



A multiscale modelling framework of coastal flooding events for global to local flood hazard assessments

Irene Benito¹, Jeroen C. J. H. Aerts^{1,2}, Philip J. Ward^{1,2}, Dirk Eilander^{1,2}, and Sanne Muis^{1,2}

¹Institute for Environmental Studies (IVM), Vrije Universiteit, Amsterdam, the Netherlands

²Deltares, Delft, the Netherlands

Correspondence: Irene Benito (i.benito.lazaro@vu.nl)

Received: 7 May 2024 – Discussion started: 24 May 2024

Revised: 7 February 2025 – Accepted: 21 April 2025 – Published: 11 July 2025

Abstract. Tropical and extratropical cyclones, which can cause coastal flooding, are among the most devastating natural hazards. Understanding coastal flood risk better can help to reduce their potential impacts. Global flood models play a key role in this process. In recent years, global models and methods for flood hazard simulation have improved, but they are still limited in the actionable information that they can provide at local scales. One notable limitation is the insufficient resolution of global models, which cannot accurately capture the complexities of storms and the topography of specific regions. Additionally, most large-scale hazard assessments tend to focus solely on either offshore water level simulations or overland flooding, often relying on static flood modelling approaches. In this study, we introduce the MOSAIC (MOdelling Sea level And Inundation for Cyclones) framework, a flexible Python-based framework designed to dynamically simulate both offshore water levels and coastal flooding. MOSAIC provides a multiscale modelling approach to automatically generate and nest high-resolution local models within a coarser global model. This approach seeks to simulate more accurate water levels, thereby enhancing coastal boundary conditions for dynamic flood modelling. We showcase the potential of MOSAIC using three historical storm events, with the aim of assessing the effects of temporal- and spatial-resolution refinements and bathymetry data. Our findings indicate that the importance of model refinements is linked to the topography of the study area and the storm characteristics. For instance, refining the temporal output resolution has a significant impact on small and rapidly intensifying tropical cyclones but is less critical for extratropical cyclones. Additionally, the refinement of spatial output locations is particularly relevant in re-

gions where water levels exhibit high spatial heterogeneity along the coast. In regions with complex topography, grid refinement and higher-resolution bathymetry play a more significant role. MOSAIC provides an automated approach to provide flood maps at a local scale. Our results confirm the proof of concept that the automated approach of MOSAIC can be used to provide high-resolution flood maps without the need for calibration or other manual steps. As such, MOSAIC provides a bridge between fully global and fully local modelling approaches. In future work, further validation could be carried out to explore the optimal settings for different regions in more detail.

1 Introduction

Coastal flood events can have devastating impacts on societies, economies, and the environment when they affect densely populated and low-lying coastal areas (Wadey et al., 2015). Tropical cyclones (TCs) and extratropical cyclones (ETCs) are the cause of the most severe coastal flooding events (Douris et al., 2021; Haigh et al., 2016; UNDRR, 2020; Wahl et al., 2017). For example, Hurricane Harvey in 2017 is one of the costliest storms in the United States' history, with estimated damage of USD 125 billion. Typhoon Idai in Mozambique in 2019 caused around 600 deaths and economic damage of USD 770 million (Nhamo and Chikodzi, 2021; Sebastian et al., 2021). In 1953, an ETC was the cause of the most severe coastal flood event in northwest Europe, resulting in more than 2000 deaths (Wadey et al., 2015). More recently, in 2010, ETC Xynthia hit the Atlantic

coast of France, causing 47 deaths and EUR 1.2 billion worth of economic damage (CGEDD, 2010).

Coastal flood events are driven by extreme sea levels, resulting from a combination of mean sea level variations, tides, storm surges, and waves (Kirezci et al., 2020; Marcos et al., 2019; Vousdoukas et al., 2017, 2018b; Wahl, 2017). In recent years, several studies have applied global hydrodynamic models to simulate coastal water levels (Dullaart et al., 2021; Muis et al., 2016; Pringle et al., 2021; Vousdoukas et al., 2016b; Wang and Bernier, 2023). Subsequently, these water levels have been used to derive extreme-water-level values for various return periods. These extreme water levels have then been used as input for global overland flood models (Wing et al., 2024), and the resulting flood hazard maps have been used to assess flood exposure and risk (Vousdoukas et al., 2016a). While these global studies have greatly improved our understanding of large-scale coastal flood risks, they do not yet have the accuracy to provide actionable information about coastal flood events at local scales.

The accuracy of large-scale hazard assessments is limited by several factors related to the quality of the input data and assumptions underlying the modelling approaches. Until now, the vast majority of large-scale hazard assessments have primarily been focused on either modelling extreme water levels or modelling overland floods. Each model component has its own limitations. We identify here three main methodological limitations of large-scale hazard assessments. First, coastal geometry strongly influences extreme sea levels (Mori et al., 2014; Woodruff et al., 2023), with large variability at the local scale. Consequently, in regions with complex morphologies, such as estuaries, semi-enclosed bays, or barrier systems, global models lack the resolution required to accurately resolve the extreme sea levels (Bunya et al., 2010; Dietrich et al., 2010; Islam et al., 2021). Grid refinement and nesting of local high-resolution models within coarser global models can result in improved coastal boundary conditions. Pelupessy et al. (2017) used a multiscale approach to obtain realistic boundary conditions by nesting a high-resolution barotropic model inside of a global circulation model. Similarly, the Coastal Storm Modeling System (CoSMoS) combines global climate models and oceanographic models that are dynamically downscaled to assess compound flooding and coastal changes at the regional to local scale (Barnard et al., 2014, 2019, 2025; Nederhoff et al., 2024). Camus et al. (2011) used a dynamic downscaling approach to translate global wave data into higher-spatiotemporal-resolution waves for the Spanish coast. Second, the accuracy of input datasets such as the meteorological forcing and the bathymetry has a large influence on the total water levels. Coarse meteorological forcings – in terms of both spatial and temporal resolution – might not be able to resolve intense storms (Hodges et al., 2017; Murakami, 2014; Thomas et al., 2021), while errors in the bathymetric datasets will propagate to the modelling of storm surge levels (Woodruff et al., 2023). Third, coastal flooding is a dy-

namic process where flood duration and physical processes play a key role. However, given the high computational costs associated with using hydrodynamic flood models, their use has been limited to local applications. Most large-scale hazard assessments have used static flood modelling methods, which neglect flood dynamics (Hinkel et al., 2014; Muis et al., 2016; Ramirez et al., 2016; Vafeidis et al., 2019; Vousdoukas et al., 2016a). Additionally, large-scale hazard assessments typically focus on a single flood driver (Alfieri et al., 2017; Hirabayashi et al., 2021; Tiggelev et al., 2020; Vousdoukas et al., 2018a; Ward et al., 2020). However, TC and ETC events often produce precipitation, river discharge, storm surges, and waves, all of which can contribute to flooding. When these drivers occur in combination, they can significantly amplify flood hazards and risks. This is demonstrated by the modelling of, for example, hurricane Florence, which hit the USA in 2018 (Gori et al., 2020). Few large-scale studies have analysed the effects and interactions of multiple flood drivers. While Bates et al. (2021) performed a combined risk assessment of fluvial, pluvial, and coastal flooding for the continental USA, Eilander et al. (2023) introduced a globally applicable compound flood modelling framework that accounts for precipitation, river discharge, and storm tides.

In this study, we present the open-source MOSAIC (Modelling Sea level And Inundation for Cyclones) framework to simulate TC and ETC water levels and coastal flooding events. Coastal flooding is dynamically modelled by coupling two existing modelling approaches: (1) to simulate water levels generated by storm surges and tides, it couples the hydrodynamic Global Tide and Surge Model (GTSM) and Delft3D Flexible Mesh software and (2) to dynamically simulate overland flooding, it couples the simulated water levels with the Super-Fast INundation of CoastS model (SFINCS). MOSAIC is based on Python and global datasets and, as such, provides a globally applicable and reproducible approach that can automatically build and process Delft3D Flexible Mesh and SFINCS models. As such, it is well suited to a model comparison study to test different model setups.

Here we showcase the potential of the MOSAIC framework by applying it to three case studies where large storm surges caused catastrophic flooding events, namely the historical storm events of TC Irma, TC Haiyan, and ETC Xynthia (see Fig. 1; Bertin et al., 2012; Cangialosi et al., 2018; Lapidez et al., 2015). For each of these storms, we simulate the coastal water levels and flood depths using automatically built uncalibrated models. Where available, we evaluate the model performance by comparing it to observed water levels and flood maps. Moreover, we perform a sensitivity analysis of different modelling settings. This includes the effects of the model resolution, output resolution, and improvements in bathymetry.

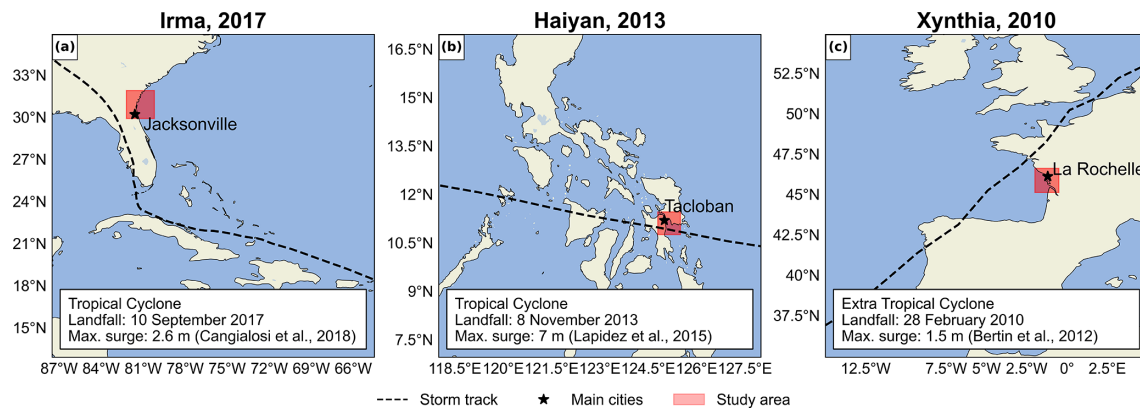


Figure 1. Case studies analysed in this paper. (a) Tropical cyclone Irma, (b) tropical cyclone Haiyan, and (c) extratropical cyclone Xynthia. The red area indicates the modelling domain of the flood analysis.

2 The MOSAIC modelling framework

The MOSAIC modelling framework, shown in Fig. 2, is a Python-based framework that integrates different packages, models, and software. It consists of two main components: (1) the simulation of global coastal boundary conditions with the Global Tide and Surge Model (GTSM) (Sect. 2.1), including the dynamic downscaling with a local high-resolution model (Sect. 2.1.3), and (2) the overland flood hazard simulations using the SFINCS model (Sect. 2.2). Python scripts that enable adjustments to the GTSM settings are used to generate different model configurations. For the flood hazard simulations, MOSAIC uses Hydro Model Tools (HydroMT) to prepare and postprocess SFINCS model input and output data.

2.1 Derivation of coastal boundary conditions

2.1.1 Meteorological forcing

The meteorological forcing datasets used in this study vary per storm. For ETC Xynthia and TC Irma, we use mean sea level pressure and the 10 m meridional and zonal wind components from the ERA5 reanalysis dataset at a horizontal resolution of 0.25° and a 1 h temporal resolution (Hersbach et al., 2019). Because TC Haiyan is not well resolved in ERA5 (see Fig. A1), we use pressure and wind from tropical cyclone track data merged with data from ERA5. The tropical cyclone track data are retrieved from the Joint Typhoon Warning Center at 6-hourly intervals (Naval Meteorology and Oceanography Command, 2022) and are converted to a polar grid with 36 radial bins, 375 angular bins, and a radius of 350 km using the Holland parametric wind model (Holland et al., 2010). Following the methodology of Dullaart et al. (2021) and Lin and Chavas (2012), we apply an anticlockwise rotation angle of $\beta = 20^\circ$ and set the storm translation to a surface background wind reduction factor of $\alpha = 0.55$. Additionally, we use an empirical surface wind re-

duction factor (SWRF) of 0.85 (Batts et al., 1980) and convert 1 min average winds to 10 min averages using a factor of 0.915 (Harper et al., 2010). The Holland model's output provides a file that defines a polar grid containing pressure and wind fields. To extend the pressure and wind fields beyond the Holland model's defined TC boundary, we linearly interpolate these fields on the outermost 75 % to align with the ERA5 background data (Deltares, 2024).

2.1.2 Global storm surge and tide model

MOSAIC uses GTSM v4.1 to simulate water levels resulting from tides and storm surges, ignoring baroclinic and wave contributions. GTSM is a global depth-averaged hydrodynamic model based on Delft3D Flexible Mesh (Kernkamp et al., 2011). It has a spatially varying resolution of 25 km deep in the ocean and 2.5 km along the coasts (1.25 km for Europe) (Dullaart et al., 2020; Muis et al., 2020). The spatially varying resolution makes it computationally efficient for the simulation of water levels at large scales. The bathymetry in the model is the 15 arcsec resolution European Marine Observation and Data Network bathymetry dataset for Europe (EMODnet; Consortium EMODnet Bathymetry, 2018) and the 30 arcsec General Bathymetric Chart of Oceans 2019 dataset for the rest of the globe (GEBCO, 2014). Tides are generated internally with tide generating forces, while storm surges originate from external forcing using pressure and wind fields (Sect. 2.1.1; Muis et al., 2020). A constant Charnock coefficient of 0.041 is applied to translate wind speeds from the external forcing into wind drag, and a background pressure of 101 325 Pa is considered. GTSM has been successfully validated using different meteorological datasets and has been shown to provide accurate extreme sea levels (Dullaart et al., 2020; Muis et al., 2016, 2020). Version 4.1 is a calibrated version of the model with improved parameterisations for internal tides and bottom-friction coefficients (Deltares, 2021; Wang et al., 2022b). GTSM provides

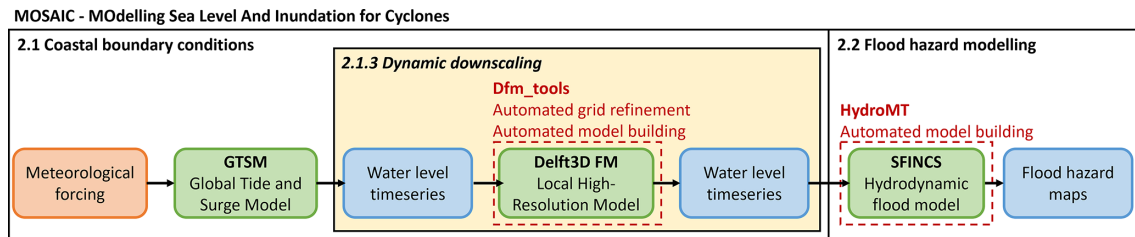


Figure 2. Flowchart showing the input (in orange), models (in green), outputs (in blue), Python packages (in red), and the optional dynamic downscaling feature (in yellow) of MOSAIC.

as output water level time series over a grid in the ocean and for locations along every ~ 5 km of the coast.

To validate the coastal component of our modelling framework, we compare water levels from GTSM to observed water levels from tide gauge stations of the Global Extreme Sea Level Analysis (GESLA) dataset (Haigh et al., 2023). This comparison is made for case studies where the GTSM output locations are found near tide gauge stations from GESLA (see Fig. 3). GTSM output is referenced to mean sea level (MSL). We translate the reference level of the GESLA water levels to the MSL by removing the annual average water level for each year and subsequently by removing the mean over the 1985–2005 period from the detrended time series. To assess the accuracy of GTSM, we calculate Pearson’s correlation coefficient and the root-mean-squared error (RMSE; see Table A1). Figures 4 and 5 show the time series of water levels at different tide gauge stations during the landfalls of TC Irma and ETC Xynthia, respectively. Pearson’s correlation between the GTSM-simulated and observed water levels is high for both events, indicating good agreement. For TC Irma, the average correlation across the nine stations is 0.93, with a standard deviation of 0.06 m. For ETC Xynthia, the average correlation across the six stations is 1.00, with a standard deviation of 0.01. Additionally, TC Irma has an average RMSE of 0.28 m with a standard deviation of 0.09 m. ETC Xynthia has an RMSE of 0.22 m, with a standard deviation of 0.08 m. The stations performing worse are those located in enclosed harbours or behind the barrier islands. The RMSE values of GTSM for both storms show results comparable to other large-scale studies that have used hydrodynamic models to simulate storm tides of storm events. Marsooli and Lin (2018) and Gori et al. (2023), for example, used the ADvanced CIRCulation model (ADCIRC) to simulate storm tides, with an average RMSE over stations of 0.31 and 0.29 m, respectively. Vogt et al. (2024) used the GeoCLaw solver and reported an average RMSE of 0.24 m over 213 tide gauge stations but with Pearson’s correlation of 0.5, showing less agreement with observed storm tides than the MOSAIC model setup presented in this study.

