



1 Accounting for Uncertainties in Forecasting Tropical Cyclone- 2 Induced Compound Flooding

3 Kees Nederhoff^{1,2,5}, Maarten van Ormondt¹, Jay Veeramony³, Ap van Dongeren^{2,4}, José A. Á. Antolínez⁵,
4 Tim Leijnse^{4,6}, Dano Roelvink^{2,4}

5 ¹ Deltares USA, 8601 Georgia Ave, Silver Spring, MD 20910, USA

6 ² CURR, UNESCO-IHE Institute for Water Education, P.O. BOX 3015, 2601 DA Delft, The Netherlands

7 ³ Naval Research Lab, Stennis Space Center, MS 39529, USA

8 ⁴ Marine and Coastal Management, Deltares, Boussinesqweg 1, Delft, 2629 HV, The Netherlands

9 ⁵ Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

10 ⁶ Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, The
11 Netherlands.

12 *Correspondence to:* Kees Nederhoff (kees.nederhoff@deltares-usa.us)

13 **Abstract.** Tropical cyclone impacts can have devastating effects on the population, infrastructure, and on natural habitats.
14 However, predicting these impacts is difficult due to the inherent uncertainties in the storm track and intensity. In addition,
15 due to computational constraints, both the relevant ocean physics and the uncertainties in meteorological forcing are only
16 partly accounted for. This paper presents a new method, called the Tropical Cyclone Forecasting Framework (TC-FF), to
17 probabilistically forecast compound flooding induced by tropical cyclones, considering uncertainties in track, forward speed,
18 and wind speed/intensity. The open-source method accounts for all major relevant physical drivers, including tide, surge, and
19 rainfall, and considers TC uncertainties through Gaussian error distributions and autoregressive techniques. The tool creates
20 temporally and spatially varying wind fields to force a computationally efficient compound flood model, allowing for the
21 computation of probabilistic wind and flood hazard maps for any oceanic basin in the world, as it does not require detailed
22 information on the distribution of historical errors. A comparison of TC-FF and JTWC operational ensembles, both based on
23 DeMaria et al. (2009), revealed minor differences of <10%, suggesting that TC-FF can be employed as an alternative, for
24 example, in data-scarce environments. The method was applied to Cyclone Idai in Mozambique. The underlying physical
25 model showed reliable skill in terms of tidal propagation, reproducing the storm surge generation during landfall and flooding
26 near the city of Beira (success index of 0.59). The method was successfully applied to forecast the impact of Idai with different
27 lead times. The case study analyzed needed at least 200 ensemble members to get reliable water levels and flood results three
28 days before landfall (<1% flood probability error and <20 cm sampling errors). Results showed the sensitivity of forecasting,
29 especially with increasing lead times, highlighting the importance of accounting for cyclone variability in decision-making
30 and risk management.

31



32 1 Introduction

33 Tropical cyclone (TC) induced compound flooding, which occurs when storm surge, heavy rainfall, high tide, and river
34 discharge coincide, can have devastating impacts on coastal communities (Wahl et al., 2015). This type of flooding is
35 particularly concerning as it can result in higher water levels and increased inland flooding, leading to damage and loss of life
36 (e.g., Resio & Irish, 2015). The increased frequency and severity of compound flooding events are expected to worsen due to
37 the effects of climate change (e.g., Lin et al. 2012; Mori and Shimura, 2023) as well as on-going coastal development and
38 population growth. Mitigation and preparedness strategies require a sound toolbox for assessing TC-induced compound
39 flooding on coastal communities that enhance short to long-term decision-making.

40

41 Operational and strategic risk analyses are instrumental in analyzing and mitigating potential environmental risks. Operational
42 risk analysis, typically associated with short-term forecasting (~several days), provides immediate response and preparedness
43 for imminent disasters, ensuring the safety and protection of people and property (Roy & Kovordányi, 2012). Conversely,
44 strategic risk analysis focuses on long-term climate variability assessments, delivering insights into hazards and their socio-
45 economic and environmental impacts, thus facilitating informed policy decisions and adaptation strategies (e.g., Nederhoff et
46 al., 2021). Though distinctly different, both perspectives are critical for comprehensive climate risk management, as they offer
47 different scales and timeframes for prevention, preparedness, response, and recovery.

48

49 Forecasting agencies such as the National Hurricane Center (NHC) have significantly improved operational meteorological
50 risk analysis, credited to gains made in numerical weather prediction models (McAdie et al., 2000, Cangialosi et al., 2020).
51 Despite advancements, operational forecast errors remain significant enough to necessitate considering the inherent
52 uncertainties in these forecasts for informed preparedness decision-making (Lamers et al., 2023). A common probabilistic
53 approach is to represent the resulting uncertainty in track prediction by a cone envelope as a graphical representation that
54 illustrates the possible track variation of the TC center (NHC, 2023). The shape of the cone can be derived from the historical
55 error data of the forecast and typically represents a 66.7% probability that the track will be within the cone (i.e., 33.3% chance
56 the track falls outside the cone). The cone increases in size with lead time as the errors in the prediction accumulate. While the
57 cone gives valuable insight into the potential range of TC variability of the core, it can be easily misinterpreted as the
58 corresponding impacted area, which can be substantially larger. Quantification of the uncertainty in track prediction can be
59 computed with several methods. For example, De Maria et al. (2009) introduced a Monte Carlo method to generate 1,000
60 realizations by randomly sampling from historical error distribution functions from the past 5 yr for both the track and intensity.
61 De Maria et al. (2013) improved their method so that the track uncertainty is estimated on a case-by-case basis using the Goerss
62 predicted consensus error (GPCE; Goerss, 2007), where the uncertainty is estimated based on the spread of a dynamical model
63 ensemble instead of historical averages. Other methods exist. For example, Chen et al. (2023) introduced a deep-learning
64 ensemble approach for predicting tropical cyclone rapid intensification. However, these methods were all derived to provide



65 insight, before landfall, into the uncertainty of the wind speeds and not designed to force hydrodynamic or wave models, and
66 can thus result in too erratic forcing conditions.

67

68 Early Warning Systems (EWS) for coastal compound flooding are sensitive to uncertainties in the TC, including nonlinear
69 interactions between the TC size, forward speed, location of landfall, tides, rainfall, and infiltration. However, often EWS for
70 coastal flooding use physics-based and, due to computational constraints, deterministic approaches in which the best track is
71 used to force a hydrological & hydrodynamic model that computes the storm surge and the complex interactions between
72 coastal, fluvial, and pluvial processes (e.g., Global Storm Surge Information System; GLOSSIS based on Delft3D Flexible
73 Mesh; Kernkamp et al., 2011 and Coastal Emergency Risks Assessment; CERA based on ADCIRC; Luetlich et al., 1992).

74

75 Several examples of probabilistic coastal flood methods capture uncertainty in forcing. For example, the Global Flood
76 Awareness System (GloFAS; Alfieri et al., 2013) is a modeling chain run by the European Centre for Medium-Range Weather
77 Forecasts (ECMWF) based on the LISFLOOD hydrological model forced by 51 ensemble members. While GloFAS is an
78 excellent resource for communities worldwide, it operates at a large scale with a relatively coarse resolution of 0.1 degrees
79 (~10 km), and is thus not designed explicitly for TCs that require high spatial resolutions (Roberts et al., 2020), and neither
80 account for relevant coastal processes such as tides. Higher resolution and the inclusion of coastal processes can be found in
81 several regional applications. For example, the Stevens Flood Advisory System (SFAS; Ayyad et al., 2022) is an ensemble-
82 based probabilistic forecasting of tide, surge, and riverine flow across the US Mid-Atlantic and Northeast coastline and runs
83 for 96 different atmospheric forcing datasets. Other examples include forecasting systems from the UK Met Office (Flowerdew
84 et al., 2010) and the Royal Netherlands Meteorological Institute (de Vries, 2009). All these systems rely on coarser numerical
85 forecasting products, focus on mid-latitude regions, and are thus not explicitly designed to forecasts hazards related to TCs.

86

87 Probabilistic modeling systems for TC-induced coastal flooding for operational risk analyses in the US and Japan include P-
88 Surge (Taylor and Glahn, 2008; Gonzalez and Taylor, 2021), which uses data from the NHC to create a set of synthetic storms
89 by perturbing the storm's position, size, and intensity based on past errors of the advisories. Subsequently, the Sea, Lake,
90 Overland, Surge from Hurricanes model (SLOSH; Jelesnianski 1992) is run and forecasts storm surge in real-time when a
91 hurricane is threatening. However, SLOSH does not account for several relevant (coastal) processes (e.g., tides, waves, rainfall,
92 infiltration) and thus lack their interactions. The Japan Meteorological Agency (JMA) does use a dynamic tide and storm surge
93 model (Higaki et al., 2009) but only accounts for a limited number of 11 ensemble members (Hasegawa et al., 2015). Moreover,
94 both methods are created with a specific region in mind and are not easily transferable to other locations.

95

96 Besides probabilistic physics-based techniques, statistical machine-learning techniques (e.g., Lecacheux et al., 2021 or Nguyen
97 & Chen, 2020) are becoming increasingly popular in reducing the computational expense of forecasting compound flooding.



98 However, these machine learning downscaling methods lack nonlinear interactions between relevant coastal processes driving
99 compound flooding. Hybrid methods focus on reducing the number of tracks simulated and proved capable of accurately
100 representing a larger set of scenarios (Bakker et al., 2022).

101

102 As introduced by Suh et al. (2015), the constraints in real-time forecasting for operational risk analysis are around both
103 'accuracy' and 'promptness'. In other words, the time constraints associated with forecasting dictate that some modeling systems
104 use a purely deterministic approach or a limited number of ensemble members to perform more detailed compound flooding
105 predictions and thus simplify the meteorological uncertainty (e.g., GLOSSIS, CERA, JMA). On the other hand, probabilistic
106 approaches for meteorology with a large number of ensemble members use simplified hydrodynamics or have an insufficient
107 resolution for TCs and thus do not account for the processes needed to forecast TC-induced coastal compound flooding (e.g.,
108 GloFAS, SFAS, NHC).

109

110 To address the limitations listed, we propose a method to generate probabilistic wind and compound flood hazard maps by
111 using, for the first time, ensembling techniques via statistical emulation of TCs combined with physics-driven modeling for
112 coastal compound flooding. The workflow emulates the TC evolution using an autoregressive technique in combination with
113 reported mean errors in track and intensity, similar to DeMaria et al. (2009) but without the need for historical error distribution
114 functions. Next, this emulator produces an ensemble of several (herein thousands) TC members. Then, for each ensemble
115 member, a time- and spatially-varying wind field is generated and used to force a computationally efficient compound flood
116 model SFINCS (Leijnse et al. 2021). The output consists of probabilistic wind and flood hazard maps that can be forecast on
117 time with limited computational resources anywhere in the world. This paper refers to the TC forecasting framework as the
118 Tropical Cyclone Forecasting Framework, TC-FF.

119

120 The paper is structured as follows. Section 2 introduces the Monte Carlo forecasting methodology. Section 3 describes the
121 case study site and historical event of interest. The materials and methods used in this paper are described in Section 4.
122 Validation in terms of tides and storms and application of the forecasting methodology are presented in Section 5. Finally,
123 Sections 6 and 7 discuss and summarize the main conclusions of the study.

124 **2 Tropical Cyclone Forecasting Framework**

125 In this paper, we introduce the probabilistic Tropical Cyclone Forecasting Framework, TC-FF, to compute TC-induced
126 compound flooding for operational risk analysis. Our approach integrates a TC emulator using a Monte Carlo-based ensemble
127 sampling generation with an autoregressive technique, which is a simplified adaptation of DeMaria et al. (2009). The ensemble
128 members are generated around the forecasted official track, considering the average historical errors in intensity, cross-track,
129 and along-track. The ensemble members are provided as input for the fast compound flood model called SFINCS.



