis now focuses on the characterization of extreme events. Understanding where and when the largest storm surges occur is essential for quantifying present-day coastal hazard and for supporting coastal management of Mediterranean cities.

To provide an overview of the spatial and seasonal patterns of extremes, Figure 4.17 shows the seasonal maximum storm surge levels across the Mediterranean Sea over the 45-year period, while Figure 4.18 presents the mean of the annual maxima over the same period.

As expected, the most intense storm surges occur predominantly during autumn and especially winter, when meteorological conditions favor the development of deep cyclonic systems. This seasonal pattern is consistent with previous studies showing that atmospherically induced sea level extremes in the Mediterranean are mainly

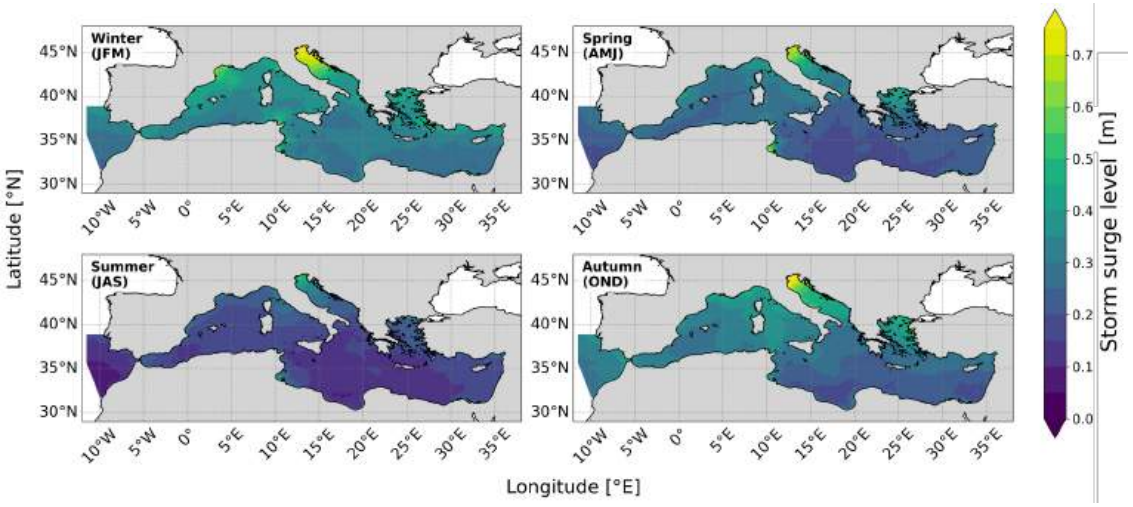

Figure 4.17: Seasonal maximum storm surge levels over the Mediterranean region from 1979 to 2023. Each panel shows the spatial distribution of the seasonal maxima for: (a) Winter (January-March), (b) Spring (April-June), (c) Summer (July-September), and (d) Autumn (October-December).

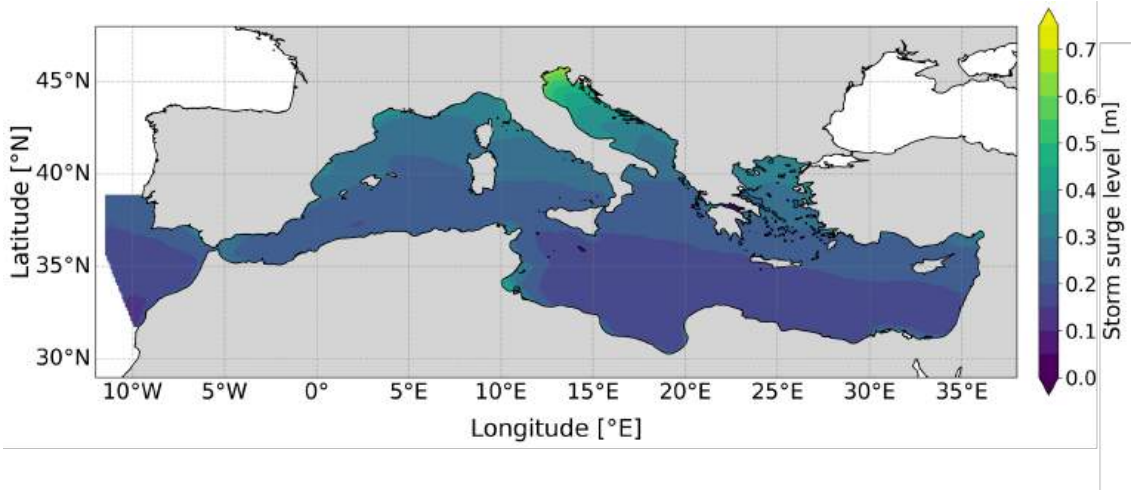

Figure 4.18: Mean of annual maximum storm surges across the entire simulation period (1979-2023).

concentrated in the cold season (e.g., Ferrarin et al., 2022; Marcos et al., 2009; Gomis et al., 2008). Conversely, storm surge activity is generally minimal during summer throughout all Mediterranean sub-basins, reflecting the less energetic atmospheric patterns typical of that season.

Among the Mediterranean sub-regions, the Adriatic Sea (particularly its north-

ern part) exhibits the highest storm surge levels. This agrees with earlier findings attributing the enhanced surge response in the Adriatic to its semi-enclosed morphology, shallow bathymetry, and exposure to southeasterly winds such as the Scirocco (Medugorac et al., 2018; Ferrarin et al., 2022; Lionello and Sanna, 2005), which can significantly amplify water level fluctuations and increase coastal flooding risk.

Beyond the marked seasonal cycle, the maps also reveal robust spatial patterns in the distribution of extremes. In all seasons, the largest storm surge levels are concentrated not only along the northern Adriatic coasts, with values locally exceeding 0.7 m in winter, but also along specific stretches of the north-western Mediterranean, such as the Gulf of Lion and parts of the Spanish coast, especially in the Gulf of Valencia. In contrast, most open-sea areas and the southern and eastern Mediterranean coasts generally experience much smaller maxima, rarely exceeding 0.3-0.4 m, even in winter.

The mean of the annual maxima provides complementary information by smoothing out the influence of single, exceptionally intense events. This field highlights those regions where relatively high storm surges occur repeatedly over the 45-year period, rather than only in isolated years. The persistently elevated values along the northern Adriatic and, to a lesser extent, the Gulf of Lion and northern Aegean Sea indicate that these areas are not only capable of producing the highest surges, but also of doing so with a relatively high frequency, which is particularly relevant for long-term coastal risk assessment.

Taken together, the seasonal maxima and the mean annual maxima provide a first, purely descriptive characterization of present-day storm surge hazard in the Mediterranean, identifying both the times of year and the regions where extreme sea level events are most likely to occur and to have significant coastal impacts.

To further characterize storm surge extremes, Figure 4.19 shows the time series of annual maximum storm surges for the representative tide gauge stations. For

each station, the series of observed (blue) and modeled (orange) annual maxima provides a compact view of how extreme events have evolved over the last decades in terms of both amplitude and temporal occurrence.

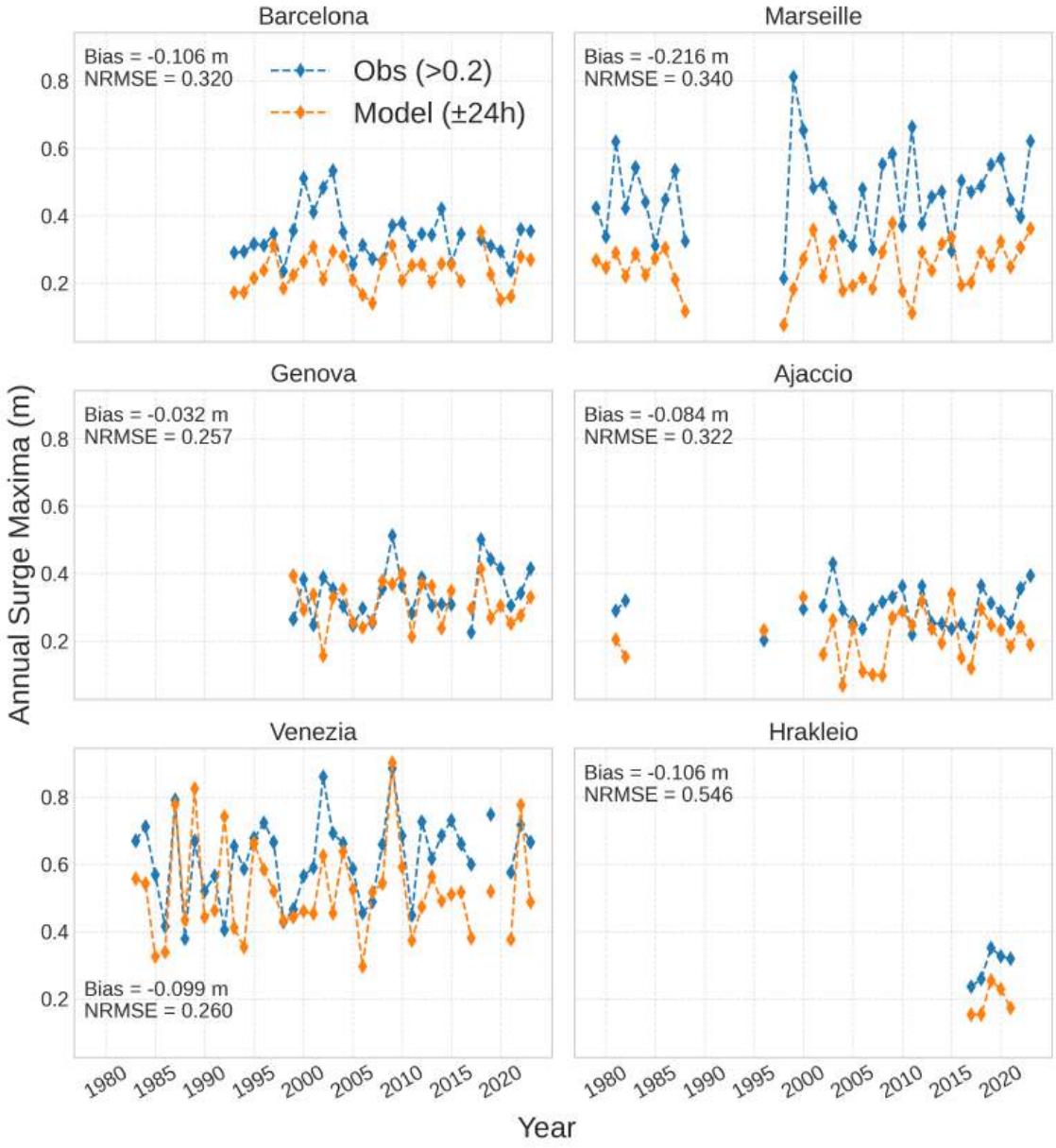

Figure 4.19: Time series of annual maximum storm surge levels for the representative tide gauges. For each station, the bias and NRMSE computed on annual maxima are also reported.

First, the time series highlight the marked differences in data availability among stations. Some gauges provide several decades of annual maxima, while others only cover a limited number of years, which directly affects how well the local statistics of rare events can be constrained.

The time series reveal that all stations experience substantial year-to-year variability in annual maxima, with clusters of particularly energetic years alternating with lower energy periods. This interannual variability is especially pronounced at sites such as Venezia and Marseille, where several winters exhibit peaks well above 0.7-0.8 m, while other years remain closer to 0.4-0.5 m. At the same time, the overall range of annual maxima tends to remain relatively stable over the full record, with no clear long-term trend emerging at these sites, which supports the use of a stationary framework for the subsequent extreme value analysis over the historical period considered.

Extreme value analysis is then performed to estimate storm surge return levels at the reference tide gauges. Figure 4.20 presents the return levels associated with 2-, 5-, 10-, 25-, 50-, and 100-year events, together with 95% confidence intervals obtained using a bootstrap procedure. Repeated bootstrap samples are drawn from the original annual maxima, the parameters of the Gumbel distribution are re-estimated for each resample, and the corresponding return levels computed. The resulting ensemble of estimates is used to derive uncertainty bands, defined as the 2.5^{th} and 97.5^{th} percentiles of the bootstrap distribution of return levels.

As expected, the uncertainty bands widen with increasing return period, reflecting the limited number of extreme events available to constrain very rare events. In all panels, return levels increase approximately linearly in semi-log space, indicating that the Gumbel distribution provides a reasonable fit to the upper tail of the surge distribution at these sites. For Marseille and Venezia, return levels approach or exceed 1 m for a 100-year event, confirming the particularly high hazard in these

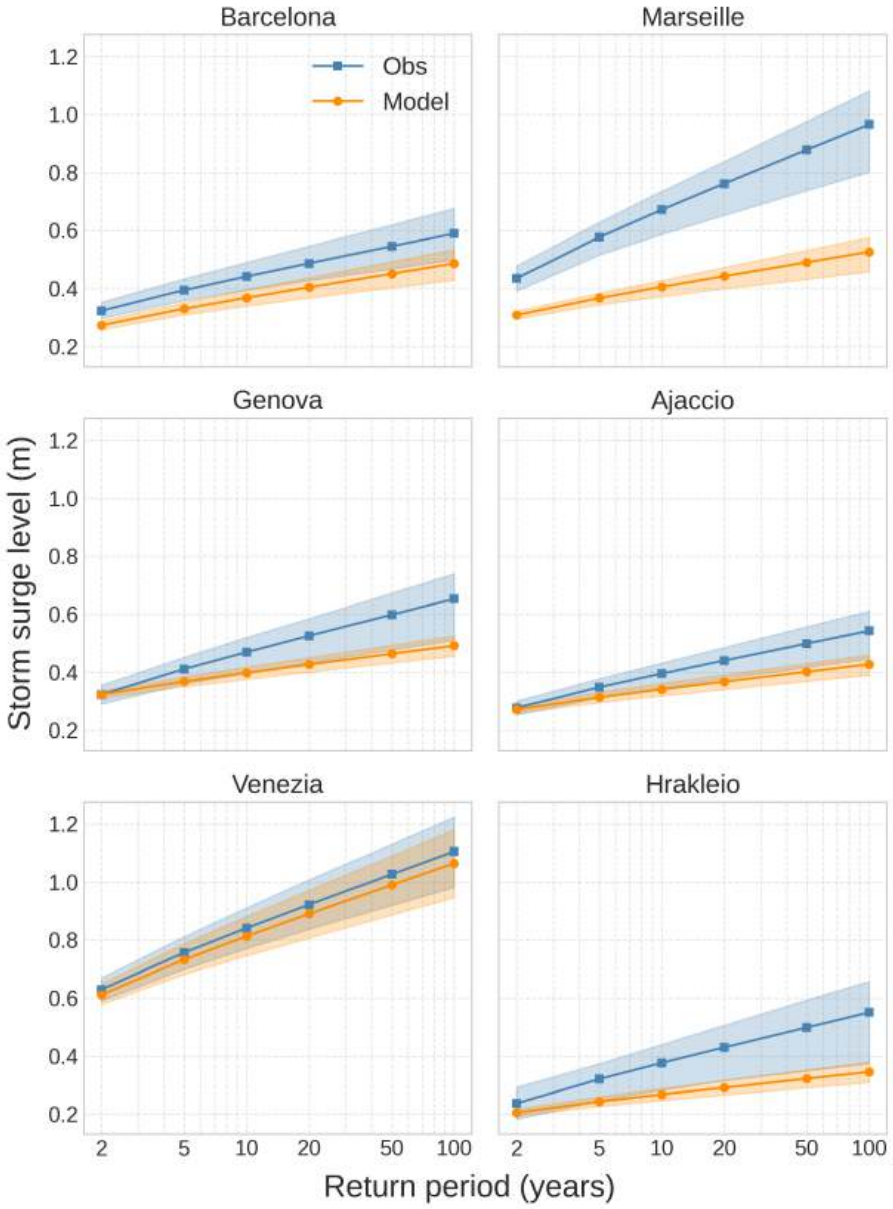

Figure 4.20: Estimated storm surge return levels for 2-, 5-, 10-, 25-, 50-, and 100-year events at six tide gauges. Shaded areas indicate 95% confidence intervals derived from a bootstrap procedure.

locations, while Ajaccio and Hrakleio exhibit markedly lower extremes, consistent with their more sheltered exposure.

In the modeling framework, return levels and their associated confidence intervals are estimated from 45 annual maxima, corresponding to the full 1979-2023 simu-

lation. In contrast, observational records often span a shorter period, resulting in fewer annual maxima. This difference in sample size has two main implications: model-based return level curves tend to be smoother and less steep, and their confidence bands are generally narrower than those derived from observations. Rather than indicating an intrinsically better behavior, this reflects the greater statistical stability of a longer record of extremes and highlights the added value of the hindcast for quantifying rare events that are poorly sampled in the tide gauge data.

Following the same procedure, return periods are then evaluated across the entire Mediterranean basin, as shown in Figure 4.21. The maps display the spatial distribution of storm surge return levels for 10-, 50-, and 100-year events, together with the values estimated at tide gauge locations, represented by the diamonds.

The maps confirm and refine the patterns already suggested by the analysis of seasonal maxima: the northern Adriatic stands out as the region with the highest return levels, with 100-year surges locally exceeding 0.7 m, while elevated values are also found along parts of the north-western Mediterranean (e.g. Gulf of Lion and Spanish coasts) and, to a lesser extent, in some sectors of the northern Aegean and Levantine Seas. In contrast, most North African and eastern Mediterranean coasts exhibit substantially lower return levels, typically below 0.3–0.4 m even for the 100-year event.

By combining point-wise return levels at tide gauges with their continuous spatial extension from the model, the figure provides a basin-scale view of present-day storm surge hazard. This information forms the basis for subsequent risk-oriented analyses and for interpreting future projections of storm surge extremes against a well-characterized historical baseline.

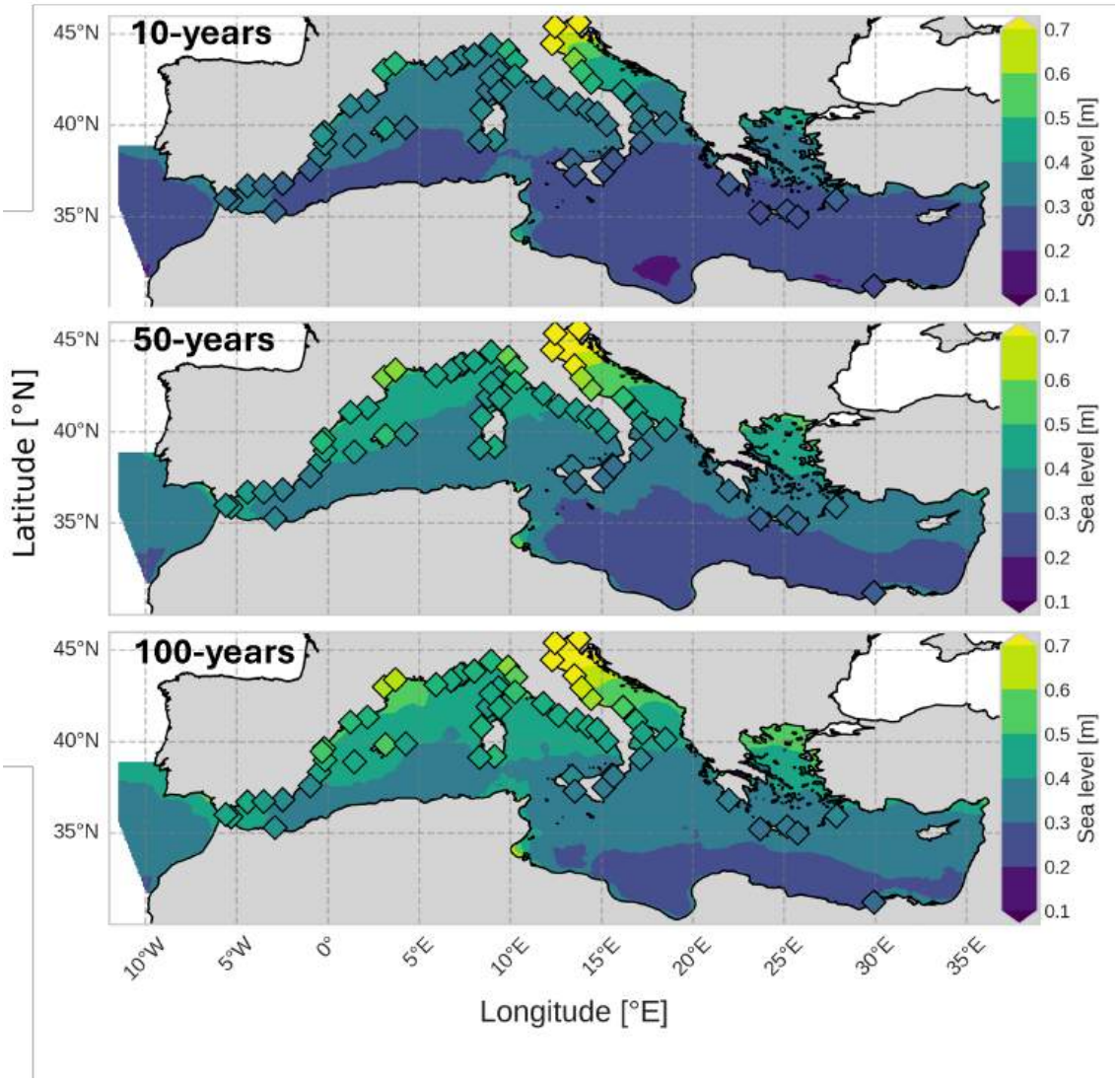

Figure 4.21: Spatial distribution of storm surge return levels for 10-, 50-, and 100-year events across the Mediterranean. Diamonds indicate tide gauge locations.

4.3 Case studies

The hydrodynamic model configuration employed for the validated hindcast is here applied within an event-based framework to simulate selected historical storms. This approach allows for the reconstruction of individual events, highlighting the spatio–temporal evolution of the surge and its sensitivity to local bathymetry and

coastal morphology, which are key factors modulating surge propagation and coastal water level response.

Although the Ligurian Sea in northwestern Italy is not typically considered among the Mediterranean regions most prone to extreme storm surges, it frequently experiences intense cyclonic systems and strong winds capable of generating significant sea level anomalies. The selected case studies, which occurred in 1955 and 2018, represent such conditions, where extreme meteorological forcing led to unusually high surge levels, causing notable damage and disruption along the Ligurian coast.

Atmospheric forcing for these simulations is derived from ERA5 reanalysis data, dynamically downscaled using the WRF model (Ferrari et al., 2020) to horizontal resolutions of 10 km and 3.3 km, enabling an assessment of the impact of forcing resolution on the simulated surge.

The use of ERA5 in this context differs from the hindcast framework, where WRF is forced by CFSv2 data. This choice is motivated by data availability constraints, as CFSv2 does not cover the 1955 event considered in this study, while ERA5 provides continuous global coverage over the entire analysis period. For consistency in the modelling framework applied to both case studies, ERA5 is therefore adopted as a common large-scale atmospheric forcing dataset for WRF downscaling.

Each case study is analyzed under varying spatial resolutions by comparing four model configurations, obtained by combining two hydrodynamic resolutions (3 km, RES_3, and 10 km, RES_10) with the two atmospheric forcing resolutions (3 km, ATM_3, and 10 km, ATM_10).