2.1.3 Dynamic downscaling

The dynamic downscaling within MOSAIC consists of two parts. First, MOSAIC generates a local high-resolution model with Delft3D Flexible Mesh using the Python package `dfm_tools` (Veenstra, 2024). `Dfm_tools` allows us to automatically create a local modelling grid with a spatially varying resolution based on the specified maximum and minimum grid cell sizes as well as Courant’s number derived from the bathymetry data provided (Veenstra, 2024). The bathymetry of the local model can be updated by interpolating a new bathymetric dataset into the newly generated grid. The settings to automatically generate the local high-resolution models used in this study can be found in Sect. 2.3. Second, MOSAIC uses an offline coupling approach to nest the local Delft3D Flexible Mesh model within GTSM. A Python script is used to first identify the boundaries of the local Delft3D Flexible Mesh model. These boundaries are then used to determine the specific locations where GTSM output should be extracted. Subsequently, GTSM provides the water level time series at the boundaries of the local model. Finally, the local high-resolution model is executed using the water levels derived from GTSM as forcing input, together with the same meteorological forcing as for GTSM.

2.2 Hydrodynamic flood hazard modelling setup

MOSAIC uses the Super-Fast INundation of CoastS (SFINCS) model to simulate overland storm surge flood depths. SFINCS is a reduced-physics hydrodynamic model developed for a more computationally efficient dynamic flooding approach than full shallow-water-equation models (Leijnse et al., 2021). It solves simplified equations of mass and momentum, similar to the LISFLOOD-FP model (Bates et al., 2010). SFINCS has been successfully applied to model compound flooding from tropical cyclone Irma in 2017 (Leijnse et al., 2021). Its modelling output results in similar results to those from full shallow-water-equation models, while reducing computational expenses by a factor of 100 (Leijnse et al., 2021). To speed up the flood model simulations, we use the subgrid schematisation from SFINCS for all the simulations (Leijnse et al., 2020).

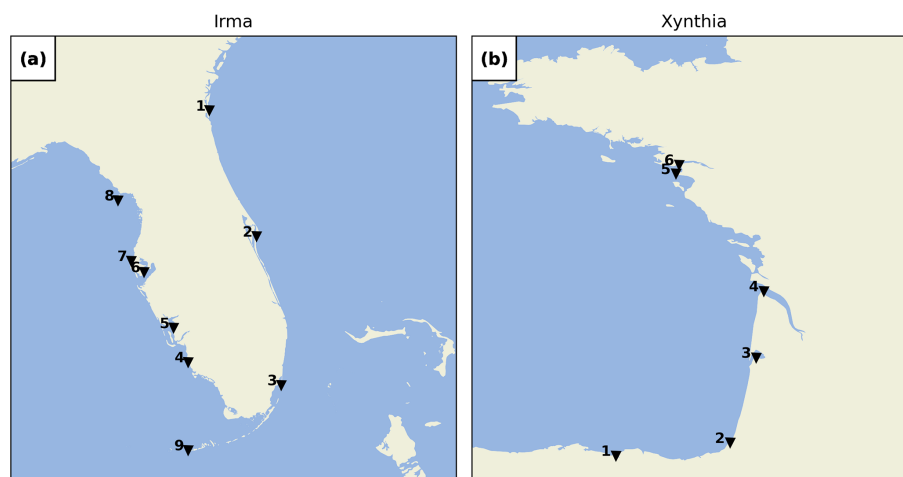


Figure 3. GESLA tide gauge stations for the Irma (a) and Xynthia (b) case studies.

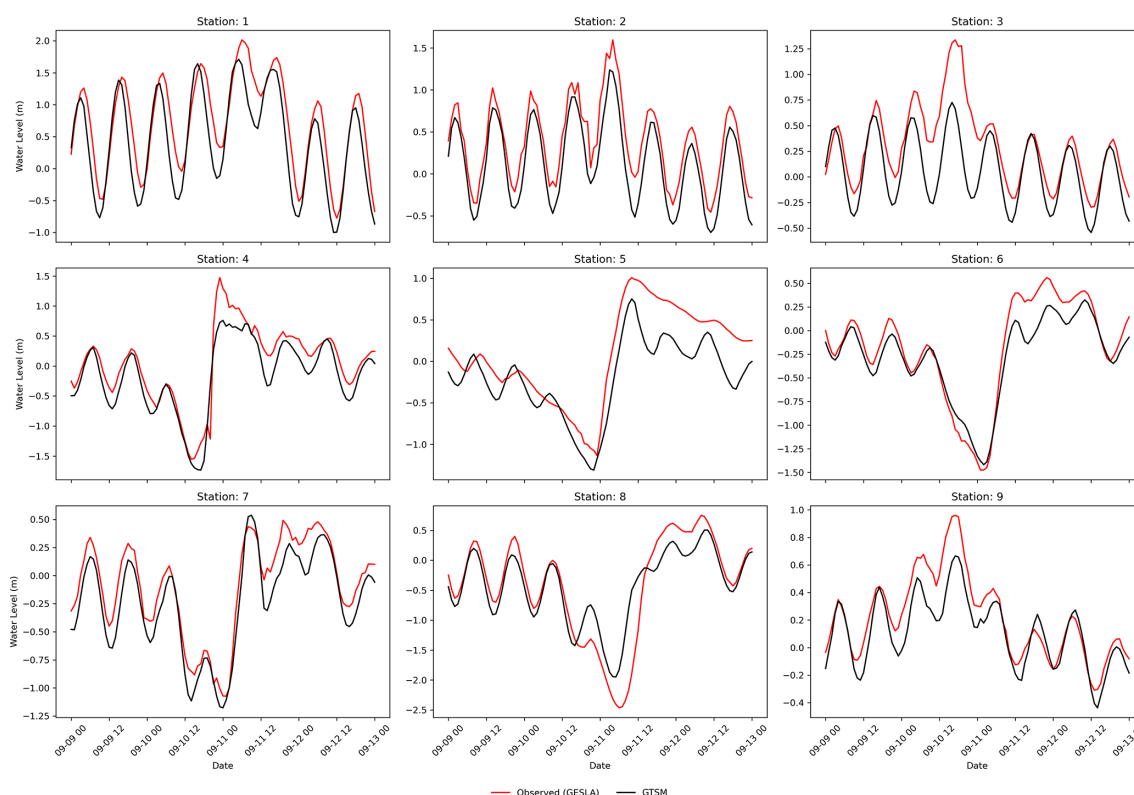


Figure 4. Validation of water levels for the Irma case study for the nine tide gauge stations depicted in Fig. 3a.

For this study, we use GEBCO 2020 (15 arcsec spatial resolution; Weatherall et al., 2020) as the input dataset for the bathymetry and FABDEM (30 m spatial resolution; Hawker et al., 2022) as the input dataset for the land elevation, except for ETC Xynthia. For ETC Xynthia, we use the 5 m resolution lidar-based digital elevation model (DEM) developed by the French National Geographic Institute (IGN) because it better represents dikes in the region, leading to better flood

estimates than with FABDEM (see Fig. A14). The spatially varying roughness coefficients used within SFINCS are derived from the land use maps of the Copernicus Global Land Service (Buchhorn et al., 2020). Within MOSAIC, SFINCS is coupled offline with water levels from GTSM at a 1 h resolution for the default settings. The Mean Dynamic Topography (DTU10MDT; Andersen and Knudsen, 2009) is used to convert the vertical reference of the water levels from mean

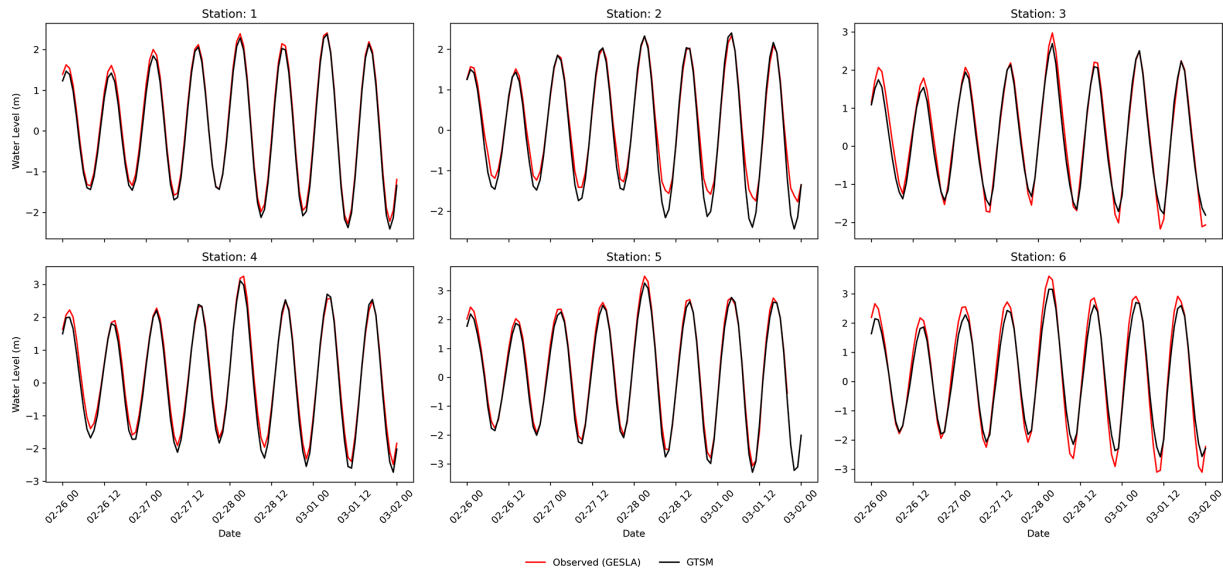


Figure 5. Validation of water levels for the Xynthia case study for the six tide gauge stations depicted in Fig. 3b.

sea level to the EGM2008 geoid. The resulting flood hazard maps have a resolution of 30 m.

To build the SFINCS models and couple them with GTSM, MOSAIC uses the HydroMTv0.7.1 (Hydro Model Tools) package (Eilander et al., 2023). HydroMT is an open-source Python package, which provides automated and reproducible model building and analysis of results. HydroMT uses a modular approach in which datasets and model setup configurations can easily be interchanged. In the MOSAIC framework presented in this paper, we take advantage of HydroMT in several ways: (1) to automatically convert the forcing files from GTSM and the other inputs into the model-specific input format, (2) to easily build a reproducible SFINCS model, and (3) to perform the analysis of the SFINCS model output. SFINCS is forced with GTSM water level time series at locations along every ~ 5 km of the coastline and provides as output the water level time series for each grid cell. Finally, flood depth maps are derived from the maximum water levels by subtracting the DEM.

To validate the hydrodynamic flood hazard modelling component of the modelling framework, we compare the modelled flood extents with observed flood extents derived from field measurements. This comparison is done for Xynthia, the only case study for which observed flood extent data are available (Breilh et al., 2013; DDTM, 2011). We measure the model skill using (1) the hit rate (H), defined as the flood area correctly simulated over the observed flooded area (Eq. 1); (2) the false-alarm ratio (F), defined as the area wrongly simulated over the observed flooded area (Eq. 2); and (3) the critical-success index (C), defined as the area correctly simulated as flooded over the union of the observed and modelled flooded areas (Eq. 3). Figure 6 shows the skill of the modelled maximum flood extents with SFINCS us-

ing the GTSM water levels as forcing. The hit rate is 0.78, correctly representing the flooding in most regions and only underestimating it in regions further inland. The false-alarm ratio of the model is 0.62. Flooding is overestimated in the north, likely due to the lack of flood protection measures included in the model that are present in reality. The critical-success index is 0.48 as a result of the areas that are simulated well and those that are over and underpredicted. While the performance of the flood model is negatively affected by the quality of the topography and the representation of local features such as dikes, we consider the performance sufficient for large-scale modelling and comparable to other studies such as Ramirez et al. (2016) and Vousdoukas et al. (2016b).

$$H = \frac{F_{\text{modelled}} \cap F_{\text{observed}}}{F_{\text{observed}}} \quad (1)$$

$$F = \frac{F_{\text{modelled}} / F_{\text{observed}}}{F_{\text{observed}}} \quad (2)$$

$$C = \frac{F_{\text{modelled}} \cap F_{\text{observed}}}{F_{\text{modelled}} \cup F_{\text{observed}}} \quad (3)$$

2.3 Sensitivity analysis

Using the MOSAIC modelling framework, we analyse the effects of refining the resolution of GTSM on the simulated water levels and assess how these propagate to the results for the flood hazard simulated by SFINCS. As described in Table 1, we categorise model configurations in two distinct groups. The first group, which contains the global model configurations (G), includes the default model configuration (G1) and configurations that modify only the global GTSM model (G2 and G3). In this group, the refinements applied are as follows: (1) the temporal output resolution, which is

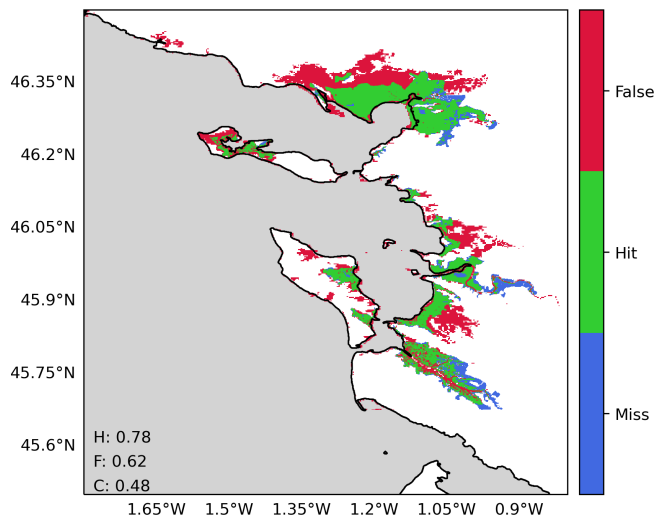


Figure 6. Validation of the flood hazard modelling component of the modelling framework for the Xynthia case study using the water levels of the default configuration of GTSM as forcing. The map compares the modelled and observed maximum flood extents, where green indicates flood areas simulated correctly, blue shows flood areas not simulated but observed, and red shows flood areas simulated but not observed. Performance indicators for the hit rate (H), false-alarm ratio (F), and critical-success index (C) are shown in the panel.

different from the implicitly calculated simulation time step of GTSM, is refined from 1 h to 10 min, allowing us to capture more changes in water levels, including the peaks of the water levels (G2), and (2) the spatial output resolution is refined from locations along the coast every ~ 5 to ~ 2 km, providing more coastal boundary conditions for the hydrodynamic flood hazard model (G3). The second group, which contains the nested model configurations (N), includes those model configurations that use a nested local model within the global model GTSM by performing dynamic downscaling. These model configurations include (1) the nesting of local high-resolution models with refined grids and GTSM (N1) and (2) the nesting of local high-resolution models with refined grids and updated bathymetry and GTSM (N2). Finally, we evaluate the combined effects of all these refinements through the “fully refined” configuration (N3), which integrates both the enhanced temporal and spatial resolutions and the nested high-resolution models and updated bathymetry. The validation of GTSM and SFINCS shows sufficient performance for all the model configurations from Table 1 and Fig. 7 (see Table A1 and Figs. A2, A3 and A15).

3 Sensitivity analysis of the model results

3.1 Multiscale storm surge modelling

Figure 8a, e, and i show the maximum water levels simulated by G1 for the three case studies and depict the maximum observed water levels for various GESLA tide gauge stations. To understand the effect of each individual refinement in the maximum water levels, Fig. 8 presents the differences in maximum water levels between each refinement and the model configuration G1. Figure 9 presents the differences in maximum water levels between the fully refined model configuration N3 and the model configuration G1.

3.1.1 Effects of higher resolution on water levels

Figure 8b, f, and j show that the refinement of the temporal output resolution of GTSM from 1 h to the 10 min intervals of G2 results in higher maximum water levels across the entire model domain for all three case studies. For TC Irma (Fig. 8b), the sensitivity of the water levels to the temporal refinement is relatively small, less than 10 cm. The small effect of the temporal refinement for TC Irma can be observed as well in Table A1 and Fig. A2, where G1 and G2 present similar time series and performance coefficients when compared to the observed water levels. For TC Haiyan (Fig. 8f), the sensitivity of the water levels is significant. Water levels increase due to the temporal refinement up to 2 m along the coastlines where TC Haiyan made landfall, showing that the 1 h resolution is too coarse to accurately capture the water level response. The cause of this is that TC Haiyan had rapid intensification, and when modelling water levels at a 1 h resolution, we overlook the storm’s peak, resulting in an underestimation of the maximum water levels. G2, however, can capture the peak of TC Haiyan more precisely (see Figs. A4 and A5). For ETC Xynthia (Fig. 8j), the sensitivity of the water levels to the temporal refinement is relatively small, less than 10 cm on average and slightly higher in enclosed basins and estuaries near La Rochelle. The small changes in water levels for ETC Xynthia are due to the inherent characteristics of ETCs, which typically have larger dimensions, lower intensity, and a slower rate of intensification compared to TCs. This means that the changes in water levels can be captured well at a 1 h resolution. The small effect of the temporal refinement for ETC Xynthia can be observed as well in Table A1 and Fig. A3, where G1 and G2 present similar time series and performance coefficients when compared to observed water levels.