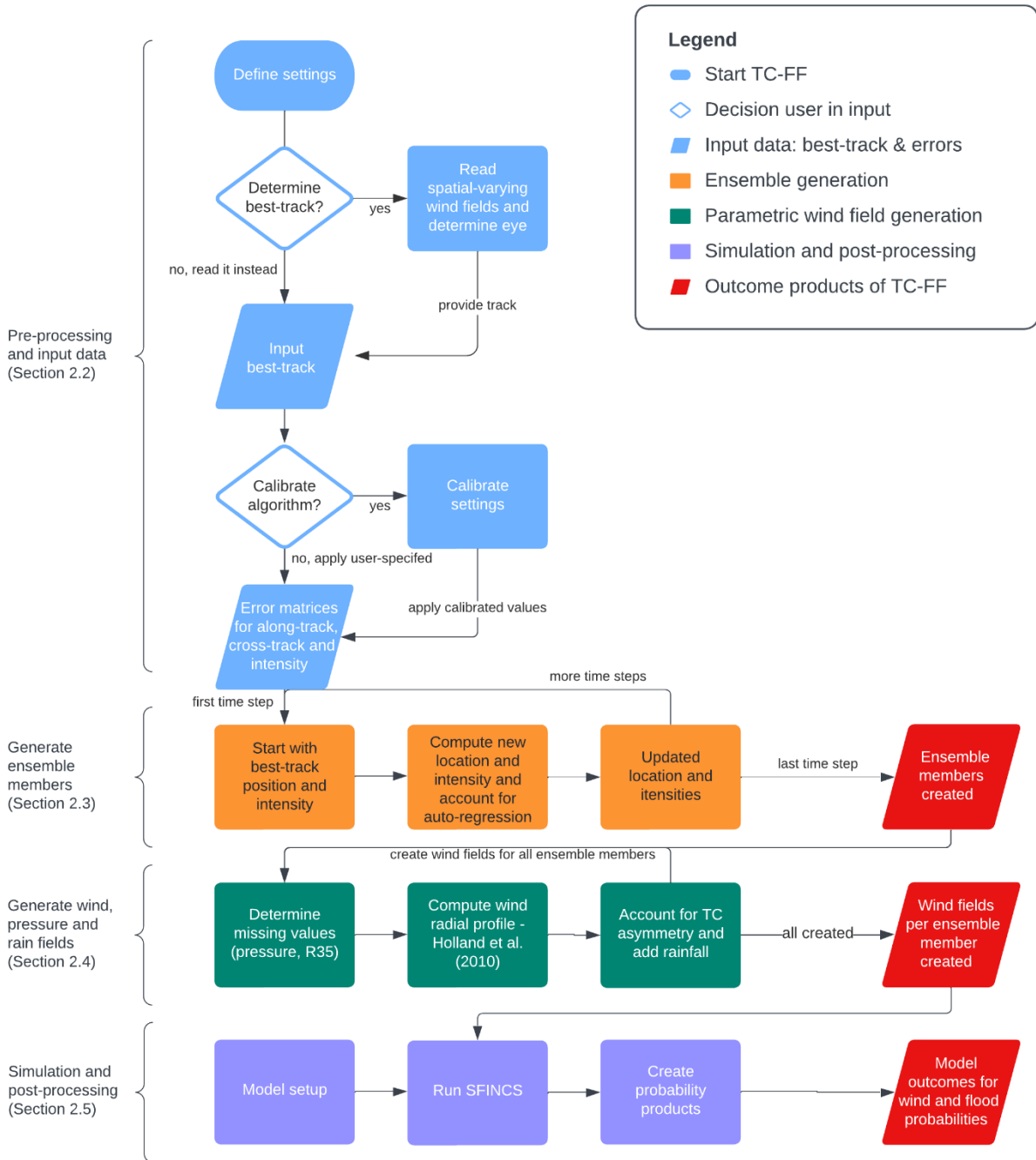
130 Additionally, TC-FF considers tidal movements, storm surge, precipitation, and infiltration. The outcomes are consolidated
131 into a unified probability product. By choice, each member has an equal likelihood of occurrence. The Python code for this
132 method is accessible on GitHub via the following link: <https://github.com/Deltares/CoastalHazardsToolkit>.

133 2.1 TC-FF flowchart

134 A compact flowchart of TC-FF used to generate the ensemble member is shown in Figure 1. The steps of this process are as
135 follows:

- 136 1. **Define settings:** The user specifies the data source, period, time step of the ensemble generation, and the number of
137 ensemble members requested.
- 138 2. **Input best track:** The code either determines the best track based on gridded time and spatial-varying wind and
139 pressure fields (e.g., COAMPS-TC; Doyle et al., 2014) or reads in the forecasted track by one of the forecasting
140 centers (e.g., NHC or other agencies).
- 141 3. **Error matrices for along-track, cross-track, and intensity:** The tool first computes random realizations based on
142 the along-track, cross-track, and intensity standard deviations imposed for the time steps requested. The imposed
143 mean absolute error is scaled with the timestep to overcome any time step dependency.
- 144 4. **Generate ensemble members:** Following the approach of DeMaria et al. (2009), a Monte Carlo method generates
145 numerous ensemble members based on error matrices of the previous step in combination with an autoregressive
146 technique for the along-track, cross-track, and intensity error.
- 147 5. **Generate wind, pressure, and rain fields:** Generate meteorological forcing conditions, i.e., the surface wind and
148 pressure fields per time step per ensemble member, based on parametric methods (e.g., Holland et al., 2010) for
149 subsequent analysis and application within numerical models. Rainfall can be included as well via intensity
150 relationships.
- 151 6. **Simulation and post-processing:** In this study, the compound flood model SFINCS is applied, but in principle, other
152 hydrodynamic models can also be applied, albeit typically at a higher computational expense. Data from the different
153 ensembles are combined into several probabilistic outputs ranging from the probability of gale force winds (wind
154 speed > 35 knots or >18 m/s), compound flooding (water depth > 15 cm) to quantile estimates (e.g., 1% exceedance
155 water level).

156



157

158 **Figure 1. Flowchart of the Tropical Cyclone Forecasting Framework (TC-FF). Pre-processing stages are represented in light blue,**
 159 **the computational core of ensemble generation is denoted in orange, the parametric wind field generation is portrayed in green, the**
 160 **hydrodynamic simulation and analysis of winds are marked in purple and outcomes in red.**



161 In the subsequent paragraphs, we describe in more detail the pre-processing, the computation of the ensemble members (track
162 and intensity variations), and the determination of time- and spatially-varying wind fields.

163 **2.2 Pre-processing and input data**

164 The pre-processing of TC-FF comprises three components.

165

166 First, one specifies the period they would like to simulate, including the total time period over which wind fields need to be
167 generated and the time period over which the ensembles need to be generated. In addition, a timestep for ensemble generation
168 (default 3 hours) needs to be specified. At this stage, one also specifies the mean absolute error and auto-regression coefficients
169 for the along-track, cross-track, and intensity. When these values are unknown, calibration needs to be performed to determine
170 them by comparing them with the reported errors of the forecast center (see calibration in Section 5.2.1). At this stage, one
171 also specifies the number of ensemble members requested. The influence of the number of ensemble members is discussed in
172 Section 5.3.2.

173

174 Second, since TC-FF creates random realizations around the best track, an input track is needed. Depending on the application,
175 TC-FF reads a forecast bulletin that generates the track or determine the best track from the output of a high-resolution regional
176 meteorological model. The determination of a track from a meteorological model is based on an algorithm that finds the
177 minimum pressure in an area of interest. It takes in grid values, u and v wind components, pressure, minimum distance for
178 clustering, and returns lists of x and y coordinates of cyclone eyes, the maximum wind speed plus pressure around each eye.

179

180 Third, before the generation of the ensemble members, TC-FF creates random errors with a normal distribution based on the
181 provided average errors. Matrices are two-dimensional, with one dimension being the number of time stamps and the other the
182 number of ensemble members. The imposed mean absolute error is scaled with the timestep to overcome any time step
183 dependency and converted into a standard deviation.

184 **2.3 Ensemble members**

185 **2.3.1 Track realizations and calibration**

186 An important component in TC-FF is the generation of track realizations (or ensemble members) from the official track
187 forecast. The official positions are interpolated with a spline function to include values at all requested times. Our approach
188 for the track realization largely follows DeMaria et al. (2009). We decompose the track error into the along-track (AT) and
189 cross-track (CT) components and account for the track error serial correlation via autoregressive regression (Equations 1 and
190 2).

$$AT_t = \alpha_t AT_{t-i} + B_{rnd} \quad \text{Equation 1}$$



$$CT_t = c_t CT_{t-i} + D_{rnd} \quad \text{Equation 2}$$

191 in which AT_t and CT_t are the AT and CT error at time steps t , a_t and c_t are constants, AT_{t-3} and CT_{t-3} are errors of the previous
192 time step (typically $i=3$ hours), and B and D are random numbers that are normal (Gaussian) distributed, scaled with the mean
193 absolute error but are limited to $\pm 2\sigma$.

194
195 Unlike DeMaria et al. (2009), we do not access the probability distributions of historical errors. Instead, we calibrate the
196 parameters (a_t , c_t , and mean absolute errors for B and D) based on the reported historical errors from the agency responsible
197 for the issued forecast (see Section 5.2.1). This is a simpler methodology and requires substantially less data (which is also
198 typically not accessible outside the forecast centers). These historical errors are routinely reported by the forecast centers (e.g.,
199 see Section 4.1.2 for information on the data sources used in this paper). Note that errors in our implementation (neither the
200 error nor the auto-regressive coefficient) vary with lead time. We calibrate a constant mean absolute error in combination with
201 a single auto-regression coefficient (see Section 5.2.1 for calibration and Section 5.2.2 for the influence of simplifications).
202 Moreover, the mean absolute error is converted into a standard deviation using a fixed relationship assuming a normal
203 distribution of the error and scaled with the applied time step to allow the user flexibility in the applied time step.

204
205 The determination of the ensemble members is subsequently based on the sum of the forecast and random error components.
206 In other words, we add the along-track and cross-track error to forecasted along- and cross-track. An example of the first 20
207 ensemble members is presented in Figure 2B. Using this procedure, 10,000 ensembles are generated for each forecast case
208 within this study; however, it is possible to use fewer ensemble members to reduce the computational cost but at larger
209 statistical uncertainty (see Section 5.3.2 for trade-offs).

210 2.3.2 Intensity realizations and calibration

211 Similar to the track realization, the maximum wind speed (intensity) at a specific interval is determined using a random
212 sampling approach. The starting point is the official forecast of intensity that is interpolated to include values at all requested
213 times, and a random error component (VE_t) is added.

$$VE_t = e_t VT_{t-3} + F_{rnd} \quad \text{Equation 3}$$

214 in which VE_t at time steps t , e_t is a constant, VE_{t-3} are errors of the previous time step (typically 3 hours) and F is random
215 numbers that are normally distributed, scaled with the mean absolute error and is limited to $\pm 2\sigma$.

216
217 The inland wind decay model adjusts the maximum intensity as a function of the distance inland, is directly based on DeMaria
218 et al. (2009) and is computed with Equation 4. If the intensity of any inland ensemble member exceeds this predetermined
219 value at any forecast time, the intensity is adjusted to match this value. Subsequently, the intensity errors are recalculated based
220 on the adjusted intensity. Additionally, if the intensity of an inland ensemble member falls <7.7 m/s (15 knots) at any point in



221 time, the TC intensity is reset to zero for all subsequent periods to overcome any unrealistic reintensifying TCs. All these
222 criteria follow DeMaria et al. (2009).

$$V_i = 20 + 120e^{0.0035D} \quad \text{Equation 4}$$

223 in which the maximum wind speed (V_i) in knots and the distance to land (D) in kilometers (with negative values indicating
224 inland cyclones) are given, the intensity of an inland cyclone can be determined.

225

226 The intensity implementation differs from DeMaria et al. (2009) in the following ways. We remove the dependency that the
227 error scales with wind intensity and bias correction. Again, the determination of the ensemble members is based on the sum of
228 the forecasted and random components computed with Gaussian mean absolute errors and an auto-regressive constant over
229 lead time. Similar to the track realization, intensity errors are scaled with the time step to overcome any time step dependency.
230 The influence of the simplifications and the difference compared to NOAA operational code based on the original DeMaria et
231 al. (2009) and DeMaria et al. (2013) implementation are discussed in Section 5.2.2.