4.3.1 The 2018 storm event in the Ligurian Sea

The first case study is the intense storm that affected the Ligurian Sea at the end of October 2018, producing severe coastal impacts and marked storm surge levels along several stretches of the coast. The event rapidly intensified on 29th October

and was characterized by a short but extremely energetic phase, which resulted in widespread damage and substantial economic losses in several coastal cities and communities. As discussed by Cavaleri et al. (2022, 2019), the exceptional impact of this storm was due to the concurrence of multiple drivers, including an intense low-pressure system, strong wind fields, and very large significant wave heights.

In this work, the event is revisited with a specific focus on the storm surge component, with the aim of complementing the primarily atmospheric and wave-oriented description provided by Cavaleri et al. (2022). In particular, the surge model allows reconstruction of the spatial structure and temporal evolution of the sea level response during the most intense phase of the storm, providing additional information on where and when the largest surge anomalies occurred within the Ligurian basin.

Figure 4.22 summarizes the synoptic-scale evolution of mean sea level pressure during the event, based on WRF simulations at 10 km resolution, highlighting the rapid deepening and the position of the cyclone over the northwestern Mediterranean. It clearly shows that the low-pressure center transits directly over the Ligurian Sea and remains in the area for several hours during 29$^{\text{th}}$ October moving northward. This persistence implies that the Ligurian basin was exposed for a sustained period to both strong pressure gradients and the associated cyclonic circulation, creating favorable conditions for a pronounced storm surge response. In particular, the passage of the pressure minimum over the basin is consistent with the rapid growth of surge levels observed during the afternoon hours, when the event reaches its peak intensity. Consistently, Figure 4.23 shows the wind field associated with the storm, based on WRF simulations at 10 km resolution. Intense winds affected the Ligurian Sea and surrounding areas on 29$^{\text{th}}$ October, creating favorable conditions for a significant wind-driven setup and a coherent basin-scale surge response.

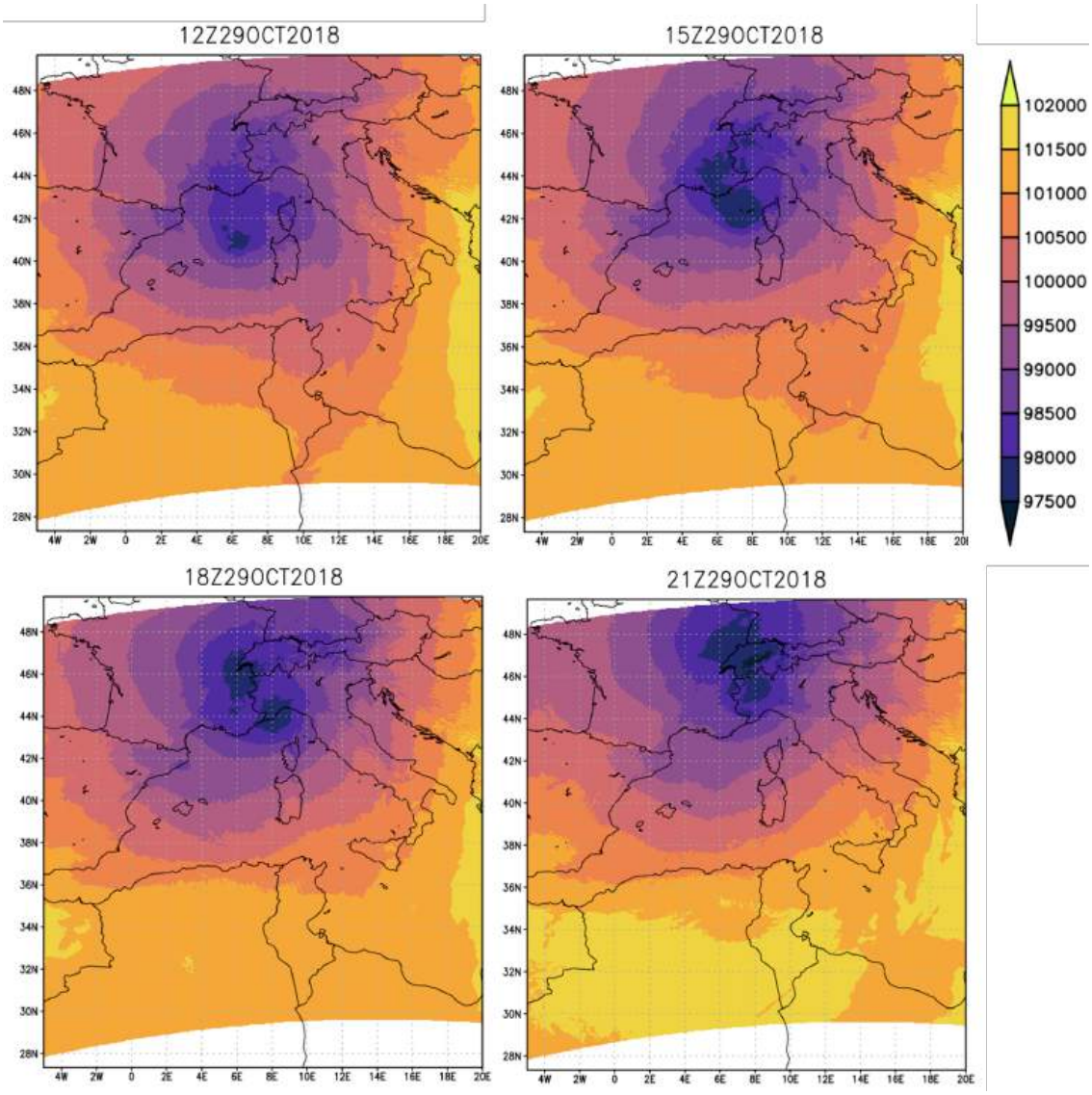

Figure 4.22: Synoptic evolution of mean sea-level pressure during the 29th October 2018 storm event, derived from ERA5 data downscaled with the WRF model at 10 km horizontal resolution, showing the position and intensification of the cyclone over the northwestern Mediterranean.

The storm surge model provides detailed information on storm surge conditions within the Ligurian basin during the peak phase of the event (Figure 4.24). The maps indicate a rapid growth of surge levels between 16:00 and 18:00 UTC on 29th October, with coastal values locally exceeding 0.5 m, followed by a partial decay by 19:00 UTC. The spatial pattern suggests that the largest surge anomalies are

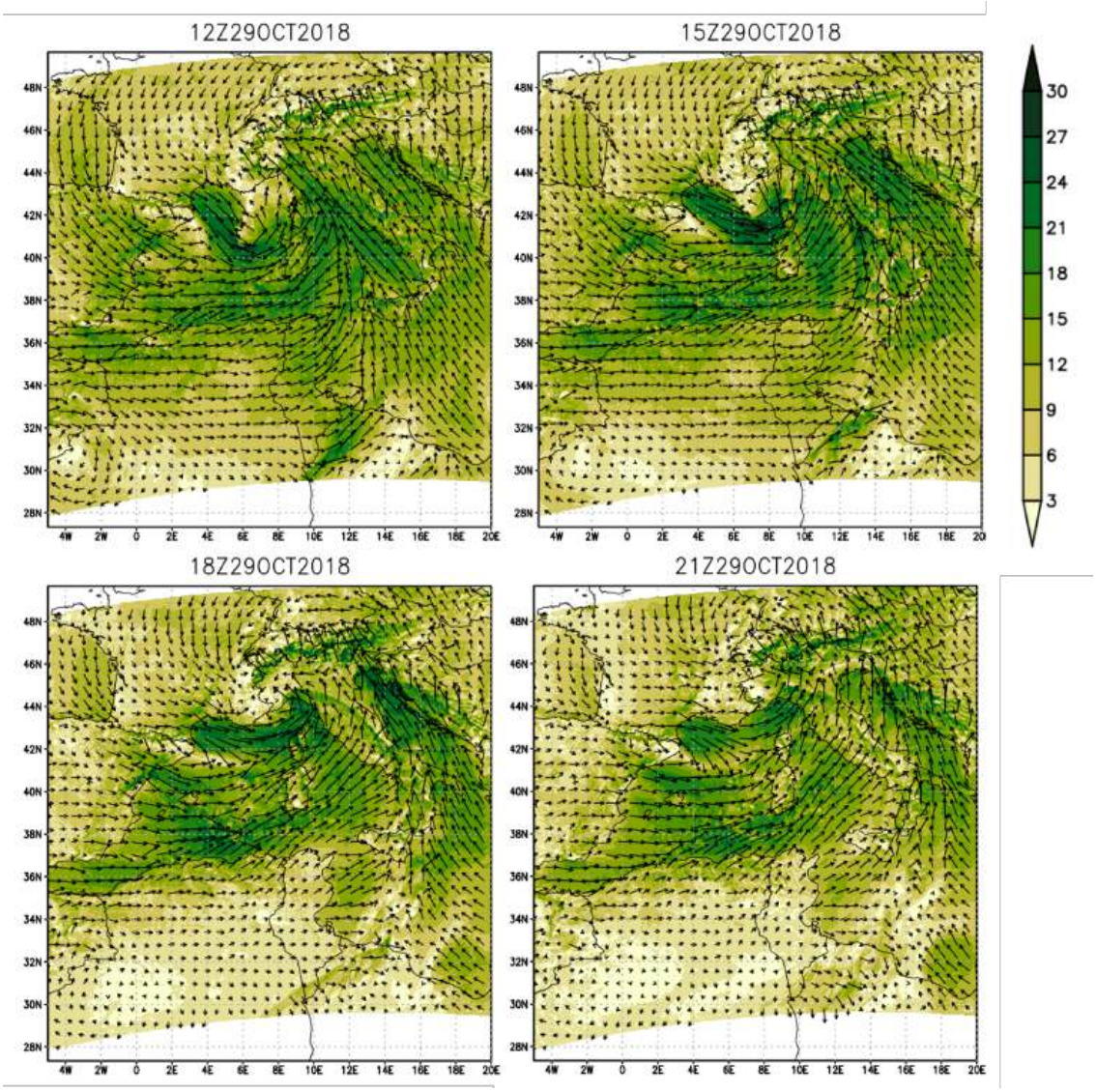

Figure 4.23: Synoptic evolution of the 10-m wind field during the 29th October 2018 storm event, derived from ERA5 data downscaled with the WRF model at 10 km horizontal resolution. Arrows indicate wind direction, while colours represent wind speed (m s^{-1}).

concentrated along the Ligurian arc, with a clear coastal intensification.

A station-based view of the event is provided in Figure 4.25, which compares modeled and observed storm surge time series at Nice, Imperia, Genova, and Livorno over October 2018, together with the corresponding scatter plots. Beyond confirming the overall consistency between modeled and observed variability during the

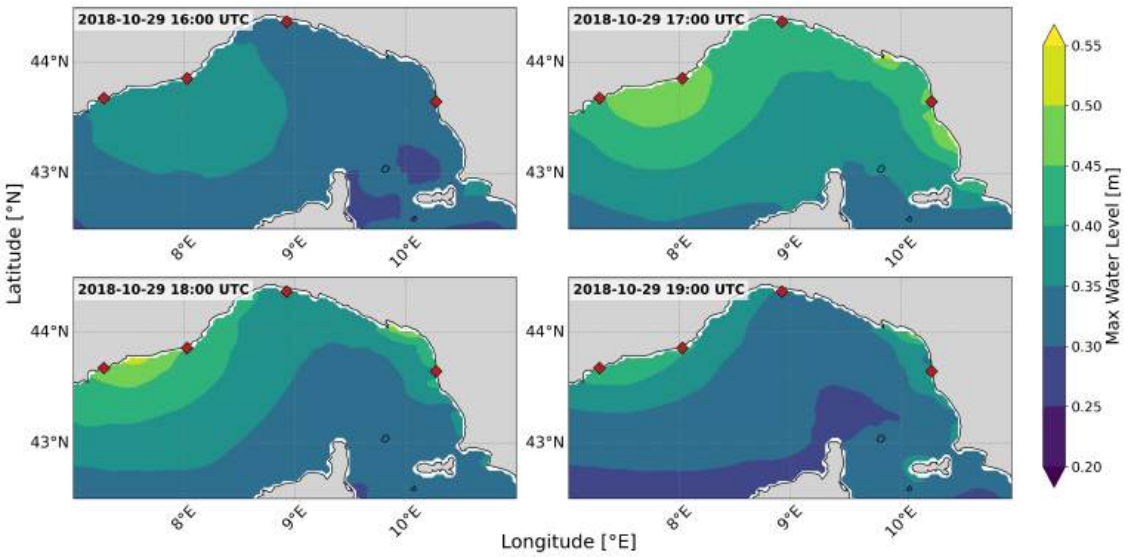

Figure 4.24: Storm surge levels during the peak phase of the 29th October 2018 Ligurian storm event. Red diamonds indicate tide gauges with available observations (Nice, Imperia, Genova, and Livorno).

storm period, these panels allow the timing and amplitude of the peak response to be interpreted in a local context. In particular, the largest surge is associated with the rapid intensification around 29th October, while the preceding and subsequent days exhibit substantially lower surge levels, highlighting the rapid evolution of the phenomenon.

The comparison reveals that the model is able to accurately reproduce the general shape and evolution of the observed sea level anomalies. This is further confirmed by the high correlation coefficients presented in the scatter plots, which exceed 0.8 for all four stations, indicating a strong agreement between the modeled and observed data during the storm period.

This good agreement suggests that the developed model can be used with a high degree of confidence to characterize storm surge events in the region. In the case of the October 2018 storm, the model output contributes to a more complete and detailed understanding of the event dynamics, complementing analyses based on atmospheric and wave data. Such reconstructions are particularly valuable for post-

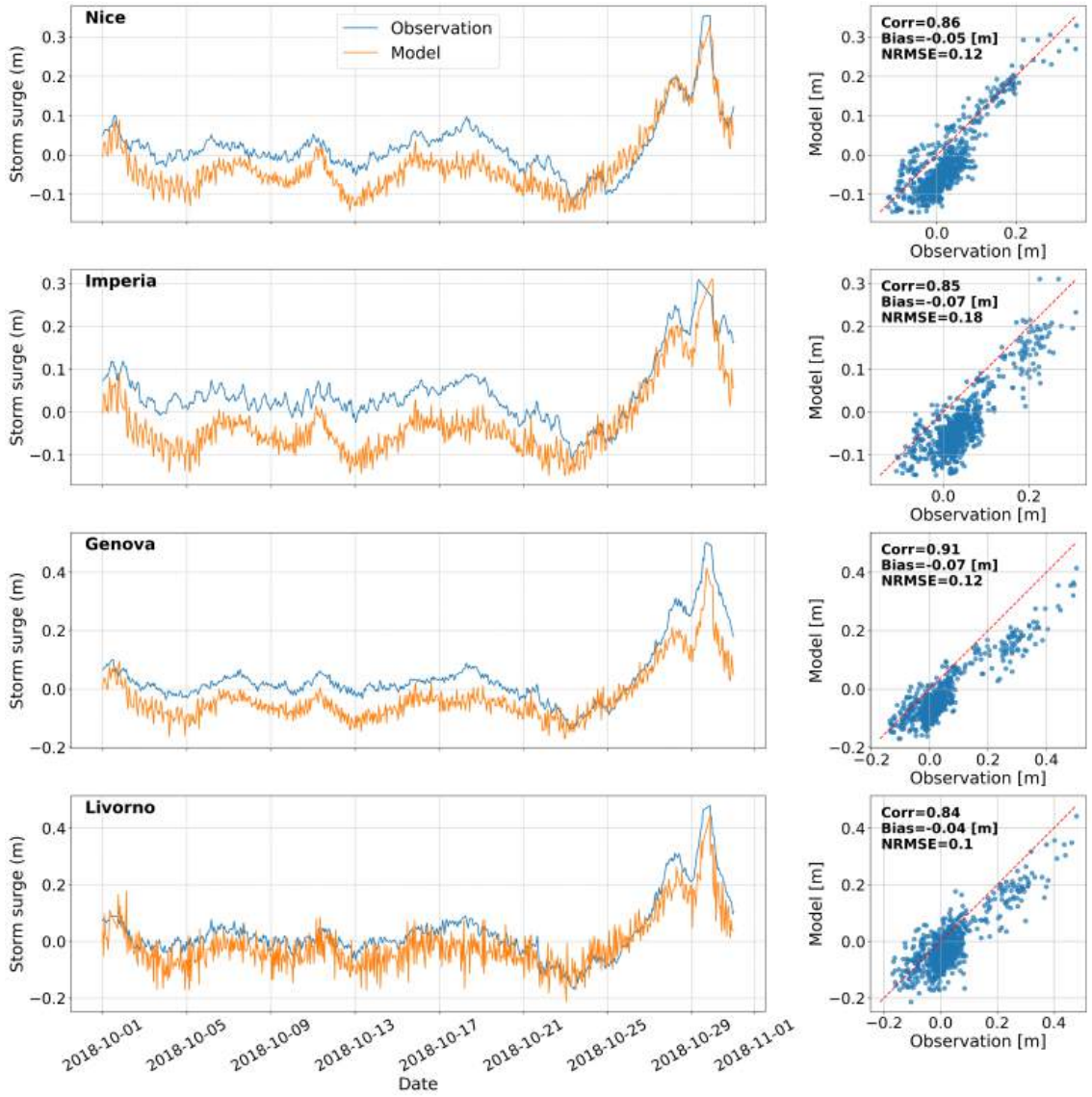

Figure 4.25: Time series of observed (blue) and modeled (orange) storm surge levels (left column) and corresponding scatter plots (right column) during the month of October 2018. In each scatter plot the statistical indexes previously mentioned are reported.

event assessments and can support improved planning and risk mitigation strategies for future coastal storms.

Finally, Figure 4.26 provides a comparative view of the simulated storm surge under the different modeling setups described previously, highlighting the influence

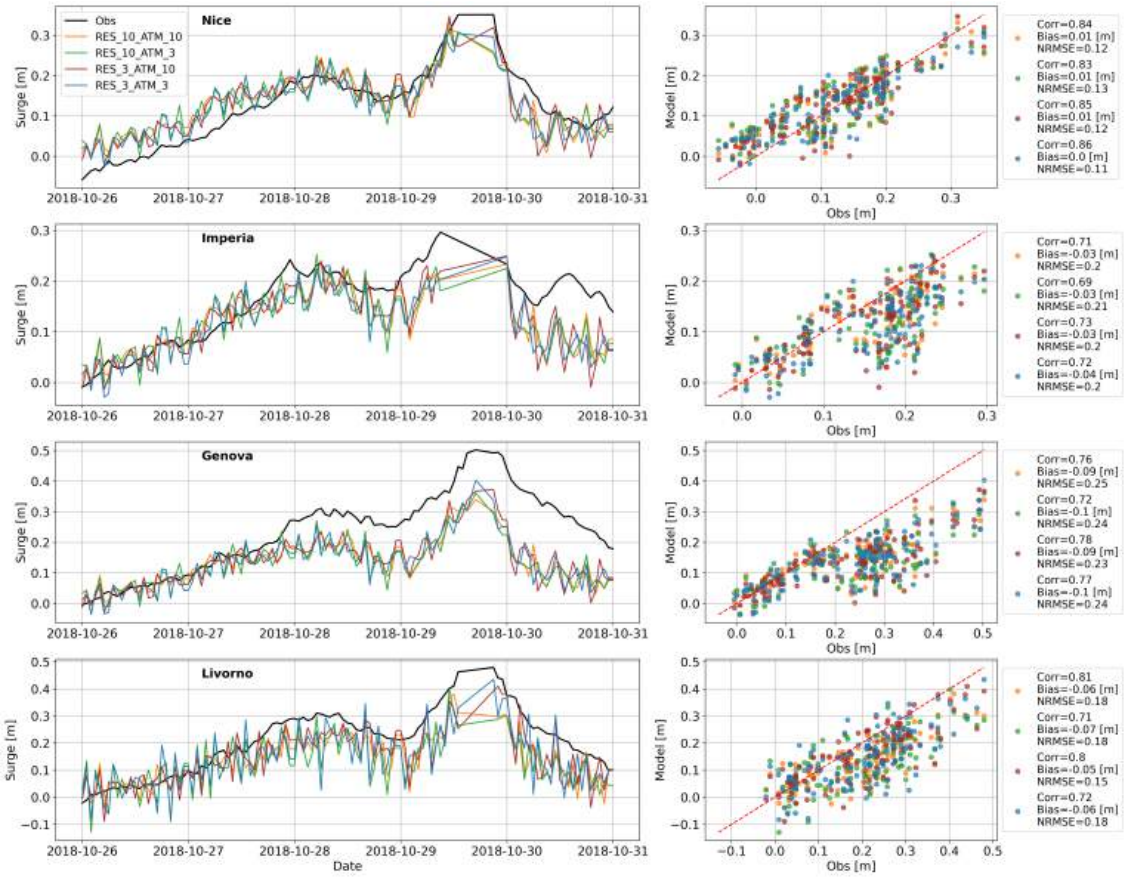

Figure 4.26: Comparison of observed storm surge time series (black) with four model configurations at selected Ligurian tide gauges during the October 2018 storm. The configurations combine two hydrodynamic resolutions (RES_3 and RES_10) with two atmospheric forcing resolutions (ATM_3 and ATM_10).

of both model and forcing resolutions on the results. For each configuration, a dedicated event-specific simulation is performed over a common 5-day period centered on the storm, encompassing the full build-up, peak, and decay phases. This approach ensures strict consistency across all configurations, as the same time window, boundary conditions, and model settings are adopted, with only the hydrodynamic and forcing resolutions varying. Consequently, differences among the four solutions can be confidently attributed to the chosen model and forcing resolutions rather than to inconsistencies in the simulation setup.

For each station, the four simulations reproduce a broadly similar temporal evolution, but with noticeable differences in peak magnitude and, in some cases, in peak timing. Overall, the configurations forced by the higher-resolution atmospheric fields tend to better capture the sharp intensification on 29th October, suggesting that the representation of wind and pressure gradients is a key factor for reproducing this event. Differences between RES_3 and RES_10 appear smaller for most of the period, indicating a weaker sensitivity to hydrodynamic grid resolution at these stations compared to the sensitivity to atmospheric forcing resolution.

Among the tested setups, RES_3_ATM_3 provides the best agreement for this specific event. However, the use of 3-km atmospheric forcing requires substantially higher computational resources and data storage and is therefore not feasible for long-term, basin-wide applications. In the present study, the 3-km forcing is obtained through a dedicated dynamical downscaling specifically targeting the October 2018 storm and is not available continuously over the full 45-year period.

Despite its coarser atmospheric resolution, the RES_3_ATM_10 configuration is still able to satisfactorily reproduce the overall temporal evolution and magnitude of the storm surge associated with the event. For this reason, RES_3_ATM_10 is selected as the final configuration for the Mediterranean-wide, long-term hindcast, representing a practical compromise between accuracy and computational cost while ensuring temporal consistency and homogeneous forcing availability over the 1979-2023 period. The higher-resolution atmospheric setup remains a valuable option for the analysis of selected extreme events or for short-term forecasting applications, where computational constraints are less limiting.

4.3.2 The event of the 1955

As a second case study, an historical storm that affected the Ligurian region on 18th February 1955 is analyzed.

The analysis of the 1955 storm is performed using the same hydrodynamic setup adopted in this study. Given the lack of direct observational constraints, the results are intended to provide a qualitative reconstruction of the associated storm surge dynamics rather than a fully validated simulation.