The model configuration G3, where the spatial output resolution is refined, is not shown in Fig. 8 because increasing the number of water level locations does not change the water level values themselves. However, this refinement becomes significant when these values are applied as coastal boundary conditions to SFINCS (see Sect. 3.2.1), as a greater number

Table 1. GTSM model configurations used in the sensitivity analysis.

Model configuration	Nomen- clature	GTSM grid resolution	Bathymetry	Spatial output resolution	Temporal output resolution
Default configuration	G1	~ 25 to 2.5/1.25 km	GEBCO2019 ¹	Original (~ 5 km)	1 h
Refined temporal output resolution	G2	~ 25 to 2.5/1.25 km	GEBCO2019 ¹	Original (~ 5 km)	10 min
Refined spatial output	G3	~ 25 to 2.5/1.25 km	GEBCO2019 ¹	Refined (~ 2 km)	1 h
Dynamic downscaling (refined grid)	N1	~ 25 to 0.45 km	GEBCO2019 ¹	Original (~ 5 km)	1 h ²
Dynamic downscaling (refined grid + updated bathymetry)	N2	~ 25 to 0.45 km	GEBCO2023	Original (~ 5 km)	1 h ²
Fully refined configuration	N3	~ 25 to 0.45 km	GEBCO2023	Refined (~ 2 km)	10 min ²

¹ EMODnet2018 for Europe (Xynthia case study). ² For the model configurations N1, N2, and N3, the temporal output resolution is also the temporal resolution of the coupling between GTSM and the local high-resolution model.

of coastal boundary conditions gives additional information to the flood model.

3.1.2 Effects of dynamic downscaling with original bathymetry on water levels

Figure 8c, g, and k show that the model configuration N1 results in significant changes in water levels for all case studies. The largest differences occur along the coasts, where the largest changes in the model grid size resolution occur. For TC Irma (Fig. 8c), the nesting of a local model at a high resolution with GEBCO2019 results in maximum water levels that are up to 0.3 m higher than G1 in the south-west of Florida. These differences between N1 and G1 gradually increase over time and are at a maximum at the peak of TC Irma (Fig. A10). While a higher grid resolution affects the tidal propagation mainly along the coast of Florida (Figs. A6 and A7), storm surge propagation is more sensitive to the bathymetry used (Figs. A8 and A9). A high resolution is needed in areas with steep bathymetry. In contrast to the coarser grid of G1, N1 resolves complex topographic features around the barrier islands better (Fig. A11), allowing water to flow more freely through these barriers. At the 10 September 2017 time step in Fig. A10, when there is a negative surge north of the barrier island, G1 produces higher water levels because water remains trapped in the north. Conversely, during the peak of TC Irma on 11 September 2017, the water levels in G1 are lower than in N1 because less water is able to travel northwards. The increased northward surge of N1 propagates further into the Gulf of Mexico, leading to higher water levels that also propagate further into the Gulf of Mexico (see Fig. A10). Water levels for nine tide gauge stations along the coast indicate that while G1 underestimates the peak of TC Irma in most locations (Fig. A2, all stations but station 7), N1 simulates higher peaks on average, resulting sometimes in overestimations (Fig. A2, station 9). The improved resolution of topographic features in the barrier island region allows stations nearby (Fig. A2, stations 4 and 9) to capture the event’s peak better compared to G1. Additionally, the performance of N1 is slightly better than

G1 for six tide gauge stations (stations 1–6), as reflected in Table A1, which shows lower RMSE values. However, for stations 7–9, G1 shows a slightly higher RMSE and Pearson’s correlation. For TC Haiyan (Fig. 8g), the differences in maximum water levels are up to 1 m higher than in G1 near the landfall regions. These differences occur due to the refinement of the grid from 2.5 km to 45 m, which results in a significant increase in the number of model grid cells that define regions of shallow bathymetry, especially around the bay near Tacloban, resulting in a more detailed representation of water levels in that region. Thanks to the increase in grid cells, the strait north of Tacloban for N1 is defined using multiple grid cells in comparison to the two-grid-cell width of G1 (see Fig. A12). Therefore, in that region, N1 allows us to better resolve the topography of the region, and water can travel more easily northwards. For ETC Xynthia (Fig. 8k), the water levels from the nested local model at a high resolution are overall lower than the water levels for G1. Near La Rochelle, those water levels are up to 0.2 m lower. When comparing the performance of N1 with G1 (Table A1 and Fig. A3), both model configurations can predict the time series pattern well, with high Pearson’s correlation coefficients. Overall, the RMSE for Xynthia is similar for most tide gauge stations, except for two stations located at the mouths of estuaries (stations 3 and 6).

3.1.3 Effects of dynamic downscaling with updated bathymetry on water levels

Figure 8d, h, and l show that the model configuration N2 results in relatively large changes in the water levels for all the case studies. The largest differences occur along the coasts and provide results similar to those from N1. For TC Irma (Fig. 8d), the nesting of a local model at a high resolution with the updated GEBCO2023 bathymetry results in maximum water levels that are 0.3 m higher than G1 in the south of Florida. Compared to N1, model configuration N2 provides slightly higher water levels south of Florida. Those differences come from differences between GEBCO2023 and GEBCO2019 in the region. N2 shows a similar performance

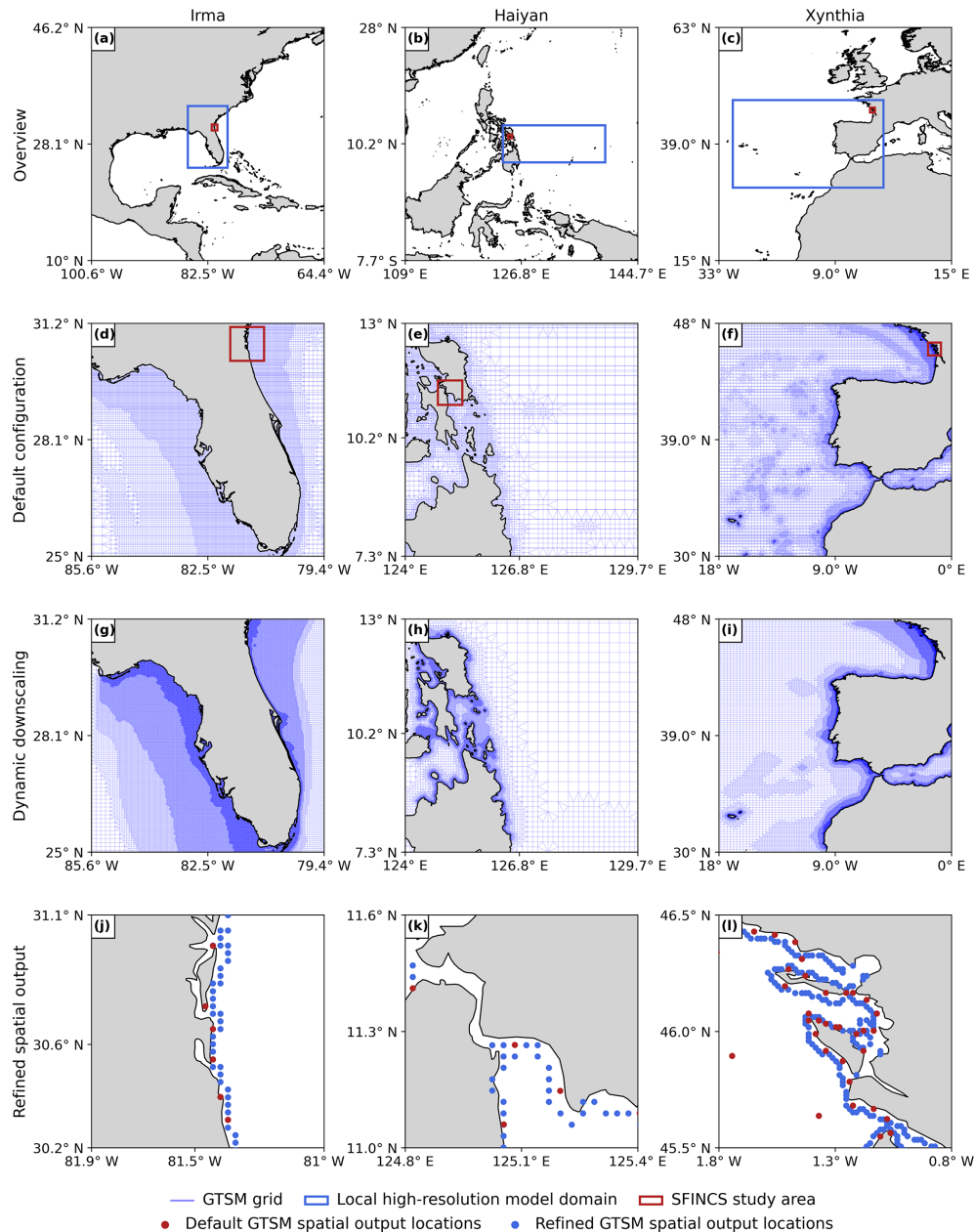


Figure 7. Overview of the model domains for the local high-resolution model and SFINCS for the three case studies (a, b, c), the default GTSM grid zoomed in (d, e, f), the local high-resolution model grid zoomed in (g, h, i), and the GTSM spatial output locations for the default configuration and the refined spatial output configuration zoomed in to the SFINCS study area (j, k, l).

to G1 and N1 across nine tide gauge stations (Table A1 and Fig. A2). For TC Haiyan (Fig. 8h), the differences in maximum water levels are up to 1 m higher than G1 at the landfall regions. Compared to N1, N2 provides, on average, higher maximum water levels, except in the Bay of Tacloban, where N1 presents higher maximum water levels on average. These differences come from the differences between GEBCO2019 and GEBCO2023. For ETC Xynthia (Fig. 8l), the water levels from the nested local model at a high resolution with GEBCO2023 are lower overall than the water levels for G1.

Compared to N1, the model configuration N2 provides a similar pattern of water level decrease; however, the maximum water level reduction compared to G1 is slightly less than for N1. The performance of N2, as shown in Table A1 and Fig. A3, is comparable to that of G1 and N2, except at two tide gauge stations (station 3 and 6) where GEBCO2023 does not accurately capture the bathymetry of the river channels in the estuaries. In contrast, EMODNET2018, the bathymetry used in model configuration N1, resolves these details better (see Fig. A13).

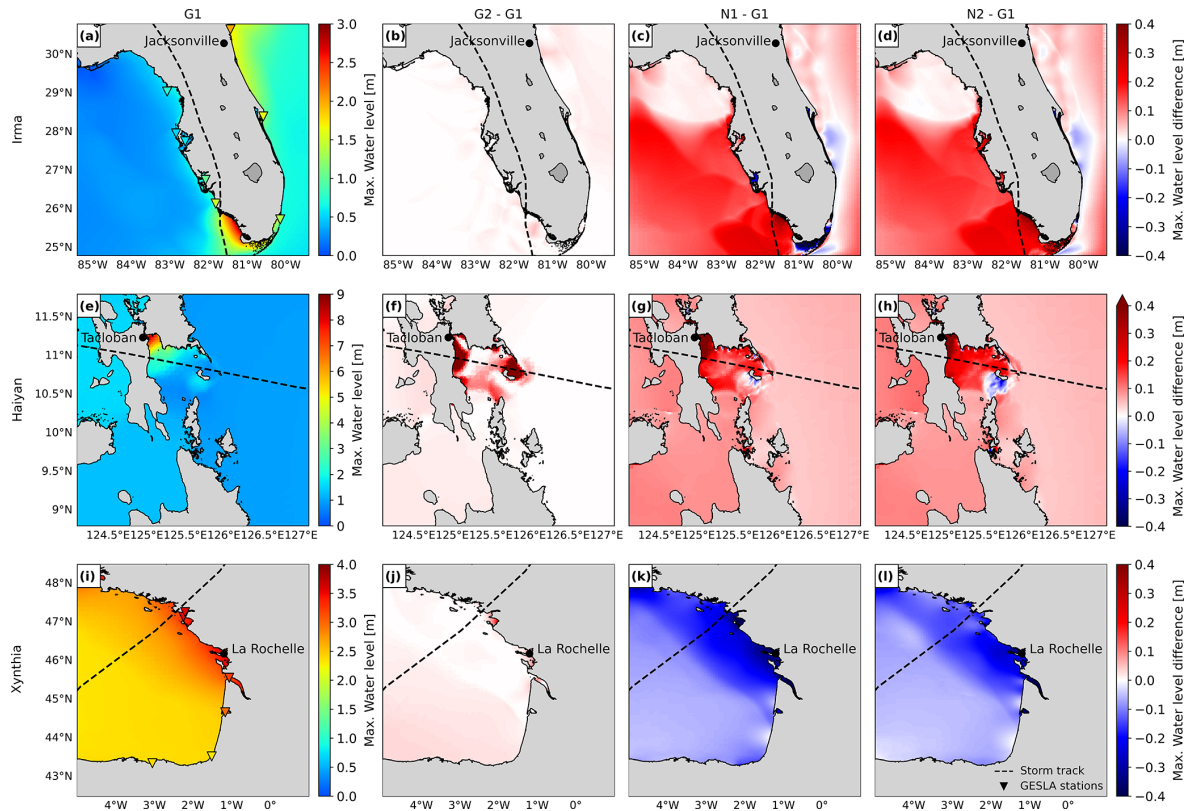


Figure 8. Panels (a), (e), and (i) show the maximum water levels for the three case studies using configuration G1, along with observed maximum water levels from GESLA tide gauge stations (inverted triangles). Panels (b), (f), and (j) show the difference in water levels between G2 and G1; panels (c), (g), and (k) show the difference between N1 and G1; and panels (d), (h), and (l) show the difference between N2 and G1. See Table 1 for configuration details

3.1.4 Effects of a fully refined model on water levels

In Fig. 9, we observe that the maximum water level differences between N3 and G1 lead to significantly different results for each case study. For TC Irma, N3 provides higher maximum water levels throughout almost the whole domain, resulting in a picture similar to that of N2 but with higher water levels along the southeast coast. The maximum differences in maximum water levels between N3 and N1 are up to 0.3 m. For TC Haiyan, N3 provides maximum water levels that resemble a combination of G2 in the regions where temporal refinement is relevant and N2 in the rest of the study area. The differences in maximum water levels between N3 and G1 for Haiyan are more than 2 m at the coast near Tacloban. Finally, for ETC Xynthia, N3 provides slightly higher maximum water levels in the south of the domain compared to G1, where the effects of G2 predominate, and lower maximum water levels in the north, where the effects of N2 are more dominant.

3.2 Hydrodynamic flood modelling

As a second step in the sensitivity analysis, we analyse how the effects of the different storm surge model configurations propagate to the SFINCS flood model. In Fig. 10, we compare the maximum flood depths of each refinement and G1. Figure 11 shows the maximum flood depth differences between N3 and G1.

