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

Irene Benito¹, Jeroen C.J.H. Aerts¹,², Philip J. Ward¹,², Dirk Eilander¹,², and Sanne Muis¹,²

¹Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, The Netherlands.
²Deltares, Delft, The Netherlands.

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

Abstract. Tropical and extratropical cyclones, which can cause coastal flooding, are among the most devastating natural hazards. Understanding better coastal flood risk 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 still present limitations to provide actionable information at local scales. In order to address some of those limitations we present MOSAIC, a novel modelling framework that couples dynamic water level and overland flood models. MOSAIC follows a multiscale modelling approach in which local models with high-resolution are nested within a coarser large-scale model to obtain higher-resolution water levels and provide better coastal boundary conditions for dynamic flood modelling. To demonstrate the capabilities of MOSAIC we simulate three historical storm events. To merit the potential of MOSAIC’s multiscale modelling approach we perform a sensitivity analysis. Our findings indicate that various model refinements influence the simulation of total water levels and flood depths. The degree of importance of each refinement is linked to the local topography of the study area, the spatial heterogeneity of the water levels and the storm characteristics. MOSAIC provides a bridge between fully global and fully local modelling approaches, paving the way towards more actionable large-scale flood risk assessments.

1 Introduction

Coastal flood events can have devastating impacts on societies, economies, and the environment when affecting 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; Dullaart et al., 2021; Haigh et al., 2016; Reduction, 2020; Wahl et al., 2017). For example, Hurricane Harvey, in 2017, is one of the costliest storms in the United States’ history, with an estimated damage of $125 billion. Typhoon Idai, in Mozambique 2019, caused around 600 deaths and economic damages of $770 million (Nhamo & 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 €1.2 billion economic damages (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 (Kirezi et al., 2020; Marcos et al., 2019; Vousdoukas et al., 2017, 2018; 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, Voukouvalas, Annunziato, et al., 2016; Wang & Bernier, 2023). Subsequently, these water levels have been used to derive extreme water level values. These extreme water levels have then been used as input into global overland flood models, and the resulting flood hazard maps have been used to assess flood exposure and risk (Vousdoukas, Voukouvalas, Mentaschi, et al., 2016). 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 concentrated on either modelling extreme water levels or overland floods. However, each model component has its own limitations. We identify here two main methodological limitations of large-scale hazard assessments. First, bathymetry and coastal geometry strongly influence extreme sea levels (Bloemendaal et al., 2019; Dullaart et al., 2020; Mori et al., 2014), with large variability at 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). Grid refinement and nesting of local high-resolution models within coarser global models can result in improved coastal boundary conditions (Pelupessy et al., 2017). Second, coastal flooding is a dynamic process where flood duration and physics play a key role. However, the high computational costs associated with using hydrodynamic flood models for large-scale hazard assessments have limited most large-scale hazard assessments to static flood modelling methods, which neglect the dynamics of flooding events for large scales (Hinkel et al., 2014; Muis et al., 2016; Ramirez et al., 2016; Vafeidis et al., 2019; Vousdoukas, Voukouvalas, Mentaschi, et al., 2016). Consequently, to date the use of hydrodynamic models has been mainly limited to local applications.

To address these two main limitations, we introduce the MOSAIC (MOdelling Sea Level And Inundation for Cyclones) modelling framework. MOSAIC is a flexible Python-based modelling framework designed to dynamically simulate TC and ETC coastal flooding events. To enhance the accuracy in complex regions, MOSAIC applies a multiscale modelling approach in which local models with high-resolution (~45 m to 25 km) are nested within a large-scale model with a coarser resolution (~2.5 km to 25 km). To enable hydrodynamic flood modelling, MOSAIC couples two main existing modelling approaches: (1) to simulate water levels generated from storm surges and tides at global to local scale it couples the hydrodynamic Global Tide and Surge Model (GTSM) and Delft3D Flexible Mesh software; and (2) to dynamically simulate overland flooding at local scale it couples the simulated water levels with the Super-Fast INunadation of CoastS model (SFINCS). We use a reproducible approach that is globally applicable and that can automatically generate local Delft3D Flexible Mesh models as well as local SFINCS models. In this study, we showcase the potential of the MOSAIC framework by applying it to three case studies where large storm surges caused catastrophic flooding events, namely historical storm events 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. Moreover, we perform a sensitivity analysis of different modelling settings with the goal to present the potential merits of the multiscale approach.