232 **2.4 Parametric wind fields**

233 After the determination of the ensemble members, the time and spatial varying wind fields are constructed and written in a
234 polar coordinate system. Several (horizontal) parametric wind profiles have been presented in the literature (e.g., Fujita, 1952;
235 Chavas et al., 2015), with the original Holland wind profile (Holland, 1980) being the most widely used due to its relative
236 simplicity. Several codes have been developed for storm surge models to provide time and spatial wind and pressure fields
237 (e.g., Hu et al., 2012 for ADCIRC). Deltares has developed the Wind Enhance Scheme (WES; Deltares, 2018) to generate TC
238 wind and pressure field around the specified location of a tropical cyclone center and given a number of TC parameters. In its
239 current implementation, information on wind radii (radius of gale force winds) can be considered in the Holland et al., (2010)
240 formulation using information either from best track-data or from the proposed relationships of Nederhoff et al. (2019), which
241 increases the accuracy of the method. Furthermore, the asymmetry of the wind field in a TC is also implemented, as delineated
242 by Schwerdt et al. (1979). Additionally, tropical cyclone-induced precipitation can be incorporated using empirical
243 relationships such as IPET (2006).

244 **2.5 SFINCS simulation and post-processing**

245 After the determination of the wind fields for all the requested ensemble members, TC-FF runs a hydrodynamic model. In this
246 study, we apply the compound flood model SFINCS (Leijnse et al., 2021), which lends itself well to a large number of
247 simulations in a reasonable amount of time due to its reduced complexity. SFINCS reads the tidal boundary conditions and
248 wind, pressure, and rainfall conditions from the wind fields. Once all the ensemble member simulations have finished,
249 probability products regarding wind and flood hazards are created. These products are created by sorting the results for each



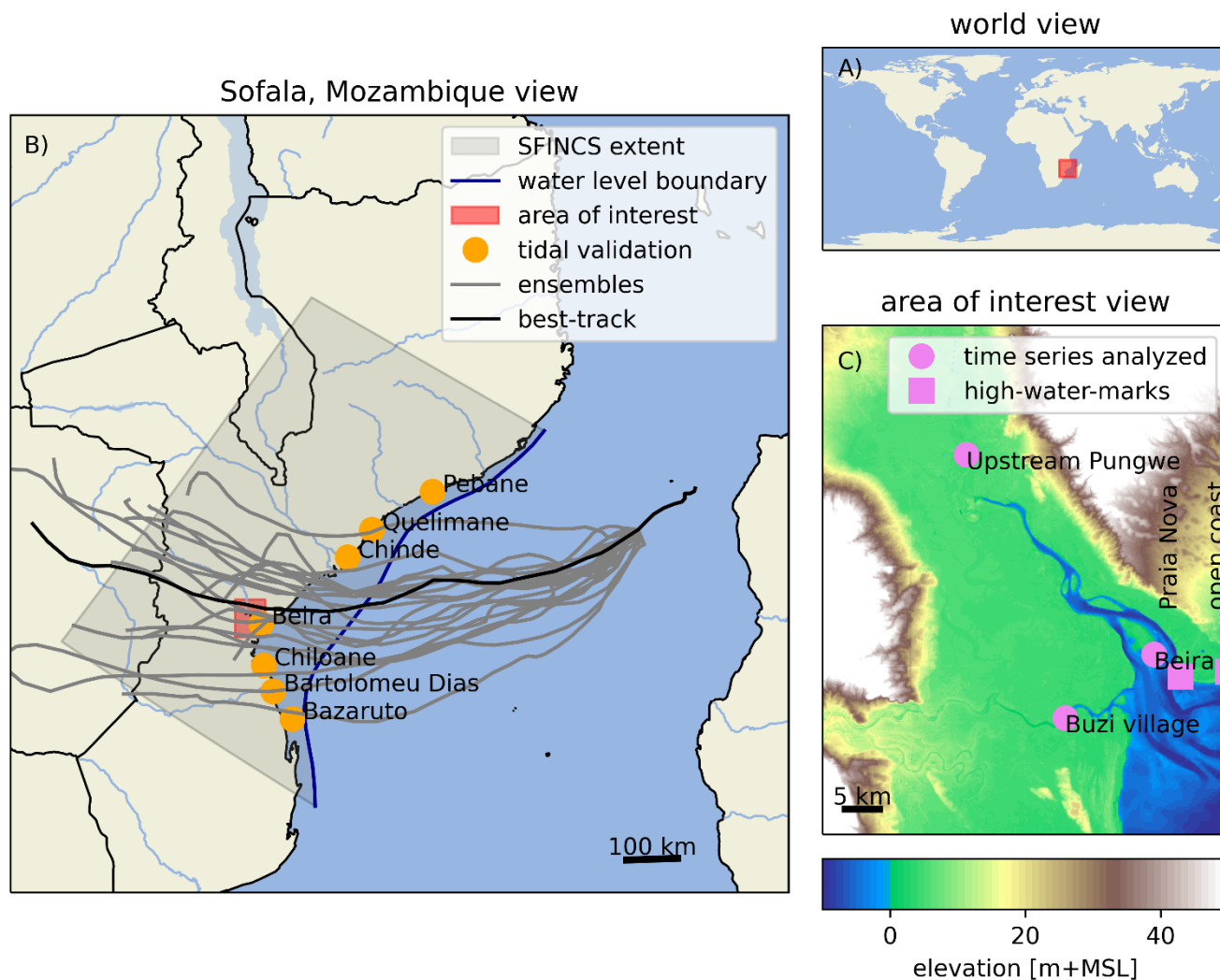
250 grid cell and providing estimates for either specific intervals (e.g. wind speeds > 35 knots or water depth > 15 cm) or quantile
251 estimates (e.g., 1% exceedance water level).

252 3 Case study

253 The TC forecasting framework is applied to a historical event that took place in Mozambique's Sofala province: Cyclone Idai,
254 in March 2019. Mozambique is a country located in southeastern Africa (Figure 2A). The country has a diverse population of
255 over 31 million people, of which 2 million live in the Sofala province in central Mozambique (National Institute of Statistics
256 of Mozambique, 2017). Sofala is primarily rural, with small communities along the Pungwe and Buzi river deltas (Emerton et
257 al., 2020). Beira is the province's largest city, home to over 500,000 people, and an important port linking the hinterland to the
258 Indian Ocean. The city is prone to flooding, particularly during the rainy season, which generally extends from October to
259 April or May. This period coincides with the cyclone season, as cyclones often bring intense rainfall to the region. The
260 vulnerability of Beira to flooding is exacerbated by factors such as climate change, rapid urbanization, and limited
261 infrastructure.

262

263 Cyclone Idai was an example of a compound flood event that affected large parts of the coastal delta of Sofala (Eilander et al.,
264 2022). The storm began as a tropical depression in the Mozambique Channel, causing extensive flooding after its first landfall
265 in early March. It later intensified as it moved back over the sea, developing into a tropical cyclone with 10-minute sustained
266 wind speeds of 165 km/h. Idai made landfall near the port city of Beira, bringing powerful winds, resulting storm and heavy
267 rains that caused widespread flooding and destruction. Large areas were flooded, first around the coast and a few days later,
268 more inland in the Buzi and Pungwe floodplains. The total rainfall across the five days from March 13-18 ranged from 250–
269 660 mm (NASA GPM, 2019). Over 112,000 houses were destroyed, and an estimated 1.85 million people were affected (UN
270 OCHA, 2019).



271

272 **Figure 2. View of the study site:** (A) Mozambique's Sofala Province is situated in the southeastern region of Africa in the Southern
273 **Hemisphere.** (B) Geographical and hydrodynamic representation of the study area. The SFINCS model extent, highlighted in Panel
274 **B,** encompasses a portion of the Sofala region, forced offshore with a water level boundary, and is validated at seven tidal stations
275 (indicated by orange circles; see Section 8.1). The best track is represented by a solid dark line, with the first 20 ensemble members
276 5 days before landfall demonstrated as gray lines. (C) The area of interest is the Pungwe estuary, situated near the city of Beira.
277 Model validation also takes place at two high-water-marks close to the city (signified by a purple box), with model outcomes depicted
278 at three diverse locations across the estuary (marked by circles).

279



280 **4 Material and methods**

281 **4.1 Materials**

282 **4.1.1 Elevation datasets**

283 Several topographic and bathymetric datasets were collected and combined to develop a merged DEM. Data includes field
284 survey data points collected during three campaigns in November-December 2020 across Beira, locally-collected LiDAR,
285 bathymetric charts, MERIT (Yamazaki et al., 2017) and GEBCO19 (IOC, IHO and BODC, 2003). Careful consideration was
286 given to prioritize specific datasets in space to ensure the most detailed, recent, and accurate datasets were used in a given
287 area. For example, survey and LiDAR data is bare earth and prioritized over the usage of MERIT and GEBCO19. The merged
288 DEM was produced on medium-resolution (50 m) regional DEM, and a fine-resolution (5 m) local DEM in Beira. For more
289 information on merging the data, one is referred to Deltares (2021).

290 **4.1.2 Forcing conditions**

291 Tidal boundary conditions were based on harmonic constituents provided by TPXO 8.0 (Egbert and Erofeeva, 2002), and tidal
292 amplitudes and phases for all available 13 components were applied. The best track data (BTD) by the Joint Typhoon Warning
293 Center (JTWC) is used throughout this study for meteorological forcing conditions (JTWC, 2022). Reported error statistics by
294 the JTWC for the 5-year average from 2016-2020 were used to inform the ensemble generation (JTWC, 2021). Ensemble
295 members from TC-FF were compared to 1,000 members produced with the code from NOAA, NHC, and JTWC based
296 DeMaria et al. (2009) and DeMaria et al. (2013) that is used operationally (Buck Sampson, personal communication; June 5,
297 2023).

298 **4.1.3 Validation data**

299 Observed tidal coefficients near the city of Beira were used for the calibration and validation of the model (van Ormondt,
300 2020; see Figure 2 for locations). The validation of the event Cyclone Idai (2019) consisted of comparing both, observed and
301 modeled flood extent in deltas of the Pungwe and Buzi rivers and high-water marks in the city of Beira. The observed flood
302 extent was derived from Sentinel-1 synthetic aperture radar data (Eilander et al., 2022); and two observed high water marks
303 (Deltares, 2021) were used, one at Praia Nova, in the western side of the city, and another one at the open coast beach in the
304 southeast (see Figure 2 for locations). Correspondingly, values of modeled flood extent and high water marks were output at
305 the same locations.



306 **4.2 Methods**

307 **4.2.1 Area schematization**

308 For this study, we employed the Super-Fast INundation of CoastS (SFINCS) model, which solves the simplified equations of
309 mass and momentum for overland flow in two dimensions (Leijnse et al. 2021). The goal was to create one continuous
310 compound flood area model that computes tidal propagation, storm surge, pluvial and fluvial flooding.

311

312 The area schematization builds upon Eilander et al. (2022) but varied in three ways. First, we extended the model alongshore
313 and in deeper water to alleviate the need to nest in a large-scale regional coastal circulation model and generate tidal
314 propagation and storm surge within the domain. Using a quadtree implementation (e.g., Liang et al. ,2008), we applied a
315 variable model resolution ranging from 8000 to 500 meters. A quadtree is a technique in which the refinement from one level
316 to another is based on the original cell but divided into 4 smaller cells with 2 times smaller grid size and allows extending the
317 model setup into deeper water without having time step restrictions in deeper water based on the explicit numerical scheme of
318 SFINCS. Second, high-resolution topo-bathymetry and land roughness were included in the native resolution utilizing sub-
319 grid lookup tables (Leijnse et al., 2020). However, the hydrodynamic computations were performed on a coarser resolution to
320 save computational time. Up to 10-meter DEM information was included in the 500-meter grid cells (i.e., factor 50 refinement).
321 Lastly, sub-grid bathymetry features were included to account for maximum dune height based on the DEM to control overflow
322 during storm conditions around Beira.

323

324 A spatially-varying roughness and infiltration was used based on land elevation. All points above mean sea level (MSL) have
325 a high Manning friction coefficient of $0.06 \text{ s/m}^{1/3}$, and an infiltration rate of 1.9 mm/hr (typical values from HSGs Group C;
326 United States Department of Agriculture, 2009), and all other points have lower friction of $0.02 \text{ s/m}^{1/3}$ to represent water and
327 do not have any infiltration. The SFINCS model was forced with tidal boundary conditions and time- and spatially-varying
328 winds, pressure, and rainfall fields. We refer to Section 8.1 for calibration of the tides, in which we show that the area model
329 reproduces tides with a median MAE of 21 cm. Wind and pressure fields were created with the Holland wind profile (Holland
330 et al. 2010) based on the BTD (see Section 2.4 for details). Rainfall for TCs was based on the Interagency Performance
331 Evaluation Task Force Rainfall Analysis (IPET, 2006) method. Comparison with the reported rainfall total revealed a
332 significant underestimation of cumulative rainfall during Idai based on IPET. Based on the magnitude of the underestimation,
333 rainfall estimates by IPET were tripled, resulting in a cumulative rainfall in the area of interest of 495 mm for the best track,
334 which is in a similar order of magnitude as observed (see Section 3).