3.2.1 Effects of a higher resolution on flood depths

Figure 10b, g, and l show that the refinement of GTSM's temporal output resolution from 1 h to the 10 min intervals for G2 provides different results for each case study. For TC Irma (Fig. 10b), the small increase in water levels as a result of the temporal output refinement (Sect. 3.1.1) also results in a small increase in flood depths. Conversely, TC Haiyan (Fig. 10g) experiences much higher water levels along the coast at a higher temporal resolution. As a result, it also experiences significantly higher flood depths, surpassing G1 by 1 m in regions near Tacloban. ETC Xynthia (Fig. 10l) experiences an increase in water levels along the coast for the 10 min temporal output resolution, especially in the study re-

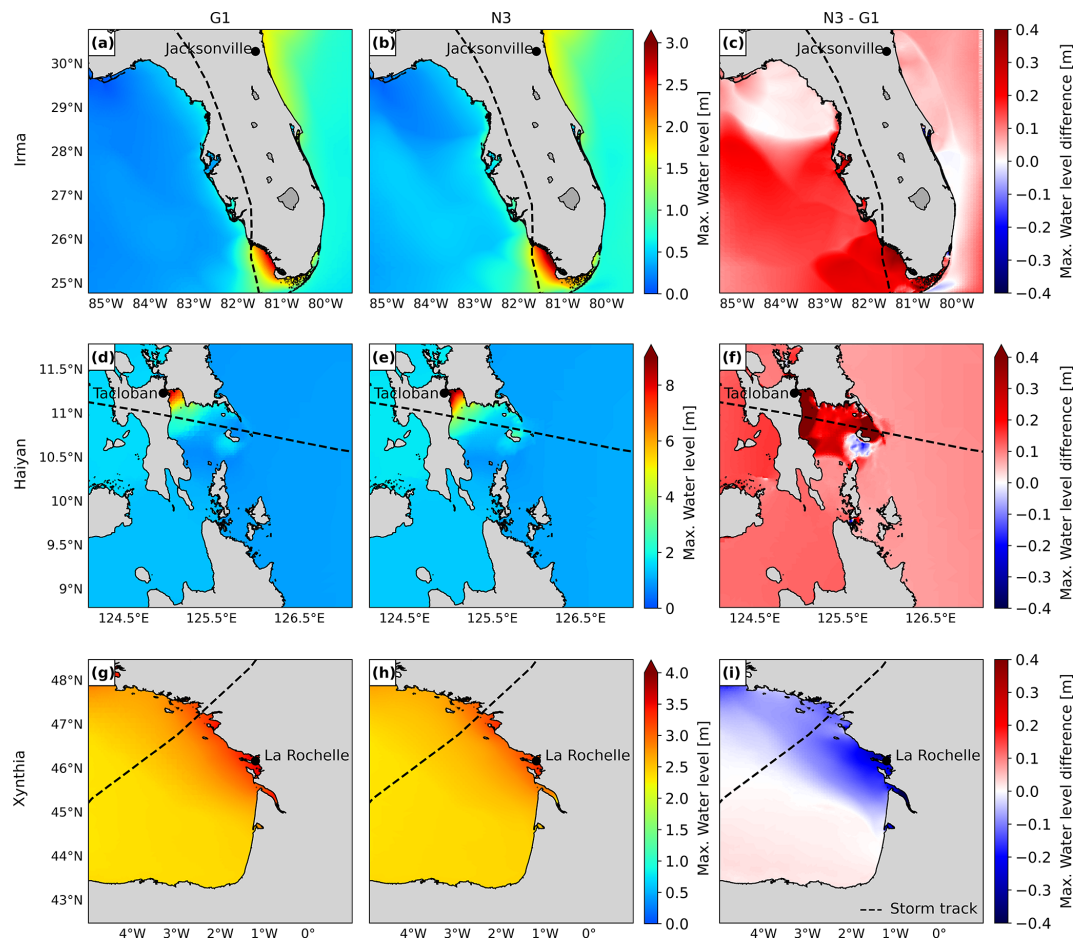


Figure 9. Panels (a), (d), and (g) show the maximum water levels for the default configuration G1. Panels (b), (e), and (h) show the maximum water levels for the fully refined configuration N3. Panels (c), (f), and (i) show the difference in maximum water levels between N3 and G1 for each case study.

gion of SFINCS. This results in an increase in flood depths of up to 0.1 m. For ETC Xynthia, G2 shows a higher hit rate and false-alarm ratio compared to G1 but the same critical-success index (see Fig. A15).

Figure 10c, h, and m show that refinement of the spatial output locations of G3 provides coastal boundary conditions for SFINCS at additional locations, thereby providing more water level input to the flood model. Figure 10c shows that this refinement results in lower flood depths north of and around Jacksonville for TC Irma. Conversely, for TC Haiyan (Fig. 10h), the increase in spatial inputs results in higher flood depths in most of the study area, in particular exceeding the G1 flood depths around Tacloban by more than 1 m. For ETC Xynthia (Fig. 10m), the refinement of spatial water level inputs leads to higher flood depths north of La Rochelle, up to 0.1 m, while south of La Rochelle, there are barely any changes compared to G1. For ETC Xynthia, G3 shows the same hit rate as G1, a higher false-alarm ratio, and the same critical-success index (see Fig. A15).

3.2.2 Effects of dynamic downscaling with original bathymetry on flood depths

Figure 10d, i, and n show that the model configuration N1 results in significant changes in the flood depths for all the case studies. For TC Irma (Fig. 10d), model configuration N1 leads to slightly higher water levels in comparison to G1. Consequently, the resulting flood depths are also larger and are more than 0.2 m above those of G1. Maximum water levels for TC Haiyan (Fig. 10i) are generally higher along the Bay of Tacloban when applying dynamic downscaling with the original bathymetry. This results, on average, in higher flood depths of more than 1 m compared to G1. Finally, ETC Xynthia (Fig. 10n) presents lower water levels for N1 compared to G1. Those lower water levels lead to lower flood depths across the whole model domain. For ETC Xynthia, N1 shows a lower hit rate and false-alarm ratio compared to G1 and the same critical-success index (see Fig. A15).

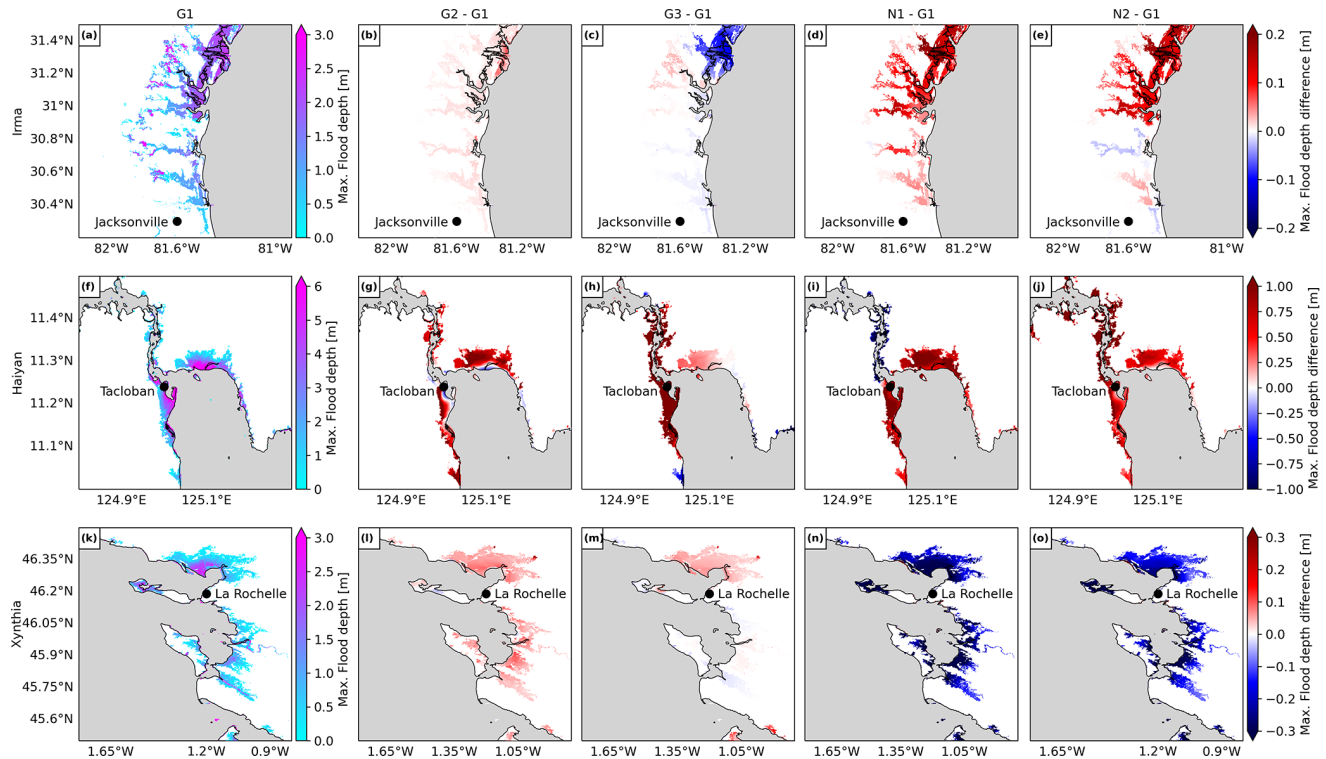


Figure 10. Panels (a), (f), and (k) show the maximum flood depth for the default configuration G1 for each case study. Panels (b), (g), and (l) show the difference between the maximum flood depth for the refined temporal output resolution configuration G2 and G1. Panels (c), (h), and (m) show the difference between the maximum flood depth for the refined spatial output configuration G3 and G1. Panels (d), (i), and (n) show the difference between the maximum flood depth for the dynamic downscaling (refined grid) configuration N1 and G1. Panels (e), (j), and (o) show the difference between the maximum flood depth for the dynamic downscaling (refined grid and updated bathymetry) configuration N2 and G1.

3.2.3 Effects of dynamic downscaling with updated bathymetry on flood depths

Figure 10e, j, and o show that the model configuration N2 results in significant changes in flood depths for all case studies. For TC Irma (Fig. 10e), model configuration N2 compared to G1 leads to higher or lower water levels depending on the region. Consequently, the resulting flood depths for N2 vary from 0.05 m lower to more than 0.2 m higher than G1. Maximum water levels for TC Haiyan (Fig. 10j) are generally higher in the Bay of Tacloban for model configuration N2 (when applying dynamic downscaling with the updated bathymetry) compared to G1. This results in larger flood depths, which, in some regions, result in a flood depth of more than 1 m higher compared to G1. However, in Tacloban Bay, N1 results, on average, in higher maximum water levels than N2, which leads to lower flood depths for N2 in comparison to N1. Finally, for ETC Xynthia (Fig. 10o) water levels are lower for N2 compared to G1. Those lower water levels lead to lower flood depths across the whole model domain. For ETC Xynthia, N2 shows a lower hit rate and false-alarm ratio compared to G1 and the same critical-success index (see Fig. A15).

3.2.4 Effects of a fully refined model on flood depths

For TC Irma, N3 provides higher water levels throughout large parts of the domain (Sect. 3.1.4) that translate into higher flood depths up to more than 0.2 m near Jacksonville. For TC Haiyan, N3 provides high water levels near Tacloban (Sect. 3.1.4), translating into high flood depths up to more than 1 m. Finally, ETC Xynthia presents lower water levels for N3 near La Rochelle (Sect. 3.1.4), which translate into lower flood depths along the coast.

To analyse the changes in flood depths over time, Fig. 12a, b, and c show the flood depth time series at the SFINCS output point locations outlined in Fig. 11 for all the model configurations. The timing and shape of the flood depth time series remain consistent across all the model configurations for all the case studies, and only slight differences in the magnitude of the flood depths are visible. Figure 12a shows that for TC Irma, all model configurations result in similar flood depths, and only N1 results in slightly higher flood depths, about 0.1 m higher than the others. Figure 12b shows that for TC Haiyan, G1 results in the lowest flood peak, while the temporal resolution of G2 plays a key role, enhancing the flood peak approximately 1 m higher than in G1. Finally,

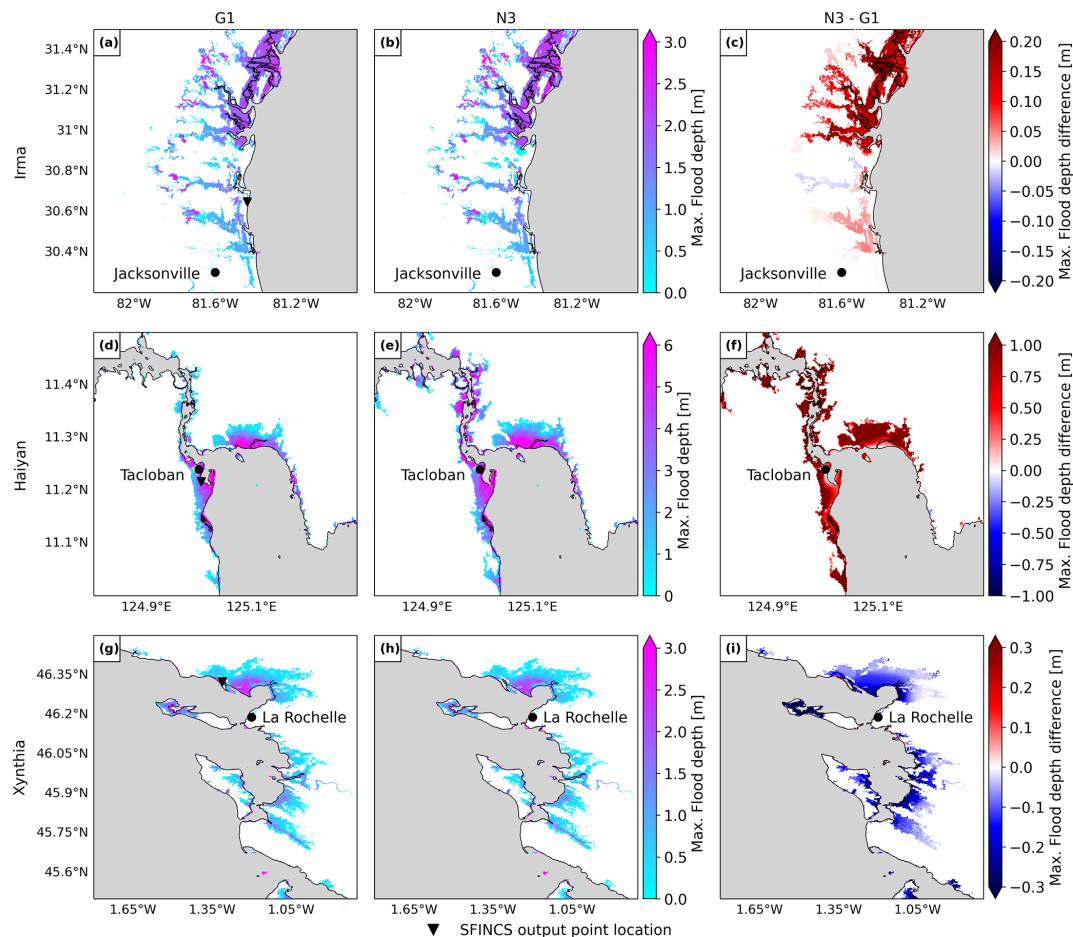


Figure 11. Panels (a), (d), and (g) show the maximum flood depth for the default configuration G1 for each case study. Panels (b), (e), and (h) show the maximum flood depth for the fully refined configuration N3. Panels (c), (f), and (i) show the difference between the maximum flood depths for N3 and G1.

Fig. 12c shows that all global model configurations (G1, G2, and G3) result in a first flood peak that is approximately 0.5 m higher than that of the nested model configurations (N1, N2, and N3). The second peak is simulated similarly by all model configurations, with N1 being the configuration that provides the lowest flood depths.

Panels a, b, and c in Fig. 12 only show the results for a single SFINCS output point location. However, the refinements might have most effect in other regions different from the SFINCS output point locations. To understand the overall effect of each refinement in the flood hazard maps, Fig. 12d, e, and f show the flood volume time series across each of the case study's model domains. While the timing and shape of the flood volume time series remains consistent across all the model configurations for all the case studies, there are differences in the magnitude of the flood volumes. Figure 12d shows that for TC Irma, the nested models lead to the highest flood volumes, with N3 being the model configuration that simulates the highest flood volume. On the other hand, the increase in spatial output of GTSM from G3 results in the

lowest flood volumes. Figure 12e shows that for TC Haiyan, N3 also leads to the highest flood volumes, while G1 results in the lowest volumes. Finally, Fig. 12f shows that for ETC Xynthia, the nested model configurations lead to the lowest flood volumes, while the global models result in higher flood volumes.

4 Discussion

4.1 Sensitivity analysis and model validation

The results of the sensitivity analysis reveal the complexity of hydrodynamic modelling and the sensitivity to specific local settings and storm characteristics. The effect of nesting higher-resolution models on water level and flood depth varies. For instance, the fully refined model configuration N3 simulates higher water levels almost everywhere for TC Irma. However, for TC Haiyan and ETC Xynthia, certain regions show higher water levels with N3, while other regions show lower water levels compared to the default global G1

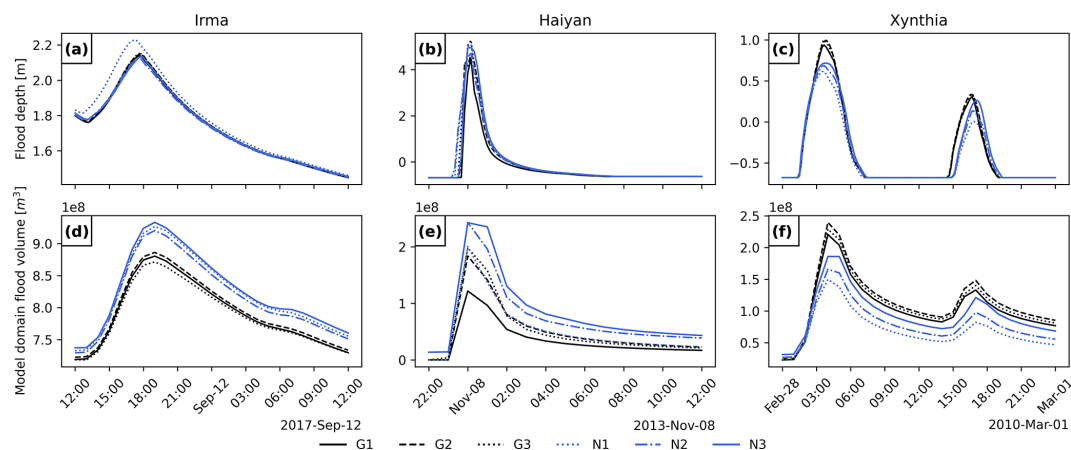


Figure 12. Flood depth time series for three observation points and flood volume time series for the SFINCS model domain of each case study and model configuration (see Table 1). The spatial location of the SFINCS output point locations can be observed in Fig. 11a, d, and g.

configuration. Similarly, flood depths around Jacksonville for TC Irma are generally higher with the refined model configuration N3, although some areas experience lower values. For TC Haiyan in Tacloban, flooding significantly increases with the refinements, whereas for ETC Xynthia, flood depths decrease notably around La Rochelle.