![Image](https://example.com/image.png)

**Figure 1.** Case studies analysed on this paper. Left: Tropical cyclone Irma; middle: Tropical cyclone Haiyan; right: 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) (Section 2.1.1), including the dynamic downscaling with a local high-resolution model (Section 2.1.3); and (2) the overland flood hazard simulations using the SFINCS model (Section 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 the 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 type. We force ETC events with mean sea level pressure and 10 m meridional and zonal wind components from the ERA5 re-analysis dataset at a horizontal resolution of 0.25 degrees and 1 hour temporal resolution (Hersbach et al., 2019). We force TC events with pressure and wind from tropical cyclone track data merged with mean sea level pressure and wind components from ERA5. The wind and pressure from tropical cyclone track data are retrieved from the National Hurricane Center from NOAA and the Joint Typhoon Warning Center at 6 hourly intervals (Naval Meteorology and Oceanography Command, 2022; NOAA, 2023) and are converted to a polar grid with 36 radial bins, 375 arcs and a radius of 500 km using the Holland parametric wind model (Holland et al., 2010). The pressure and wind fields derived from track data are linearly interpolated from their outermost 33% with the ERA5 data in the background (Deltas, 2024).

#### 2.1.2 Global storm surge and tide model

MOSAIC uses GTSMv3.0 to simulate total 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 simulating water levels at large scales. The bathymetry in the model is the 15 arcseconds resolution EMODnet bathymetry dataset for Europe (Consortium EMODnet Bathymetry, 2018), and the 30 arcseconds General Bathymetric Chart of Oceans 2014 dataset for the rest of the globe (GEBCO, 2014). Tides are generated internally with tide generating forces, while storm surges originate from external forcing with pressure and fields (Section 2.1.1; Muis et al., 2020). 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).

#### 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 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 the Courant’s number derived from the bathymetry data (Veenstra, 2024). The settings to automatically generate the local high-resolution models used in this study can be found in Section 2.3. Second, MOSAIC uses an offline coupling approach to nest the local Delft3D Flexible Mesh model within GTSM. GTSM provides the water level time series at the model boundaries of the local model. Then, 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 for tropical cyclone Irma in 2017 (Eilander et al., 2022; 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 schematization from SFINCS for all the simulations (Leijnse et al., 2020).

For this study, we use GEBCO 2020 (15 arc seconds spatial resolution; (Weatherall et al., 2020) and FABDEM (30 m spatial resolution; (Hawker et al., 2022)) as input datasets for the bathymetry and the land elevation respectively. 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 1-hourly resolution for the default settings. The Mean Dynamic Topography (DTU10MDT; (Andersen & Knudsen, 2009)) is used to convert the vertical reference of the water levels from mean sea level to the EGM2008 geoid, which is the datum of FABDEM. The resulting flood hazard maps have a resolution of 30 m.

To build the SFINCS models and couple them with GTSM, MOSAIC makes use of the HydroMTv1.0.0 (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 setups 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 input into their relevant format; (2) to easily build a complete SFINCS model; and (3) to perform the analysis of the flood model output.

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 how those propagate into the results for the flood hazard simulated by SFINCS. As described in Table 1, we define a fully refined model configuration by combining three different refinements that can lead to more accurate results than those for the default GTSM: (1) the temporal output resolution, which is different than the implicitly calculated timestep of GTSM, is refined from 1-hourly to 10-minute, allowing to capture more changes in water levels, including the peaks of the water levels; (2) the spatial output resolution is refined from locations along the coast every ~5 km to ~2 km, providing more coastal boundary conditions for the hydrodynamic flood hazard model; and (3) the dynamic downscaling is performed to nest local high-resolution models into GTSM, allowing to resolve better the water levels on areas with complex topographies. Subsequently, we compare the coastal water levels and flood depths of the “fully refined” against the results of the default configuration of GTSM. To understand how each of the three model refinements affect the output of the fully refined configuration, we also analyse the effects of each individual refinement against the results of the default configuration. These comparisons will provide insights in the potential merits of refining the global modelling approach and dynamic downscaling. Figure 3 provides a detailed overview of the different modelling configurations for the three case studies.
Table 1. GTSM model configurations used in the sensitivity analysis.