335 **4.2.2 Simulations periods**

336 The validation of the area schematization focused on two time periods. First, 3 spring-neap cycles (January 13, 2022, until
337 February 26, 2022) were used for the tidal calibration and validation in the area of interest (see Appendix 8.1). Second, Idai



338 was hindcasted forced with the JTWC BTD and compared to observational data for flood extent and high-water levels (Section
339 5.1). After validation of the area schematization, the new forecasting methodology introduced in Section 2 was applied. Various
340 lead times ranging from 1 to 5 days before the second landfall for 10,000 ensemble members were computed (Section 5.3).

341

342 Model runs were performed on the Deltares Netherlands Linux-based High-Performance Computing platform using 10 Intel
343 Xeon CPU E3-1276 v3. On average, a 7-day Idai simulation took about 4 minutes on a single core. Running all 50,000 events
344 took ~15 days.

345 4.2.3 xModel skill

346 Several accuracy metrics were calculated throughout this study: model bias, mean-absolute-error (MAE; Equation 5), root-
347 mean-square-error (RMSE; Equation 6), unbiased RMSE (uRMSE; RMSE with bias removed from the predicted value). These
348 error metrics are used for comparison in water levels, wind speed and track errors.

$$MAE = \frac{1}{N} \sum (|y_i - x_i|) \quad \text{Equation 5}$$

$$RMSE = \sqrt{\frac{1}{N} \sum (y_i - x_i)^2} \quad \text{Equation 6}$$

349 where N is the number of data points, y_i is the i -th prediction (modeled) value, x_i is the i -th measurement.

350

351 Moreover, skill is quantified by binary flood metrics (Wing et al. 2017). The model output (M) is converted to one of two
352 states: wet (1) or dry (0), using a commonly used threshold of 15 cm (e.g., Wing et al. 2017) and compared to the Sentinel
353 benchmark data (B). The Critical Success Index (C ; Equation 7) accounts for both overprediction and underprediction and can
354 range from 0 (no match between modeled and benchmark data) to 1 (perfect match between modeled and benchmark data).

$$C = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1 + M_1 B_0} \quad \text{Equation 7}$$

355

356 For the comparison of cumulative distribution functions (CDF) of cross-track, along-track and intensity, we also applied the
357 Continuous Ranked Probability Score (CRPS; Matheson & Winkler, 1976). CRPS measures how good forecasts are in
358 matching observed outcomes; where $CRPS = 0$, the forecast is wholly accurate, and $CRPS = 1$, the forecast is wholly
359 inaccurate.



$$\text{CRPS}(\mathbf{F}, \mathbf{x}) = \int_{-\infty}^{\infty} [\mathbf{F}(\mathbf{y}) - \mathbf{F}_0(\mathbf{y})]^2 d\mathbf{y} \quad \text{Equation 8}$$

360 where $F(y)$ is the CDF is associated with an empirical probabilistic reference and prediction.

361 **4.2.4 Analysis method**

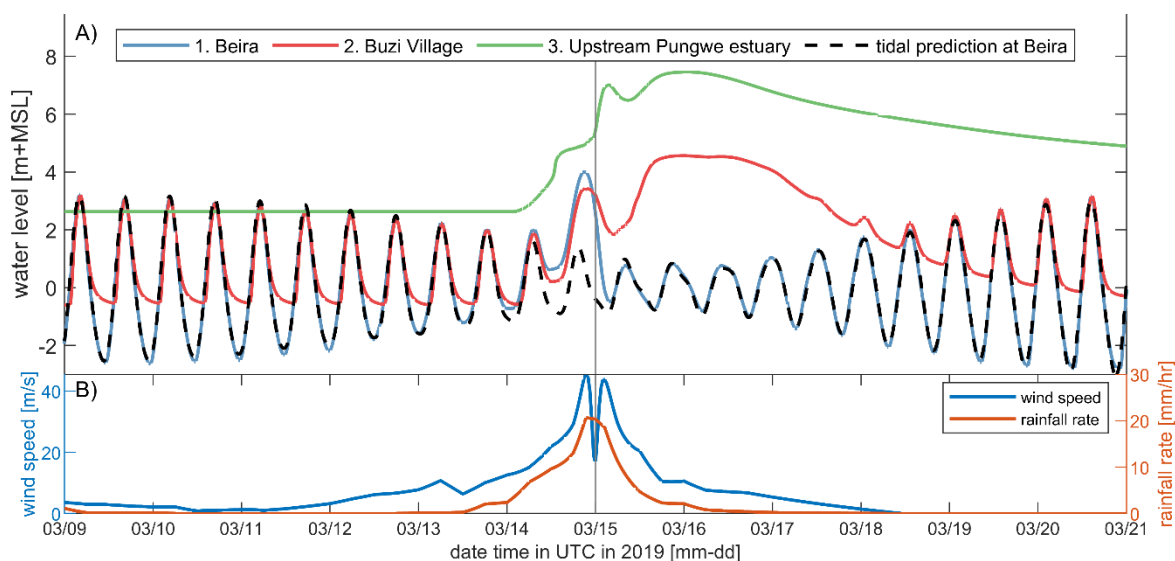
362 The analysis of forecasting results was undertaken using several methods. Initially, extreme wind speeds and water levels were
363 assessed by charting them as time-series data, inclusive of quantile estimates such as the 95% confidence interval (CI).
364 Following this, the maximum values registered during the simulation were organized into cumulative distribution functions
365 (CDFs). This process offered insights into their exceedance probability. Finally, the mean probability of flooding was
366 computed. The method to derive this value entailed counting the instances where computational cells registered a minimum of
367 15 cm of water. Only cells positioned above mean sea level (MSL) were incorporated into the area estimates.



368 5 Results

369 5.1 Verification of the numerical model for Cyclone Idai

370 Computed water levels near Beira show the strong tidal modulation and the wind-induced storm surge during the landfall of
371 the cyclone (Figure 3 – panel A; blue line for water level and vertical line for moment of landfall). Based on the difference
372 between the predicted astronomic tide and the total modeled water level, we estimate a storm surge of >3.5m due to the ~45
373 m/s wind speeds (Figure 3 – panel B). The storm surge at Beira is driven by wind setup as well as pluvial and fluvial drivers.
374 Deeper in the estuary, in the Pungwe flood plains, water levels peaked several days after landfall due to intense upstream
375 rainfall and subsequent runoff. Water levels near Buzi Village seem to be a combined result of first marine and second riverine-
376 driven water levels.



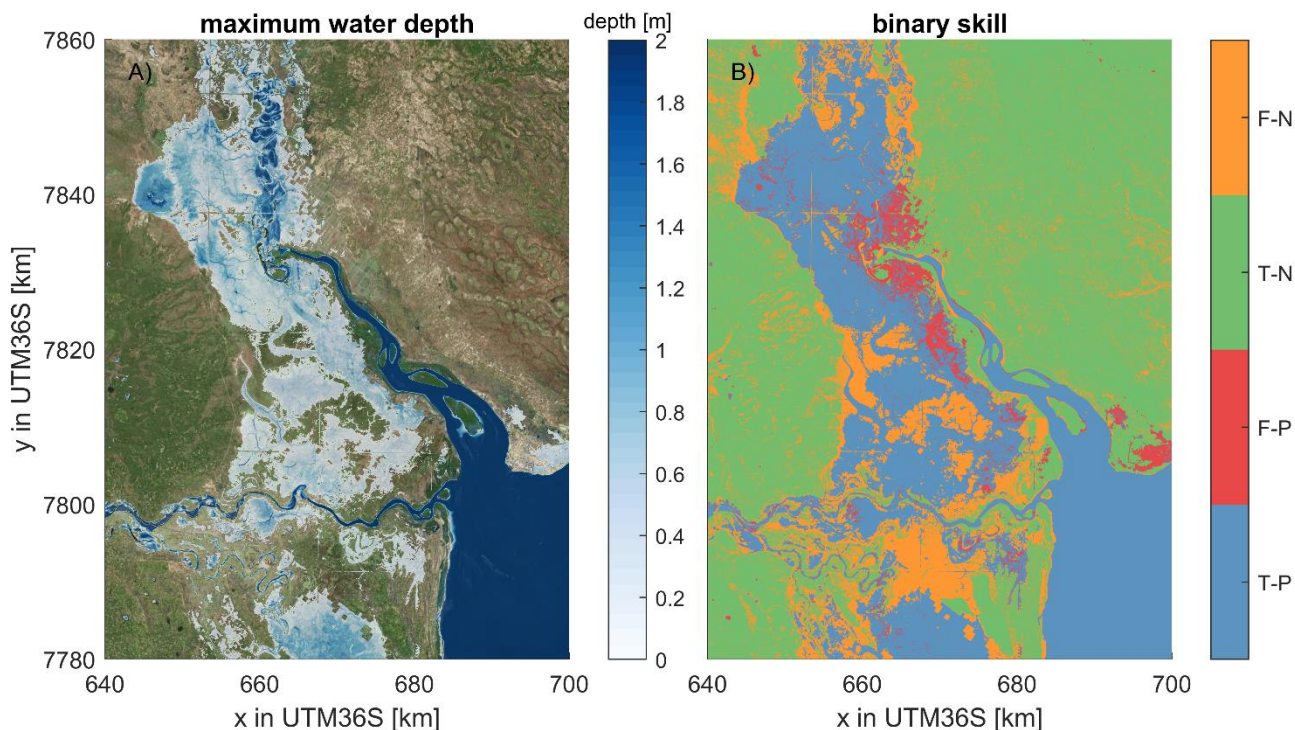
377

378 **Figure 3. Time series of water levels, wind speed, and precipitation within the study area. (A) Computed water levels at various**
379 **locations (blue for Beira, red for Buzi village, and green for upstream in the Pungwe estuary (see Figure 2C for their location) and**
380 **the black dashed line representing the astronomical prediction at Beira. (B) Simulated wind speed (blue) and rainfall rate (red) over**
381 **the same period. Idai made landfall on March 15, and its powerful winds and rainfall resulted in marine flooding at Beira and**
382 **riverine-driven flooding upstream in the estuary. The vertical line represents the moment of landfall.**

383 Validation of the SFINCS model for the observed extent (blue colors in Figure 4A) gives confidence in the ability to simulate
384 the compound flooding (Figure 4). The model can reproduce the Sentinel-1 derived extent with a Critical Success Index of
385 0.59. This skill score is comparable to previous work by Eilander et al. (2022), albeit somewhat lower. Based on the differences
386 between the modeled and satellite-derived extent, it becomes apparent that the model underestimates the flooding around the
387 Buzi River (false negative; orange colors in Figure 4B around 660-7800 km). We hypothesize this is due to the lack of river
388 inflow related to an underestimation of rainfall further upstream and/or overestimation of infiltration due to soil saturation
389 which is not considered. Moreover, the comparison with satellite-derived flood extent indicates an overestimation of the
390 flooding at Beira (false positive; red colors in Figure 4). Here, we suspect that the benchmark data might be off, and the coastal



391 flooding already receded before the Sentinel data recorded the extent. The observed high-water marks near Beira ranged from
392 3.6m within the estuary to 2.9 m + MSL at the open coast and are reproduced by SFINCS with respectively 3.8 and 3 m +
393 MSL. This difference suggests a positive bias of the model results at the coast of ~10-20 cm, similar to the tidal validation (see
394 Appendix 8.1), which revealed a median MAE of 21 cm.