Refining the temporal output resolution (model configuration G2) has a significant influence on small, rapidly intensifying TCs like Haiyan. Compared to the default global configuration G1, this results in water levels and flood depths that are 2 and 1 m higher. For ETCs, the refinement of the temporal output resolution does not lead to substantial changes in water levels or flood depths, indicating that a 1 h temporal resolution is sufficient. Refining the spatial output locations of GTSM (model configuration G3) provides more detailed coastal boundary conditions for SFINCS. This is most relevant for regions where the coastal water levels show large spatial variations. For TC Haiyan, for example, the increase in coastal output locations in the Bay of Tacloban from 4 to 20 locations (see Fig. 7) leads to flood depths 1 m higher than in G1. Furthermore, regions with more complex topographies such as the south of Florida for TC Irma or the Tacloban Bay for TC Haiyan are influenced by the grid refinement of N1, leading to larger differences compared to G1 in terms of water levels and, consequently, flooding. The updating of bathymetry also plays an important role, contributing to differences between N1 and N2 in all the case studies.

The validation of our results also highlights the complexities of hydrodynamic modelling and how each specific setting impacts overall performance. It is challenging to assess the storm surge model performance due to the limited number of tide gauge stations available, with poor spatial coverage in many regions (Haigh et al., 2023). Another source of uncertainty is the location of these tide gauge stations, which are often situated in enclosed basins or harbours, where hydrodynamic models have more difficulty simulating water

levels compared to open-sea conditions. Besides, the validation of the flood hazard models is difficult due to the contribution of other flood drivers neglected in this study. The automated, uncalibrated MOSAIC configurations tested in this study have a storm surge modelling performance with Pearson's correlations above 0.92 and an average RMSE of less than 0.3 m. These results are comparable to the well-established GTSM model (Muis et al., 2016) and to other large-scale studies (Gori et al., 2023; Marsooli and Lin, 2018; Vogt et al., 2024). Similarly, the flood hazard modelling results align with those from other studies that simulated coastal flooding from ETC Xynthia (Ramirez et al., 2016; Vousdoukas et al., 2016a). All model configuration refinements perform adequately and have similar results, making it difficult to determine which configuration consistently provides the best overall performance based on the validation. This outcome largely depends on the storm characteristics and regional topography.

4.2 Limitations

There are several limitations that need to be taken into account when using MOSAIC. Limitations that are linked to general flood hazard modelling and are not specific to MOSAIC include the following: (1) the meteorological forcing data can be a large source of uncertainty when modelling extreme water levels (Dullaart et al., 2020). MOSAIC allows us to combine the results of the Holland parametric wind model with climate reanalysis datasets to enhance the wind and pressure fields at the peripheries of the TCs. Nonetheless, the implementation of more advanced parametric wind models or high-resolution climate data could further improve the water level simulations (Emanuel and Rotunno, 2011; Hu et al., 2011). (2) The accuracy of the bathymetry has a large influence on storm surge modelling (Mori et al., 2014; Woodruff et al., 2023). Global bathymetry is rather coarse and can have large errors (Weatherall et al., 2020), but for many regions,

high-resolution and accurate bathymetry is not available. This will impact the effect of dynamic downscaling, where MOSAIC uses bathymetry data to generate the model grid. Using higher-resolution local bathymetry enables finer grid refinement and higher accuracy of local data (Consortium EMODnet Bathymetry, 2018; NOAA, 2014; NOAA National Geophysical Data Center, 2001), which can enhance the accuracy of the results (Woodruff et al., 2023). (3) The accuracy of digital elevation models (DEMs) has a large influence on flood modelling (Hawker et al., 2022). In this paper, we use FABDEM and IGN, but MOSAIC allows us to replace the DEM with any dataset, and we recommend that users of MOSAIC use the best data available for their region of interest. In addition to the effects of DEMs, the presence of flood protection structures has a substantial impact on flood hazard models. The neglect of dikes in our SFINCS model is one of the reasons why our modelling framework overestimates flooding for ETC Xynthia. MOSAIC's HydroMT component supports the implementation of levees as 1D line features in the SFINCS model, and this ability could be used when there is local information on flood protection levels.

The main limitation specific to the automated approach of MOSAIC is related to the generation of the local high-resolution models. These automatically generated models can present instabilities when refined grid cells are present at the model boundaries. Therefore, care needs to be taken when applying dynamic downscaling. To solve this problem, the first 0.3° around the model domain are not refined in this study. When changes in grid refinement are abrupt, for example due to steep bathymetry, model instabilities can also occur. The nesting of multiple models inside of each other would allow for a smoother grid transition and might solve this issue. Nevertheless, placing the model boundaries in such a way that they do not cut through topographically complex regions is recommended. Furthermore, it is to be noted that the models presented here (except G1) are uncalibrated. Although they show adequate performance, detailed calibration of the bed level, bottom friction, and roughness coefficients could improve the modelling results (Wang et al., 2022a).

Automated modelling tools like MOSAIC have the advantage of being efficient, reducing potential human errors, and being reproducible and transparent. However, they also have their limitations. Users must be aware of the underlying modelling assumptions and should carefully review the model outputs of their specific case study (Remmers et al., 2024).

4.3 Directions for future research

There are various directions to further develop and improve MOSAIC. In this study, we have implemented MOSAIC to simulate coastal flooding driven by storm surges. However, flooding typically results from a combination of various drivers. Our results underestimate flooding near estuaries and deltas due to the exclusion of precipitation and

river discharge and underestimate it near steep coasts due to the exclusion of waves and overtopping. Considering the fact that HydroMT and SFINCS can include pluvial and fluvial drivers (Eilander et al., 2023), there is potential to incorporate the modelling of compound events into MOSAIC. Waves can significantly contribute to coastal flooding and, in some regions, are the dominant driver of extreme water levels (Parker et al., 2023). However, the inclusion of wave contributions in large-scale assessments has been limited due to the computational cost of traditional wave-resolving numerical models. The development of more computationally efficient wave solvers offers an opportunity to implement dynamic wave simulations in large-scale assessments and in MOSAIC. For instance, Leijnse et al. (2024) developed an efficient solver currently being integrated into SFINCS. Furthermore, this first version of MOSAIC makes use of offline coupling for both the local high-resolution model and the SFINCS model. However, new software developments such as the Oceanographic Multi-purpose Software Environment (OMUSE; Pelulessy et al., 2017) could be used to enable online coupling, as well as to further expand MOSAIC by allowing for coupling with other models such as hydrological or ocean models. We envisage various directions for the future application of MOSAIC beyond the modelling of historical coastal floods presented here. By leveraging the flexibility of MOSAIC to modify input datasets, the modelling framework can be used to study events under historical and climate change conditions. Furthermore, taking advantage of MOSAIC's multiscale modelling approach, TC/ETC high-resolution hazard assessments can be obtained globally. When linked to impact models, such as Delft-FIAT (Slager et al., 2016), MOSAIC could also be used for risk assessments.

4.4 Added value of the MOSAIC framework

The main added value of MOSAIC is its flexibility, simulating water levels and coastal flooding anywhere in the world, with customisable datasets and resolutions that enable efficient, region-specific storm event simulations. Users of MOSAIC can easily simulate storm events in any region with this modelling framework. First, they can select the appropriate meteorological forcing. Within MOSAIC, users can choose gridded meteorological data from reanalysis datasets or climate models to simulate ETCs or TCs, provided that the data accurately capture the TC wind and pressure fields (as seen with ETC Xynthia and TC Irma in this study). Alternatively, they can select a hybrid approach that combines the Holland model with ERA5 in the background when modelling smaller TCs with rapid intensification (such as TC Haiyan in this study). Depending on the specific storm simulated and the study area, users can select different model refinements. For rapidly intensifying storms, users can choose a more refined temporal output resolution, while nested models can help resolve the topography and bathymetry in regions with complex coastlines. If the users have coastal boundary condi-

tions available, MOSAIC can automatically generate stand-alone local high-resolution Delft3D FM models without having to couple them with GTSM. Although uncalibrated, these model configurations demonstrate similar performance to the well-established global model GTSM but at a significantly lower computational cost. The hydrodynamic flood modelling part of MOSAIC offers user-defined settings as well, enabling users, for instance, to choose the most suitable DEM for their study area or implement flood protection measures through MOSAIC's HydroMT component.

5 Concluding remarks

The MOSAIC modelling framework introduced in this study allows us to dynamically simulate coastal flooding events through the coupling of dynamic water level and overland flood models, making use of a Python environment. This approach is automated and reproducible and, combined with the underlying global datasets used, makes it globally applicable. MOSAIC's flexibility allows us to easily simulate coastal flooding events globally, while also using local high-resolution models. Based on our results, we conclude that the refinement of the global modelling approach can significantly impact the simulation of coastal water levels and flood depths at the local scale, although the differences in local settings mean that there is no one-size-fits-all approach. We recommend a higher-temporal-output resolution for rapidly intensifying TCs; a spatial output refinement for regions with heterogeneous water levels; and nested local models with high-resolution bathymetry, if available, for regions with complex topography. The flexibility and ease of use of MOSAIC make it a valuable resource for users to further explore the optimal settings for their case study and region of interest.

Appendix A: Supporting tables and figures

Table A1. Validation indicators that compare the maximum total water levels and observations from GESLA for the Irma and Xynthia case studies.

Irma					Pearson's correlation [–]			
RMSE [m]								
Station	G1	G2	N1	N2	G1	G2	N1	N2
1	0.41	0.41	0.39	0.40	0.92	0.92	0.92	0.92
2	0.28	0.27	0.25	0.25	0.98	0.98	0.98	0.98
3	0.33	0.33	0.32	0.33	0.79	0.78	0.81	0.79
4	0.27	0.26	0.21	0.24	0.96	0.96	0.96	0.94
5	0.35	0.35	0.33	0.31	0.93	0.93	0.93	0.93
6	0.18	0.18	0.17	0.21	0.98	0.98	0.98	0.94
7	0.17	0.17	0.14	0.14	0.97	0.97	0.95	0.95
8	0.39	0.39	0.42	0.45	0.92	0.92	0.90	0.88
9	0.16	0.16	0.18	0.10	0.93	0.92	0.90	0.96
Average	0.28	0.28	0.27	0.27	0.93	0.93	0.93	0.92
Standard deviation	0.09	0.09	0.10	0.11	0.06	0.06	0.05	0.05
Xynthia					Pearson's correlation [–]			
RMSE [m]								
Station	G1	G2	N1	N2	G1	G2	N1	N2
1	0.12	0.13	0.13	0.13	1.00	1.00	1.00	1.00
2	0.27	0.29	0.22	0.26	0.99	0.99	0.99	0.99
3	0.21	0.20	0.47	0.61	0.99	0.99	0.95	0.91
4	0.20	0.21	0.19	0.34	1.00	1.00	1.00	0.98
5	0.18	0.18	0.24	0.25	1.00	1.00	0.99	0.99
6	0.34	0.31	0.49	0.92	0.99	0.99	0.98	0.90
Average	0.22	0.22	0.29	0.42	1.00	1.00	0.99	0.96
Standard deviation	0.08	0.07	0.15	0.29	0.01	0.01	0.02	0.04

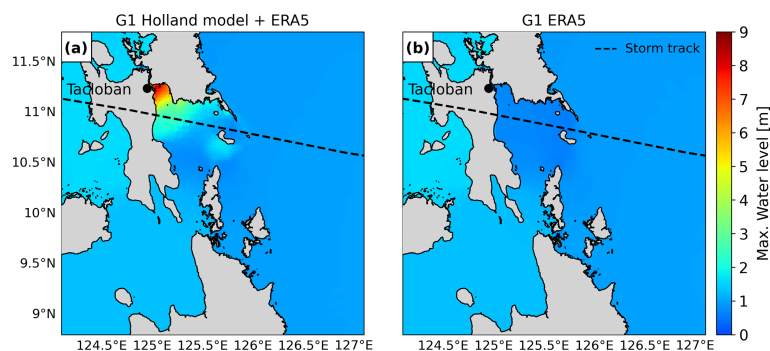


Figure A1. Maximum water level outputs of GTSM for the Haiyan case study with different meteorological forcings. **(a)** Maximum total water levels with the Holland model combined with ERA5 as forcing. **(b)** Maximum total water levels with ERA5 as forcing.

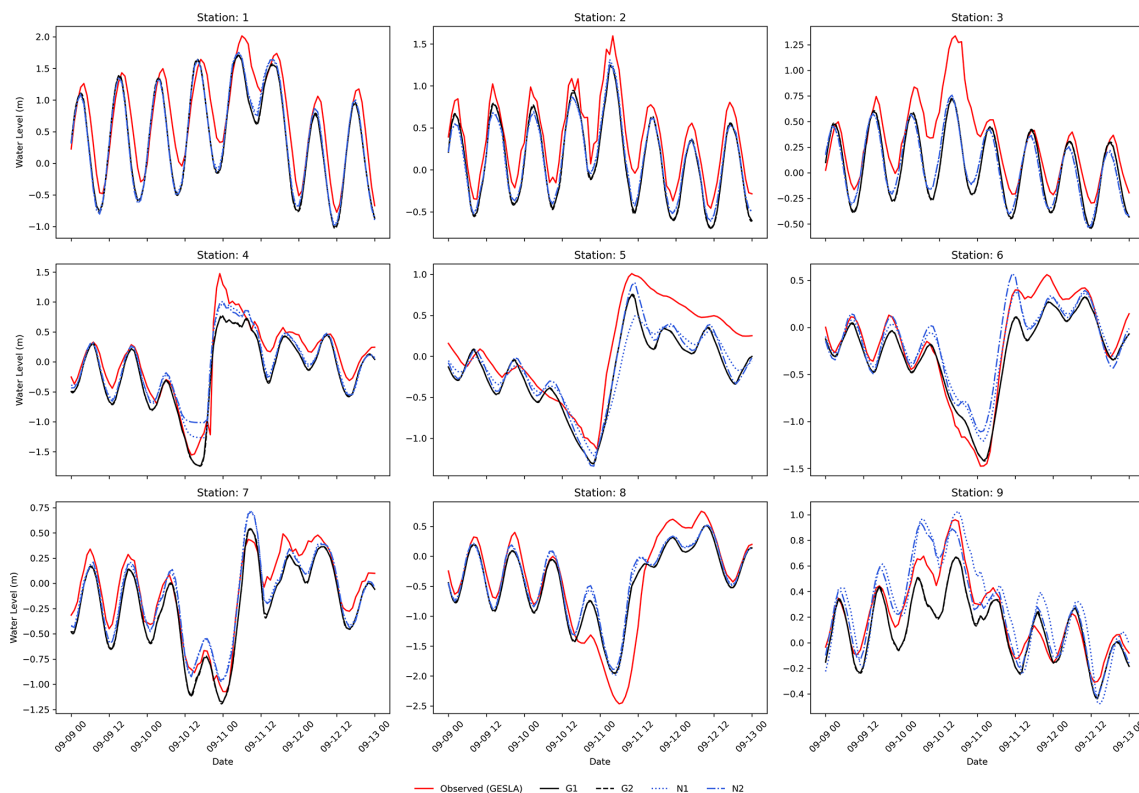


Figure A2. Validation of total water levels for the Irma case study for the nine locations depicted in Fig. 3a.

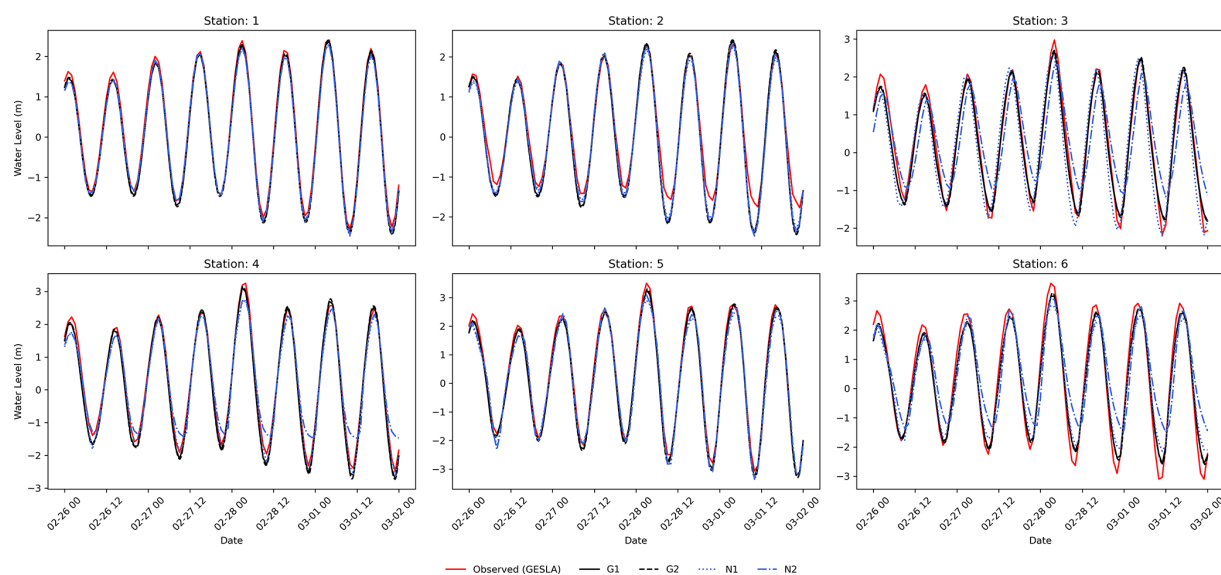


Figure A3. Validation of total water levels for the Xynthia case study for the six locations depicted in Fig. 3b.