<table>
<thead>
<tr>
<th>Model configuration</th>
<th>GTSM grid resolution</th>
<th>Bathymetry</th>
<th>Spatial output resolution</th>
<th>Temporal output resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default configuration</td>
<td>~25 to 2.5/1.25km</td>
<td>GEBCO2014*</td>
<td>Original (~5 km)</td>
<td>1h</td>
</tr>
<tr>
<td>Fully refined configuration</td>
<td>~25 to 0.45km</td>
<td>GEBCO2023</td>
<td>Refined (~2 km)</td>
<td>10min**</td>
</tr>
<tr>
<td>Refined temporal output</td>
<td>~25 to 2.5/1.25km</td>
<td>GEBCO2014*</td>
<td>Original (~5 km)</td>
<td>10min</td>
</tr>
<tr>
<td>Refined spatial output</td>
<td>~25 to 2.5/1.25km</td>
<td>GEBCO2014*</td>
<td>Refined (~2 km)</td>
<td>1h</td>
</tr>
<tr>
<td>Dynamic downscaling</td>
<td>~25 to 0.45km</td>
<td>GEBCO2023</td>
<td>Original (~5 km)</td>
<td>1h**</td>
</tr>
</tbody>
</table>

* EMODnet for Europe

** When applying dynamic downscaling, the temporal output resolution is also the temporal resolution of the coupling between GTSM and the local high-resolution model.
Figure 3. Overview of the model domains for the local high-resolution model and SFINCS, for the three case studies (panels a, b, c); default GTSM grid zoomed in (d, e, f); local high-resolution model grid zoomed in (g, h, i); and GTSM spatial output locations for the default configuration and the refined spatial output configuration, zoomed into the SFINCS study area (j, k, l).
Sensitivity analysis of the model results

3.1 Multiscale storm surge modelling

As a first step in the sensitivity analysis, we compare the maximum total water levels of the fully refined configuration and the default configuration. In Fig. 4 we observe that the maximum water level differences between the fully refined and the default configurations lead to significantly different results for each case study. For TC Irma the fully refined configuration provides higher maximum water levels throughout almost the whole domain, with maximum water levels up to 7.9 m, while the default configuration reaches 7.5 m. The maximum differences in maximum water levels are up to approximately 0.5 m. For TC Haiyan the fully refined and default configurations provide maximum water levels up to 13.5 m. However, there are differences in maximum water levels for Haiyan of more than 1 m both higher and lower than the default configuration depending on the region of the model domain. Finally, for ETC Xynthia the fully refined configuration provides lower maximum water levels across the whole domain, being up to 0.3 m lower than the default configuration.

Figure 4. Maximum water levels for the three case studies, for the default configuration (panels a, d, g) and for the fully refined configuration (panels b, e, h). Difference between the maximum water level for the fully refined model configuration and the default configuration (panels c, f, i).
To understand the contribution of each refinement in the maximum water level differences between the fully refined configuration and the default configuration, Fig. 5 presents the differences in maximum water levels between each refinement and the default configuration. Refining the temporal output resolution of GTSM from 1-hourly to 10-minute intervals (Fig. 5 panels b, e, h) results in higher maximum water levels across the entire model domain for all three case studies. For TC Irma (Fig. 5 panel b), the sensitivity of the water levels to the temporal refinement is relatively small, up to 20 cm. Water levels increase due to the temporal output refinement mostly at the location of landfall, south of Florida, while near Jacksonville the water level changes are much smaller. The cause for this might be the fact that TC Irma did not directly pass over Jacksonville, therefore the storm surge generated in that region is mainly being driven by less intense outer winds that are less sensitive to changes in temporal resolution. For TC Haiyan (Fig. 5 panel e), the sensitivity of the water levels is significant. Water levels increase due to the temporal refinement up to about 1 m along the coasts where TC Haiyan made landfall, showing that 1-hourly resolution is too coarse to accurately capture the water levels response. The cause for this might be that TC Haiyan had a rapid intensification, and when modelling water levels at 1-hourly resolution we might overlook the storm’s peak, resulting in an underestimation of the maximum water levels. For ETC Xynthia (Fig. 5 panel h), the sensitivity of the water levels to the temporal refinement is relatively small, up to 10 cm. The small changes in water levels for ETC Xynthia are likely 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 meteorological forcing can be well captured at a 1-hourly resolution.