395
396 **Figure 4. Maximum computed water depth (Panel A) and binary skill of flood extents for Idai (Panel B).** Water depths are
397 downscaled from the model resolution to the 10x10 meter resolution of the topo-bathymetry. The binary skill evaluation (Panel B)
398 assists in determining the model's accuracy and dependability, and the Sentinel-1 radar data is used as a reference to determine
399 skill. A true-positive (T-P) outcome denotes a correct flood prediction by the model compared to Sentinel-1 derived extent, whereas
400 a false-positive (F-P) occurs when the model forecasts a non-existent flood. In contrast, a false-negative (F-N) indicates where the
401 model overlooks an actual flood, and a true-negative (T-N) result occurs when the model accurately predicts the lack of a flood event.
402 The model produces large-scale flooding, which is largely also observed in the data, but local differences of over- and
403 underestimation exist. The coordinate system of this figure is WGS 84 / UTM 36 S (EPSG 32736). © Microsoft.

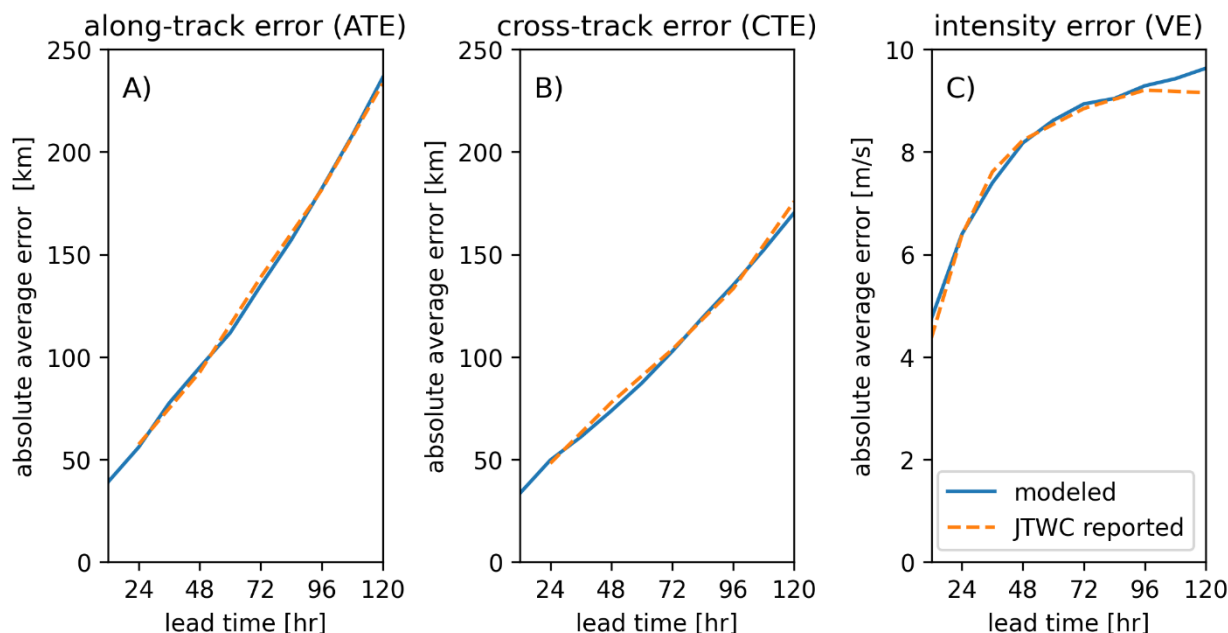
404



405 5.2 Calibration and influence of simplifications of TC-FF

406 5.2.1 Calibration of TC-FF: mean absolute error and auto-regression

407 This study used JTWC-reported errors for the along-track, cross-track, and intensity for the Southern Hemisphere to calibrate
408 our methodology (JTWC, 2021). For other case studies, for example, based on different forecasting agencies or in other ocean
409 basins), these reported errors can be used instead. Calibration is performed by minimizing the square-root difference between
410 computed and reported mean absolute values for various lead times using the Nelder-Mead method. This effort resulted in
411 mean absolute errors for B and D of 68.5 and 55.3 km and autoregression coefficients a_t , c_t , of 1.214 and 1.181 (Figure 5A
412 and B) for the along-track and cross-track. Moreover, we calibrated the mean absolute error and regression coefficients for the
413 intensity, which resulted in mean absolute errors for F of 9.28 m/s and autoregression coefficient e_t of 0.624 (Figure 5C).



414

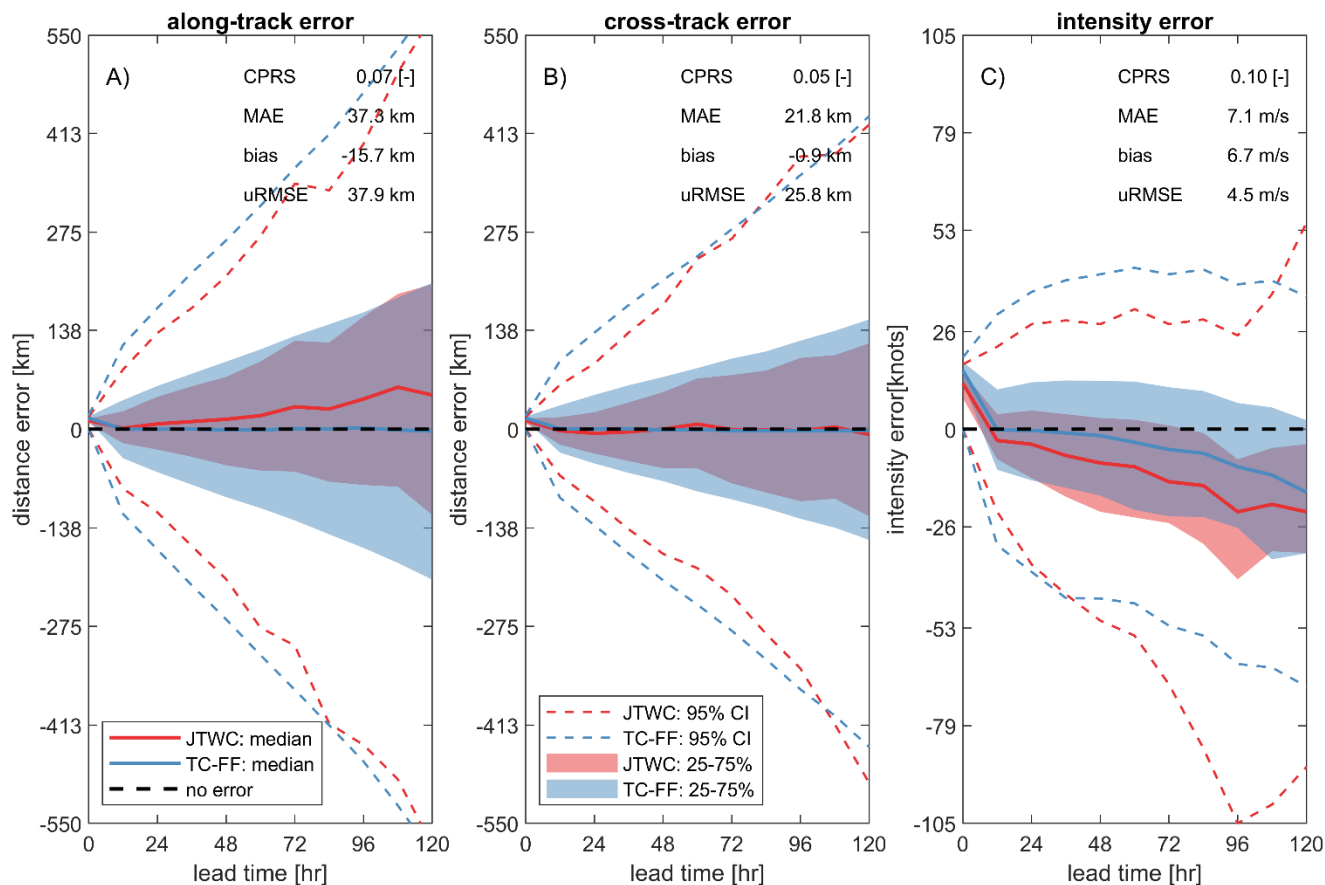
415 **Figure 5. Comparison of calibration results for the probabilistic forecasting method TC-FF (solid blue line) and the Joint Typhoon**
416 **Warning Center (JTWC) reported error statistics based on the 5-year average (2016-2020) in the Southern Hemisphere (dashed**
417 **orange line). Panel A represents the along-track error, Panel B demonstrates the cross-track error, and Panel C exhibits the wind**
418 **speed or intensity error. Modeled errors are based on 1,000 ensemble members. Modeled absolute average errors are similar to**
419 **JTWC.**

420 5.2.2 Comparison of TC-FF with operational forecast products

421 Errors produced by TC-FF are compared to the implementation from NOAA, NHC and JTWC that are used operationally.
422 Minor differences between the TC-FF and full implementation based on DeMaria et al. (2009) and De Maria et al. (2013) exist
423 and are attributed to the simplifications used in the error distribution (including the lack of GPCE) and lack of bias conditions.
424 The distribution in along-track, cross-track, and intensity error is typically in the same order (Figure 6), which is confirmed by



425 a median CPRS over various lead times from 0 to 120 hours of 0.07, 0.05, 0.10 and median MAE of 37 km, 21 km, and 7 m/s
426 of for respectively the along-track, cross-track, and intensity. At the same time, TC-FF has by design no bias corrections,
427 whereas the operational system does, leading to the positive median along-track error in red compared to the blue line in Figure
428 6A and a median bias of -16 km. Besides the median estimates, the interquartile range (25-75%) and 95% CI match relatively
429 well for the along-track and cross-track errors. Larger differences are found for the intensity error. In general, the wind intensity
430 error looks visually erratic and doesn't start at zero for no lead time, which is the result of the inland wind decay model. Both
431 JTWC and TC-FF have a negative bias due to the effect of land, but TC-FF does have a median bias of +6.7 m/s compared to
432 JTWC, suggesting that TC-FF overestimates. However, more substantial differences are found for the interquartile range and
433 95% CI. These findings for the along-track, cross-track, and intensity are supported by a more detailed analysis of the CDF
434 for the different parameters as a function of lead time (Figure 12, Figure 13, Figure 14). For the along-track and cross-track,
435 we observe an increase in the MAE and uRMSE as a function of lead time but a decrease in the CPRS. The increasingly larger
436 error distribution influences this pattern. Moreover, TC-FF produces Gaussian-distributed errors while the JTWC error
437 distribution differs since it is based on historical error distribution and adjusted based on the GPCE. Similar to Figure 6, larger
438 differences are found for the intensity error, which is influenced by the bias correction that increases with lead times.



439

440 **Figure 6.** Comparison of validation results for the probabilistic forecasting method TC-FF (blue line) and the Joint Typhoon
 441 Warning Center (JTWC) operational product (red line). Panel A represents the along-track error, Panel B demonstrates the cross-
 442 track error, and Panel C exhibits the wind speed or intensity error. Errors are computation for both the TC-FF and JTWC are
 443 based on 1,000 ensemble members. Solid lines are median estimats, shaded areas the interquartile range (25-75% CI) and dashed
 444 line the 95% CI. TC-FF and JTWC produce broadly similar error distributions for different lead times.