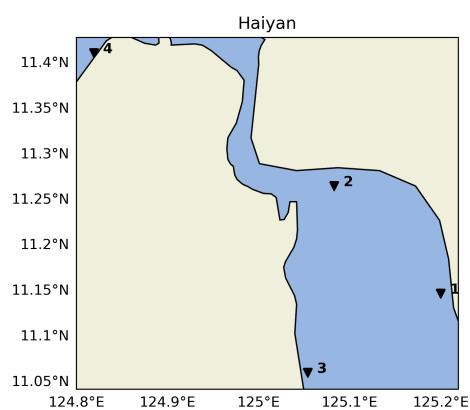


Figure A4. GTSM output locations for the Haiyan case study.

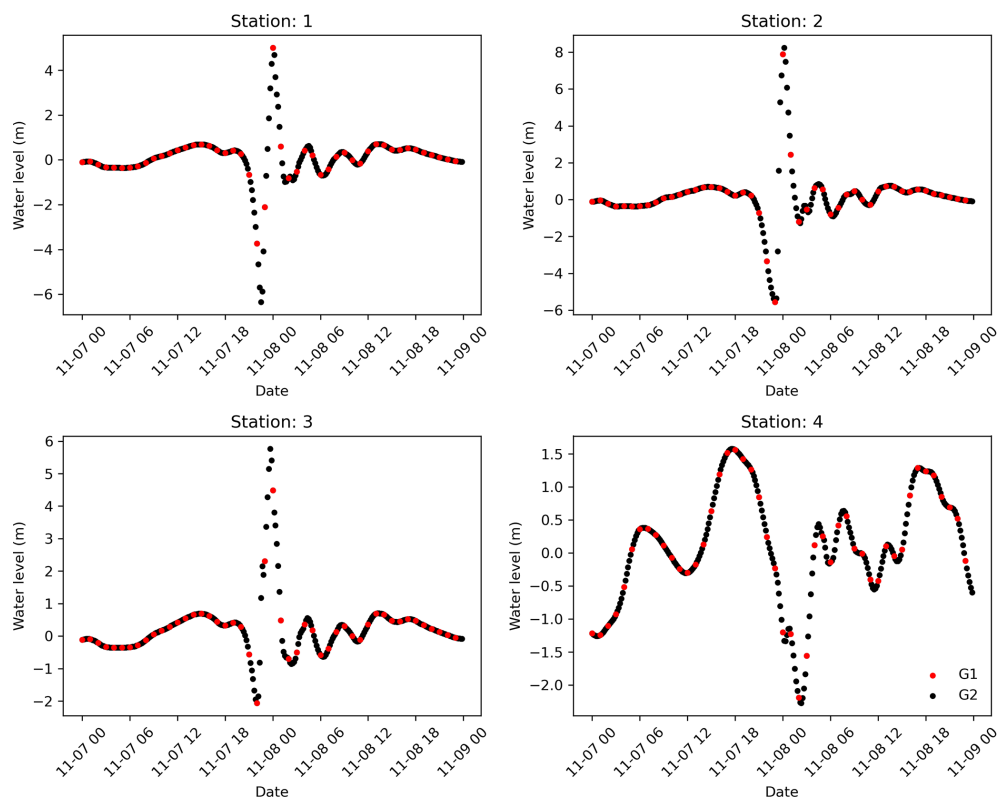


Figure A5. Haiyan total water level time series for the GTSM output locations provided in Fig. A4. Time series for the default configuration (G1) and the refined temporal output resolution configuration (G2).

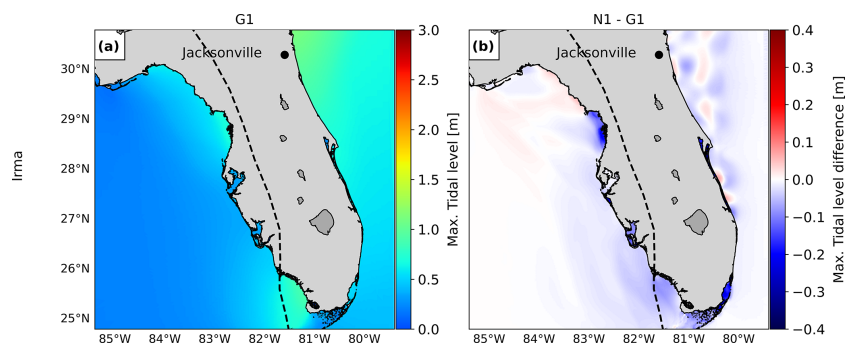


Figure A6. Maximum water levels for the tide-only simulation of G1 (a). The difference between the maximum water level for the tide-only simulations of N1 and G1 (b).

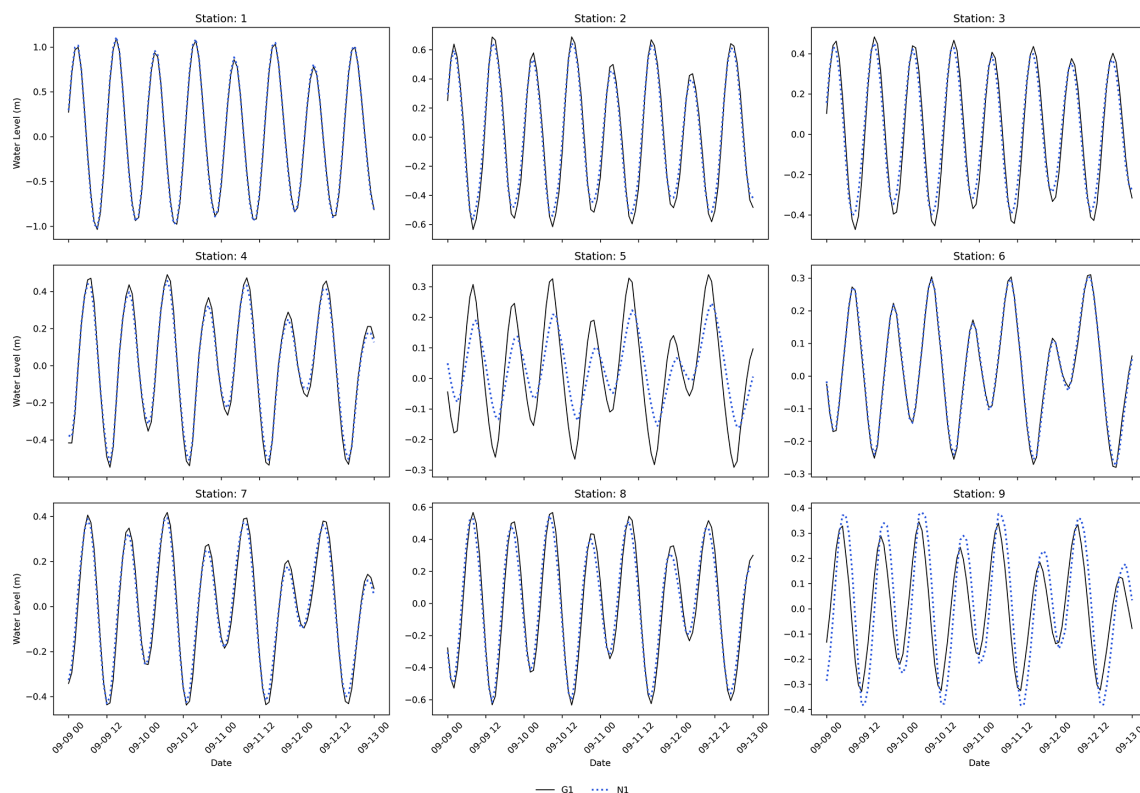


Figure A7. Water levels for the tide-only simulations for the Irma case study model configurations G1 and N1 for the nine locations depicted in Fig. 3a.

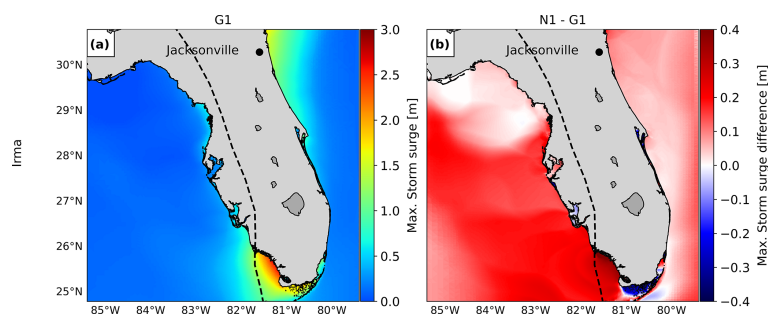


Figure A8. Maximum water levels for the storm-surge-only simulation of G1 (a). The difference between the maximum water level for the tide-only simulations of N1 and G1 (b).

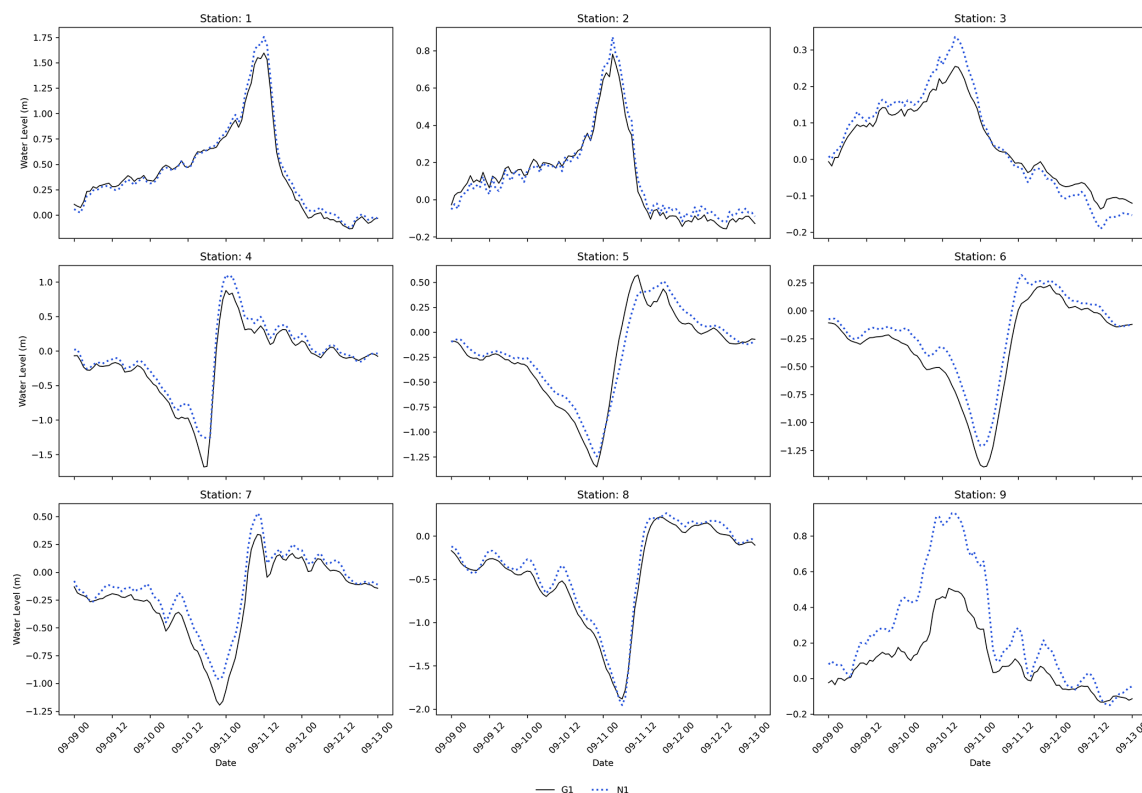


Figure A9. Water levels for the storm-surge-only simulations for the Irma case study model configurations G1 and N1 for the nine locations depicted in Fig. 3a.

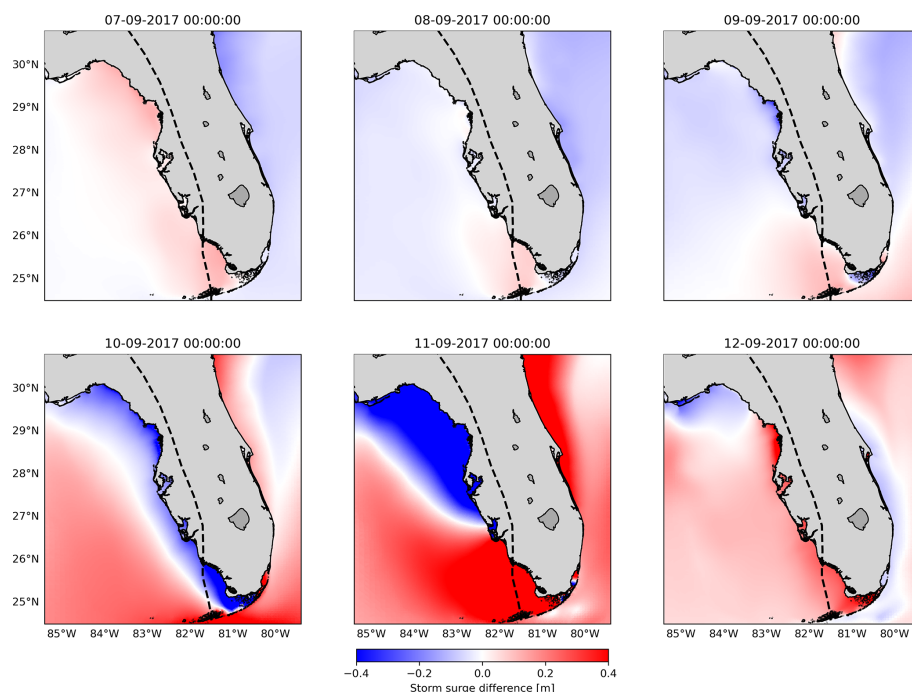


Figure A10. The difference in water levels for the storm-surge-only simulations of N1 and G1 for different time steps: before TC Irma made landfall (7 until 9 September 2017), during the peak (between 10 and 11 September 2017), and after the peak (12 September 2017).

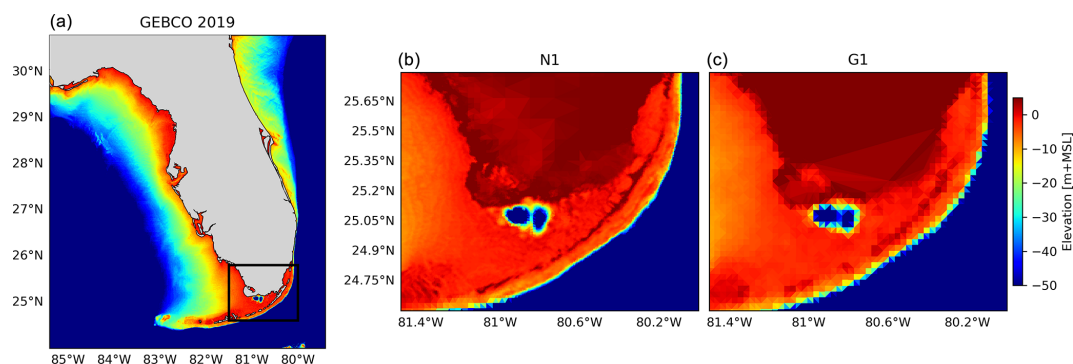


Figure A11. (a) GEBCO2019 for the study area; the black rectangle shows the barrier island region from the middle and right panels. (b) Bathymetry in the barrier island interpolated to the grid of the model configuration N1. (c) Bathymetry in the barrier island interpolated to the grid of the model configuration G1.

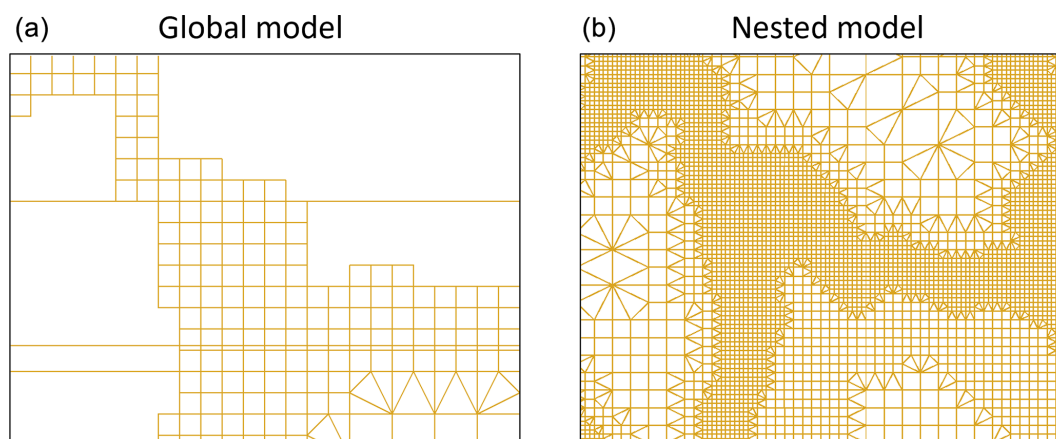


Figure A12. A close look at the unstructured grid of the global GTSM model with a grid resolution up to 2.5 km along the coast (a) and the nested grid of dynamic downscaling with a grid resolution up to 0.45 km along the coast (b) for the Haiyan case study.