The model is run for a dedicated 5-day simulation window centered on the storm, using boundary conditions and numerical settings consistent with the reference configuration. Simulations are performed adopting different spatial resolutions for both the hydrodynamic model and the atmospheric forcing, following the experimental design described in the previous sections, in order to assess the sensitivity of the reconstructed storm surge to model and forcing resolution.

Figures 4.27 and 4.28, both based on the 10 km atmospheric forcing, indicate a developed synoptic phenomenon over northern Italy and the northwestern Mediterranean, associated with a strong pressure gradient and intense winds over the basin. Compared with the 2018 case study, the 1955 storm appears to be characterized by a broader-scale pressure anomaly and a more spatially extended wind field, suggesting a more basin-wide forcing rather than a sharply localized peak.

As for the 2018 storm, the sensitivity to model configuration is evaluated by comparing simulations performed with different hydrodynamic grid resolutions and atmospheric forcing resolutions (Figure 4.29). In this case, the four configurations produce very similar surge evolution at Nice, Imperia and Genova, with only modest differences in peak magnitude, suggesting that the event is primarily controlled by the large-scale atmospheric forcing and that resolution effects play a secondary role for this specific storm. In contrast, larger differences emerge at Livorno, where higher short-term variability indicates a stronger influence of local coastal geometry or the representation of nearshore processes.

The spatial distribution of the simulated storm surge during the core phase of the event is shown in Figure 4.30. Surge levels are elevated across most of the Ligurian

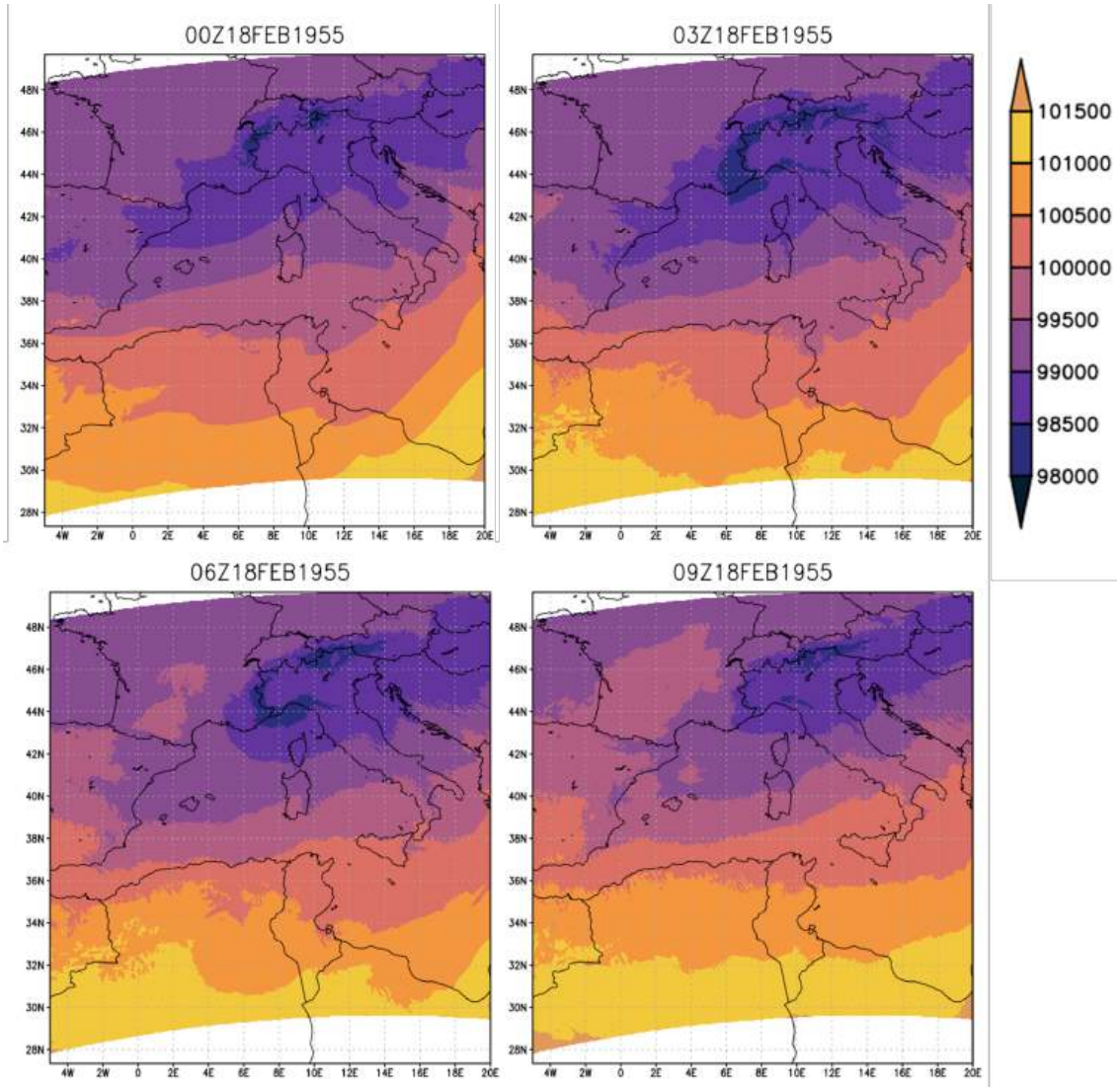

Figure 4.27: Synoptic evolution of mean sea-level pressure during the 18 February 1955 storm event, derived from atmospheric forcing at 10 km spatial resolution.

basin at 00:00-06:00 UTC, reaching values around 0.30-0.40 m along portions of the coast, and then decreasing by 09:00 UTC. Compared with the October 2018 event, the February 1955 case exhibits lower peak surge values over the Ligurian arc and a less localized coastal maximum. At the same time, elevated surge levels appear to persist for a longer time window, as relatively high values are maintained over several consecutive snapshots, suggesting a more sustained basin-wide response

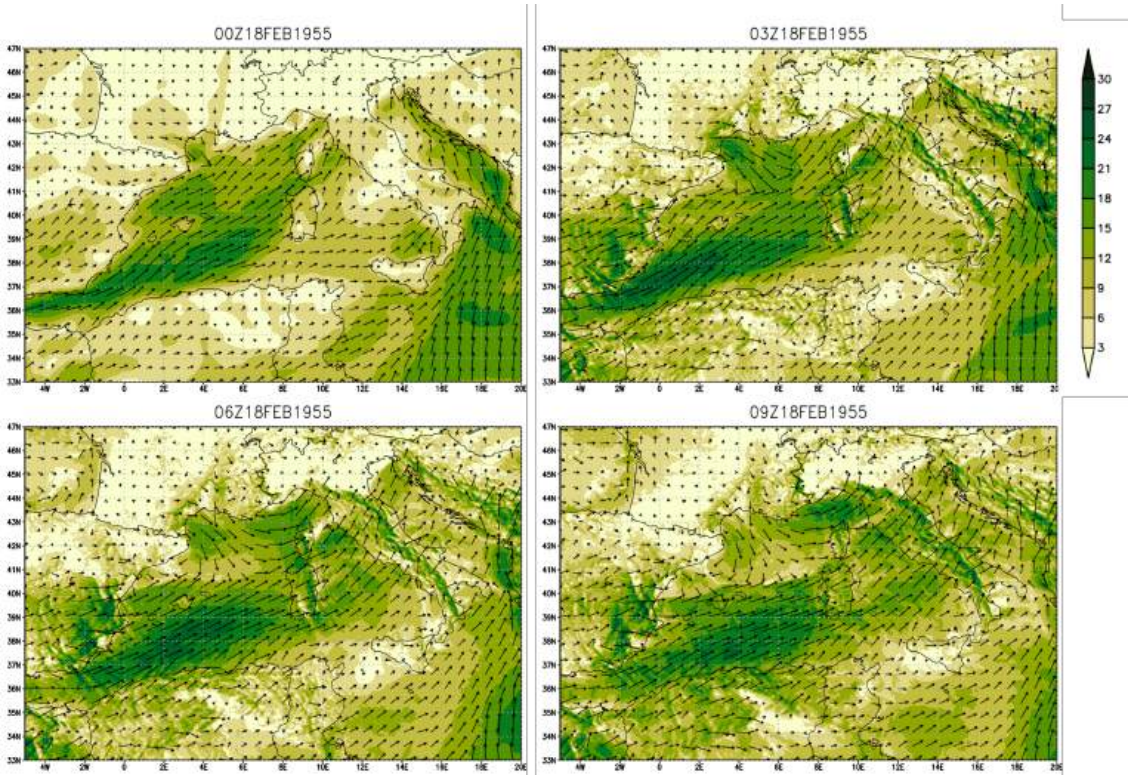

Figure 4.28: Synoptic evolution of the 10-m wind field during the 18 February 1955 storm event, derived from atmospheric forcing at 10 km spatial resolution. Arrows indicate wind direction, colours represent wind speed (m s^{-1}).

rather than a short-lived, sharply peaked event.

Overall, this second case study complements the 2018 event by illustrating how different synoptic settings can lead to distinct storm surge responses in the Ligurian Sea, in terms of both peak amplitude and spatial structure. In addition, it highlights the flexibility of the modeling framework: event-scale reconstructions can be performed consistently with the validated hindcast setup even outside the 1979-2023 period, provided that suitable atmospheric forcing is available.

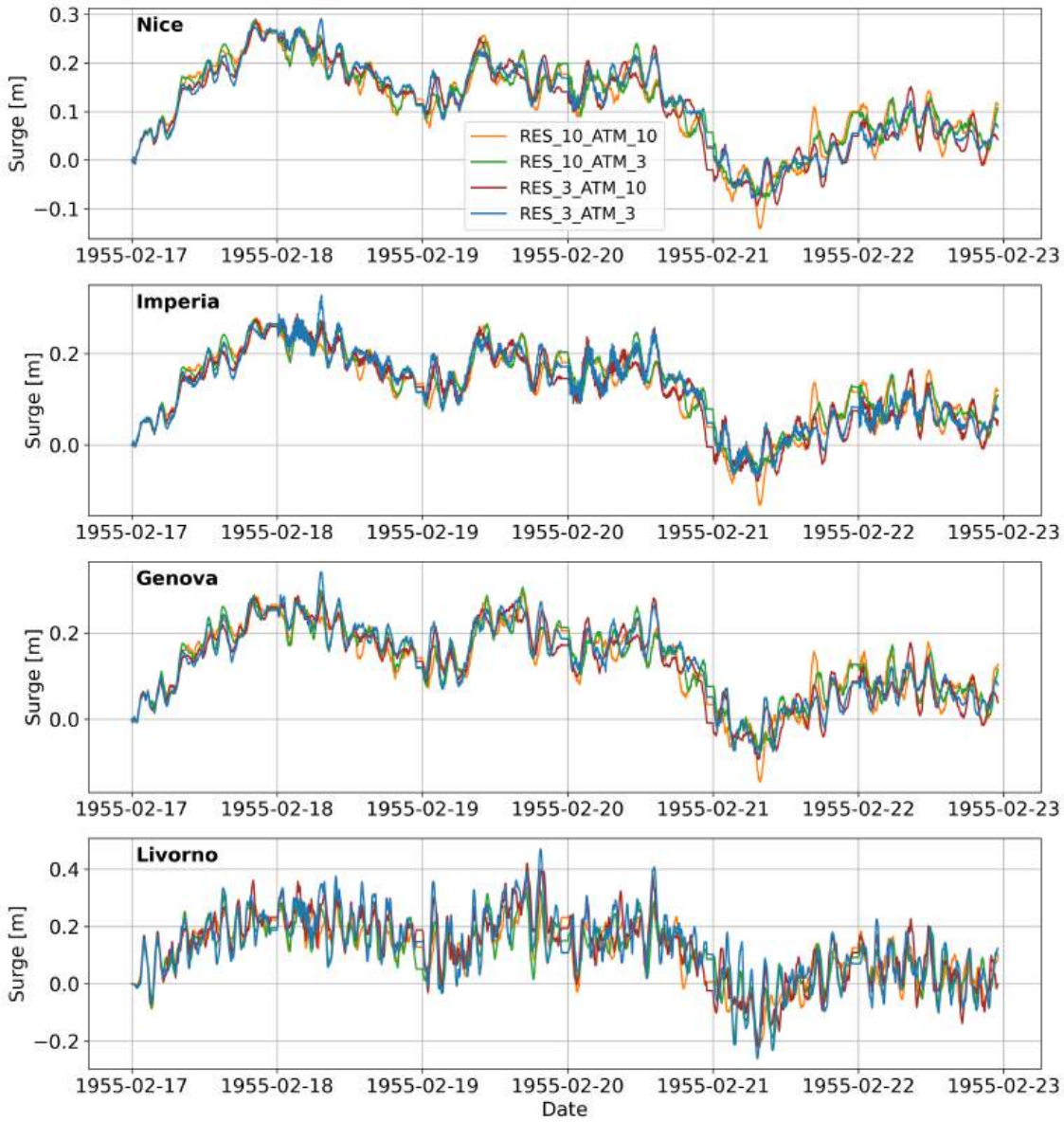

Figure 4.29: Modeled storm surge time series at selected Ligurian and nearby coastal locations during the February 1955 storm event, comparing the four RES/ATM sensitivity configurations.

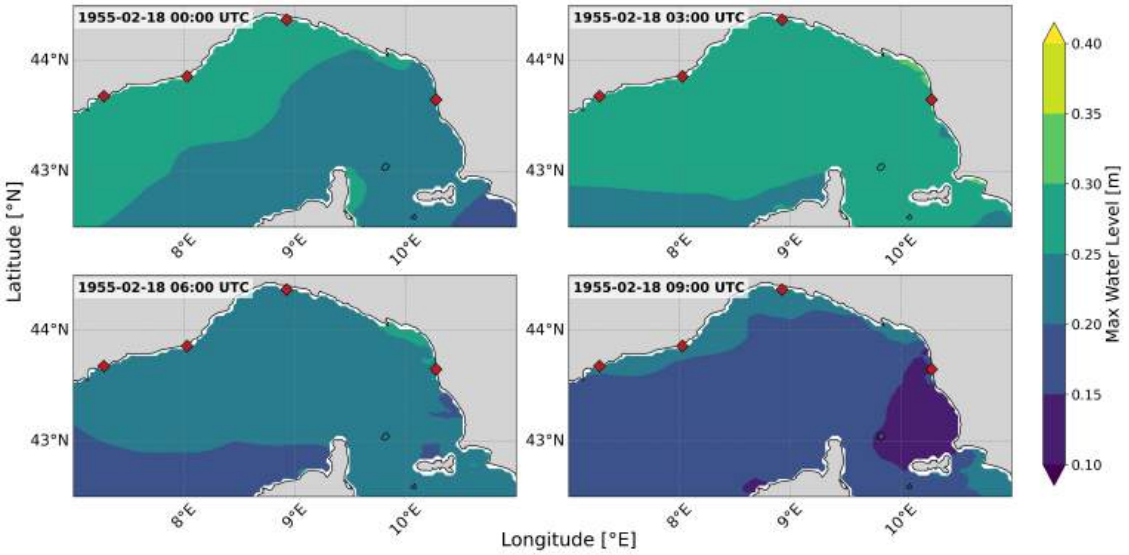

Figure 4.30: Maps of storm surge levels during the storm event that hit the Ligurian region on 18 February 1955. The red diamonds indicate the locations of the cities of Nice, Imperia, Genoa, and Livorno.

5

Results - Projections

5.1 Projection framework and raw outputs

The development of an extensive dataset of future storm surge projections represents a key outcome of this PhD research. In order to investigate how storm surge levels may evolve under different future atmospheric conditions, a total of 17 distinct meteorological forcings are employed, yielding 17 corresponding storm surge projections up until the end of the century.

As introduced in the methodological framework, all projections are based on the RCP8.5 climate change scenario and include two main temporal segments: a baseline (historical) period spanning from 1970 until 2005 and a future RCP8.5 period covering from 2030 until 2100. Although the available climate forcing would have allowed the projected simulations to be extended continuously from 2006 onward, the computational cost associated with running the projections was already considerable

so, at this stage, priority was given to increasing the number of projections included in the ensemble and to characterize future climate from mid-century and onward, rather than extending each simulation without interruption. The current temporal window starting in 2030 still supports analyses at mid-century and end-century horizons (approximately 2050 and 2100) which are presented in the following sections, while the completion of the missing years will be addressed in future developments.

A prerequisite for quantifying future changes is the availability of a reliable reference dataset. For this purpose, the hindcast dataset developed and validated in the previous chapters provides the baseline against which the baseline simulations of each climate projection can be evaluated and, after correction, against which future changes can be expressed. Using the same hydrodynamic model framework for both hindcast and projections ensures methodological consistency and reduces the risk that projected differences arise from the hydrodynamic model configuration inconsistencies rather than from changes in atmospheric forcing.

In the projections analysis, particular emphasis is placed on annual maxima storm surge levels, because extremes are the most relevant for coastal hazard assessment and for the design of protection and adaptation measures. Figure 5.1 shows, for each of the 17 projections, the spatial distribution of the mean annual maxima over the end-century period (2074-2100).

The maps are arranged such that projections in the first column share the same RCM (RCA4) while differing in the driving GCM. Projections in the second column are based on the HIRHAM5 RCM, whereas those in the third column employ either the COSMO-crCLIM-v1-1 or the CCLM4-8-17 RCMs. For clarity, each panel is labeled with the corresponding GCM–RCM combination, following the nomenclature and model pairings illustrated in Figure 3.3 and detailed in Appendix A. This layout allows a qualitative assessment of the influence of the large-scale climate signal from the GCM compared to the downscaling effects introduced by the RCM.

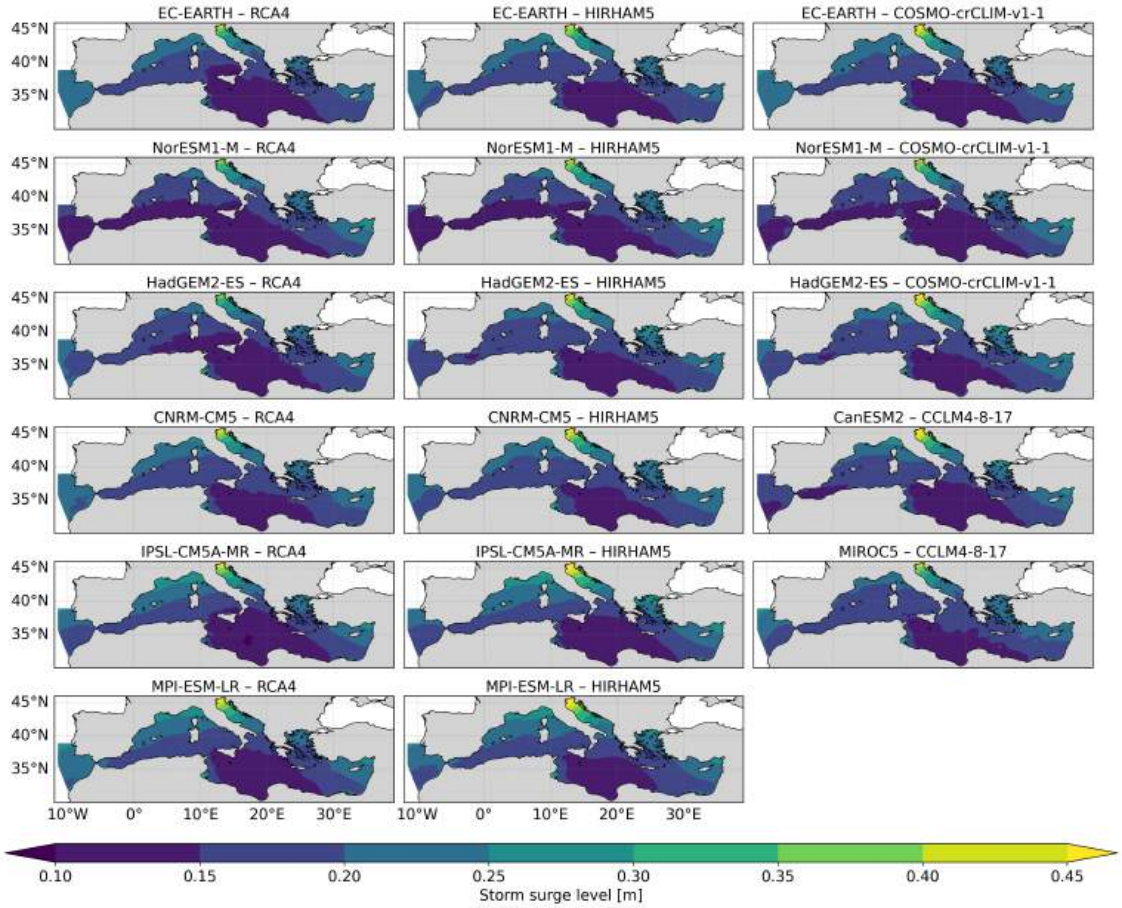

Figure 5.1: Mean annual maxima storm surge levels for the end-century period (2074-2100) for each of the 17 EURO-CORDEX driven projections. Each map is labeled with the corresponding GCM-RCM combination, following the nomenclature and model pairings illustrated in Figure 3.3 and detailed in Appendix A.

Overall, Figure 5.1 reveals consistent spatial distribution between the different projections in terms of the mean annual maxima patterns and magnitude. In all cases, the largest values (up to approximately 0.5 m) occur in the northern Adriatic, consistently indicating this region as the primary hot-spot for storm surge extremes within the Mediterranean basin following the same spatial patterns of the hindcast data (Figure 4.18). Other areas, which reach values of the order of 0.3 m, include sectors of the Ligurian Sea, North-Tyrrhenian Sea, Gulf of Lion, and North-Levantine Sea, while most of the North African coast exhibits the lowest maxima

values, with the notable exception of the Gulf of Gabes where extreme surges appear more pronounced (reaching 0.25-0.3 m).

To better highlight the projected changes associated with each GCM-RCM combination, Figure 5.2 shows the difference between the end-century mean annual maxima (2074-2100) and the corresponding historical values (1979-2005), computed separately for each projection. These difference maps allow identification of areas where annual maxima increase or decrease locally, and whether consistent patterns emerge across the basin and the different models.