445 5.3 Forecasting of Idai using the TC-FF

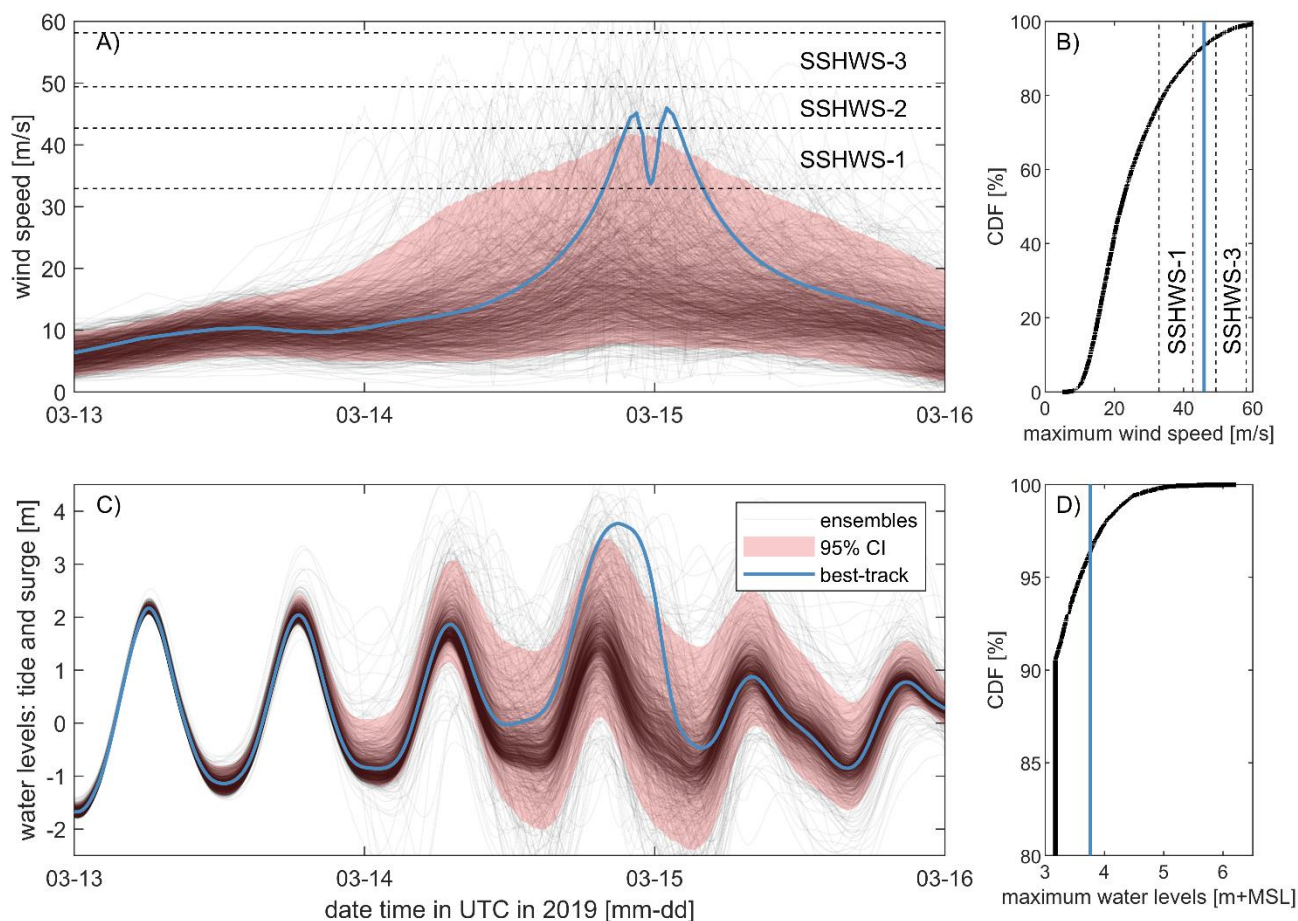
446 This section presents the application of forecasting Idai using the TC-FF.

447 5.3.1 Uncertainty three days before landfall

448 The TC-FF method with 10,000 ensemble members is applied to the case of Cyclone Idai. The results reveal that accounting
 449 for the uncertainty of the TC track and intensity of eye three days before landfall results in considerable uncertainty regarding
 450 wind speeds and water levels near Beira (Figure 7) or the region (Figure 8). In particular, the wind speeds show a 95% CI of
 451 about 7-40 m/s at the moment of landfall (Figure 7A) versus ~45 m/s or a Saffir-Simpson Hurricane Wind Scale (SSHWS) of
 452 2 of the best track. Moreover, TC category 1 wind speeds could occur as early as March 15 at 21:30 UTC or as late as March



453 16 at 19:50 UTC. This spread of possible maximum wind speeds at Beira results from the large uncertainties in intensity and
454 a difference in landfall location and time. Based on the same model simulations, the empirical cumulative distribution function
455 (CDF) of the maximum wind speed at Beira ranges from 8.8 to 59.2 with a median wind speed of 25.5 m/s, while the best
456 track has a 5.9% exceedance probability (Figure 7B). Consequently, water levels vary greatly (Figure 7C). For example,
457 ensemble members can exhibit a sizeable wind-driven setup due to TC wind blowing from offshore into the estuary, pushing
458 water up in the estuary and at Beira. For landfall locations west of the estuary, the wind blows offshore, resulting in a large
459 set-down. Note that Beira is in the Southern Hemisphere, and due to the Coriolis effect, TCs spin clockwise. The highest water
460 levels occur when high tide and wind-driven setup coincide, which explains the three peaks in the 95% CI water level given
461 the semi-diurnal tide and the highest possible wind speed for ~1.5 days (Figure 7C). The maximum water levels are dominated
462 by the tide except in the situation of cyclone impact (see the CDF in Figure 7D and the minimum value of ~3.5m+MSL around
463 90%, which is influenced by the tide and time window over which it is determined). The specific track of Idai resulted in
464 relatively extreme conditions compared to other possible combinations (both for winds and water levels). A similar pattern
465 can be observed in the spatial maps shown in Figure 8. The average probability of flooding in the area is 26%, with higher
466 probabilities of flooding found in the lower-lying portions of the estuary (note we are excluding points below MSL; Figure
467 8A). The 1% exceedance flood depth threshold shows a large extent and is quite similar to the computed extent due to Idai
468 (see Figure 4A for comparison with Figure 8B). The main difference is that there is more flooding near the city of Beira and
469 somewhat less near Buzi Village. The match between the 1% exceedance flood depth and the best track with Idai suggests that
470 the event was relatively severe and implies that even though many other potential scenarios could have unfolded, they likely
471 would not have resulted in the same extensive flooding caused by Idai.



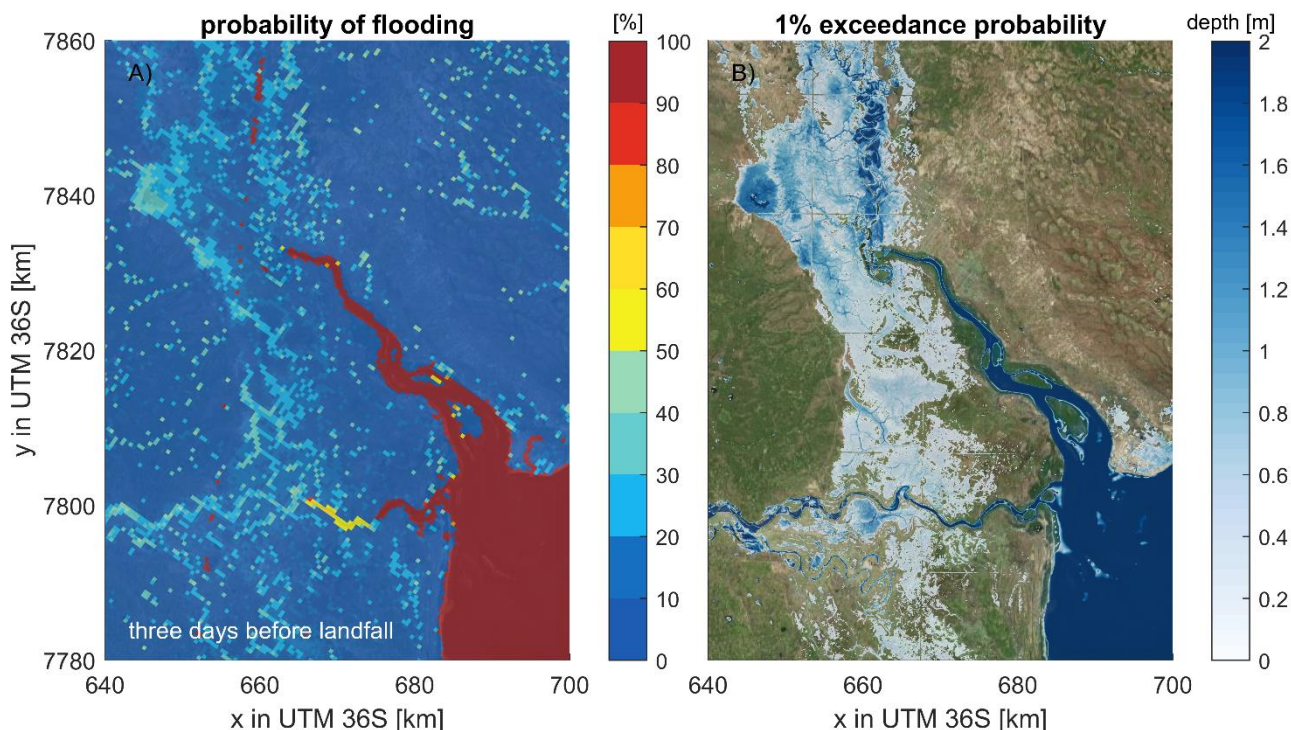
472

473 **Figure 7. Multi-panel Analysis of wind and water Levels three days before landfall: (A) time series of wind speeds, (B) maximum**
474 **wind speeds, (C) time series of water levels near Beira, and (D) maximum water levels. Data is derived from 10,000 ensemble**
475 **members (black transparent line; every 10th plotted) with red shading representing the 95% CI. The best track (blue line) and the**
476 **Saffir-Simpson Hurricane Wind Scale are included for comparison (panels A and B only). There is substantial uncertainty in wind**
477 **speeds and water levels near Beira three days prior to landfall.**

478

479

480



481
482

483 **Figure 8. Probabilistic flood analysis for Cyclone Idai three days before landfall: (A) Spatial distribution of flooding probability; (B)**
484 **Corresponding 1% exceedance water depth estimates, highlighting areas at most significant hazard. Results in panel A are**
485 **determined from 10,000 ensemble members on the original 500-meter model resolution, while water depth in panel B is downscaled**
486 **to the original 10x10-meter bathymetry resolution. Higher probabilities of flooding are found in the lower-lying portions of the**
487 **estuary. The coordinate system of this figure is WGS 84 / UTM 36 S (EPSG 32736). © Microsoft.**

488 5.3.2 Influence of sampling size

489 As described by Cashwell and Everett (1959) and DeMaria et al. (2009), the precision of Monte Carlo techniques is
490 proportional to the number of ensemble members (N). The convergence rate typically shows a slower progression than $1/N$,
491 constituting a limitation intrinsic to all Monte Carlo methods. To investigate the convergence rate and the error induced by
492 employing a finite number of ensemble members, the Idai forecasting case three days prior to landfall is used, analogous to
493 the preceding section, albeit with a variable number of ensemble members. Additionally, bootstrapping is employed to
494 approximate convergence rates and the accompanying uncertainty.

495

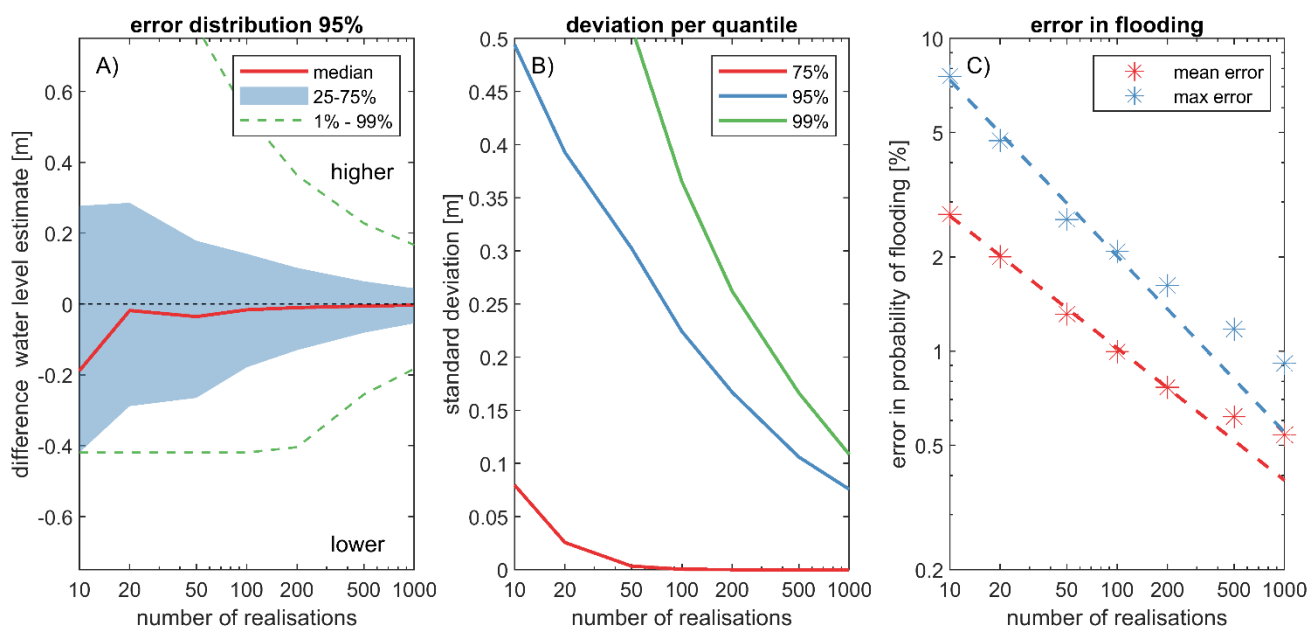
496 The estimation of the 95% exceedance maximum water levels in proximity to Beira exhibits convergence with the number of
497 ensemble members, albeit with considerable deviations compared to a fully converged solution with 10,000 members when
498 implementing a low number of ensemble members (Figure 9A). For instance, employing merely 50 ensemble members results
499 in an interquartile range (25-75%) of -0.28 to +0.10 m. Increasing the number of ensemble members reduces this sampling
500 uncertainty to a range of -0.09 to +0.06 m for 200 ensemble members.