Refining the spatial output resolution is not shown in Fig. 5 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 Section 3.2), as a greater number of coastal boundary conditions offer additional information for the flood model.

Dynamic downscaling (Fig. 5 panels c, f, i) results in relatively large changes in the water levels for all the case studies. The largest differences occur along the coasts. For TC Irma (Fig. 5 panels c), the nesting of a local model at high-resolution results in maximum water levels that are about 0.5 m higher than the default configuration in the south of Florida. Those differences might be due to changes in the bathymetric data from GEBCO2014 to GEBCO2023, which can resolve better the shallow regions around the landfall location and due to the grid refinement of the region near landfall, which provides more accuracy around the barrier islands, allowing to resolve better the water levels. Near Jacksonville, approximately 500 km distance from the landfall location, the difference in maximum water levels instead are approximately 0.1 m higher. For TC Haiyan (Fig. 5 panels f), the differences in maximum water levels are approximately up to 0.4 m higher and lower than the default configuration near the landfall regions. These differences might occur due to the update in bathymetry, where certain areas are shallower and certain areas deeper for GEBCO2023 in comparison to the default GEBCO2014 dataset. Furthermore, the refinement in the grid from 2.5 km to 45 m 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 resolution of water levels in that region. Thanks to the increase on grid cells, on the strait north of Tacloban GTSM is able to better resolve the water levels, and therefore, water can travel northwards. For ETC Xynthia (Fig. 5 panels i), the water levels from the nested local model at high-resolution are overall lower than water levels for the default configuration. Near La Rochelle, those water levels are approximately 0.3 m lower. The reason for this might be due to differences in the bathymetry in the region around La Rochelle, and due to the fact that the refined grid has more grid cells that resolve the water travelling through the estuaries in the region.
Figure 5. Maximum water levels for the three case studies, for the default configuration (panels a, d, g). Difference between the maximum water level for each specific model configuration and the default configuration. Difference in water levels for the temporal output resolution refined configuration (panels b, e, h) and the dynamic downscaling configuration (panels c, f, i).

3.2 Hydrodynamic flood modelling

As a second step in the sensitivity analysis, we compare the maximum flood depth of the fully refined configuration and the default configuration. In Fig. 6 we observe that the maximum flood depth differences between the fully refined and the default configuration lead to different results for each case study. For TC Irma the fully refined configuration provides higher water levels throughout the whole domain (Section 3.1) that translate into higher flood depths up to more than 0.2 m. For TC Haiyan, on the one hand the fully refined configuration provides lower water levels near Tacloban (Section 3.1), translating into lower flood depths, while on the other hand, regions south of Tacloban and north of the bay experience higher water levels that translate into higher flood depths as compared to the default configuration. Finally, ETC Xynthia presents lower water levels for the fully refined configuration (Section 3.1) that translate into lower flood depths along the coast, up to approximately 0.3 m across the whole domain. Further inland, however, flood depths are higher than for the default configuration.
Figure 6. Panels a, d, g show the maximum flood depth for the default configuration for each case study. Panels b, e, h show the maximum flood depth for the fully refined configuration. Panels c, f, i show the difference between the maximum flood depth for the fully refined and the default configuration.