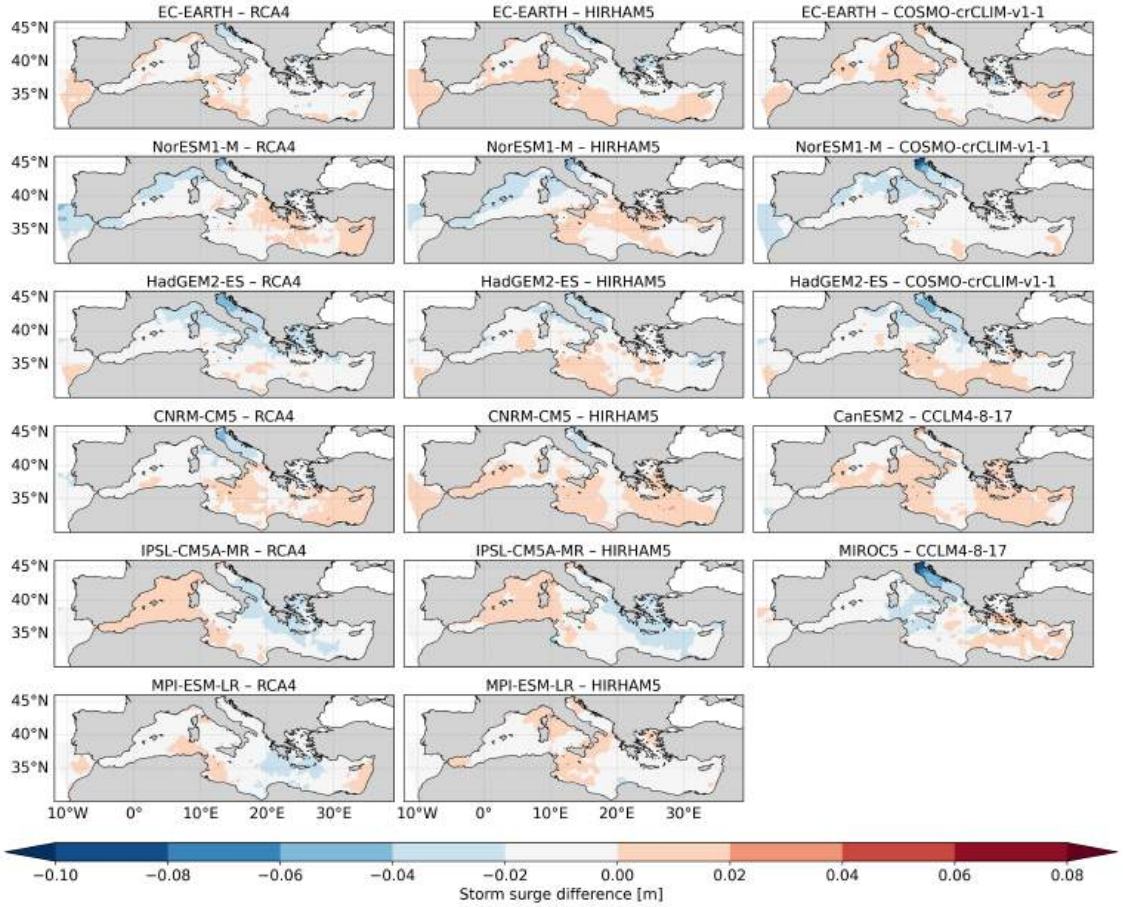

Figure 5.2: Projected change in mean annual maximum storm surge levels for each of the 17 projections, computed as the difference between 2074-2100 and 1979-2005. Each map is labeled with the corresponding GCM-RCM combination, following the nomenclature and model pairings illustrated in Figure 3.3 and detailed in Appendix A.

No basin-wide uniform trend is evident: none of the projections indicates a spatially homogeneous increase or decrease across the entire Mediterranean Sea. Instead, changes are regionally heterogeneous, with different sub-basins exhibiting opposite tendencies within the same projection. Importantly, the magnitude of the projected changes in mean annual maxima is generally limited, typically within approximately ± 0.10 m.

Despite the generally modest magnitude of the projected changes, Figure 5.2 highlights notable differences among projections. For example, the projection driven by the CanESM2 GCM tends to exhibit predominantly positive anomalies across several regions, whereas the projection driven by MIROC5 shows predominantly negative anomalies, particularly in the Adriatic Sea. Since these projections share the same RCM but differ in the GCM, this contrast suggests that the driving GCM may exert a stronger influence on the resulting storm surge changes than the RCM used for downscaling.

This interpretation is further supported by the general tendency of projections from the same GCM to display more similar spatial patterns than projections sharing the same RCM. For instance, both projections driven by IPSL-CM5A-MR suggest an increase in storm surge maxima in the western Mediterranean and a decrease in the eastern basin, largely independent of whether RCA4 or HIRHAM5 RCM is used. Likewise, the three projections driven by HadGEM2-ES consistently show an increase in the southern part of the basin and a decrease in the northern sectors, again largely independent of the selected RCM.

Besides assessing the differences of the developed projections, since they all refer to a single future scenario, it is useful to analyze them jointly, in order to summarize the ensemble-mean behavior and to assess whether coherent and spatially consistent signals of change emerge. To this end, mean annual maximum estimated by each projection, treated as an independent realization of the future climate, are combined

following an ensemble-based approach. This is conceptually consistent with the assessment frameworks adopted in recent climate assessments (e.g., IPCC AR6) when synthesizing information from multiple projections. Figure 5.3 shows the multi-model ensemble-mean change in mean annual maximum storm surge levels, computed as the unweighted average of the 17 projection-specific differences (2074–2100 minus 1979–2005). The ensemble-mean pattern is weak over most of the basin and does not indicate a coherent Mediterranean-wide increase or decrease, with only localized areas exhibiting small positive or negative anomalies. The limited amplitude of the ensemble-mean signal is consistent with the heterogeneous and often opposite-signed responses across individual projections shown in Figure 5.2, which partly compensate when averaged.

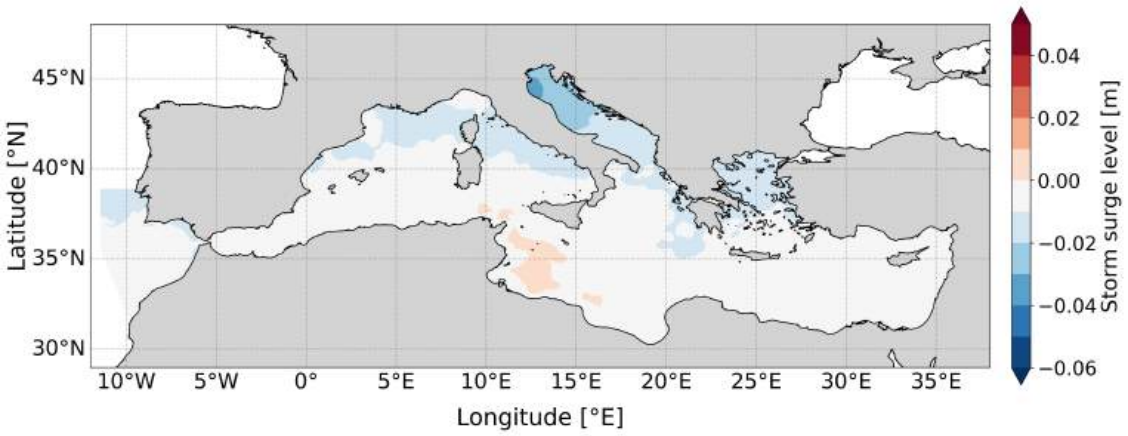

Figure 5.3: Multi-model ensemble-mean projected change in mean annual maximum storm surge levels, computed as the average of the 17 projection differences (2074–2100 minus 1979–2005).

5.2 Bias adjustment: Gumbel vs GEV

The changes discussed above describe differences between the historical and end-century segments within each climate projection, as well as in the ensemble mean, but they do not yet quantify variations relative to the validated hindcast reference.

This step is essential because each GCM-RCM combination may exhibit systematic biases in the historical period, which would affect the interpretation of projected changes and any subsequent extreme value analysis.

Figure 5.4 shows the bias in annual maxima between the historical simulations of each GCM-RCM combination and the hindcast dataset over their common period (1979-2005), computed as the difference between each projections minus the hindcast. The dominance of negative values indicates that, overall, the projections underestimate annual maxima with respect to the hindcast reference over the same period. The bias is generally modest (within approximately ± 0.05 m across most of the basin), but larger negative biases occur in some localized regions and for specific projections, notably in the Adriatic Sea, the northern Aegean, and in some cases the Alboran Sea.

In order to evaluate future changes consistently with respect to the hindcast reference and to ensure comparability across all projections, it is therefore necessary to correct these systematic differences. For this reason, a bias correction is applied following the quantile-mapping approach described in Section 3.3.1.

As described in the methodology, bias correction is applied only to the annual maximum storm surge values, and it is performed independently at each grid point and for each climate projection. Consequently, all post-correction analyses are carried out on annual maxima only. An example of the bias adjustment procedure is shown in Figure 5.5. For a randomly selected projection (NorESM1-M - RCA4; projection n.7 in Appendix A) and an arbitrarily chosen grid point located in the Ligurian Sea, the figure illustrates how annual maxima are corrected through a quantile-mapping approach.

Bias correction is carried out by matching the distribution of annual maximum storm surge values from the historical segment of each climate projection with the corresponding distribution from the hindcast reference. This is done by building the

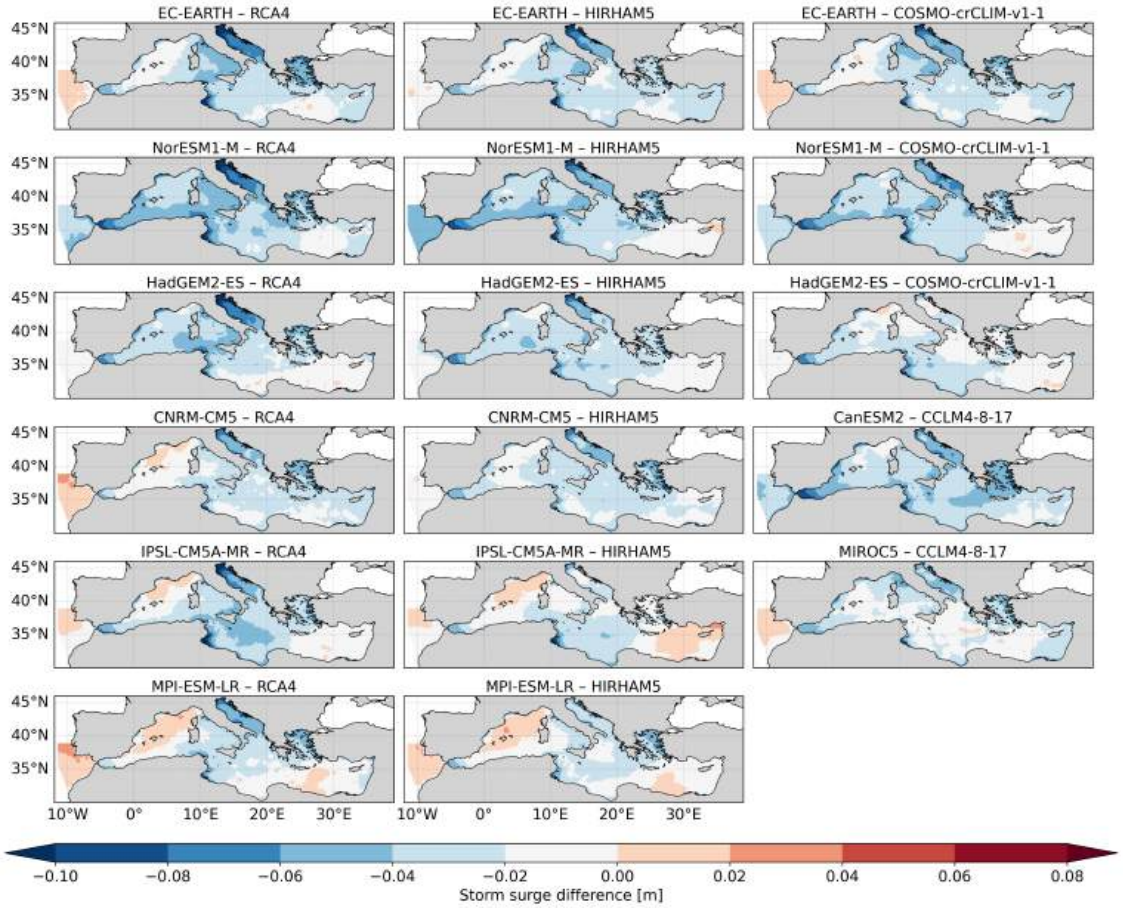

Figure 5.4: Bias in mean annual maximum storm surge levels between the historical segment of each GCM-RCM combination and the hindcast reference over the common period 1979-2005 (computed as historical minus hidncast).

empirical CDF of the historical annual maxima from a given projection and then fitting it with two alternative parametric extreme-value distributions (Gumbel and Generalized Extreme Value, GEV) in order to assess the differences arising from the choice of distribution. The same fitting procedure is applied to the hindcast annual maxima, so that both datasets are represented by smooth, comparable cumulative distributions. Once these two distributions are available, each historical annual maximum from the projection is translated into its percentile with respect to the fitted historical distribution. The corrected value is obtained by taking the same

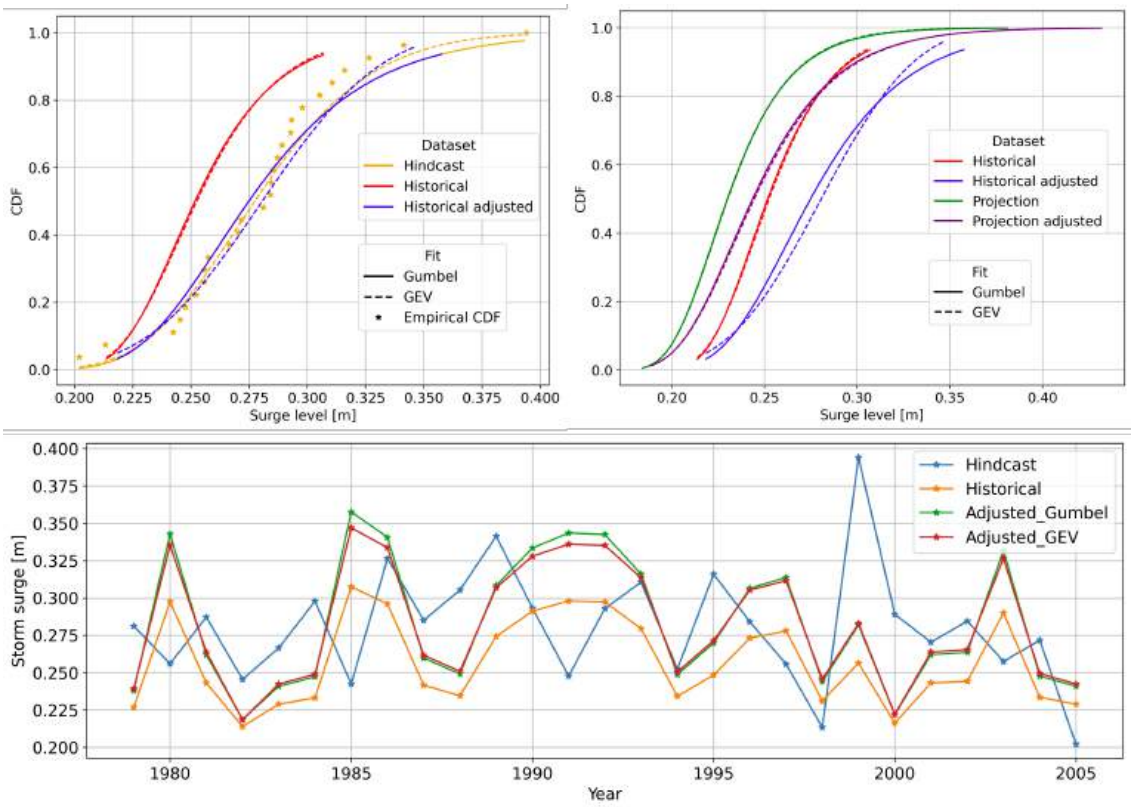

Figure 5.5: Example of the annual maximum bias adjustment (quantile-mapping) at a representative grid point in the Ligurian Sea for a selected EURO-CORDEX-driven projection (NorESM1-M–RCA4; n.7 in Appendix A). Top row: cumulative distribution functions (CDFs) of annual maximum storm surge levels. The left panel compares the hindcast reference with the model historical segment before and after adjustment using both Gumbel and GEV parametric fits. The right panel illustrates the transfer of the estimated bias from the historical to the future segment under the assumption of bias stationarity. Bottom: annual-maximum time series for 1979-2005 comparing hindcast, raw historical simulation, and bias-adjusted historical series obtained using the Gumbel- and GEV-based corrections.

percentile in the fitted hindcast distribution and reading the corresponding storm surge level. Repeating this operation for all historical annual maxima produces a bias adjusted historical series whose distribution is aligned with the hindcast reference, while preserving the rank structure of the original values.

For the future projections, the same transfer function is applied under the assumption of bias stationarity, i.e., that the bias estimated in the historical period remains valid under future conditions.

It is worth noting that the fitted CDFs are built using all available annual maxima within each time window. In the example of Figure 5.5, the hindcast CDF is based on 45 values (1979-2023), the historical projection CDF on 36 values (1970-2005), and the future projection CDF on 71 values (2030-2100). Because these sample sizes differ, the fitted CDFs (and associated Gumbel/GEV parameters) are affected by different levels of sampling uncertainty, which is typically more pronounced for shorter windows and particularly for the upper tail. This should be considered when interpreting differences between fitted distributions and the resulting bias-adjusted extremes.

Finally, the comparison of time series over 1979-2005 highlights that the bias corrected historical series follows the same year-to-year variability as the raw historical projection, while being shifted so that its distribution matches the hindcast reference.

After completing the bias adjustment procedure over the full model domain and for all 17 projections, it becomes possible to reassess end-century changes in storm surge extremes against a validated reference. In this section, changes are evaluated in terms of the mean of annual maxima, consistently with the bias correction strategy, which is applied only to annual maximum values at each grid point. Figure 5.6 shows, for each projection, the end-century minus historical change in the mean annual maximum storm surge after bias correction, where the annual maxima distributions are fitted using a Gumbel model.

Overall, the magnitude of the projected change remains modest across the Mediterranean basin, with values typically fluctuating by only a few centimeters around zero. More pronounced negative changes are localized in specific areas, most notably the northern Adriatic in several projections. Similarly to the pre-correction differences, the signal is not spatially uniform across the basin: increases and decreases coexist in different sub-regions within the same projection, highlighting a

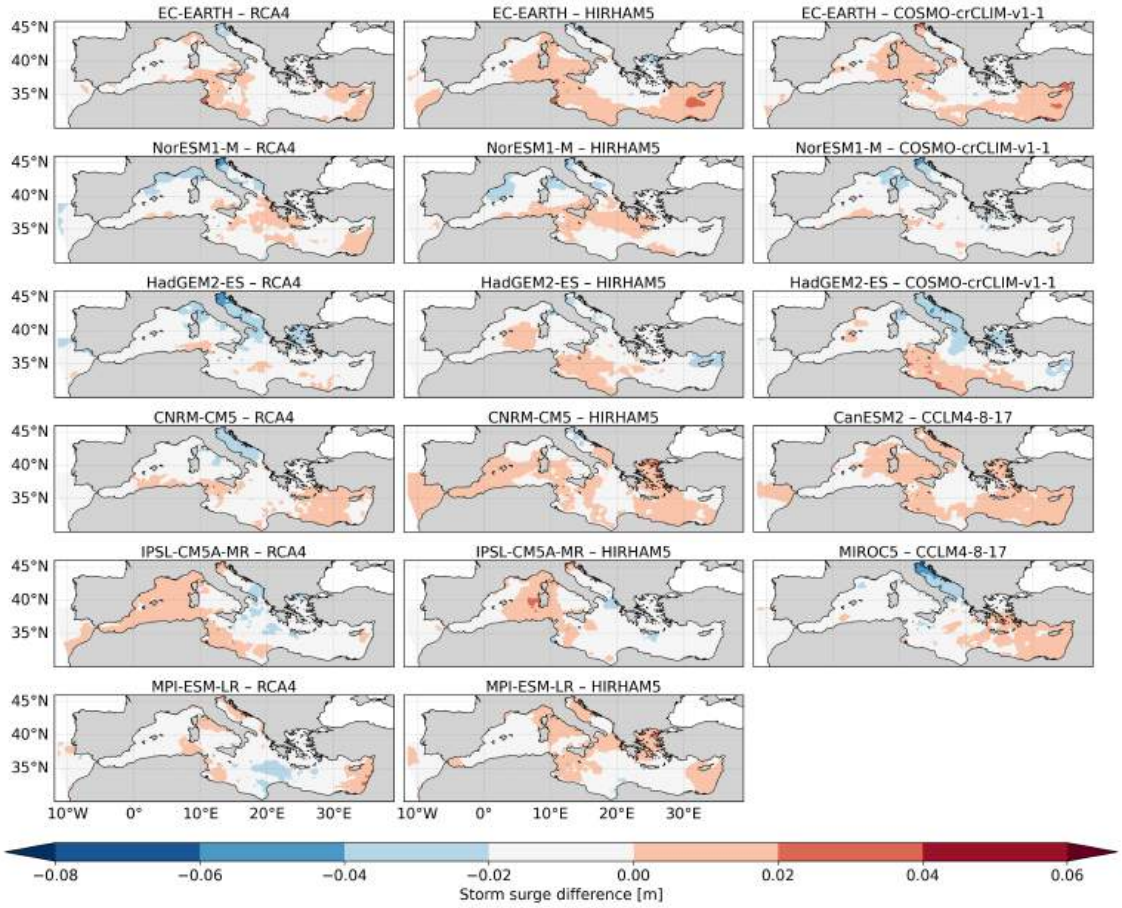

Figure 5.6: Gumbel-bias-corrected changes in mean annual maximum storm surge (end-century minus historical in meters) for each projection.

regional dependence of the projected response.

The same analysis is repeated using a GEV distribution to represent annual maxima during the bias adjustment step. Figure 5.7 indicates that the spatial patterns of end-century changes are very similar to those obtained with the Gumbel fit, suggesting that the large scale structure of the projected differences is not strongly affected by the choice of parametric extreme value model. Minor differences can be observed in the local amplitude of the change, with the GEV-based correction sometimes producing slightly larger positive or negative anomalies in localized areas; nevertheless, the overall scale of variability remains small.

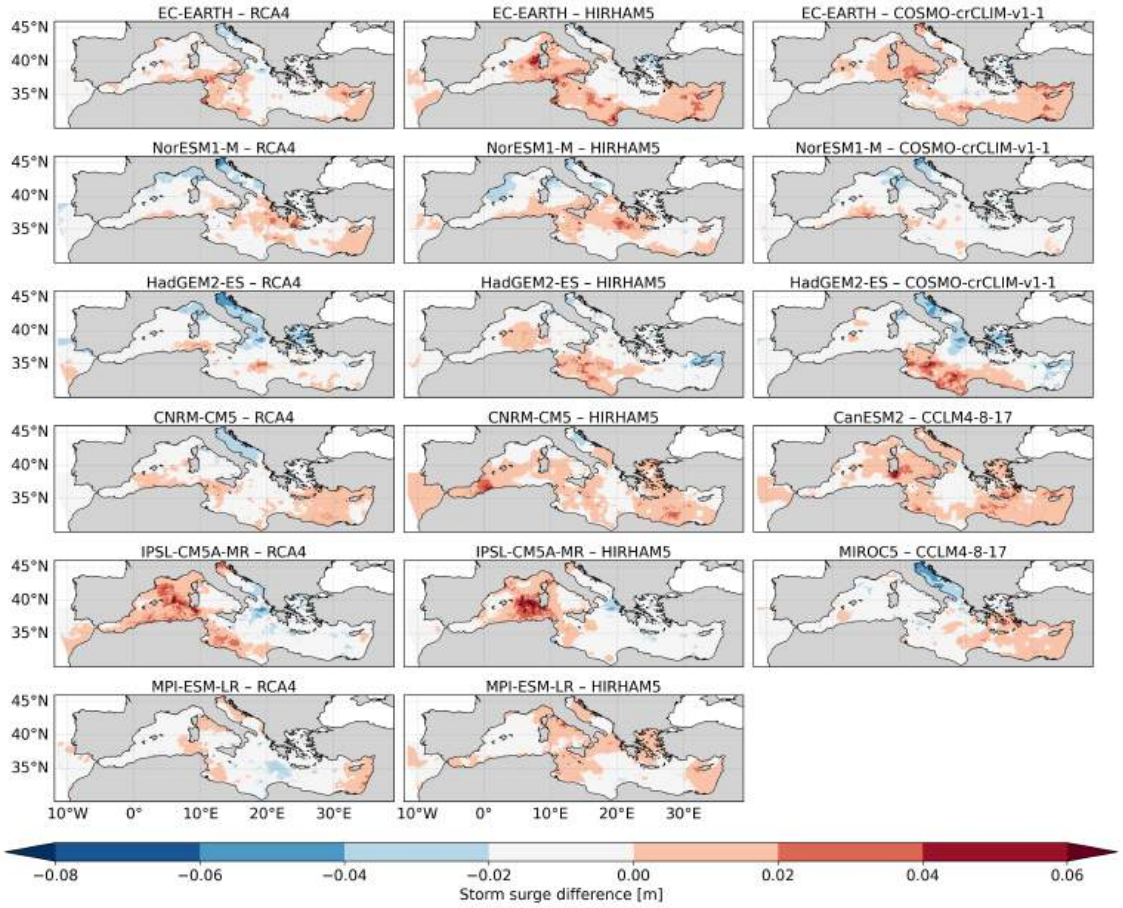

Figure 5.7: Same as Figure 5.6, but using a Generalized Extreme Value (GEV) distribution to fit annual maxima within the quantile-mapping procedure.