501 Similarly, the standard deviation for several quantiles of maximum water level estimates at Beira reduces with more ensemble
502 members. It exhibits a similar pattern from higher to lower quantiles (Figure 9B). In essence, estimating rare events necessitates
503 executing more ensemble members to attain comparable convergence. This study found that the 95% exceedance maximum
504 water level at Beira when utilizing 200 ensemble members has a standard deviation of 21 cm (blue line Figure 9B). This level
505 of convergence seems acceptable since it is in a similar order as the skill of the hydrodynamic model (see Section 5.1).

506

507 The probability of error in flood potential is expressed as a function of N on a log-log plot (Figure 9C). Compared to a fully
508 converged solution with 10,000 members, for N=200, the mean error constitutes 0.95%, and the maximum error amounts to
509 1.53%. Note that this estimate is without considering the model error. In the log-log diagram, the errors exhibit near-linear
510 correlations with N and could serve as a basis for determining the number of ensemble members needed for a specified
511 confidence level. For instance, to achieve a maximum error of 1% in flood probability, it would be necessary to utilize 500
512 ensemble members.



513

514 **Figure 9. Sampling size effects on flood estimation accuracy. (A) Quantiles of sampling error for the 5% exceedance water level. (B)**
515 **Standard deviation of 75%, 95%, and 99% quantiles, illustrating the uncertainty in estimation. (C) Comparison of maximum and**
516 **average error in flood probability predictions. All panels were generated using 10,000 ensemble members and a 1000-bootstrap**
517 **resampling approach. Using more ensemble members reduces the sampling uncertainty.**

518 5.3.3 Importance of lead time

519 Thus far, the probabilistic TC forecasting framework has been implemented three days prior to the landfall of Idai.
520 Nevertheless, the forecast's results fluctuate with lead times, consequently influencing the associated evaluations of water
521 levels (Figure 10) and flood probabilities (Figure 11).

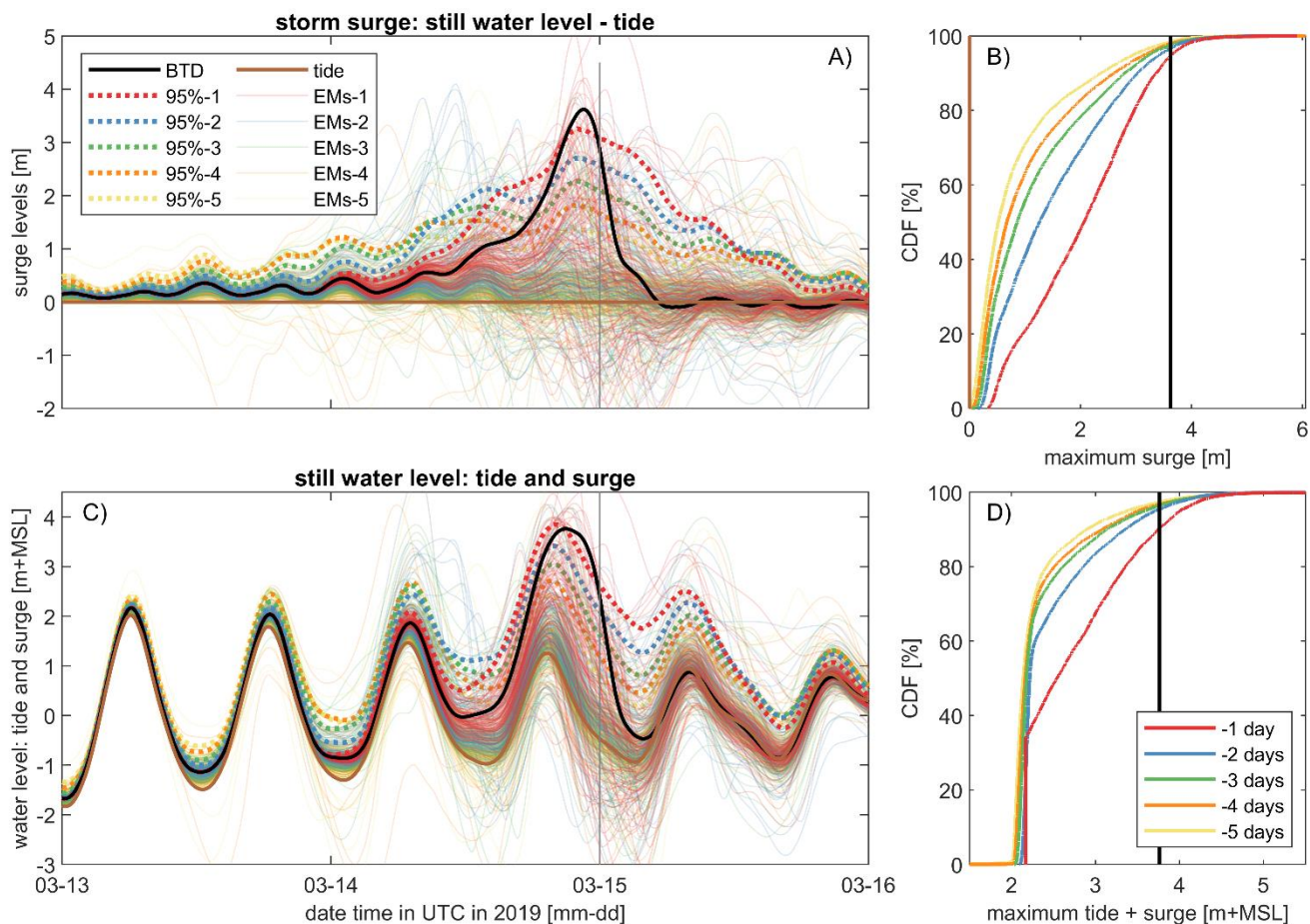


522

523 The predicted water levels (tide + surge) vary with lead times (Figure 10A and C). Specifically, at a lead time of five days
524 before landfall, an (unsurprisingly) larger spread between the ensemble members is observed compared to lead times of, for
525 example, one or three days. Moreover, as landfall approaches, the time series converges since increasing ensemble members
526 produce highly similar predictions. For example, notice how individual ensemble members 1 day before landfall show similar
527 storm surges and still water levels (i.e., the concentration of lines which becomes more apparent in Figure 10). Moreover, the
528 5% and 95% exceedance values become less spread out and more peaked around landfall (dashed lines in Figure 10). This
529 convergence is more apparent for the storm surge. The CDF of the maximum storm surge levels increases with reducing lead
530 time (Figure 10B). For example, the median storm surge increases from 0.5m five days before landfall to 0.9 and 2.0m for lead
531 times of three days and one day, respectively (notice the increasing median estimate in the CDF plot from 5 to 1 day in Figure
532 10B). This increase in maximum storm surge shows the increasing certainty that the TC will land near Beira. However, for
533 other locations, the opposite may occur as the landfall shifts away from it. The still water levels are influenced by both tidal
534 motions and the influence of the TC (Figure 11C). This strongly influences the maximum computed still water level (Figure
535 11D). For instance, the lowest maximum water level for all simulations is around ~2 m above MSL, resulting from the
536 maximum tidal range rather than the TC itself. The 95th quantile of the maximum still water level is 3.4 m + MSL five days
537 prior to landfall, which increases to 3.6 and 4.0 m+MSL for lead times of three days and one day, respectively. The best track
538 of Idai is included as a reference and estimated to have a 9% probability of exceedance 1 day before landfall.

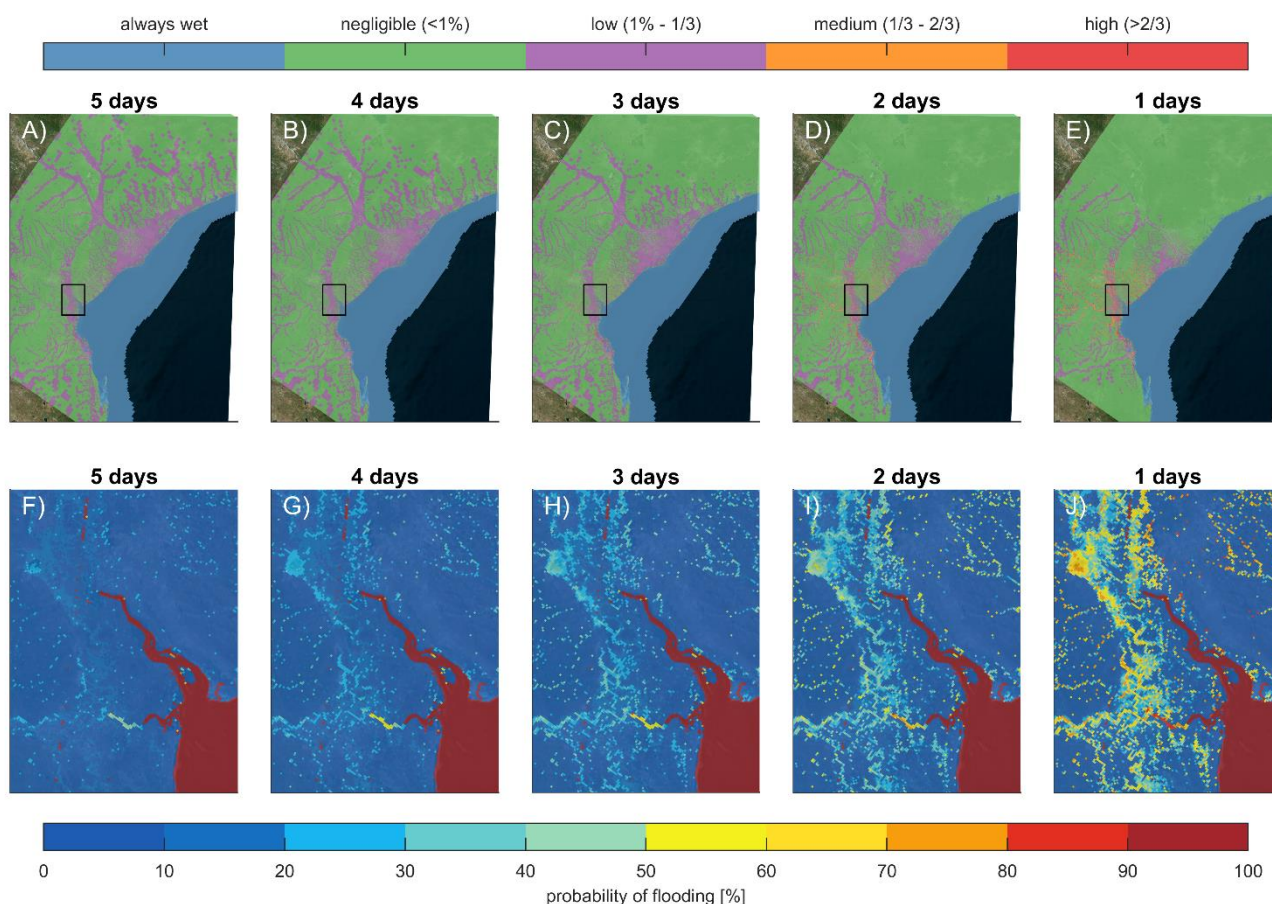
539

540 A large portion of the Sofala province faces a minor flood risk five days before the actual landfall. The flood probability for
541 the estuary near Beira increases as lead times reduce (Figure 11B). In particular, the average probability of flooding five days
542 before landfall is 15%, increasing to 17 and 24% for lead times of three and one day, respectively. Conversely, for the entire
543 model domain, a probability of greater than 1% flooding declines from 97 to 94 and 64 km² for lead times of five, three, and
544 one day (Figure 10A). In other words, five days before landfall, less confidence in predictions translates into more spatial
545 variability on flooding probability tied to a larger impact area. Closer to the actual landfall, there is more certainty over which
546 area will be affected.