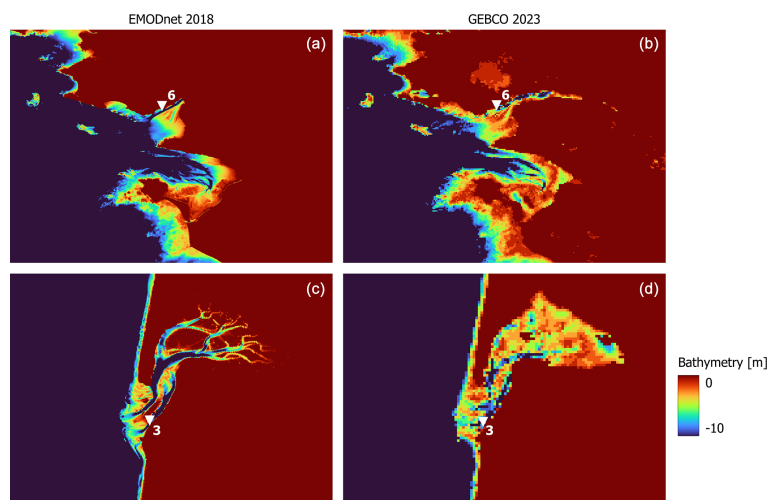


Figure A13. A close look at the bathymetry of two stations (**a, b**: station 6 and **c, d**: station 3) that provide lower performance with updated bathymetry for the Xynthia case study. (**a, c**) Bathymetric map of EMODNet2018. (**b, d**) Bathymetric map of GEBCO2023.

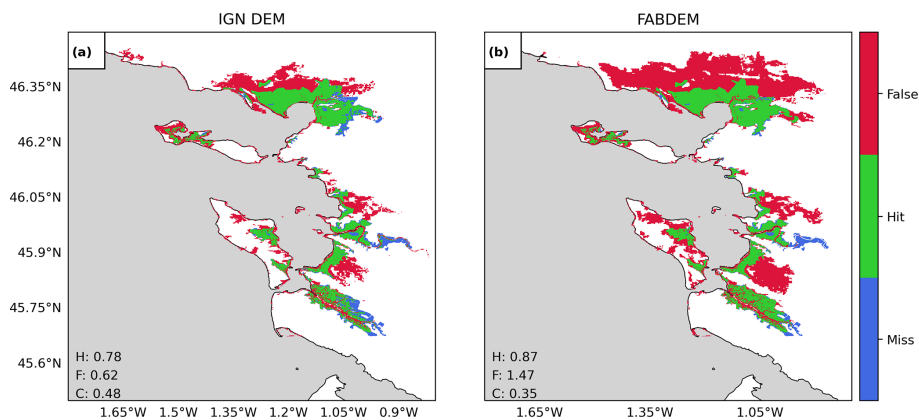


Figure A14. Validation of flood extents for the Xynthia case study against observed flood extents. The maps compare the modelled and observed maximum flood extents for a SFINCS model generated with ING's DEM (**a**) and FABDEM (**b**), where green indicates flood areas simulated correctly, blue shows flood areas not simulated but observed, and red shows flood areas simulated but not predicted. Performance indicators for the hit rate (H), false-alarm ratio (F), and critical-success index (C) are shown in each panel.

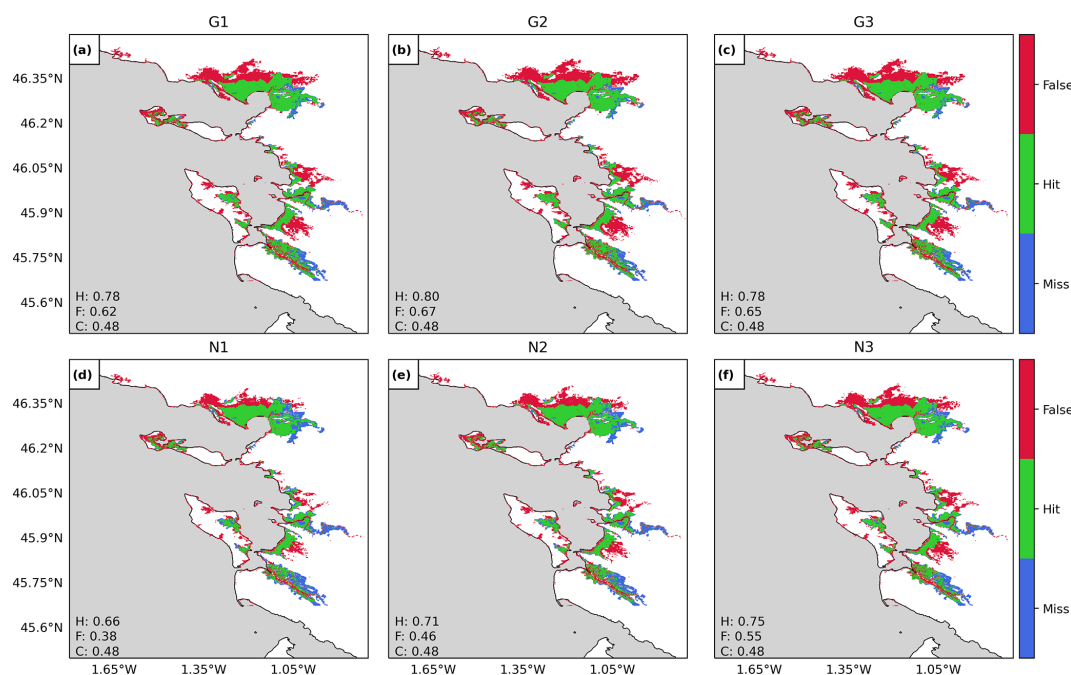


Figure A15. Validation of flood extents for the Xynthia case study against observed flood extents. The maps compare the modelled and observed maximum flood extents for each model configuration; see Table 1, where green indicates flood areas simulated correctly, blue shows flood areas not simulated but observed, and red shows flood areas simulated but not predicted. Performance indicators for the hit rate (H), false-alarm ratio (F), and critical-success index (C) for each configuration are shown in each panel.

Code availability. The underlying code for this study is available at <https://doi.org/10.5281/zenodo.15296397> (Benito, 2025a).

Data availability. The datasets compiled and/or analysed during the current study are available on Zenodo: <https://doi.org/10.5281/zenodo.15268382> (Benito, 2025b).

Author contributions. IB: conceptualisation, investigation, methodology, modelling, visualisation, analysis, writing – original draft. JCJHA: conceptualisation, investigation, methodology, writing – review and editing, supervision. PJW: conceptualisation, investigation, methodology, writing – review and editing, supervision. DE: conceptualisation, investigation, methodology, modelling, writing – review and editing, supervision. SM: conceptualisation, investigation, methodology, modelling, writing – review and editing, supervision.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Natural Hazards and Earth System Sciences*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. The authors would like to thank the SURF Co-operative for support in using the Dutch national e-infrastructure under grant nos. EINF-2224 and EINF-5779.

Financial support. This research has been supported by the European Research Council, H2020 European Research Council (grant no. 884442,) and the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (grant no. ASDI.2018.036).

Review statement. This paper was edited by Mauricio Gonzalez and reviewed by two anonymous referees.

References

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., and Feyen, L.: Global projections of

- river flood risk in a warmer world, *Earth's Future*, 5, 171–182, <https://doi.org/10.1002/2016EF000485>, 2017.
- Andersen, O. B. and Knudsen, P.: DNSC08 mean sea surface and mean dynamic topography models, *J. Geophys. Res.-Oceans*, 114, <https://doi.org/10.1029/2008JC005179>, 2009.
- Barnard, P. L., Van Ormondt, M., Erikson, L. H., Eshleman, J., Hapke, C., Ruggiero, P., Adams, P. N., and Foxgrover, A. C.: Development of the Coastal Storm Modeling System (CoSMoS) for predicting the impact of storms on high-energy, active-margin coasts, *Nat. Hazards*, 74, 1095–1125, <https://doi.org/10.1007/s11069-014-1236-y>, 2014.
- Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A. F., Limber, P., O'Neill, A. C., van Ormondt, M., Vitousek, S., Wood, N., Hayden, M. K., and Jones, J. M.: Dynamic flood modeling essential to assess the coastal impacts of climate change, *Sci. Rep.*, 9, 4309, <https://doi.org/10.1038/s41598-019-40742-z>, 2019.
- Barnard, P. L., Befus, K. M., Danielson, J. J., Engelstad, A. C., Erikson, L. H., Foxgrover, A. C., Hayden, M. K., Hoover, D. J., Leijnse, T. W. B., Massey, C., McCall, R., Nadal-Caraballo, N. C., Nederhoff, K., O'Neill, A. C., Parker, K. A., Shirzaei, M., Ohenhen, L. O., Swarzenski, P. W., Thomas, J. A., van Ormondt, M., Vitousek, S., Vos, K., Wood, N. J., Jones, J. M., and Jones, J. L.: Projections of multiple climate-related coastal hazards for the US Southeast Atlantic, *Nat. Clim. Change*, 15, 101–109, <https://doi.org/10.1038/s41558-024-02180-2>, 2025.
- Bates, P. D., Horritt, M. S., and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, *J. Hydrol.*, 387, 33–45, <https://doi.org/10.1016/j.jhydrol.2010.03.027>, 2010.
- Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., Savage, J., Olcese, G., Neal, J., Schumann, G., Gius-tarini, L., Coxon, G., Porter, J. R., Amodeo, M. F., Chu, Z., Lewis-Gruss, S., Freeman, N. B., Houser, T., Delgado, M., Hamidi, A., Bolliger, I., E. McCusker, K., Emanuel, K., Ferreira, C. M., Khalid, A., Haigh, I. D., Couasnon, A., E. Kopp, R., Hsiang, S., and Krajewski, W. F.: Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates, *Water Resour. Res.*, 57, e2020WR028673, <https://doi.org/10.1029/2020WR028673>, 2021.
- Batts, M. L., Cordes, M., Russell, L., Shaver, J., and Simiu, E.: Hurricane Wind Speeds in the United States, *Natl. Bur. Stand. Build. Sci. Ser.*, 106, <https://doi.org/10.1061/JSDEAG.0005541>, 1980.
- Benito, I.: Coastal flooding events developed following a multiscale flood modelling approach (TC Irma, TC Haiyan and ETC Xynthia), Zenodo [code], <https://doi.org/10.5281/zenodo.15296397>, 2025a.
- Benito, I.: Coastal flooding events developed following a multiscale flood modelling approach (TC Irma, TC Haiyan and ETC Xynthia), Zenodo [data set], <https://doi.org/10.5281/zenodo.15268382>, 2025b.
- Bertin, X., Bruneau, N., Breilh, J. F., Fortunato, A. B., and Karpytchev, M.: Importance of wave age and resonance in storm surges: The case Xynthia, Bay of Biscay, *Ocean Model.*, 42, 16–30, <https://doi.org/10.1016/j.ocemod.2011.11.001>, 2012.
- Breilh, J. F., Chaumillon, E., Bertin, X., and Gravelle, M.: Assessment of static flood modeling techniques: application to contrasting marshes flooded during Xynthia (western France), *Nat. Hazards Earth Syst. Sci.*, 13, 1595–1612, <https://doi.org/10.5194/nhess-13-1595-2013>, 2013.
- Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsend-bazar, N.-E., Herold, M., and Fritz, S.: Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2018: Globe, Zenodo, <https://doi.org/10.5281/zenodo.3518038>, 2020.
- Bunya, S., Dietrich, J. C., Westerink, J. J., Ebersole, B. A., Smith, J. M., Atkinson, J. H., Jensen, R., Resio, D. T., Luettich, R. A., Dawson, C., Cardone, V. J., Cox, A. T., Powell, M. D., Westerink, H. J., and Roberts, H. J.: A High-Resolution Coupled Riverine Flow, Tide, Wind, Wind Wave, and Storm Surge Model for Southern Louisiana and Mississippi. Part I: Model Development and Validation, *Mon. Weather Rev.*, 138, 345–377, <https://doi.org/10.1175/2009MWR2906.1>, 2010.
- Camus, P., Mendez, F. J., and Medina, R.: A hybrid efficient method to downscale wave climate to coastal areas, *Coast. Eng.*, 58, 851–862, <https://doi.org/10.1016/j.coastaleng.2011.05.007>, 2011.
- Cangialosi, J. P., Latta, A. S., and Berg, R.: Tropical cyclone report: hurricane Irma, National Hurricane Center, Miami, 2018.
- CGEDD: Tempête Xynthia: Retour d'expérience, évaluation et propositions d'action, 2010.
- Consortium EMODnet Bathymetry: EMODnet Digital Bathymetry (DTM 2018), <https://sextant.ifremer.fr/geonetwork/srv/api/records/18ff0d48-b203-4a65-94a9-5fd8b0ec35f6> (last access: 21 June 2022), 2018.
- DDTM: Éléments de mémoire sur la tempête Xynthia du 27 et 28 février 2010, <https://www.charente-maritime.gouv.fr/Actions-de-l-Etat/Environnement-risques-naturels-et-technologiques/Risques-naturels-et-technologiques/Generalites-sur-la-prevention-des-risques-naturels/Elements-de-memoire-Xynthia/Elements-de-memoire-sur-la-tempete-Xynthia-du-27-et-28-fevrier-2010> (last access: 16 October 2024), 2011.
- Deltares: Model description and development – Global Tide and Surge Model – Deltares Public Wiki, 649, <https://publicwiki.deltares.nl/display/GTSM/Model+description+and+development> (last access: 10 July 2024), 2021.
- Deltares: D-Flow Flexible Mesh User Manual, 2024.
- Dietrich, J. C., Bunya, S., Westerink, J. J., Ebersole, B. A., Smith, J. M., Atkinson, J. H., Jensen, R., Resio, D. T., Luettich, R. A., Dawson, C., Cardone, V. J., Cox, A. T., Powell, M. D., Westerink, H. J., and Roberts, H. J.: A High-Resolution Coupled Riverine Flow, Tide, Wind, Wind Wave, and Storm Surge Model for Southern Louisiana and Mississippi. Part II: Synoptic Description and Analysis of Hurricanes Katrina and Rita, *Mon. Weather Rev.*, 138, 378–404, <https://doi.org/10.1175/2009MWR2907.1>, 2010.
- Douris, J., Kim, G., Abrahams, J., Lapitan Moreno, J., Shumake-Guillemot, J., Green, H., and Murray, V.: WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970–2019) (WMO-No. 1267), WMO, Geneva, 1–90 pp., COI: 20.500.12592/b3d1n6, 2021.
- Dullaart, J. C. M., Muis, S., Bloemendaal, N., and Aerts, J. C. J. H.: Advancing global storm surge modelling using the new ERA5 climate reanalysis, *Clim. Dynam.*, 54, 1007–1021, <https://doi.org/10.1007/s00382-019-05044-0>, 2020.
- Dullaart, J. C. M., Muis, S., Bloemendaal, N., Chertova, M. V., Couasnon, A., and Aerts, J. C. J. H.: Accounting for tropical cyclones more than doubles the global population exposed to low-