To quantify the sensitivity to the chosen distribution, Figure 5.8 reports the difference between the Gumbel-based and GEV-based results. The maps confirm that the two bias correction choices lead to very similar end-century change estimates: differences are generally close to zero and remain within a narrow range of only a few centimeters. This indicates that, for the present application, uncertainty associated with choosing Gumbel versus GEV in the bias correction step is small compared with the spatial variability across the basin and the inter-model spread across projections.

As done previously, in addition to examining the behavior of each model individ-

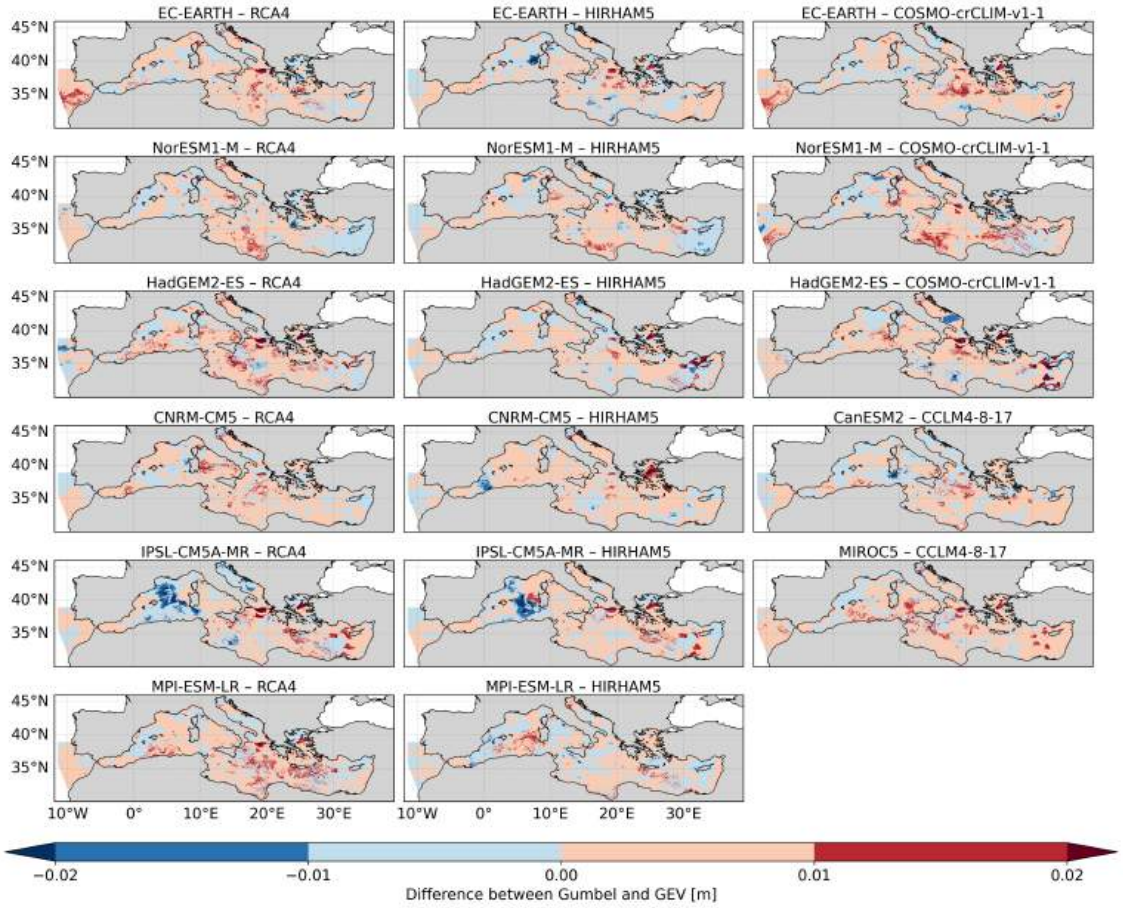

Figure 5.8: Difference between the bias-corrected end-century changes obtained with Gumbel and GEV fits, computed as $\Delta_{\text{Gumbel}} - \Delta_{\text{GEV}}$ (Gumbel minus GEV).

ually, the overall expected change by adopting an ensemble-mean approach is also considered. Figure 5.9 shows the multi-model ensemble-mean bias-corrected change in mean annual maximum storm surge levels, computed as the unweighted average of the 17 projection-specific differences between the end-century and the historical periods, using the Gumbel-based quantile-mapping approach to perform the bias adjustment.

In line with the results related to the changes before the bias correction procedure is applied shown in Figure 5.3, the amplitude of the ensemble-mean change in storm surge maxima is limited, consistent with the spatially heterogeneous and

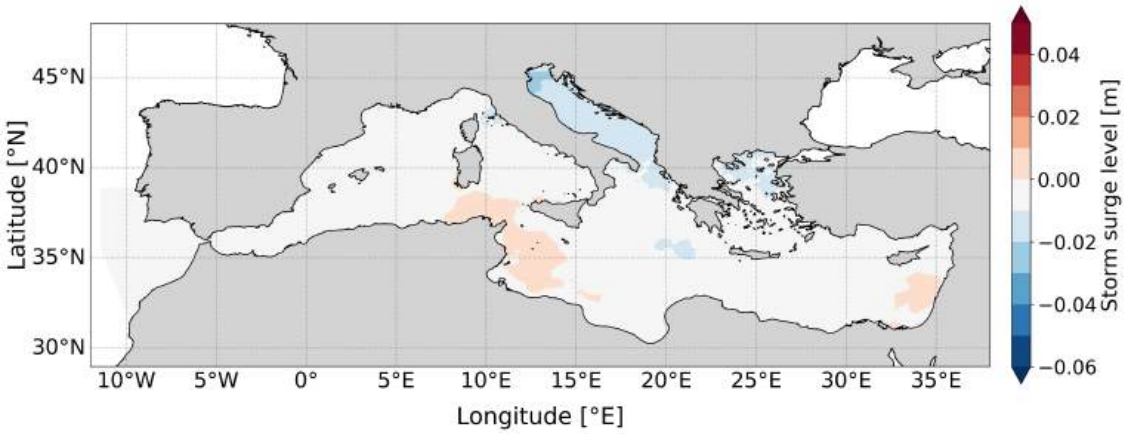

Figure 5.9: Multi-model ensemble-mean bias-corrected change in mean annual maximum storm surge levels, computed as the unweighted average of the 17 projection-specific differences between the end-century (2074-2100) and historical (1979-2005) periods. The bias adjustment is performed using the Gumbel-based quantile-mapping approach, and changes are expressed as end-century minus historical.

often opposite-signed responses across individual projections. Overall, the expected change remains weak across the Mediterranean basin, with a tendency towards a decrease in the northern sectors, particularly in the Adriatic Sea, and also over parts of the northern Aegean and northern Tyrrhenian, while a very modest increase is projected along the Tunisian coast.

Since Gumbel- and GEV-based bias corrections yield nearly identical estimates of the end-century change (Figure 5.8), presenting separate ensemble-mean change maps for both choices would be redundant. Accordingly, Figure 5.9 reports the ensemble-mean bias-corrected change using the Gumbel-based adjustment as a representative example, as very similar spatial patterns and magnitudes are expected for the GEV-based case. Given this limited sensitivity, the subsequent return level analysis is carried out using a single distribution choice for consistency. In the following, return levels are therefore computed based on the GEV formulation, as it provides a more flexible description of the upper tail through its shape parameter and is therefore well suited for extrapolation to long return periods, which are central to the present hazard assessment.

5.3 Future changes in storm surge return levels

Estimating return levels associated with specific return periods is a widely used approach to quantify the magnitude of both relatively frequent and rare events. A return period is defined as the average time between two exceedances of a given level.

This analysis, complementary to the results presented in the previous sections, allows the identification of Mediterranean areas where storm surge levels associated with selected return periods are expected to increase or decrease in the future, thus providing an assessment of potential changes in coastal hazard. In line with the spatially heterogeneous patterns discussed above, no basin-wide uniform signal is expected; instead, sub-regional responses emerge depending on the specific climate projection considered. In particular, where the projected change (future minus historical) is positive in Figures 5.6 and 5.7, the return levels for a given return period are higher than in the historical period, whereas negative changes indicate lower return levels and a potential reduction in hazard.

Figure 5.10 refers to the single climate projection obtained from the combination GCM:CanESM2 and RCM:CCLM4-8-17. It shows storm surge return levels for 10-, 25-, 50-, and 100-year return periods for the historical period of the projection (first column), together with the projected changes at mid-century (second column) and end-century (third column) for the same return periods.

The historical maps (first column) indicate that, for a 10-year return period, storm surge levels are generally below 0.30 m across most of the Mediterranean basin, with a prominent exception in the northern Adriatic, where values locally reach or exceed approximately 0.70 m. Additional areas with relatively higher levels are visible in the northern Aegean, the Gulf of Gabes, and along portions of

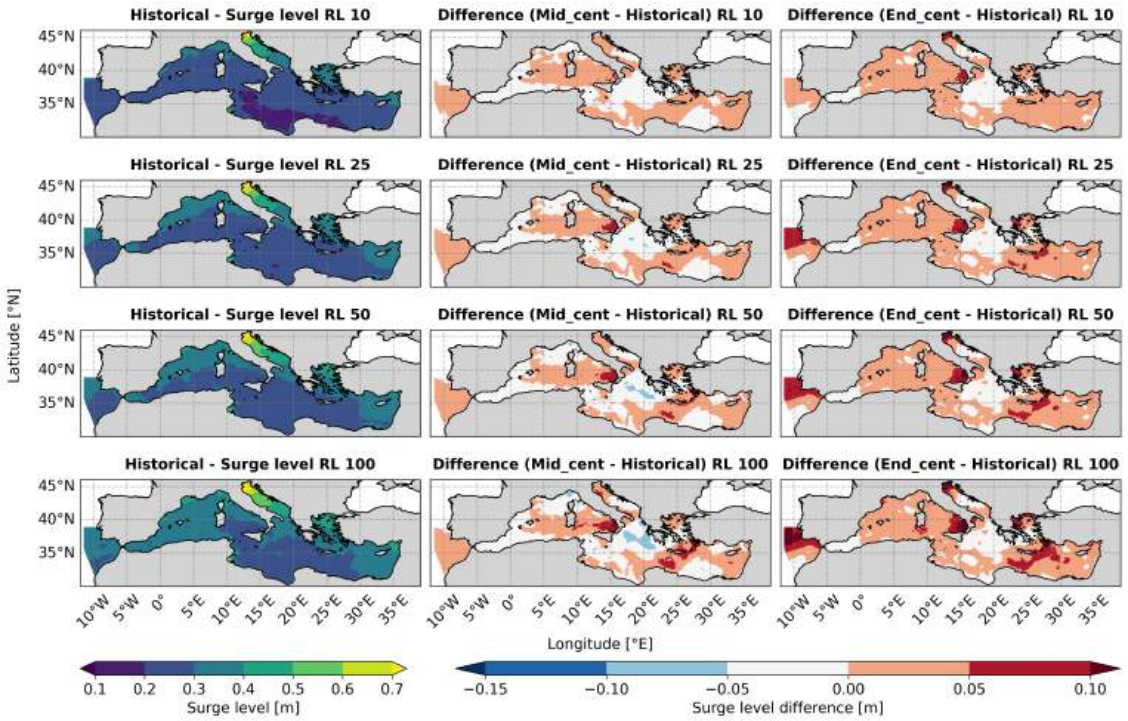

Figure 5.10: Storm surge return levels for 10-, 25-, 50-, and 100-year return periods (rows) for the historical period (first column) of the projection based on GCM:CanESM2 and RCM:CCLM4-8-17, and projected changes at mid-century (second column) and end-century (third column).

the Ligurian and adjacent French coasts.

As expected, return levels increase with the return period while broadly preserving the same geographical pattern. For the 100-year return period, the northern Adriatic remains the main hot-spot and the affected area becomes more spatially extensive. Elevated levels are also evident in the Gulf of Gabes and parts of the eastern Mediterranean, where values locally approach and in some areas exceed about 0.50 m.

Projected changes (second and third columns) are shown in red where return levels increase and in blue where they decrease relative to the historical baseline. For this specific projection, changes are predominantly positive across large portions of the basin. At mid-century, differences are generally modest (often on the order of a few centimeters) and more spatially patchy. By end-century, the signal becomes

stronger and more widespread, and several hot-spots show increases that locally exceed about 0.10 m, particularly in the northern Adriatic, the northern Aegean, and along parts of the southern Tyrrhenian coast (e.g., Calabria and Campania).

Considering another projection, namely the only additional simulation produced with the same RCM (CCLM4-8-17) but driven by the GCM MIROC5 (Figure 5.11), the projected pattern of change appears markedly different from that obtained for the CanESM2-driven projection.

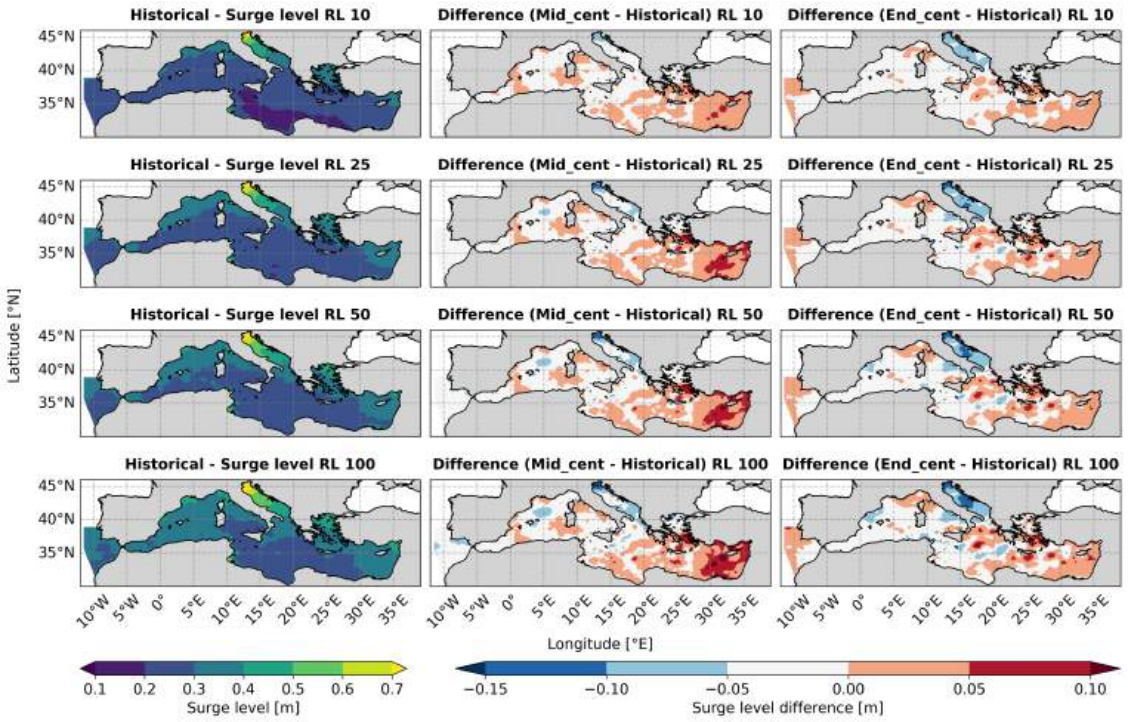

Figure 5.11: Same as Figure 5.10, but for the projection based on GCM:MIROC5 and RCM:CCLM4-8-17.

When focusing on projected changes in return levels, similar general considerations apply, but with a different temporal and spatial signature. In contrast to the previous projection, which showed a more coherent basin-wide increase becoming stronger towards end-century, this MIROC5-driven projection exhibits more pronounced and spatially structured changes already at mid-century, with relevant

increases in parts of the eastern Mediterranean that were not highlighted in the previous case.

By end-century, the signal over the Adriatic Sea shifts towards decreases, which become particularly evident for longer return periods. For instance, for the 100-year return level, the Adriatic region shows projected reductions locally reaching approximately 0.15 m, indicating an opposite behavior compared to the end-century increases found in the CanESM2-driven projection.

Overall, this comparison further confirms that the choice of the driving GCM can have a substantial impact on the projected changes in storm surge return levels, leading to projections that may differ not only in magnitude but also in sign across key Mediterranean sub-regions.

Two additional examples are provided in Figures 5.12 and 5.13, which are based on the same driving GCM (IPSL-CM5A-MR) but employ two different RCMs (RCA4 and HIRHAM5, respectively).

Despite the use of different regional models, the two projections show broadly consistent spatial patterns of change in storm surge return levels. In both cases, changes tend to be predominantly positive over large portions of the basin, with the magnitude of the signal generally increasing for longer return periods and becoming more pronounced towards end-century.

This agreement between Figures 5.12 and 5.13, obtained with the same GCM but different RCMs, further supports the idea that the driving GCM can exert a stronger influence on the projected storm surge signal than the choice of the RCM, at least for the configurations considered here.

In addition, both projections highlight areas of increasing return levels that were not emphasized in the previous examples, pointing to additional potential hot-spots of future storm surge hazard. For instance, a marked strengthening is visible in parts of the northwestern Mediterranean (including the Ligurian and Provençal sectors

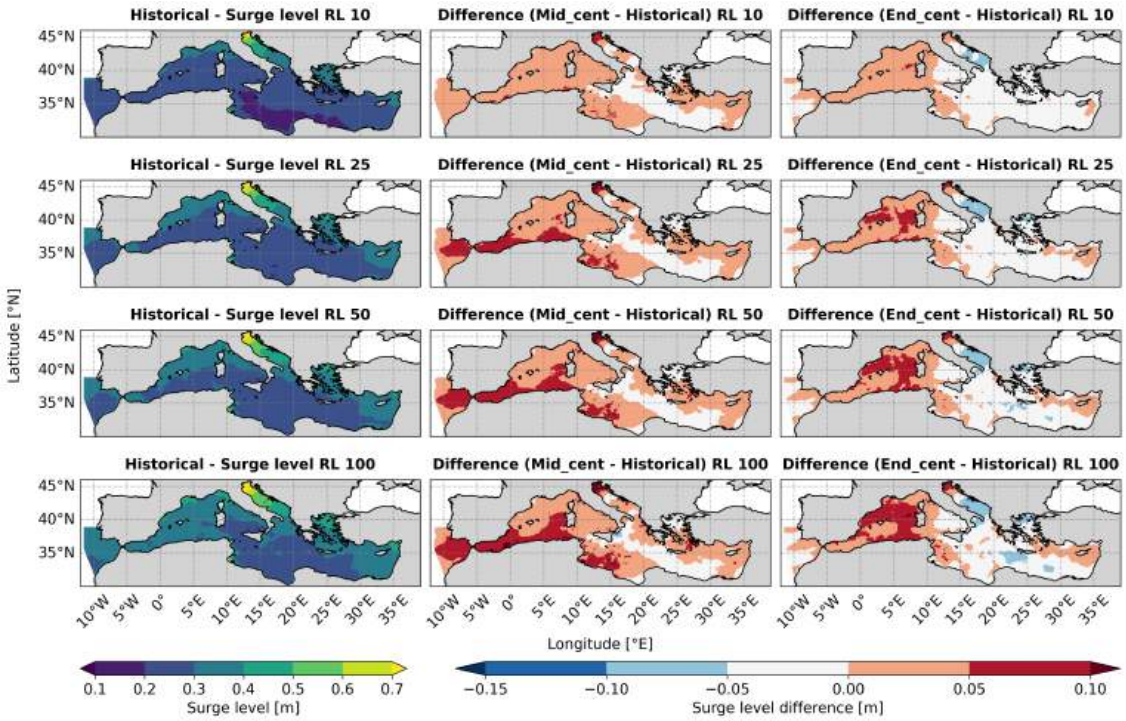

Figure 5.12: Same as previous figures, but for the projection based on GCM:IPSL-CM5A-MR and RCM:RCA4.

as well as the Alboran Sea and the the northwestern African coast) and in localized coastal stretches of the central Mediterranean, especially for the 50- and 100-year return periods.

These emerging differences, both in terms of the regions experiencing change and the sign of the change, naturally raise the question of what the actual direction of future storm surge evolution will be. At the same time, it is important to note that the projected variations discussed here remain overall limited, typically within a range of approximately ± 0.15 m.

This motivates the need for an analysis based on a large ensemble of climate projections, which can provide diversified yet complementary information on the plausible future evolution of the phenomenon. In contrast, an assessment based on only a small subset of projections would be unable to fully capture the spread of

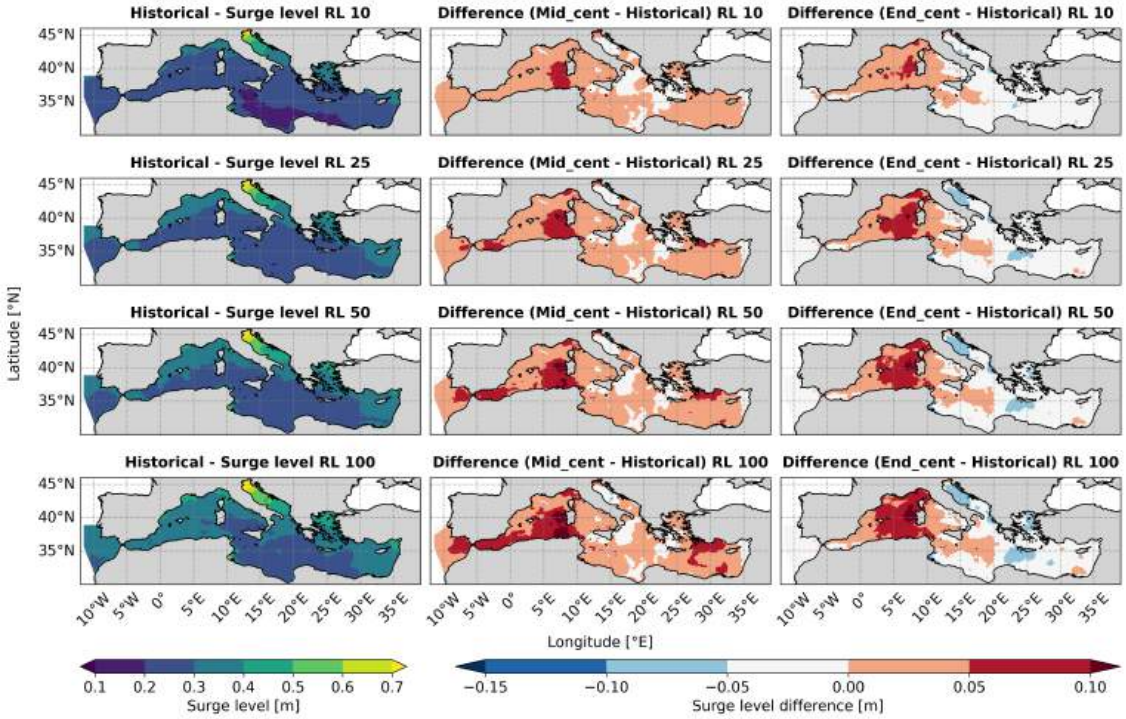

Figure 5.13: Same as previous figures, but for the projection based on GCM:IPSL-CM5A-MR and RCM:HIRHAM5.

possible outcomes and could therefore lead to conclusions that are overly dependent on the specific choice of GCM-RCM combination.