To understand the contribution of each refinement in the maximum flood depth differences between the fully refined configuration and the default configuration, Fig. 7 presents the differences in maximum flood depths between each refinement and the default configuration.

Refining GTSM’s temporal output resolution from 1-hourly to 10-minute intervals (Fig. 7 panels b, f, j) provides different results for each case study. For TC Irma (Fig. 7 panel b), the small increase in water levels as a result of the temporal output refinement (Section 3.1) also results in a small increase in flood depths. Conversely, TC Haiyan (Fig. 7 panel f) experiences much higher water levels at higher temporal resolution. As a result, it also experiences significantly higher flood depths, surpassing the default configuration by 1 m in regions near Tacloban. ETC Xynthia (Fig. 7 panel j) experiences an increase in water levels along the coast for the 10-minute temporal output resolution, which results in a general increase in flood depths of approximately 0.1 m.

Refining the spatial output locations from GTSM provides coastal boundary conditions to SFINCS at additional locations, thereby improving the accuracy of the water levels input to the flood model. Figure 7 panel c shows that this refinement results in lower flood depths north of Jacksonville for TC Irma. Conversely, for TC Haiyan (Fig. 7 panel g), the increase in spatial inputs results in higher flood depths in most of the study area, particularly exceeding 1 m the default configuration flood depths.
around Tacloban. For ETC Xynthia (Fig. 7 panel k) the refinement of spatial water level inputs overall leads to higher flood depths, being north of La Rochelle approximately 0.1 m higher.

Dynamic downscaling for TC Irma (Fig. 7 panel d) leads to higher water levels in comparison to the default configuration. Consequently, the resulting flood depths are also higher, exceeding approximately 0.2 m above those of the default configuration. Model results for TC Haiyan (Fig. 7 panel h) generally result in lower water levels in the bay of Tacloban when applying dynamic downscaling, resulting in lower flood depths. On the strait north of Tacloban, GTSM is able to better resolve the water levels, and therefore, water can travel north, leading to higher flood depths along the strait. Finally, ETC Xynthia (Fig. 7 panel l) has lower water levels due to the dynamic downscaling. Those lower water levels lead to lower flood depths across the whole model domain.

Figure 7. Panels a, e, i show the maximum flood depth for the default configuration for each case study. Panels b, f, j show the difference between the maximum flood depth for the temporal refined configuration and the default configuration. Panels c, g, k show the difference between the maximum flood depth for the refined spatial output configuration and the default configuration. Panels d, h, l show the difference between the maximum flood depth for the dynamic downscaling configuration and the default configuration.

To analyse the changes of flood depths over time, Fig. 8 panels a, c show the flood depth timeseries at the SFINCS output point locations outlined in Fig. 6, for the default and fully refined configurations as well as the three individual refinements. While the timing and shape of the flood depth timeseries remains consistent across all the model configurations for all the case studies, there are differences in the magnitude of the flood depths. Figure 8 panel a shows that for TC Irma the fully refined configuration leads to higher flood depths than those for the default configuration. It can be observed that the cause for that increase in flood depths is mostly triggered by the dynamic downscaling, which results in flood depths more than 0.5 m higher than those with the default configuration. On the contrary, the temporal and spatial output refinements do not have much influence on the flood depths at this SFINCS output point location. Figure 8 panel b shows that for TC Haiyan the fully refined configuration leads to similar flood depths than those for the default configuration. It can be observed that refining the temporal...
and spatial output resolution results in flood depths approximately 1 m higher than the default configuration. However, the dynamic downscaling produces lower flood depths, compensating the other refinements and leading to similar results for the fully refined and default configurations at the peak of the event. A smaller flood peak occurs around 05:00 on the 8th of November, where the timing of the fully refined configuration occurs earlier, probably as an effect of the temporal output resolution refinement. Finally, Fig. 8 panel c shows that for ETC Xynthia the fully refined configuration leads to lower flood depths than those for the default configuration. The results of the fully refined configuration seem mostly influenced by the dynamic downscaling but with higher flood depths at the flood peaks of the event, which could be the result of the temporal and spatial output refinements.