- probability coastal flooding, *Commun. Earth Environ.*, 2, 1–11, <https://doi.org/10.1038/s43247-021-00204-9>, 2021.
- Eilander, D., Couasnon, A., Leijnse, T., Ikeuchi, H., Yamazaki, D., Muis, S., Dullaart, J., Haag, A., Winsemius, H. C., and Ward, P. J.: A globally applicable framework for compound flood hazard modeling, *Nat. Hazards Earth Syst. Sci.*, 23, 823–846, <https://doi.org/10.5194/nhess-23-823-2023>, 2023.
- Emanuel, K. and Rotunno, R.: Self-Stratification of Tropical Cyclone Outflow. Part I: Implications for Storm Structure, *J. Atmos. Sci.*, 68, 2236–2249, <https://doi.org/10.1175/JAS-D-10-05024.1>, 2011.
- GEBCO: General bathymetric chart of the oceans (GEBCO) 2014 Grid, <https://www.gebco.net/> (last access: 21 June 2022), 2014.
- Gori, A., Lin, N., and Smith, J.: Assessing Compound Flooding From Landfalling Tropical Cyclones on the North Carolina Coast, *Water Resour. Res.*, 56, e2019WR026788, <https://doi.org/10.1029/2019WR026788>, 2020.
- Gori, A., Lin, N., Schenkel, B., and Chavas, D.: North Atlantic Tropical Cyclone Size and Storm Surge Reconstructions From 1950-Present, *J. Geophys. Res.-Atmos.*, 128, e2022JD037312, <https://doi.org/10.1029/2022JD037312>, 2023.
- Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., Horsburgh, K., and Gouldby, B.: Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK, *Sci. Data*, 3, 1–14, <https://doi.org/10.1038/sdata.2016.107>, 2016.
- Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., Arns, A., Bradshaw, E., and Thompson, P.: GESLA Version 3: A major update to the global higher-frequency sea-level dataset, *Geosci. Data J.*, 10, 293–314, <https://doi.org/10.1002/gdj3.174>, 2023.
- Harper, B. A., Kepert, J. D., and Ginger, J. D.: Guidelines for Converting Between Various Wind Averaging Periods in Tropical Cyclone Conditions, WMO, 54 pp., 2010.
- Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., and Neal, J.: A 30 m global map of elevation with forests and buildings removed, *Environ. Res. Lett.*, 17, 024016, <https://doi.org/10.1088/1748-9326/ac4d4f>, 2022.
- Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, *ECMWF Newsl.*, 17–24, <https://doi.org/10.21957/vf291hehd7>, 2019.
- Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J., Marzeion, B., Fettweis, X., Ionescu, C., and Levermann, A.: Coastal flood damage and adaptation costs under 21st century sea-level rise, *P. Natl. Acad. Sci. USA*, 111, 3292–3297, <https://doi.org/10.1073/pnas.1222469111>, 2014.
- Hirabayashi, Y., Alifu, H., Yamazaki, D., Imada, Y., Shiogama, H., and Kimura, Y.: Anthropogenic climate change has changed frequency of past flood during 2010–2013, *Prog. Earth Planet. Sci.*, 8, 36, <https://doi.org/10.1186/s40645-021-00431-w>, 2021.
- Hodges, K., Cobb, A., and Vidale, P. L.: How Well Are Tropical Cyclones Represented in Reanalysis Datasets?, *J. Climate*, 30, 5243–5264, <https://doi.org/10.1175/JCLI-D-16-0557.1>, 2017.
- Holland, G. J., Belanger, J. I., and Fritz, A.: A revised model for radial profiles of hurricane winds, *Mon. Weather Rev.*, 138, 4393–4401, <https://doi.org/10.1175/2010MWR3317.1>, 2010.
- Hu, K., Chen, Q., and Kimball, S. K.: Consistency in hurricane surface wind forecasting: an improved parametric model, *Nat. Hazards*, 61, 1029–1050, <https://doi.org/10.1007/s11069-011-9960-z>, 2011.
- Islam, M. R., Lee, C.-Y., Mandli, K. T., and Takagi, H.: A new tropical cyclone surge index incorporating the effects of coastal geometry, bathymetry and storm information, *Sci. Rep.*, 11, 16747, <https://doi.org/10.1038/s41598-021-95825-7>, 2021.
- Kernkamp, H. W. J., Van Dam, A., Stelling, G. S., and de Goede, E. D.: Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf, *Ocean Dynam.*, 61, 1175–1188, <https://doi.org/10.1007/s10236-011-0423-6>, 2011.
- Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D., and Hinkel, J.: Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century, *Sci. Rep.*, 10, 1–12, <https://doi.org/10.1038/s41598-020-67736-6>, 2020.
- Lapidez, J. P., Tablazon, J., Dasallas, L., Gonzalo, L. A., Caba-caba, K. M., Ramos, M. M. A., Suarez, J. K., Santiago, J., Lagmay, A. M. F., and Malano, V.: Identification of storm surge vulnerable areas in the Philippines through the simulation of Typhoon Haiyan-induced storm surge levels over historical storm tracks, *Nat. Hazards Earth Syst. Sci.*, 15, 1473–1481, <https://doi.org/10.5194/nhess-15-1473-2015>, 2015.
- Leijnse, T., Nederhoff, K., Van Dongeren, A., McCall, R. T., and Van Ormondt, M.: Improving Computational Efficiency of Compound Flooding Simulations: the SFINCS Model with Subgrid Features, 2020, NH022-0006, 2020.
- Leijnse, T., Van Ormondt, M., Nederhoff, K., and Van Dongeren, A.: Modeling compound flooding in coastal systems using a computationally efficient reduced-physics solver: Including fluvial, pluvial, tidal, wind- and wave-driven processes, *Coast. Eng.*, 163, 103796, <https://doi.org/10.1016/j.coastaleng.2020.103796>, 2021.
- Leijnse, T. W. B., van Ormondt, M., van Dongeren, A., Aerts, J. C. J. H., and Muis, S.: Estimating nearshore infragravity wave conditions at large spatial scales, *Front. Mar. Sci.*, 11, <https://doi.org/10.3389/fmars.2024.1355095>, 2024.
- Lin, N. and Chavas, D.: On hurricane parametric wind and applications in storm surge modeling, *J. Geophys. Res.-Atmos.*, 117, 1–19, <https://doi.org/10.1029/2011JD017126>, 2012.
- Marcos, M., Rohmer, J., Voudoukas, M. I., Mentaschi, L., Le Cozannet, G., and Amores, A.: Increased Extreme Coastal Water Levels Due to the Combined Action of Storm Surges and Wind Waves, *Geophys. Res. Lett.*, 46, 4356–4364, <https://doi.org/10.1029/2019GL082599>, 2019.
- Marsooli, R. and Lin, N.: Numerical Modeling of Historical Storm Tides and Waves and Their Interactions Along the U.S. East and Gulf Coasts, *J. Geophys. Res.-Oceans*, 123, 3844–3874, <https://doi.org/10.1029/2017JC013434>, 2018.
- Mori, N., Kato, M., Kim, S., Mase, H., Shibutani, Y., Takemi, T., Tsuboki, K., and Yasuda, T.: Local amplification of storm surge by Super Typhoon Haiyan in Leyte Gulf, *Geophys. Res. Lett.*, 41, 5106–5113, <https://doi.org/10.1002/2014GL060689>, 2014.
- Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H., and Ward, P. J.: A global reanalysis of storm surges and extreme sea levels, *Nat. Commun.*, 7, 11969, <https://doi.org/10.1038/ncomms11969>, 2016.

- Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., Yan, K., and Verlaan, M.: A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections, *Front. Mar. Sci.*, 7, 1–15, <https://doi.org/10.3389/fmars.2020.00263>, 2020.
- Murakami, H.: Tropical cyclones in reanalysis data sets, *Geophys. Res. Lett.*, 41, 2133–2141, <https://doi.org/10.1002/2014GL059519>, 2014.
- Naval Meteorology and Oceanography Command: Naval Oceanography Portal, Best Track Archive, <https://www.metoc.navy.mil/jtwc/jtwc.html?best-tracks> (last access: 10 October 2024), 2022.
- Nederhoff, K., Leijnse, T. W. B., Parker, K., Thomas, J., O'Neill, A., van Ormondt, M., McCall, R., Erikson, L., Barnard, P. L., Foxgrover, A., Klessens, W., Nadal-Caraballo, N. C., and Massey, T. C.: Tropical or extratropical cyclones: what drives the compound flood hazard, impact, and risk for the United States Southeast Atlantic coast?, *Nat. Hazards*, 120, 8779–8825, <https://doi.org/10.1007/s11069-024-06552-x>, 2024.
- Nhamo, G. and Chikodzi, D.: Cyclones in Southern Africa: Volume 1: Interfacing the Catastrophic Impact of Cyclone Idai with SDGs in Zimbabwe, Springer International Publishing, Cham, <https://doi.org/10.1007/978-3-030-72393-4>, 2021.
- NOAA: Continuously Updated Digital Elevation Model (CUD-DEM) – Ninth Arc-Second Resolution Bathymetric-Topographic Tiles files, https://coast.noaa.gov/htdata/raster2/elevation/NCEI_ninth_Topobathy_2014_8483/ (last access: 3 February 2025), 2014.
- NOAA National Geophysical Data Center: U.S. Coastal Relief Model Vol.3 – Florida and East Gulf of Mexico, <https://doi.org/10.7289/V5W66HPP>, 2001.
- Parker, K., Erikson, L., Thomas, J., Nederhoff, K., Barnard, P., and Muis, S.: Relative contributions of water-level components to extreme water levels along the US Southeast Atlantic Coast from a regional-scale water-level hindcast, *Nat. Hazards*, 117, 2219–2248, <https://doi.org/10.1007/s11069-023-05939-6>, 2023.
- Pelupessy, I., van Werkhoven, B., van Elteren, A., Viebahn, J., Candy, A., Portegies Zwart, S., and Dijkstra, H.: The Oceanographic Multipurpose Software Environment (OMUSE v1.0), *Geosci. Model Dev.*, 10, 3167–3187, <https://doi.org/10.5194/gmd-10-3167-2017>, 2017.
- Pringle, W. J., Wirasaet, D., Roberts, K. J., and Westerink, J. J.: Global storm tide modeling with ADCIRC v55: unstructured mesh design and performance, *Geosci. Model Dev.*, 14, 1125–1145, <https://doi.org/10.5194/gmd-14-1125-2021>, 2021.
- Ramirez, J. A., Lichter, M., Coulthard, T. J., and Skinner, C.: Hyper-resolution mapping of regional storm surge and tide flooding: comparison of static and dynamic models, *Nat. Hazards*, 82, 571–590, <https://doi.org/10.1007/s11069-016-2198-z>, 2016.
- Remmers, J., Teuling, A., Dahm, R., Dam, A. van, and Melsen, L.: Power to the programmer: Modeller's perspective on automating the setup of hydrodynamic models for Dutch water authorities, *Socio-Environ. Syst. Model.*, 6, 18657–18657, <https://doi.org/10.18174/sesmo.18657>, 2024.
- Sebastian, A., Bader, D. J., Nederhoff, C. M., Leijnse, T. W. B., Bricker, J. D., and Aarninkhof, S. G. J.: Hindcast of pluvial, fluvial, and coastal flood damage in Houston, Texas during Hurricane Harvey (2017) using SFINCS, *Nat. Hazards*, 109, 2343–2362, <https://doi.org/10.1007/s11069-021-04922-3>, 2021.
- Slager, K., Burzel, A., Bos, E., De Bruikn, K., Wagenaar, D. J., and Winsemius, H. C.: User Manual Delft-FIAT version 1, 2016.
- Thomas, S. R., Nicolau, S., Martínez-Alvarado, O., Drew, D. J., and Bloomfield, H. C.: How well do atmospheric reanalyses reproduce observed winds in coastal regions of Mexico?, *Meteorol. Appl.*, 28, e2023, <https://doi.org/10.1002/met.2023.2021>, 2021.
- Tiggeloven, T., de Moel, H., Winsemius, H. C., Eilander, D., Erkens, G., Gebremedhin, E., Diaz Loaiza, A., Kuzma, S., Luo, T., Iceland, C., Bouwman, A., van Huijstee, J., Ligtoet, W., and Ward, P. J.: Global-scale benefit–cost analysis of coastal flood adaptation to different flood risk drivers using structural measures, *Nat. Hazards Earth Syst. Sci.*, 20, 1025–1044, <https://doi.org/10.5194/nhess-20-1025-2020>, 2020.
- UNDRR: The human cost of disasters: an overview of the last 20 years (2000–2019) | UNDRR, <https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019> (last access: 27 September 2022), 2020.
- Vafeidis, A. T., Schuerch, M., Wolff, C., Spencer, T., Merckens, J. L., Hinkel, J., Lincke, D., Brown, S., and Nicholls, R. J.: Water-level attenuation in global-scale assessments of exposure to coastal flooding: a sensitivity analysis, *Nat. Hazards Earth Syst. Sci.*, 19, 973–984, <https://doi.org/10.5194/nhess-19-973-2019>, 2019.
- Veenstra, J.: dfm_tools: A Python package for pre- and post-processing D-FlowFM model input and output files, Zenodo, <https://doi.org/10.5281/zenodo.10633862>, 2024.
- Vogt, T., Treu, S., Mengel, M., Frieler, K., and Otto, C.: Modeling surge dynamics improves coastal flood estimates in a global set of tropical cyclones, *Commun. Earth Environ.*, 5, 1–19, <https://doi.org/10.1038/s43247-024-01707-x>, 2024.
- Vousdoukas, M. I., Voukouvalas, E., Mentaschi, L., Dottori, F., Giardino, A., Bouziotas, D., Bianchi, A., Salamon, P., and Feyen, L.: Developments in large-scale coastal flood hazard mapping, *Nat. Hazards Earth Syst. Sci.*, 16, 1841–1853, <https://doi.org/10.5194/nhess-16-1841-2016>, 2016a.
- Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., and Feyen, L.: Projections of extreme storm surge levels along Europe, *Clim. Dynam.*, 47, 3171–3190, <https://doi.org/10.1007/s00382-016-3019-5>, 2016b.
- Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., and Feyen, L.: Extreme sea levels on the rise along Europe's coasts, *Earth's Future*, 5, 304–323, <https://doi.org/10.1002/2016EF000505>, 2017.
- Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., and Feyen, L.: Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard, *Nat. Commun.*, 9, 1–12, <https://doi.org/10.1038/s41467-018-04692-w>, 2018a.
- Vousdoukas, M. I., Bouziotas, D., Giardino, A., Bouwer, L. M., Mentaschi, L., Voukouvalas, E., and Feyen, L.: Understanding epistemic uncertainty in large-scale coastal flood risk assessment for present and future climates, *Nat. Hazards Earth Syst. Sci.*, 18, 2127–2142, <https://doi.org/10.5194/nhess-18-2127-2018>, 2018b.
- Wadey, M. P., Haigh, I. D., Nicholls, R. J., Brown, J. M., Horsburgh, K., Carroll, B., Gallop, S. L., Mason, T., and Bradshaw, E.: A comparison of the 31 January–1 February 1953 and 5–6 December 2013 coastal flood events around the UK, *Front. Mar. Sci.*, 2, <https://doi.org/10.3389/fmars.2015.00084>, 2015.

- Wahl, T.: Sea-level rise and storm surges, relationship status: Complicated!, *Environ. Res. Lett.*, 12, 111001, <https://doi.org/10.1088/1748-9326/aa8eba>, 2017.
- Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., and Slangen, A. B. A.: Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis, *Nat. Commun.*, 8, 1–12, <https://doi.org/10.1038/ncomms16075>, 2017.
- Wang, P. and Bernier, N. B.: Adding sea ice effects to a global operational model (NEMO v3.6) for forecasting total water level: approach and impact, *Geosci. Model Dev.*, 16, 3335–3354, <https://doi.org/10.5194/gmd-16-3335-2023>, 2023.
- Wang, X., Verlaan, M., Veenstra, J., and Lin, H. X.: Data-assimilation-based parameter estimation of bathymetry and bottom friction coefficient to improve coastal accuracy in a global tide model, *Ocean Sci.*, 18, 881–904, <https://doi.org/10.5194/os-18-881-2022>, 2022a.
- Wang, X., Verlaan, M., Apecechea, M. I., and Lin, H. X.: Parameter estimation for a global tide and surge model with a memory-efficient order reduction approach, *Ocean Model.*, 173, 102011, <https://doi.org/10.1016/j.ocemod.2022.102011>, 2022b.
- Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, G. A., Schäfer, A., Stanley, T., Veldkamp, T. I. E., and Winsemius, H. C.: Review article: Natural hazard risk assessments at the global scale, *Nat. Hazards Earth Syst. Sci.*, 20, 1069–1096, <https://doi.org/10.5194/nhess-20-1069-2020>, 2020.
- Weatherall, P., Tozer, B., Arndt, J. E., Bazhenova, E., Brिंगensparr, C., Castro, C., Dorschel, B., Ferrini, V., Hehemann, L., Jakobsson, M., Johnson, P., Ketter, T., Mackay, K., Martin, T., McMichael-Phillips, J., Mohammad, R., Nitsche, F., Sandwell, D., and Viquerat, S.: The GEBCO_2020 Grid – a continuous terrain model of the global oceans and land, <https://doi.org/10.5285/a29c5465-b138-234d-e053-6c86abc040b9>, 2020.
- Wing, O. E. J., Bates, P. D., Quinn, N. D., Savage, J. T. S., Uhe, P. F., Cooper, A., Collings, T. P., Addor, N., Lord, N. S., Hatchard, S., Hoch, J. M., Bates, J., Probyn, I., Himsworth, S., Rodríguez González, J., Brine, M. P., Wilkinson, H., Sampson, C. C., Smith, A. M., Neal, J. C., and Haigh, I. D.: A 30 m Global Flood Inundation Model for Any Climate Scenario, *Water Resour. Res.*, 60, e2023WR036460, <https://doi.org/10.1029/2023WR036460>, 2024.
- Woodruff, J., Dietrich, J. C., Wirasaet, D., Kennedy, A. B., and Bolster, D.: Storm surge predictions from ocean to subgrid scales, *Nat. Hazards*, 117, 2989–3019, <https://doi.org/10.1007/s11069-023-05975-2>, 2023.