5.4 Robustness of projected changes

Since projected trends from individual simulations are not spatially uniform and exhibit substantial inter-model differences, it is important to assess the robustness of the projected changes in terms of consistency across the different climate model chains.

Accordingly, projected changes are interpreted not only through their ensemble mean, but also through a robustness metric that combines inter-model agreement on the sign of change and the magnitude of the signal relative to natural variability as

already introduced in Chapter 3. This approach follows the IPCC AR6 robustness framework, which classifies changes based on thresholds for model agreement (80% agreement in sign) and for signal emergence above a fixed threshold.

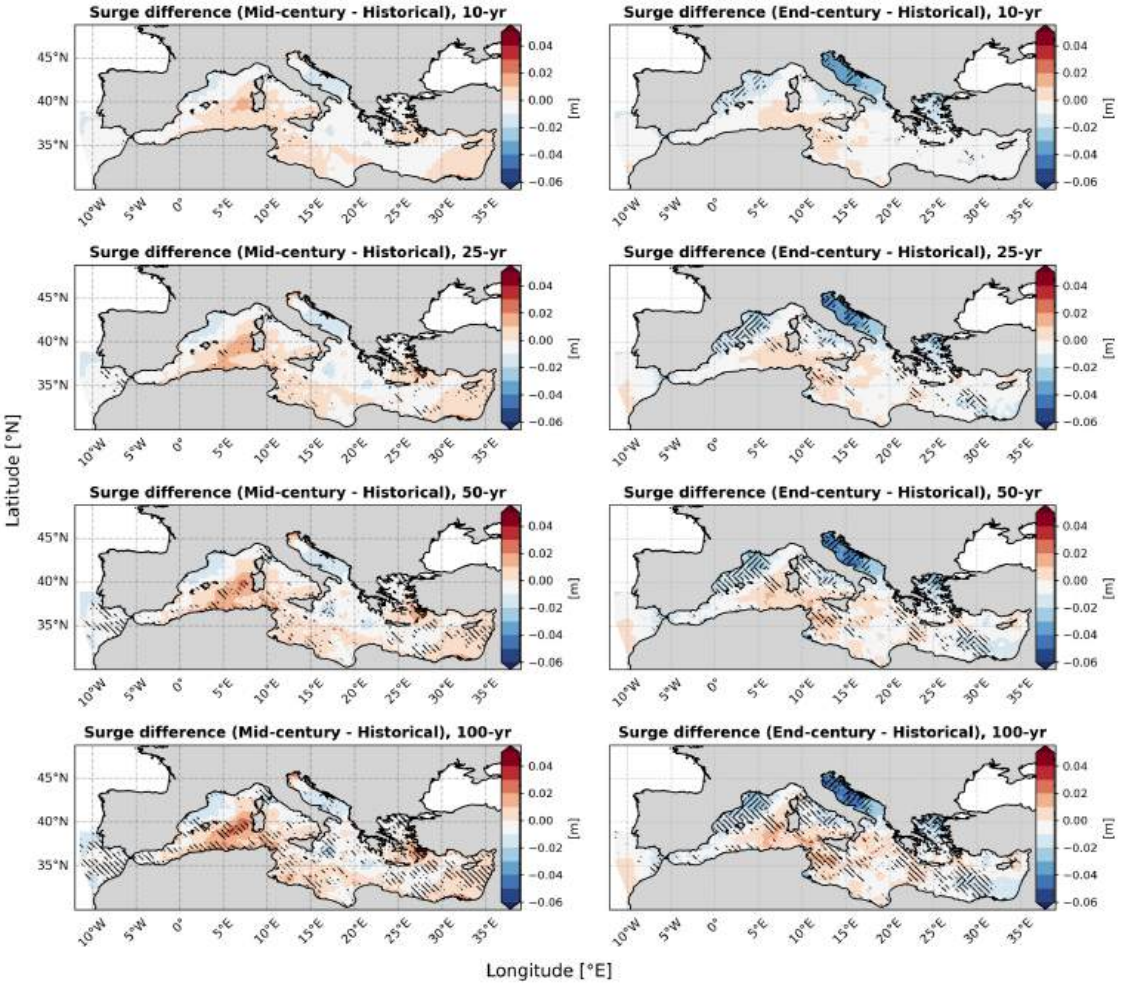

Figure 5.14: Ensemble-mean of storm surge return levels for 10-, 25-, 50- and 100-year return periods at mid-century (left column) and end-century (right column), obtained by averaging the 17 projection members. Regions with // hatching indicate robust change, regions with \\ hatching indicate conflicting change, and unhatched areas denote no robust change, according to the criteria described in Section 3.3.3. Note that this hatching convention is the opposite of that used in IPCC AR6, where robust change is typically shown as an absence of hatching; here, the choice was inverted to improve the visual identification of robust and non-robust areas in the Mediterranean domain.

Figure 5.14 shows the ensemble-mean changes in storm surge return levels at mid-century and end-century, together with their robustness across the ensemble.

The ensemble-mean changes remain generally small over most of the Mediterranean basin and for all considered return periods, typically within a few centimeters relative to the historical baseline. This further confirm that storm surges are not expected to vary substantially in the future relative to the historical period.

At mid-century, robust signals emerge only in a few localized areas and primarily for rare events (50- and 100-year return periods), while large portions of the basin are classified as no robust change or conflicting change. By end-century, the spatial pattern of robust change becomes more pronounced in the northern Adriatic, while robustness remains limited over most of other sub-basins, such as the northern Aegean, the Tyrrhenian Sea along the Campania coast, and the Gulf of Lion. Over the majority of the basin, conflicting or no robust changes persist, indicating that the projected changes in storm surge extremes are generally small compared with internal variability and the spread among models.

5.5 Pooled-ensemble return levels

Using a large ensemble of climate projections, especially when all simulations are produced within a consistent modeling framework as in this thesis, provides an additional key advantage: individual projections can be aggregated into a single, longer dataset. On the one hand, this aggregation yields information that is less dependent on the behavior of any single model chain and is therefore more representative of the ensemble as a whole. On the other hand, by increasing the number of available extreme values, it improves the statistical robustness of return level estimates, particularly for long return periods.

In practice, the return levels shown in the previous section were estimated from annual maxima extracted over 27-year time windows (mid-century: 2034-2060; end-century: 2074-2100). By pooling the annual maxima from all 17 projections, the

sample size increases to $17 \times 27 = 459$ annual maxima for each future period (and similarly for the historical reference period).

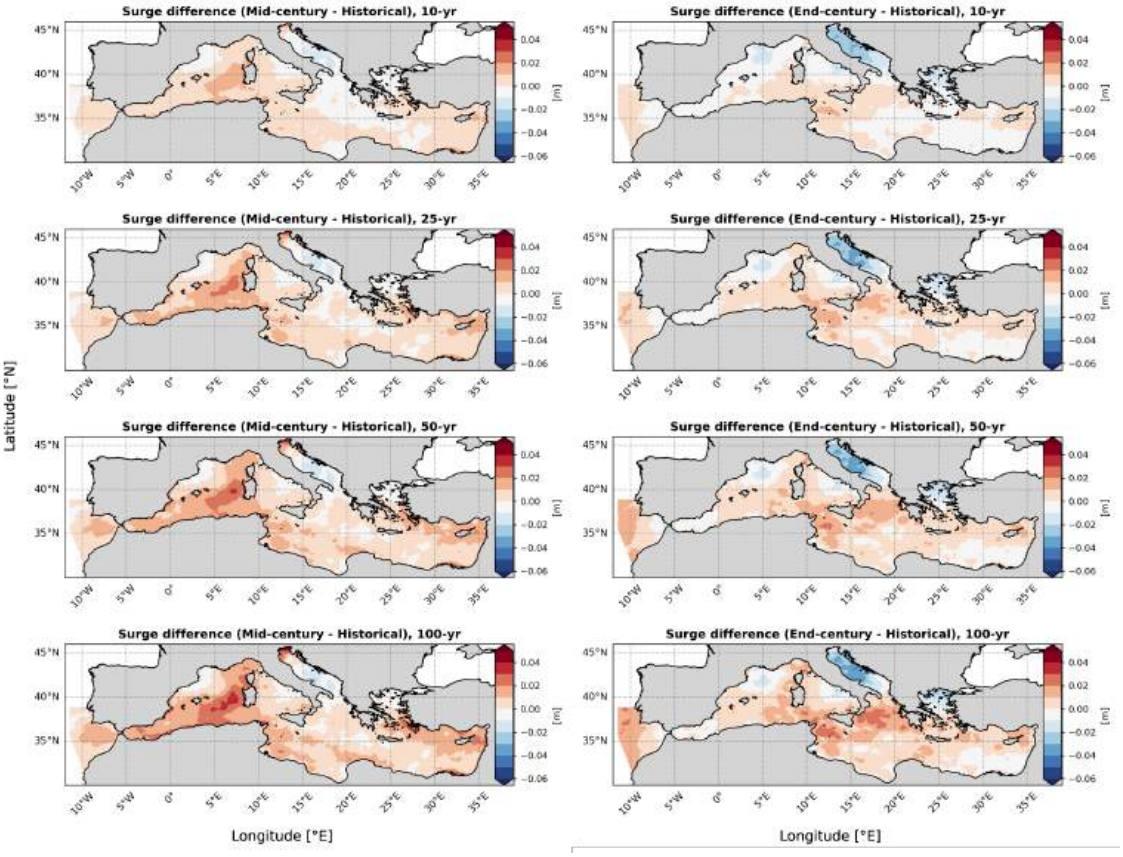

Figure 5.15: Projected changes in storm surge return levels estimated from the pooled ensemble of annual maxima (17 projections) relative to the historical baseline. Left column: mid-century (2034-2060); right column: end-century (2074-2100). Rows correspond to 10-, 25-, 50-, and 100-year return periods.

The aggregated annual maxima are then used to estimate storm surge return levels for 10-, 25-, 50-, and 100-year return periods across the Mediterranean basin. Figure 5.15 shows the projected changes relative to the historical baseline at mid-century (left column) and end-century (right column). Overall, the projected changes are small, remaining within approximately ± 0.10 m across most of the basin for all considered return periods. At mid-century, a modest increase is visible in parts of the Mediterranean, including the northern Adriatic and several localized areas

in the western basin and the eastern Mediterranean. By end-century, the signal becomes weaker over most sub-basins, while the largest coherent change appears in the Adriatic Sea, where return levels tend to decrease relative to the historical period.

These results suggest that, when pooling extremes across the full 17-member ensemble, large basin-wide changes in storm surge return levels are not expected, and the projected signal remains modest compared to the spatial variability of present-day extremes.

The maps related to the end-century pooled-ensemble return levels are also compared in Figure 5.16 to those obtained by the ensemble-mean procedure, i.e. estimated separately for each projection and then averages the resulting return levels across the 17 members (same maps reported in Figure 5.14 neglecting the hatches identifying the robustness). The comparison isolates the methodological differences between pooling extremes before fitting and averaging return-level estimates after fitting, and highlights how these choices affect the magnitude and spatial structure of projected return levels.

Overall, both approaches depict a weak basin-scale signal, with broadly consistent spatial patterns across all return periods. The most coherent feature is a decrease in the northern Adriatic, which becomes more evident for longer return periods (50–100 yr) in both panels. Differences between methods are generally subtle: the pooled-ensemble maps tend to exhibit slightly more spatially continuous anomalies and locally larger amplitudes, whereas the ensemble-mean maps appear more damped, consistent with partial balance when averaging member-wise estimates.

Moving from a basin-scale and largely qualitative perspective to a more local and quantitative illustration, Figure 5.17 shows the return level curves at a single grid point in the Ligurian Sea, next to the city of Chiavari. Here, results are again derived using the pooled-ensemble approach. This location is used here purely for

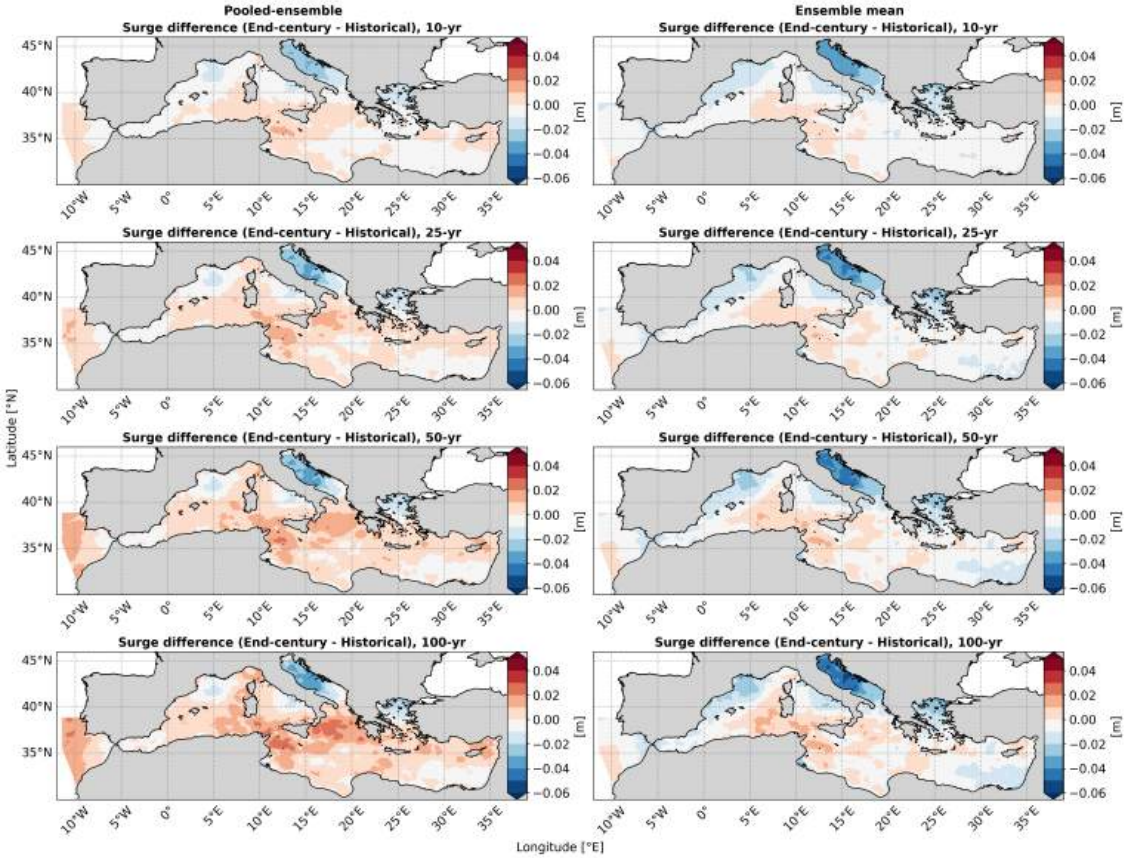

Figure 5.16: Comparison of end-century storm surge return levels obtained with two ensemble approaches. Left: pooled-ensemble return levels estimated by fitting the extreme-value model to annual maxima pooled across all 17 projections (2074-2100). Right: ensemble-mean return levels obtained by averaging member-wise return level estimates.

illustrative purposes and does not represent a specific hot-spot, but it was selected as a convenient reference point for additional analyses not central to the present thesis.

The figure shows three return level curves (historical, mid-century, and end-of-century), all estimated from the pooled ensemble annual maxima. In each panel, empirical estimates derived from the sorted annual maxima are shown together with a GEV fit, which is then used to estimate return levels for 10-, 25-, 50-, and 100-year return periods. At this specific location, differences between the historical and future periods are very small and, for shorter return periods, can be below

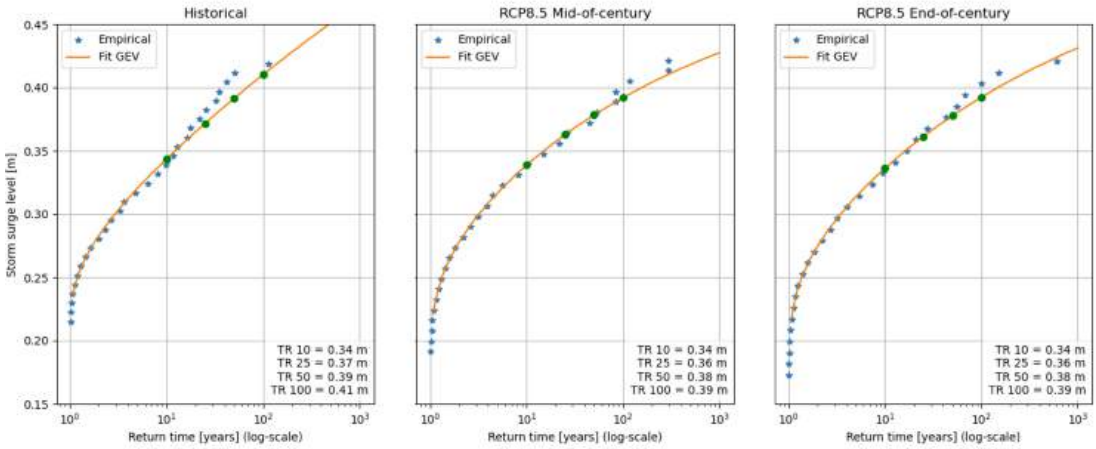

Figure 5.17: Return level curves for the historical, mid-century and end-of-century periods at a single model grid point in the Ligurian Sea (Chiavari), estimated from the ensemble of projections.

the centimeter level (e.g., the 10-year return level is essentially unchanged across the three periods). For the 100-year return level, the historical estimate is about 0.41 m, decreasing to approximately 0.39 m in both mid-century and end-of-century, indicating a slight reduction of extremes in the ensemble signal at this site.

In summary, the combined use of pooled-ensemble and ensemble-based robustness approaches (presented in this section and in the previous one respectively) provides complementary insights into future storm surge extremes. While the pooled-ensemble analysis emphasizes the average behavior across projections and benefits from enhanced statistical sampling of extremes, the ensemble-based robustness assessment reveals the substantial inter-model uncertainty and cautions against over-interpreting ensemble-mean signals, particularly at basin scale.

6

Discussion

The research carried out within the framework of this PhD thesis can be broadly divided into two main components. On the one hand, it focuses on the development and validation of a storm surge modeling framework; on the other hand, it addresses the characterization of the storm surge phenomenon and the associated coastal risk in the past and a quantitative assessment of its possible future evolution until the end of the 21st century under a high-emissions climate change scenario.

Despite addressing different objectives, these two components are strongly interconnected and share several common methodological aspects. Among these, the choice of the numerical model and the selection of its specific parameterizations play a central role, as they directly affect both the reliability of the hindcast simulations and the robustness of the future projections.

As discussed in Section 2.1, numerous studies have previously investigated storm surge processes across different regions and spatial scales of the Mediterranean Sea.

However, only a limited number of these studies have employed the Delft3D modeling suite, which is more commonly used in other coastal and hydrodynamic applications, while alternative modeling frameworks have often been preferred. In this context, the selection of the spatial resolution adopted in this thesis is guided by a careful balance between model performance and computational cost. A key requirement is methodological consistency: both the hindcast dataset and the climate projections are developed using a single hydrodynamic model configuration, in order to ensure that the differences identified between historical and future conditions can be attributed primarily to changes in atmospheric forcing rather than to inconsistencies in model setup.

Several sensitivity and comparative analyses are therefore carried out, including the case study presented in Section 4.3.1. These analyses supports the selection of a spatial resolution of 3 km, which is found to provide the best compromise between agreement with observations and computational efficiency.

A similar rationale guides the choice of atmospheric forcing for the hindcast simulations. The accurate representation of extreme sea level events is known to depend strongly on the quality of atmospheric forcing, particularly on the ability of weather reanalyses to reproduce extreme wind speeds and the evolution of low-pressure systems. Although ERA5 represents one of the most advanced global reanalysis products currently available, it has been shown to underestimate extremes, especially extreme winds, particularly during the earlier decades of the reanalysis period (Mentaschi et al., 2023; Campos et al., 2022). For this reason, the hindcast simulations are forced using wind and atmospheric pressure fields derived from a dynamical downscaling using the WRF model at a spatial resolution of 10 km (Ferrari et al., 2020), which provides a finer representation of mesoscale atmospheric processes than ERA5.

As shown in the results presented in Section 4.1, the model demonstrates a good

overall ability to reproduce observed storm surge variability. Nevertheless, several limitations of the validation process must be acknowledged. One important limitation, which likely contributes to locally lower correlation values, is related to the spatial representativeness of the tide gauge observations. Many of the stations used for validation are located within harbors, estuaries, or other semi-enclosed coastal environments, whose fine-scale morphology is not fully resolved by the model grid. While the adopted coastal resolution is adequate for open-sea processes, it may not capture local bathymetric features, breakwaters, or complex geometries that significantly influence water levels in these settings. Such spatial mismatches can lead to localized discrepancies between observed and simulated values, particularly under low-energy conditions or in areas strongly affected by anthropogenic modifications.

Additional discrepancies may also arise from limitations in the observational datasets themselves. As discussed previously, several tide gauge records are fragmented, sometimes available only at hourly resolution, and not always referenced consistently over time. Although extensive quality control procedures are applied first through the GESLA project flags and subsequently within the framework of this study, these corrections and classifications remain a potential source of uncertainty in the validation process.

Furthermore, the scarcity of tide gauge observations in parts of the eastern and southern Mediterranean introduces additional uncertainty regarding model performance in these regions. Nevertheless, it is assumed that the available validation results provide a representative assessment of the model's overall capability to simulate storm surge dynamics across the basin.

Overall, the results obtained in this study are consistent with those reported in previous literature, as discussed in Section 4.1.3, taking into account differences in model configurations, spatial resolution, and atmospheric forcing. A clear seasonal dependence of model performance is also evident: the model reproduces storm

surge variability more accurately during autumn and winter, when surges are more intense, whereas skill decreases in periods characterized by weak surge activity. During summer, when the synoptic wind- and pressure-driven surge signal is comparatively small, observed sea level variability is increasingly influenced by processes not represented in this modeling setup, including baroclinic and steric changes associated with seasonal temperature and salinity stratification and internal circulation variability, as well as nearshore contributions such as wave setup. As a result, the model explains a smaller fraction of the observed variability during summer, leading to lower correlation values.