Panels a, b, c in Fig. 8 only show the results for a single SFINCS output point location. However, the refinement that has the most effect on the fully refined configuration might be different when looking at other SFINCS output point locations. For TC Haiyan, for example, when looking at Fig. 6 and Fig. 7 we observe that the water levels and resulting flood depths for the fully refined configuration are similar to those for the dynamic downscaling near Tacloban, while south of Tacloban the temporal and spatial output refinements seem to be more dominant. For ETC Xynthia, when looking at Fig. 6 and Fig. 7 we observe that the flood depths for the fully refined configuration are similar to those for the dynamic downscaling along the coasts, but further inland the temporal and spatial output refinements seem more dominant. To understand the overall effect of each refinement in the fully refined configuration, Fig. 8 panels d, e, f show the flood volume timeseries across each of the case study’s model domain. While the timing and shape of the flood volume timeseries remains consistent across all the model configurations for all the case studies, there are differences in the magnitude of the flood volumes. Figure 8 panel d shows that for TC Irma the fully refined configuration leads to higher flood volumes than those for the default configuration. It can be observed that the cause for that increase in flood volume is mostly triggered by the dynamic downscaling, which results in flood volumes of more than 2x10^7 m^3 than those with the default configuration. On the contrary, the temporal output refinement does not have much influence on the flood volume, and the spatial output refinement leads to lower flood volumes than the default scenario. Figure 8 panel e shows that for TC Haiyan the fully refined configuration leads to significantly higher flood volumes than those for the default configuration. At the peak of the event, the fully refined configuration leads to approximately 2x10^8 m^3 more than the default. This increase in volume seems to be the result of a combination of all the refinements, with the increase in temporal output resolution being the refinement that leads to most similar results. Finally, Fig. 8 panel f shows that for ETC Xynthia the fully refined configuration presents similar flood volumes as the default configuration. At the first flood peak, the flood volume provides similar results as the default, due to the lower flood volumes caused by the dynamic downscaling that seem to compensate the higher volumes caused by the temporal and spatial output refinements. In the second flood peak, however, the fully refined configuration shows lower flood depths than those for the default configuration. The reason for this could be the fact that compared to the first peak, the dynamic downscaling in this peak presents lower flood volumes with respect to the default configuration.
Figure 8. Flood depth timeseries for three observation points and flood volume timeseries for the SFINCS model domain of each case study. The spatial location of the SFINCS output point locations can be observed in Fig. 6.

4 Discussion and Conclusions

The MOSAIC modelling framework introduced in this study allows to dynamically simulate coastal flooding events. MOSAIC enables the coupling of dynamic water level and overland flood models making use of a Python environment. As such, the approach is automated and reproducible, and combined with the hydrodynamic models used, this makes the approach globally applicable. It allows us to easily simulate coastal flooding events globally, and its multiscale modelling approach also allows us to enhance the simulation at local scale by providing refined water levels. MOSAIC provides a bridge between fully global and fully local modelling approaches, and thereby paves the way for more actionable large-scale flood risk assessments.