In terms of spatial distribution, the regions most strongly affected by storm surges are consistent with expectations based on previous studies and with the known influence of coastal morphology on surge generation. The Adriatic Sea clearly emerges as the most surge-prone region of the Mediterranean. Its shallow bathymetry and elongated geometry make it particularly sensitive to strong Sirocco winds, which generate a large effective fetch and favor the development and amplification of storm surge within the basin. Other regions experiencing intense surge events include the Gulf of Gabes, the Nile Delta, the Gulf of Lion, and the northern Aegean Sea. These areas share common features, such as shallow coastal waters, low-lying coastlines, semi-enclosed basin geometries, and exposure to persistent winds from preferred directions.

Particular attention is devoted to extreme events during the validation process, with the aim of quantifying differences between observed and modeled surges in terms of both magnitude and timing. Extreme surge levels are of primary interest for coastal impact assessment and for estimating return periods associated with rare events. The analysis of selected case studies indicates that the bias associated with peak events is generally larger than that computed over the full time series. This behavior is expected, as extreme peaks are rare and therefore poorly sampled,

making the estimates more sensitive to systematic errors and resolution limitations. This effect is even more evident when the bias is computed using annual maxima only, since the sample contains just one value per year and is thus particularly sensitive to any mismatch in the reproduction of the observed peak. Conversely, when considering the full dataset, positive and negative errors tend to partially compensate, resulting in a smaller overall bias.

Once the validation process has quantified the model performance, the hindcast dataset is adopted as the reference representation of observed conditions for the analysis of the future projection dataset. This implies that the historical segment of each projection is bias-adjusted over the same reference period covered by the hindcast dataset, and that future changes are evaluated relative to this baseline.

To summarize, the main findings regarding future storm surge changes in the Mediterranean can be outlined as follows. First, projected changes in storm surge extremes are generally small in magnitude, typically within ± 0.10–0.15 m even by the end of the century. Second, these changes exhibit strong spatial heterogeneity, with no coherent basin-wide increase or decrease. Third, most of the Mediterranean basin is characterized by non-robust signals according to the IPCC AR6 framework, indicating that projected changes are not statistically distinguishable from natural variability. Only limited regions, such as the Adriatic Sea, the Gulf of Lion, and the northern Aegean Sea, show a consistent tendency toward decreasing extremes. Finally, the spread among projections is largely controlled by the driving GCM, highlighting the dominant role of large-scale atmospheric circulation changes in determining future storm surge behavior.

The storm surge projections presented in this thesis indicate that future changes across the Mediterranean basin are characterized by strong spatial heterogeneity but limited magnitude. No projection suggests a uniform basin-wide increase or decrease in storm surge extremes; instead, increases and decreases coexist within

the same projection, both at end-century and over nearer-term horizons.

To assess whether these signals are meaningful beyond natural variability and inter-model spread, projected changes are evaluated using the IPCC AR6 robustness framework. Under this definition, for an ensemble-based analysis such as that carried out in this thesis, a change is considered robust when it is both consistent across the ensemble (at least 80% of ensemble members agree on the sign of change) and sufficiently large to emerge from natural variability (at least 66% of the models simulate a change exceeding a threshold representative of interannual variability)

Accordingly, most of the Mediterranean basin is characterized by projected non-robust changes, indicating that the projections do not reach the 80% agreement threshold on the sign of change. A smaller portion of the basin is classified as exhibiting conflicting change, meaning that there is some tendency toward change, but it does not exceed the threshold associated with interannual variability. Only limited areas are identified as experiencing a robust change, namely the Adriatic Sea, the Gulf of Lion, and the northern Aegean Sea, where the projections consistently indicate a slight decrease in storm surge extremes. In practical terms, return levels associated with long return periods (e.g., 50- and 100-year) are projected to decrease in these regions relative to the historical baseline.

Despite the diversity of the climate projections considered, the magnitude of the projected changes in mean annual maxima and return levels is generally modest, typically remaining within ± 0.10–0.15 m even by the end of the century. This suggests that, under the RCP8.5 scenario, changes in storm surge extremes alone are unlikely to represent a dominant driver of increased coastal flood hazard at the basin scale.

However, the relatively small magnitude of the storm surge signal should not be interpreted as negligible. Even changes of a few centimeters can become relevant when combined with other contributors to extreme sea levels, such as mean sea

level rise, tidal effects, or wave setup. In particular, IPCC AR6 sea level rise projections for the Mediterranean indicate a likely increase of about 0.15–0.33 m by 2050 (relative to 1995–2014), and about 0.3–0.6 m under a very low-emissions scenario (SSP1-1.9) up to about 0.6–1.1 m under a very high-emissions scenario (SSP5-8.5) by 2100. Since coastal flooding is driven by the total water level, storm surge changes should therefore be regarded as an additive component to a rising baseline, and assessments should consider compound contributions rather than storm surge alone.

From a regional perspective, the Adriatic Sea, especially its northern sector, remains the Mediterranean sub-basin most prone to storm surge extremes due to its geomorphological and bathymetric setting. Nevertheless, the projections considered here consistently indicate a modest decrease in extreme levels in this area. A similar, though weaker, tendency toward decreasing extremes is also found in the Gulf of Lion. Elsewhere across the Mediterranean, projected changes are generally small and spatially fragmented, and they do not point to the emergence of new, coherent hot-spots.

These findings are consistent with previous studies on storm surge projections in the Mediterranean Sea (Section 2.3). At both European and basin scales, several works have reported limited changes in storm surge extremes under future climate scenarios. For instance, Vousdoukas et al. (2016) and subsequent studies (Vousdoukas et al., 2017, 2018) found minimal variations in storm surge levels across European coasts, including the Mediterranean, with projected changes generally small compared to present-day variability. Similarly, basin-scale analyses such as Makris et al. (2023) indicate a tendency towards reduced storminess and highlight strong regional differences in projected surge extremes.

More recent studies based on CMIP6 projections further support these findings. Irazoqui Apecechea et al. (2025) report a robust decrease in extreme surge levels

over parts of the Mediterranean by the end of the century, while Muis et al. (2023) identify a general tendency toward reduced storm surge magnitudes at the global scale, including the Mediterranean basin. The results obtained in this work are therefore in good agreement with the existing literature, both in terms of the limited magnitude of projected changes and their strong spatial heterogeneity.

Overall, the consistency between the present results and previous studies strengthens the conclusion that future changes in storm surge extremes in the Mediterranean are likely to remain modest, with no clear basin-wide intensification signal. Instead, changes in total water levels are expected to be primarily driven by mean sea level rise and by compound effects with other processes.

A key outcome of this work is the identification of the dominant influence of the driving GCM on projected storm surge changes, compared to the more limited impact of the RCM. Projections sharing the same GCM but employing different RCMs generally display similar spatial patterns of change, whereas projections driven by different GCMs can produce contrasting signals, sometimes even differing in sign in the same region. This finding highlights the importance of large-scale atmospheric circulation changes inherited from the GCM in shaping future storm surge behavior. Since storm surges are primarily driven by synoptic-scale wind and pressure fields, uncertainties at the global scale propagate directly into regional projections. Consequently, the inclusion of multiple GCMs is essential to capture the full range of plausible future storm surge responses.

The bias adjustment applied in this study represents a necessary step to ensure consistency between climate projections and the validated hindcast reference. The comparison between Gumbel- and GEV-based quantile-mapping approaches indicates that the choice of the extreme value distribution has only a minor influence on the estimated changes in storm surge extremes, suggesting that this source of uncertainty is small relative to the spread introduced by different climate model

chains. Nevertheless, the bias correction relies on the assumption of bias stationarity, which remains a potential source of uncertainty, particularly under strong climate change scenarios.

The analysis of return levels from individual projections confirms that projected changes can differ substantially across GCM-RCM combinations, both in magnitude and spatial distribution. By aggregating annual maxima across the full ensemble, a more robust picture emerges: ensemble-based return levels show that, when averaged across all projections, future changes remain small and do not indicate a substantial basin-wide intensification of storm surge extremes. This ensemble signal is consistent with the notion that internal variability and model uncertainty dominate over a clear forced trend in storm surge extremes within the Mediterranean.

These contrasting outcomes provided by the various projections highlight the limitations of inferring future storm surge behavior from a small subset of projections, as model-dependent differences can lead to substantially different signals, sometimes even of opposite sign in the same region.

Several limitations should be considered when interpreting the projected storm surge changes presented in this thesis. The analysis is based on a single emission scenario (RCP8.5), does not account for potential future changes in tidal dynamics, and relies on the assumption of stationarity in both model bias and extreme value statistics. Future work should address these aspects by including multiple emission scenarios, extending the temporal coverage of simulations, and exploring non-stationary approaches to extreme value analysis.

Overall, the datasets developed here provide a solid basis for basin-scale assessments of storm surge hazard, for identifying vulnerable coastal stretches, and for integration into broader coastal flood risk frameworks that combine storm surge with additional contributors to total water level. Future applications may include coupling with wave projections, compound-event analyses, and the development of

nested coastal models for local hot-spots where urban exposure and complex morphology require higher nearshore resolution.

7

Conclusions

This PhD thesis developed and validated a basin-wide, high-resolution storm surge modeling framework for the entire Mediterranean Sea, producing a consistent dataset for both historical reconstruction and future projections using Delft3D-FLOW, representing the first application of this modeling system at such spatial extent over the Mediterranean basin. The hindcast, spanning 1979–2023 and forced by high-resolution atmospheric fields from the WRF model, provides spatially continuous information on storm surge variability and extremes in a region characterized by heterogeneous observational coverage.

A central methodological outcome of this work is the demonstrated importance of accurately treating the Atlantic open boundary. Accounting for the Inverse Barometer Effect at the boundary leads to a measurable and systematic improvement in model performance compared with configurations neglecting this contribution, confirming that physically consistent barometric forcing at the basin connection is

essential for realistic Mediterranean storm surge simulations.

Validation against tide gauge derived surge residuals indicates overall good model performance, with pronounced seasonal variability. Performance is highest during autumn and winter, when storm surges are strongest and most relevant for coastal hazard, and decreases during summer, when surge amplitudes are small and local-scale processes not explicitly resolved by the model become relatively more influential. The hindcast enables basin-scale characterization of extremes through spatial diagnostics of maxima and return levels, while targeted case studies, such as the October 2018 Ligurian storm, demonstrate the model's ability to reproduce the timing, spatial footprint, and intensity of high-impact surge events.

Model credibility is further supported by localized inter-comparisons with storm surge simulations developed in previous studies, whose outputs were made available by the respective authors. These comparisons indicate overall consistency in surge representation across different modeling approaches, reinforcing confidence in the robustness of the developed framework.

While some limitations remain, particularly in the precise timing and magnitude of peak surge events, potentially related to the spatial resolution of the model not detailed enough to capture the near-coast tide gauge locations, and unresolved physical processes, the hindcast dataset can be regarded as a reliable reference for assessing past storm surge conditions across the Mediterranean basin.

Building on this validated historical baseline, a large ensemble of seventeen storm surge projections was produced up to 2100 using EURO-CORDEX climate forcing under the AR5 RCP8.5 climate change scenario. A dedicated bias-adjustment methodology enabled coherent comparisons between projections and the hindcast reference, providing a structured basis to investigate future changes in storm surge statistics and extremes while explicitly quantifying the uncertainty associated with using multiple combinations of global and regional climate models.

Ensemble analysis reveals that simulations driven by the same Global Climate Model tend to exhibit more similar spatial patterns of projected surge change than simulations sharing the same Regional Climate Model, indicating that large-scale atmospheric forcing exerts a stronger influence on projected surge variability than the regional downscaling step. At the same time, individual projections can yield locally contrasting signals, highlighting the importance of large ensembles for capturing the range of plausible future outcomes and their uncertainty.

Consistent with previous literature, projected changes in Mediterranean storm surge extremes are generally modest, suggesting that regions currently most exposed to surge-related hazard are likely to remain the most vulnerable in the future. However, storm surge represents only one component of extreme sea levels driving coastal flooding. Future coastal risk in the Mediterranean will be predominantly influenced by sea level rise, which is a direct consequence of climate change but lies beyond the scope of the present thesis.

Overall, the datasets and methodologies developed in this work provide a robust foundation for basin-scale assessments of present and future storm surge hazard in the Mediterranean, with direct relevance for both scientific research and coastal risk management applications.

7.1 Research outlook

While this dissertation has contributed significantly to the understanding and modeling of Mediterranean storm surges, several limitations and open questions remain, offering meaningful avenues for future research to build upon the findings presented here.

Flexible Mesh Refinement — In particularly vulnerable regions, or within the framework of specific projects, the model resolution can be locally refined us-

ing unstructured grids, which enable a more accurate representation, not only of coastal morphology, but also of civil structures in harbors and other anthropized areas. A promising direction involves the use of flexible-mesh modeling systems such as Delft3D Flexible Mesh (Delft3D-FM) software (Deltares, 2014), which offers advanced capabilities for simulating hydrodynamics on unstructured grids and handling complex coastal geometries. A natural next step is to employ a two-way nested modeling approach in which regular-grid simulations provide boundary conditions to higher-resolution local domains. Such an approach would enhance the accuracy of surge propagation near the coast, improve the representation of small-scale bathymetric features, and strengthen the capability to assess local impacts. This refinement would also facilitate more targeted studies supporting coastal management, early warning systems, and climate adaptation planning.

Extension to new Climate Change Scenarios — The climate change scenario explored in this dissertation represent the current state-of-the-art in terms of high-resolution atmospheric input. As discussed, the analyses presented here rely on the RCP8.5 scenario from the CMIP5 ensemble. New socioeconomic pathways and updated climate projections, introduced with CMIP6, continue to emerge and are expected to be updated in the future; however, high-resolution atmospheric projections derived from these newer scenarios are not yet fully available for the Mediterranean region. Future studies may incorporate these updated scenarios and refined regional climate model ensembles as they become accessible. Doing so would allow for capturing a broader range of possible futures and evaluating the sensitivity of storm surge projections to alternative global warming trajectories.

Beyond these methodological advancements, the research carried out in this dissertation has produced a large and valuable dataset that can serve as a foundation for multiple further applications and scientific explorations.

Extreme Sea Level analysis — Future work should also focus on a fully inte-

grated analysis involving storm surges, wind waves, and tidal dynamics. Although this dissertation has centered primarily on the surge component, extreme sea levels in the Mediterranean are the result of a combination of several interacting factors. Coupling hindcast and projected storm surge datasets with wave climate and tidal variability would provide a more physically complete view of coastal hazards. Such combined assessments are crucial to identify hot-spots of compound flooding risk, evaluate the statistical dependence between different forcings, and quantify the contributions of each driver to extreme events. This perspective is especially relevant in the context of climate change, where shifts in wave climate or modified tidal regimes could further modulate surge impacts. Moreover, integrating sea level rise projections within probabilistic frameworks could enhance long-term risk assessments, providing more actionable information for decision-makers.

SFINCS integration — A possible practical application of this work is the use of the storm surge time series generated in this dissertation as input for SFINCS (Leijnse et al., 2021), a fast and flexible inundation model developed by Deltares. This would allow simulated sea level variations to be translated into high-resolution flood maps, providing detailed estimates of inundation depth, spatial extent, and timing under both historical and projected conditions. Such an application represents a direct exploitation of the research outputs, supporting coastal risk management, engineering design, and adaptation planning. By coupling hydrodynamic simulations with flood impact modeling, SFINCS enables actionable assessments that can inform decision-makers and emergency preparedness strategies, bridging scientific results with real-world applications.

Appendix A: Atmospheric Projections Models

The details of the seventeen GCM–RCM combinations of atmospheric projections, based on the RCP8.5 climate change scenario and used to develop the storm surge dataset through the end of the 21st century, are provided here.

All projections use EUR-11 resolution (0.11°, ∼12.5 km) and 6-hourly temporal resolution (6hr).

1) EUR-11_CLMcom-CCCma-CanESM2_r1i1p1-CCLM4-8-17_v1_6hr

GCM: CanESM2 (Canadian Earth System Model 2)

GCM developed by: Canadian Centre for Climate Modelling and Analysis (CCCma)

RCM: CCLM4-8-17

RCM developed by: Climate Limited-area Modeling Community (CLMcom)

2) EUR-11_CLMcom-MIROC-MIROC5_r1i1p1-CCLM4-8-17_v1_6hr

GCM: MIROC5

GCM developed by: Atmosphere and Ocean Research Institute (University of Tokyo) and Japan Agency for Marine-Earth Science and Technology (JAMSTEC)

RCM: CCLM4-8-17

RCM developed by: Climate Limited-area Modeling Community (CLMcom)

3) EUR-11_SMHI-CNRM-CERFACS-CNRM-CM5_r1i1p1-RCA4_v1_6hr

GCM: CNRM-CM5

GCM developed by: Centre National de Recherches Météorologiques / CERFACS (CNRM-CERFACS)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

4) EUR-11_SMHI-IPSL-IPSL-CM5A-MR_r1i1p1-RCA4_v1_6hr

GCM: IPSL-CM5A-MR

GCM developed by: Institut Pierre-Simon Laplace (IPSL)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

5) EUR-11_SMHI-MOHC-HadGEM2-ES_r1i1p1-RCA4_v1_6hr

GCM: HadGEM2-ES

GCM developed by: Met Office Hadley Centre (MOHC, UK)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

6) EUR-11_SMHI-MPI-M-MPI-ESM-LR_r1i1p1-RCA4_v1a_6hr

GCM: MPI-ESM-LR

GCM developed by: Max Planck Institute for Meteorology (MPI-M, Germany)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

7) EUR-11_SMHI-NCC-NorESM1-M_r1i1p1-RCA4_v1_6hr

GCM: NorESM1-M

GCM developed by: Norwegian Climate Centre (NCC)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

8) EUR-11_SMHI-ICHEC-EC-EARTH_r1i1p1-RCA4_v1_6hr

GCM: EC-EARTH

GCM developed by: EC-Earth Consortium (Europe-wide)

RCM: RCA4

RCM developed by: Swedish Meteorological and Hydrological Institute (SMHI)

9) EUR-11_DMI-CNRM-CERFACS-CNRM-CM5_r1i1p1-HIRHAM5_v2_6hr

GCM: CNRM-CM5

GCM developed by: Centre National de Recherches Météorologiques / CERFACS (CNRM-CERFACS)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

10) EUR-11_DMI-IPSL-IPSL-CM5A-MR_r1i1p1-HIRHAM5_v1_6hr

GCM: IPSL-CM5A-MR

GCM developed by: Institut Pierre-Simon Laplace (IPSL)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

11) EUR-11_DMI-MOHC-HadGEM2-ES_r1i1p1-HIRHAM5_v2_6hr

GCM: HadGEM2-ES

GCM developed by: Met Office Hadley Centre (MOHC, UK)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

12) EUR-11_DMI-MPI-M-MPI-ESM-LR_r1i1p1-HIRHAM5_v1_6hr

GCM: MPI-ESM-LR

GCM developed by: Max Planck Institute for Meteorology (MPI-M, Germany)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

13) EUR-11_DMI-NCC-NorESM1-M_r1i1p1-HIRHAM5_v3_6hr

GCM: NorESM1-M

GCM developed by: Norwegian Climate Centre (NCC)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

14) EUR-11_DMI-ICHEC-EC-EARTH_r1i1p1-HIRHAM5_v1_6hr

GCM: EC-EARTH

GCM developed by: EC-Earth Consortium (Europe-wide)

RCM: HIRHAM5

RCM developed by: Danish Meteorological Institute (DMI)

15) EUR-11_CLMcom-ETH-MOHC-HadGEM2-ES_r1i1p1-COSMO-crCLIM-v1-1_v1_6hr

GCM: HadGEM2-ES

GCM developed by: Met Office Hadley Centre (MOHC, UK)

RCM: COSMO-crCLIM v1-1

RCM developed by: ETH Zurich / Climate Limited-area Modeling Community (CLMcom)

16) EUR-11_CLMcom-ETH-NCC-NorESM1-M_r1i1p1-COSMO-crCLIM-v1-1_v1_6hr

GCM: NorESM1-M

GCM developed by: Norwegian Climate Centre (NCC)

RCM: COSMO-crCLIM v1-1

RCM developed by: ETH Zurich / Climate Limited-area Modeling Community (CLMcom)

17) EUR-11_CLMcom-ETH-ICHEC-EC-EARTH_r1i1p1-COSMO-crCLIM-v1-1_v1_6hr

GCM: EC-EARTH

GCM developed by: EC-Earth Consortium (Europe-wide)

RCM: COSMO-crCLIM v1-1

RCM developed by: ETH Zurich / Climate Limited-area Modeling Community (CLMcom)

Appendix B: Tide gauges from GESLA v4 dataset for model validation

Note. The provider acronyms used in the table are defined as follows:

- CMEMS = Copernicus Marine Environment Monitoring Service;
- CV = City of Venice;
- IEO = Oceanographic Spanish Institute;
- ISPRA = Italian Institute for Environmental Protection and Research;
- REFMAR = Reference Network of Tide Gauges (France);
- UHSLC = University of Hawaii Sea Level Center;

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Ajaccio	FRA	CMEMS, REFMAR	from 1981	1 hour
Alcudia	ESP	CMEMS	from 2009	1 hour
Alexandria	EGY	UHSLC	2009 - 2016	1 hour
Algeciras	ESP	CMEMS, IEO	from 1980	1 hour
Alicante	ESP	CMEMS	from 1979	1 hour
Almeria	ESP	CMEMS	from 2006	1 hour
Ancona	ITA	ISPRA	from 1986	1 hour until 1998, then 10 minutes

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Anzio	ITA	ISPRA	from 2010	10 minutes
Barcelona	ESP	CMEMS	from 1993	1 hour
Bari	ITA	ISPRA	from 1979	1 hour until 1998, then 10 minutes
Cagliari	ITA	ISPRA	from 1986	1 hour until 1998, then 10 minutes
Carboneras	ESP	CMEMS	from 2013	1 hour
Carloforte	ITA	ISPRA	from 1999	10 minutes
Cartagena	ESP	CMEMS	from 2019	10 minutes
Catania	ITA	ISPRA	from 1979	1 hour until 1998, then 10 minutes
Centuri	FRA	CMEMS, REFMAR	from 2010	1 hour
Ceuta	ESP	IEO, UHSLC	from 1979	1 hour
Civitavecchia	ITA	ISPRA	from 1981	1 hour until 1998, then 10 minutes
Colonia Sant Pere	ESP	CMEMS	from 2015	1 hour
Crotone	ITA	ISPRA	from 1991	1 hour until 1998, then 10 minutes
Fos sur mer	FRA	CMEMS, REFMAR	from 2006	1 hour
Gaeta	ITA	ISPRA	from 2010	10 minutes

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Gandia	ESP	CMEMS	from 2007	1 hour
Genova	ITA	ISPRA	from 1998	10 minutes
Hrakleio	GRE	CMEMS	from 2016	10 minutes
Ibiza	ESP	CMEMS	from 2003	1 hour
Ierapetra	GRE	CMEMS	from 2016	1 10 minutes
Ile Rousse	FRA	CMEMS, REFMAR	from 2013	1 hour
Imperia	ITA	ISPRA	from 1986	1 hour until 1998, then 10 minutes
Koroni	GRE	CMEMS	from 2013	10 minutes
La Spezia	ITA	ISPRA	from 2010	10 minutes
La Figueirette	FRA	CMEMS, REFMAR	from 2011	1 hour
Livorno	ITA	ISPRA	from 1979	1 hour until 1998, then 10 minutes
Mahon	ESP	CMEMS	from 2009	1 hour
Malaga	ESP	CMEMS, IEO	from 1992	1 hour
Marina di Campo	ITA	ISPRA	from 2011	10 minutes
Marseille	FRA	CMEMS, REFMAR, UHSLC	from 1979	1 hour
Melilla	ESP	CMEMS	from 2007	1 hour

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Messina	ITA	ISPRA	from 1979	1 hour until 1998, then 10 minutes
Monaco	MCO	CMEMS, REFMAR	from 1981	1 hour
Motril	ESP	CMEMS	from 2005	1 hour
Napoli	ITA	ISPRA	from 1986	1 hour until 1998, then 10 minutes
Nice	FRA	CMEMS, REFMAR	from 1981	1 hour
Ortona	ITA	ISPRA	from 1996	1 hour until 1998, then 10 minutes
Otranto	ITA	ISPRA	from 1987	1 hour until 1998, then 10 minutes
Paleochora	GRE	CMEMS	from 2013	1 hour
Palermo	ITA	ISPRA	from 1992	1 hour until 1998, then 10 minutes
Palinuro	ITA	CMEMS, ISPRA	from 1987	1 hour until 1998, then 10 minutes
Palma de Mallorca	ESP	CMEMS, IEO	from 2009	1 hour
Plimiri	GRE	CMEMS	from 2015	1 hour
Port Ferreol	FRA	REFMAR	from 2012	1 hour

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Porto Empedocle	ITA	ISPRA	from 1999	10 minutes
Porto Torres	ITA	ISPRA	from 1985	1 hour until 1998, then 10 minutes
Ravenna	ITA	ISPRA	from 1986	1 hour until 1998, then 10 minutes
Reggio Calabria	ITA	ISPRA	from 1999	10 minutes
Sagunto	ESP	CMEMS	from 2007	1 hour
S. Benedetto del Tronto	ITA	CMEMS, ISPRA	from 2010	1 hour
Salerno	ITA	ISPRA	from 1991	1 hour until 1998, then 10 minutes
Sciacca	ITA	ISPRA	from 2012	10 minutes
Sete	FRA	CMEMS, REFMAR	from 1984	1 hour
Solenzara	FRA	CMEMS, REFMAR	from 2002	1 hour
Taranto	ITA	ISPRA	from 1993	1 hour until 1998, then 10 minutes
Tarifa	ESP	CMEMS, IEO	from 2009	1 hour
Tarragona	ESP	CMEMS	from 2011	1 hour

Tide gauge	Country	Abbreviated institution	Available observation	Temporal resolution
Toulon	FRA	CMEMS, REFMAR	from 1981	1 hour
Trieste	ITA	ISPRA, UHSLC	from 1988	1 hour until 1998, then 10 minutes
Valencia	ESP	CMEMS	from 1992	1 hour
Venezia	ITA	CV, ISPRA	from 1986	1 hour until 1998, then 10 minutes
Vieste	ITA	ISPRA	from 1990	1 hour until 1998, then 10 minutes

Table 1: Details of the Mediterranean tide gauges used in this study, extracted from the GESLA-4 dataset. The table lists, for each station, country, managing institution, period of available observations, and temporal resolution.