The results of the sensitivity analysis conducted in this study reveal the complexity of hydrodynamic modelling and the sensitivity to specific local settings and storm characteristics. When comparing the fully refined and the default configurations, the behaviour of each case study differs in terms of changes in water levels and flood depth, both, spatially and in magnitude. For TC Irma the fully refined configuration results in higher water levels and flood depths, while for TC Haiyan it results in region with higher and regions with lower water levels and flood depths, and for ETC Xynthia water levels become lower overall, and flood depths result lower along the coast but higher further inland. The effects of refining the temporal output resolution on the one hand seem to have large influence for TCs, resulting in water levels and flood depths more than 0.5 m and 1 m higher for TC Haiyan, respectively. On the other hand, for ETCs, the refining of the temporal output resolution does not change the water levels and flood depths significantly, being 1-hourly temporal resolution enough. Refining the spatial output locations of GTSM provides more coastal boundary conditions for SFINCS, resulting always in more accurate flood depth results. However, for regions where the water levels have more spatial heterogeneity along the coast, this refinement becomes most relevant. For TC Haiyan, for example, this refinement leads to flood depths 1 m higher than the default configuration. Furthermore, changes in water levels due to dynamic downscaling are notably affected by the bathymetry used to generate the local high-resolution models. This refinement leads to maximum water levels more than 0.5 m higher for TC Irma and 0.3 m lower for ETC Xynthia, in comparison to the default configuration. Looking at the effects of each refinement, we observe that while for TC Irma the refinement that most affects the model results is the dynamic downscaling, for TC Haiyan and ETC Xynthia the fully refined configuration seems to be a combination of all the refinements. Based on these results, we conclude that refinement of the global modelling approach can significantly impact the simulation of coastal water levels and flood depths at local scale. This highlights the potential merit of a multiscale modelling approach within local refinement as applied in MOSAIC.
There are several limitations that need to be taken into account when using MOSAIC. Limitations that are linked to general flood hazard modelling and not specific to MOSAIC include the following: (1) the meteorological forcing data can be a big source of uncertainty when modelling extreme water levels. MOSAIC allows to combine the results of the Holland model with climate reanalysis datasets in the background to enhance the wind and pressure fields at the peripheries of the TCs. Nonetheless, the implementation of more advanced wind parametric models could further improve the water level simulations (Emanuel & Rotunno, 2012; Hu et al., 2011). (2) the accuracy of the bathymetry has a large influence on storm surge modelling (Bloemendaal et al., 2019; Dullaart et al., 2020; Mori et al., 2014). When performing dynamic downscaling, MOSAIC uses bathymetry data to generate the model grid and subsequently simulate total water levels. Using higher-resolution local bathymetry enables finer grid refinement and enhances the accuracy of the results. However, such high-resolution bathymetry is not always available. MOSAIC is set up to allow the substitution of bathymetric data with alternative datasets, to adjust the grid resolution and refinement, and to define the desired domain of the local high-resolution model. (3) digital elevation models (DEMs) can have a large influence on flood model simulations, affecting the flood hazard depth map results. In this paper we use the FABDEM dataset, but MOSAIC allows to replace the DEM with higher resolution local datasets, when available.

MOSAIC’s main limitation lies in the generation of the local high-resolution models for dynamic downscaling. These automatically generated local high-resolution 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 degrees around the model domain are not being refined in this study. When changes in grid refinement are abrupt, model instabilities can also occur. The nesting of multiple models in each other would allow for a smoother grid transition and might solve this issue.

In this study, we have implemented MOSAIC to simulate TC and ETC coastal flooding as a result of storm surges. Future research on TCs and ETCs may further develop MOSAIC and include other drivers such as waves, rainfall and discharge. Considering that HydroMT and SFINCS are capable of handling compound flooding induced by pluvial and fluvial drivers (Eilander et al., 2023), there is potential for future enhancements of MOSAIC to incorporate the modelling of compound events. Furthermore, MOSAIC currently makes use of offline coupling for both the local-high resolution model and the SFINCS model. However, new software developments such as Oceanographic Multi-purpose Software Environment (OMUSE; Pelupessy et al., 2017) could be used in the future to move from offline to online coupling, and to further expand MOSAIC by allowing for coupling with other models such as hydrological or ocean models. The flexibility of MOSAIC to modify the input datasets could be leveraged 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, MOSAIC could also be used for risk assessments.

**Data availability**
The datasets compiled and/or analysed during the current study are available on Zenodo. *Note: to be published with Doi upon acceptance of the paper.*

**Code availability**
The underlying code for this study is available on Zenodo. *Note: to be published with Doi upon acceptance of the paper.*

**References**


### Acknowledgements

This work was carried out in the EU-ERC COASTMOVE project nr 884442 and the NWO MOSAIC project nr ASDI.2018.036. The authors would like to thank the SURF Cooperative for the support in using the Dutch national e-infrastructure under grant no. EINF-2224 and EINF-5779.

### Author contributions


### Competing interests

One of the (co-)authors is a member of the editorial board of Natural Hazards and Earth System Sciences.