References

Agulles, M., Marcos, M., Amores, A., and Toomey, T. (2024). Storm surge modelling along european coastlines: The effect of the spatio-temporal resolution of the atmospheric forcing. *Ocean Modelling*, 192:102432.

Amores, A., Marcos, M., Carrió, D. S., and Gómez-Pujol, L. (2020). Coastal impacts of storm gloria (january 2020) over the north-western mediterranean. *Natural Hazards and Earth System Sciences*, 20.

Androulidakis, Y., Makris, C., Mallios, Z., Pytharoulis, I., Baltikas, V., and Krestenitis, Y. (2023). Storm surges and coastal inundation during extreme events in the mediterranean sea: the ianos medicane. *Natural Hazards*, 117(1):939–978.

Azzopardi, B., Balzan, M. V., Cherif, S., Doblas-Miranda, E., dos Santos, M., Dobrinski, P., Falder, M., Hassoun, A. E. R., Giupponi, C., Koubi, V. V., et al. (2020). Climate and environmental change in the mediterranean basin, current situation and risks for the future. first mediterranean assessment report.

Bajo, M., Međugorac, I., Umgiesser, G., and Orlić, M. (2019). Storm surge and seiche modelling in the adriatic sea and the impact of data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 145.

Bajo, M., Zampato, L., Umgiesser, G., Cucco, A., and Canestrelli, P. (2007). A finite element operational model for storm surge prediction in venice. *Estuarine, Coastal and Shelf Science*, 75(1-2):236–249.

Bouwer, L. and Bakker, P. (2019). Extremes, abrupt changes and managing risks.

Campos, R. M., Gramcianinov, C. B., de Camargo, R., and da Silva Dias, P. L. (2022). Assessment and calibration of era5 severe winds in the atlantic ocean using satellite data. *Remote Sensing*, 14(19):4918.

Campos-Caba, R., Alessandri, J., Camus, P., Mazzino, A., Ferrari, F., Federico, I., Vousdoukas, M., Tondello, M., and Mentaschi, L. (2024). Assessing storm surge model performance: what error indicators can measure the model's skill? *Ocean Science*, 20(6):1513–1526.

Cassola, F., Ferrari, F., and Mazzino, A. (2015). Numerical simulations of mediterranean heavy precipitation events with the wrf model: A verification exercise using different approaches. *Atmospheric Research*, 164:210–225.

Causio, S., Shirinov, S., Federico, I., De Cillis, G., Clementi, E., Mentaschi, L., and Coppini, G. (2025). Coupling ocean currents and waves for seamless cross-scale modeling during medicane ianos. *Ocean Science*, 21(3):1105–1123.

Cavaleri, L., Bajo, M., Barbariol, F., Bastianini, M., Benetazzo, A., Bertotti, L., Chiggiato, J., Davolio, S., Ferrarin, C., Magnusson, L., et al. (2019). The october 29, 2018 storm in northern italy–an exceptional event and its modeling. *Progress in oceanography*, 178:102178.

Cavaleri, L., Bajo, M., Barbariol, F., Bastianini, M., Benetazzo, A., Bertotti, L., Chiggiato, J., Ferrarin, C., Trincardi, F., and Umgiesser, G. (2020). The 2019 flooding of venice and its implications for future predictions. *Oceanography*, 33(1):42–49.

Cavaleri, L., Barbariol, F., Bertotti, L., Besio, G., and Ferrari, F. (2022). The 29 october 2018 storm in northern italy: Its multiple actions in the ligurian sea. *Progress in Oceanography*, 201:102715. Available online 18 December 2021.

Cid, A., Castanedo, S., Abascal, A. J., Menéndez, M., and Medina, R. (2014). A high resolution hindcast of the meteorological sea level component for southern europe: the gos dataset. *Climate dynamics*, 43(7):2167–2184.

Codiga, D. (2025). Utide unified tidal analysis and prediction functions. `https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions`. MATLAB Central File Exchange.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

Corvaro, S., Marini, F., Rocchi, S., and Lorenzoni, C. (2025). Extreme storm surge and wave height analysis in the adriatic sea. *Coastal Engineering Proceedings*, 38(38):papers.52.

de Vries, H., Breton, M., de Mulder, T., Krestenitis, Y., Proctor, R., Ruddick, K., Salomon, J. C., Voorrips, A., et al. (1995). A comparison of 2d storm surge models applied to three shallow european seas. *Environmental Software*, 10(1):23–42.

Dee, D. P., Uppala, S., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al. (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597.

Defina, A. (2003). Numerical experiments on bar growth. *Water Resources Research*, 39(4).

Deltares (2014). *Delft3D-FLOW: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. User Manual*. Deltares, Delft, The Netherlands.

Deltares, S. (2014). D-flow flexible mesh, technical reference manual.

Donatini, L., Lupieri, G., Contento, G., Pedroncini, A., Cusati, L. A., and Crosta, A. (2015). Mwm: A 35 years wind & wave high resolution hindcast dataset and an operational forecast service for the mediterranean sea. DHI Italia & HyMOLab, University of Trieste.

Egbert, G. D. and Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic technology*, 19(2):183–204.

Favaretto, C., Ruol, P., and Martinelli, L. (2025). Analysis of the vulnerability of the lagoon levees of the po delta to coastal flooding in a changing climate. *Estuarine, Coastal and Shelf Science*, 313:109089.

Fernández-Montblanc, T., Vousdoukas, M. I., Mentaschi, L., and Ciavola, P. (2020). A pan-european high resolution storm surge hindcast. *Environment International*, 135:105367.

Ferrari, F., Besio, G., Cassola, F., and Mazzino, A. (2020). Optimized wind and wave energy resource assessment and offshore exploitability in the mediterranean sea. *Energy*, 190:116447.

Ferrarin, C., Lionello, P., Orlić, M., Raicich, F., and Salvadori, G. (2022). Venice as a paradigm of coastal flooding under multiple compound drivers. *Scientific reports*, 12(1):5754.

Ferrarin, C., Pantillon, F., Davolio, S., Bajo, M., Miglietta, M. M., Avolio, E., Carrió, D. S., Pytharoulis, I., Sanchez, C., Patlakas, P., González-Alemán, J. J., and Flaounas, E. (2023). Assessing the coastal hazard of medicane ianos through ensemble modelling. *Natural Hazards and Earth System Sciences*, 23:2273–2294.

Fortelli, A., Fedele, A., Natale, G. D., Matano, F., Sacchi, M., Troise, C., and Somma, R. (2021). Analysis of sea storm events in the mediterranean sea: The case study of 28 december 2020 sea storm in the gulf of naples, italy. *Applied Sciences*, 11(23):11460.

Gomis, D., Ruiz, S., Sotillo, M. G., Álvarez Fanjul, E., and Terradas, J. (2008). Low frequency mediterranean sea level variability: The contribution of atmospheric pressure and wind. *Global and Planetary Change*, 63(2):215–229. Mediterranean climate: trends, variability and change.

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., Arns, A., Bradshaw, E., and Thompson, P. (2023). Gesla version 3: A major update to the global higher-frequency sea-level dataset. *Geoscience Data Journal*, 10(3):293–314.

Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., Horsburgh, K., and Gouldby, B. (2016). Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the uk. *Scientific Data*, 3.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hilmi, N., Ali, E., Carnicer Cols, J., Cramer, W., Georgopoulou, E., Le Cozannet, G., and Tirado, C. (2022). Ipcc ar6 wgii cross-chapter paper 4: Mediterranean region. In *EGU General Assembly Conference Abstracts*, pages EGU22–10590.

Intergovernmental Panel on Climate Change (IPCC) (2022). Summary for Policymakers. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Irazoqui Apecechea, M., Melet, A., Menendez, M., Lobeto, H., and Valle-Rodriguez, J. B. (2025). Projections of changes in extreme storm surges for european coasts using statistical downscaling. *EGUsphere*, 2025:1–29.

Ismail, H. (2018). Climate change and its impact on coastal cities: A case study from alexandria.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional Environmental Change*, 14(2):563–578.

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., Benestad, R., Boberg, F., Buonomo, E., Cardoso, R. M., et al. (2020). Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community. *Regional environmental change*, 20(2):1–20.

Ji, T., Li, G., Liu, Y., Liu, R., and Zhu, Y. (2021). Spatiotemporal features of storm surge activity and its response to climate change in the southeastern coastal area of china in the past 60 years. *Journal of Geophysical Research: Atmopsheres*, 126.

Krestenitis, Y. N., Androulidakis, Y. S., Kontos, Y. N., and Georgakopoulos, G. (2011). Coastal inundation in the north-eastern mediterranean coastal zone due to storm surge events. *Journal of Coastal Conservation*, 15(3):353–368.

Le Provost, C. et al. (2023). pytide: Harmonic tidal analysis and prediction in python. Available at `https://github.com/CNES/pangeo-pytide`.

Leijnse, T., van Ormondt, M., Nederhoff, K., and van Dongeren, A. (2021). Modeling compound flooding in coastal systems using a computationally efficient reduced-physics solver: Including fluvial, pluvial, tidal, wind-and wave-driven processes. *Coastal Engineering*, 163:103796.

Lin-Ye, J., Garcia-Leon, M., Gracia, V., Ortego, M. I., Lionello, P., Conte, D., Perez-Gomez, B., and Sanchez-Arcilla, A. (2020). Modeling of future extreme storm surges at the nw mediterranean coast (spain). *Water*, 12(2):472.

Lionello, P., Galati, M., and Elvini, E. (2012). Extreme storm surge and wind wave climate scenario simulations at the venetian littoral. *Physics and Chemistry of the Earth, Parts A/B/C*, 40:86–92.

Lionello, P. and Sanna, A. (2005). Mediterranean wave climate variability and its links with nao and indian monsoon. *Climate Dynamics*, 25(6):611–623.

Lira-Loarca, A., Cáceres-Euse, A., De-Leo, F., and Besio, G. (2022). Wave modeling with unstructured mesh for hindcast, forecast and wave hazard applications in the mediterranean sea. *Applied Ocean Research*, 122:103118.

MacFerrin, M., Amante, C., Carignan, K., Love, M., and Lim, E. (2025). The earth topography 2022 (etopo 2022) global dem dataset. *Earth System Science Data*, 17(5):1835–1849.

Makris, C. V., Tolika, K., Baltikas, V. N., Velikou, K., and Krestenitis, Y. N. (2023). The impact of climate change on the storm surges of the mediterranean sea: Coastal sea level responses to deep depression atmospheric systems. *Ocean Modelling*, 181:102149.

Marcos, M., Tsimplis, M. N., and Shaw, A. G. P. (2009). Sea level extremes in southern europe. *Journal of Geophysical Research: Oceans*, 114(C1).

Medugorac, I., Mirko, O., Ivica, J., Zoran, P., and Miroslava, P. (2018). Adriatic storm surges and related cross-basin sea-level slope. *Journal of Marine Systems*, 181:79–90.

Mejjad, N., Rossi, A., and Pavel, A. B. (2022). The coastal tourism industry in the mediterranean: A critical review of the socio-economic and environmental pressures & impacts. *Tourism Management Perspectives*, 44:101007.

Menemenlis, D., Fukumori, I., and Lee, T. (2007). Atlantic to mediterranean sea level difference driven by winds near gibraltar strait. *Journal of physical oceanography*, 37(2):359–376.

Mentaschi, L., Vousdoukas, M. I., García-Sánchez, G., Fernández-Montblanc, T., Roland, A., Voukouvalas, E., Federico, I., Abdolali, A., Zhang, Y. J., and Feyen, L. (2023). A global unstructured, coupled, high-resolution hindcast of waves and storm surge. *Frontiers in Marine Science*, 10:1233679.

Muis, S., Aerts, J. C., Á. Antolínez, J. A., Dullaart, J. C., Duong, T. M., Erikson, L., Haarsma, R. J., Apecechea, M. I., Mengel, M., Le Bars, D., et al. (2023).

Global projections of storm surges using high-resolution cmip6 climate models. *Earth's Future*, 11(9):e2023EF003479.

Muis, S., Apecechea, M. I., Álvarez, J. A., Verlaan, M., Yan, K., Dullaart, J., Aerts, J., Duong, T., Ranasinghe, R., le Bars, D., et al. (2022). Global sea level change time series from 1950 to 2050 derived from reanalysis and high resolution cmip6 climate projections. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 10.

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C., and Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature communications*, 7(1):11969.

Nicholls, R. J., Wong, P. P., Burkett, V., Codignotto, J. O., Hay, J., Ragoonaden, S., Woodroffe, C. D., et al. (2007). Coastal systems and low-lying areas.

on Climate Change (IPCC), I. P. (2014). Summary for policymakers. in: Climate change 2014: Mitigation of climate change. contribution of working group iii to the fifth assessment report of the intergovernmental panel on climate change.

Oppenheimer, M., Glavovic, B. C., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., Deconto, R. M., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., and Sebesvari, Z. (2019). *Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities*, chapter 4. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Pena, F., Obeysekera, J., Jane, R., Nardi, F., Maran, C., Cadogan, A., de Groen, F., and Melesse, A. (2023). Investigating compound flooding in a low elevation coastal karst environment using multivariate statistical and 2d hydrodynamic modeling. *Weather and Climate Extremes*, 39:100534.

Pérez-Gómez, B., Manzano, F., Alvarez-Fanjul, E., González, C., Cantavella, J. V., and Schindelé, F. (2016). Lessons derived from two high-frequency sea level events in the atlantic: implications for coastal risk analysis and tsunami detection. *Frontiers in Marine Science*, 3:206.

Pérez Gómez, B., Vilibić, I., Šepić, J., Međugorac, I., Ličer, M., Testut, L., Fraboul, C., Marcos, M., Abdellaoui, H., Álvarez Fanjul, E., et al. (2022). Coastal sea level monitoring in the mediterranean and black seas. *Ocean science*, 18(4):997–1053.

Ponte, R. M. (2006). Low-frequency sea level variability and the inverted barometer effect. *Journal of Atmospheric and Oceanic Technology*, 23(4):619–629.

Pugh, D., Woodworth, P. L., and Woodworth, P. (2014). *Sea-level science: understanding tides, surges, tsunamis and mean sea-level changes*. Cambridge university press.

Reistad, M., Breivik, ., Haakenstad, H., Aarnes, O. J., Furevik, B. R., and Bidlot, J.-R. (2011). A high-resolution hindcast of wind and waves for the north sea, the norwegian sea, and the barents sea. *Journal of Geophysical Research: Oceans*, 116.

Rizzi, J., Gallina, V., Torresan, S., Critto, A., Gana, S., and Marcomini, A. (2016). Regional risk assessment addressing the impacts of climate change in the coastal area of the gulf of gabes (tunisia). *Sustainability Science*, 11(3):455–476.

Romero-Martín, R., Sanuy, M., and Jiménez, J. A. (2025). Unveiling the role of storm surges as a driver of flooding on the western mediterranean: a case study of the ebro delta. *Natural hazards*, 121(4):4961–4984.

Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaounas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., Lombardi, E., et al. (2016). Med-

cordex initiative for mediterranean climate studies. *Bulletin of the American Meteorological Society*, 97(7):1187–1208.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E. (2014). The ncep climate forecast system version 2. `https://doi.org/10.5065/D6RJ49WQ`. NOAA/NWS/NCEP, doi:10.5065/D6RJ49WQ.

Šepić, J., Vilibić, I., Rabinovich, A. B., and Monserrat, S. (2015). Widespread tsunami-like waves of 23-27 june in the mediterranean and black seas generated by high-altitude atmospheric forcing. *Scientific reports*, 5(1):11682.

Shankar, C. G. and Behera, M. R. (2021). Improved wind drag formulation for numerical storm wave and surge modeling. *Dynamics of Atmospheres and Oceans*, 93:101193.

Shchepetkin, A. F. and McWilliams, J. C. (2005). The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean modelling*, 9(4):347–404.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G. (2008). A description of the advanced research wrf version 3. Technical Report NCAR/TN-475+STR, National Center for Atmospheric Research. Technical Note.

Small, C. and Nicholls, R. J. (2003). A global analysis of human settlement in coastal zones. *Journal of coastal research*, pages 584–599.

Smith, G. A., Hemer, M., Greenslade, D., Trenham, C., Zieger, S., and Durrant, T. (2021). Global wave hindcast with australian and pacific island focus: From past to present. *Geoscience Data Journal*.

Sterl, A. and Caires, S. (2005). Climatology, variability and extrema of ocean waves: The web-based knmi/era-40 wave atlas. *International journal of climatology*, 25(7):963–978.

Tay, C., Lindsey, E. O., Chin, S. T., McCaughey, J. W., Bekaert, D., Nguyen, M., Hua, H., Manipon, G., Karim, M., Horton, B. P., et al. (2022). Sea-level rise from land subsidence in major coastal cities. *Nature Sustainability*, 5(12):1049–1057.

Tebaldi, C., Ranasinghe, R., Vousdoukas, M., Rasmussen, D., Vega-Westhoff, B., Kirezci, E., Kopp, R. E., Sriver, R., and Mentaschi, L. (2021). Extreme sea levels at different global warming levels. *Nature Climate Change*, 11(9):746–751.

Toledo, I., Pagán, J. I., López, I., Olcina, J., and Aragonés, L. (2024). Storm surge in spain: Factors and effects on the coast. *Marine Geology*, 476:107373. Open Access under CC BY-NC-ND.

Toomey, T., Amores, A., Marcos, M., and Orfila, A. (2022). Coastal sea levels and wind-waves in the mediterranean sea since 1950 from a high-resolution ocean reanalysis. *Frontiers in Marine Science*, 9:991504.

Toomey, T., Lira-Loarca, A., Marcos, M., Besio, G., and Orfila, A. (2025). Future wave climate in the mediterranean sea and associated uncertainty from an ensemble of 31 gcm-rcm wave simulations. *Earth's Future*, 13(4):e2024EF004992.

Torresan, S., Furlan, E., Critto, A., Michetti, M., and Marcomini, A. (2020). Egypt's coastal vulnerability to sea-level rise and storm surge: Present and future conditions. *Integrated Environmental Assessment and Management*, 16(5):761–772.

Toth, Z., Kalnay, E., Tracton, S. M., Wobus, R., and Irwin, J. (1997). A synoptic evaluation of the ncep ensemble. *Weather and forecasting*, 12(1):140–153.

Umgiesser, G., Bajo, M., Ferrarin, C., Cucco, A., Lionello, P., Zanchettin, D., Papa, A., Tosoni, A., Ferla, M., Coraci, E., et al. (2021). The prediction of floods in venice: methods, models and uncertainty. *Natural hazards and earth system sciences*, 21(8):2679–2704.

Umgiesser, G., Canu, D. M., Cucco, A., and Solidoro, C. (2004). A finite element model for the venice lagoon. development, set up, calibration and validation. *Journal of Marine Systems*, 51(1-4):123–145.

Vecchio, A., Anzidei, M., and Serpelloni, E. (2024). Sea level rise projections up to 2150 in the northern mediterranean coasts. *Environmental Research Letters*, 19(1):014050.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., and Feyen, L. (2017). Extreme sea levels on the rise along europe's coasts. *Earth's Future*, 5(3):304–323.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., and Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nature communications*, 9(1):2360.

Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., and Feyen, L. (2016). Projections of extreme storm surge levels along europe. *Climate Dynamics*, 47(9):3171–3190.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, volume 100. Academic press.

Wood, M., Haigh, I. D., Le, Q. Q., Nguyen, H. N., Tran, H. B., Darby, S. E., Marsh, R., Skliris, N., Hirschi, J. J.-M., Nicholls, R. J., and Bloemendaal, N.

(2023). Climate-induced storminess forces major increases in future storm surge hazard in the south china sea region. *Natural Hazards and Earth System Sciences*, 23:2475–2504.

Yu, C. S., Decouttere, C., and Berlamont, J. (1998). Storm surge simulations in the adriatic sea. In Gambolati, G., editor, *CENAS*, volume 28 of *Water Science and Technology Library*, pages 225–237. Springer, Dordrecht.

Zhang, Y. J., Ye, F., Stanev, E. V., and Grashorn, S. (2016). Seamless cross-scale modeling with schism. *Ocean Modelling*, 102:64–81.