547

548 **Figure 10. Forecasted water levels in Beira for 1-5 day lead times: temporal evaluation and cumulative distribution. Panel A and C:**
 549 **Time series illustrating the forecasted water levels in proximity to Beira with lead times ranging from 1 to 5 days prior to landfall**
 550 **showcasing both individual ensemble members (solid transparent lines; every 100th plotted), tide-only (brown), best-track (black)**
 551 **and quantile estimates (95% dashed lines). Panels A and C use the same colors and line styles. Panel B and D: Cumulative**
 552 **distribution function (CDF) showing the maximum water levels in ascending order for all ensemble members, providing insights**
 553 **into the probability of occurrence for various water level thresholds. Panels B and D use the same colors. Panels A and B show the**
 554 **storm surge levels (computed still water levels minus predicted tidal levels), while Panels C and D present the still water level (tide**
 555 **and surge).**



556

557 **Figure 11. Evolution of the flood probability prior to landfall: Panels A-E depict the spatial distribution of flooding probabilities at**
558 **5, 4, 3, 2, and 1 day(s) before landfall, respectively. Color gradients represent the varying probability. The top panels focus on the**
559 **entire area simulated and the bottom panels on the Pungwe and Buzi river deltas. With decreasing lead time, the area that could be**
560 **affected decreases while there is an increased probability of flooding near Beira. The coordinate system of this figure is WGS 84 /**
561 **UTM 36 S (EPSG 32736). © Microsoft.**

562



563 6 Discussion

564 This paper describes a new probabilistic method to forecast TC-induced coastal compound flooding by tide, surge, and rainfall
565 using Monte Carlo sampling. Due to the limited number of observations on TC evolution, for short-term operational analyses,
566 an autoregressive technique that imposes potential errors on top of the forecasted track is preferred over those parametric
567 sampling techniques used for long-term and strategic risk assessments based on historical records (e.g., Nederhoff et al., 2021).
568 In addition, for the same scarcity of observation, there is limited knowledge of the underlying joint distribution between TC
569 and ocean characteristics, which makes Monte Carlo sampling preferred compared to sampling techniques that are highly
570 efficient for complex multivariate patterns such as cluster analysis (e.g., Choi et al., 2009) and MDA methods (e.g., Bakker et
571 al., 2022). However, exploring the possibility of increasing efficiency via the aforementioned methods is important, especially
572 since the error space increases as a function of lead time, and estimating these events requires increasing amounts of ensemble
573 members (Figure 9B). However, this is a topic that requires an in-depth analysis and is beyond the scope of the present study.
574

575 Compared to the implementation of DeMaria et al. (2009) and DeMaria et al. (2013), the ensemble generation is simplified by
576 removing bias corrections, applying a single normal error distribution calibrated on historical errors (Figure 5), and does not
577 account for GPCE. While we acknowledge these simplifications, this method does make it possible to account for TC
578 forecasting errors for any ocean basin based on reported average historical errors alone. Nevertheless, the behavior of a specific
579 tropical cyclone (TC) does not necessarily conform to the "average" pattern, and differences between the operational JTWC
580 model were found (Figure 6). The implications of these assumptions on the precision and predictive proficiency of our
581 approach for coastal compound flooding remain undetermined. Our implementation has been recently integrated into an
582 operational system tailored for the contiguous United States. Verification of the reliability of this operational system is
583 currently pending. Regardless, TC-FF compares well with the predictions provided by ECMWF of Idai that showed a
584 probability of 50 to 90% of severe flooding four to one day before landfall (Figure 10).
585

586 In the introduced methodology, we apply the compound flooding model SFINCS. While the validation gave confidence that
587 the hydrodynamic model reproduces the main tidal motions and flooding during Idai, differences did exist compared to the
588 data (Figure 4). The model skill could be improved by including additional wind radii information in the parametric wind
589 model (e.g. radius of gale force winds along different quadrants) and more accurately resolving on-land winds, rainfall, and
590 infiltration processes. For example, Done et al. (2020) present a methodology to account for terrain effects by adjusting winds
591 from a parametric wind field model by using a numerical boundary layer model. Here, we applied the IPET empirical
592 relationship that relates pressure drop to rainfall intensity. Initially, deployment showed the necessity to triple the rainfall rate
593 due to severe underestimation of the total rainfall and associated flooding. Improvement (stochastic parametrization) of TC
594 rainfall could overcome this limitation. Moreover, SFINCS was run with a constant infiltration rate and does not account for
595 drainage systems and flood protection measures besides the frontal levee. It is also unknown how the topo-bathymetry that



596 was collected before Idai influenced results. Lastly, the effects of waves (e.g., setup, runup, overtopping) and morphological
597 change were not considered. All these limitations affect the model skill and could explain some mismatches observed compared
598 to Sentinel-1 data and high-water marks at Beira. However, the computational efficiency of SFINCS allowed us to run
599 thousands of ensemble members on limited computational resources. We accept the loss of some model accuracy with this
600 gain of speed.

601

602 The focus of the development of TC-FF has been geared to the computation of overland flooding. However, TCs pose
603 significant hazards through both water *and* wind. A study by Rappaport (2014) indicated that from 1963 to 2012 in the United
604 States, approximately 90% of fatalities associated with tropical cyclones were due to water-related incidents. The wind-related
605 fatalities were about 8%. This does not provide insight into the cause of damage associated with landfalling TCs, nor does it
606 provide insight into how these ratios vary across the globe. Regardless, TC-FF does provide the possibility to estimate extreme
607 wind speeds and link this to potential damage as an additional data product. Including wind damage as part of our framework
608 is something we are planning to work on in the future.

609 **7 Conclusions**

610 A new method and highly flexible open-source tool was developed to perform probabilistic forecasting of tropical cyclone-
611 induced coastal compound flooding. The Tropical Cyclone Forecasting Framework, TC-FF, computes a set of ensemble
612 members based on a simplified DeMaria et al. (2009) method. In particular, TC-FF uses gridded time- and spatially-varying
613 wind and pressure fields or forecasted tracks and combines this with historical observed error on the along-track, cross-track,
614 and intensity. Subsequently, the tool creates a temporally and spatially varying wind field, including rainfall, to force a
615 computationally efficient compound flood model. This approach allows for the inference of probabilistic wind and flood hazard
616 maps calibrated to any ocean basin in the world with limited computational resources. In contrast to the current practice, TC-
617 FF allows uncertainty analysis using large ensembles produced with physics-based models, narrowing down confidence bands
618 on forecasting coastal compound flooding focused on operational TC risk analyses.

619

620 The validation of the quadtree SFINCS model for Mozambique's Sofala province showed reliable skill in terms of tidal
621 propagation in the area of interest (median MAE of 21 cm), including good skill in reproducing the observed flood extent for
622 the case of the flooding caused by Cyclone Idai (2021). The model was able to reproduce the storm surge generation during
623 landfall and flooding near the city of Beira, including the subsequent compound flooding resulting from rainfall runoff in the
624 Pungwe estuary (critical success index of 0.59). Moreover, the model runs efficiently with a wall clock time of 4 minutes for
625 a 7-day event allowing it to be deployed in probabilistic operational assessments when using multiple cores.

626



627 TC-FF was calibrated with the average reported errors for the southern hemisphere via the Nelder-Mead method to determine
628 the mean absolute errors and autoregression coefficients. A comparison between TC-FF and JTWC (based on the complete
629 implementation of DeMaria et al., 2009) and DeMaria et al., 2013) revealed minor differences. In particular, for various lead
630 times from 0 to 120 hours, a median Continuous Ranked Probability Score (CRPS) of 0.07, 0.05, 0.10 and median MAE of 37
631 km, 21 km, and 7 m/s for respectively the along-track, cross-track, and intensity error were found. These findings give
632 confidence that the TC-FF, including the simplified DeMaria et al. (2009) implementation, can be used for more generalized
633 applications in data-scarce environments.

634

635 TC-FF provides valuable insights into the uncertainty of wind speeds, water levels, and potential flooding due to Idai, revealing
636 the impacts of track and intensity uncertainties. This is demonstrated in the wide array of possible maximum wind speeds and
637 significant fluctuations in water levels, which are primarily affected by tidal influences and the cyclone. For instance, even
638 just three days prior to landfall, there's a broad spread in the predicted flood areas. This suggests that there is still a significant
639 chance that Idai may not hit the anticipated area or may not generate a substantial storm surge.

640

641 The precision of forecasts is directly related to the number of ensemble members used. A mean error in flood probability of
642 less than 1% and <20 cm sampling errors for the 1% exceedance water level at Beira required 200 members. Based on that,
643 we determine that at least 200 ensemble members are needed to get reliable water levels and flood results three days before
644 landfall. A higher number of ensemble members reduces sampling uncertainty and increases the accuracy of water level and
645 flood potential estimates.

646

647 The lead time before landfall has a considerable impact on the forecast's precision. As the lead time decreases, the variability
648 of forecasts diminishes, and the forecasts converge to similar predictions. Similarly, the probability of flooding in certain areas,
649 such as the estuary near Beira, increases as the lead time shortens, providing more certainty over the areas that will be affected
650 by the event.

651

652 TC-FF offers a significant advancement compared to the current status quo of a single deterministic simulation when
653 forecasting tropical cyclone compound flooding hazards. This approach facilitates a comprehensive understanding of complex
654 interdependencies and uncertainties. By quantifying the likelihood of various outcomes, probabilistic methods enable
655 stakeholders to make more informed decisions, allocate resources better, and enhance preparedness and resilience in the face
656 of these catastrophic natural phenomena.

657

658 *Code and data availability.*



659 The code and data are freely available to other researchers and consultants. The Python code for this method is accessible on
660 GitHub via the following link: <https://github.com/Deltares/CoastalHazardsToolkit>.

661

662 *Author contributions.*

663 KN and MvO developed the method and the outline for the manuscript. KN wrote the initial manuscript, with editorial
664 comments by JV, AvD, JA, T and DR.

665

666 *Competing interests.*

667 The authors declare that they have no conflict of interest.

668

669 *Acknowledgments and financial support*

670 The authors thank the Deltares research programs 'Natural Hazards' and 'Risk Analysis and Management', which have provided
671 financing to develop and write this paper. Moreover, we want to thank Buck Sampson for input and data regarding the
672 operational wind field probabilities.

673



674 **8 Appendices**

675 **8.1 Tidal calibration and validation**

676 A tidal calibration was performed on the SFINCS computed tidal constituents compared to the tidal constituents at Beira.
 677 Constituents with an amplitude of more than 5 cm (M2, S2, N2, K2, and K1) were adjusted in terms of amplitude
 678 (multiplication) and phase (addition). Amplitude changes varied between 0.84 and 1.07 while phase difference changed on
 679 average by 40°. These calibration steps of adjusting the tidal constituents substantially reduced tidal errors at the Beira from a
 680 MAE of 43 to 17 cm. Secondly, model skill in reproducing tidal amplitudes and phases is assessed at 7 tide stations across the
 681 area of interest (including the calibration station of Beira). The SFINCS model reproduces tide with a median MAE of 21 cm,
 682 median RMSE of 25 cm, and median difference in M2 and S2 amplitude and phase of respectively -10 and -1 cm and -10 to -
 683 12° (median values computed over the different stations). Our hypothesis is that the reduction in tidal error observed at Beira
 684 throughout the calibration process might be due to a misalignment in the amplitudes and phases of the TPXO model which
 685 were used to generate the tidal boundary conditions (see Section 3.1.2). Presumably, the bathymetry contributes to the error
 686 observed in the validation process.

687

688 **Table A1. Evaluation of model proficiency in replicating tides near the Sofala province. Stations are ordered south to north. Columns**
 689 **one and two present the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), respectively, as error metrics for the**
 690 **comparison between observed and simulated tidal time series. The final four columns display the discrepancy (Δ) in amplitude (A)**
 691 **and phase difference (ϕ) for the two most prominent tidal constituents in the area (M2 and S2), where Δ is calculated as the difference**
 692 **between observed and simulated values.**

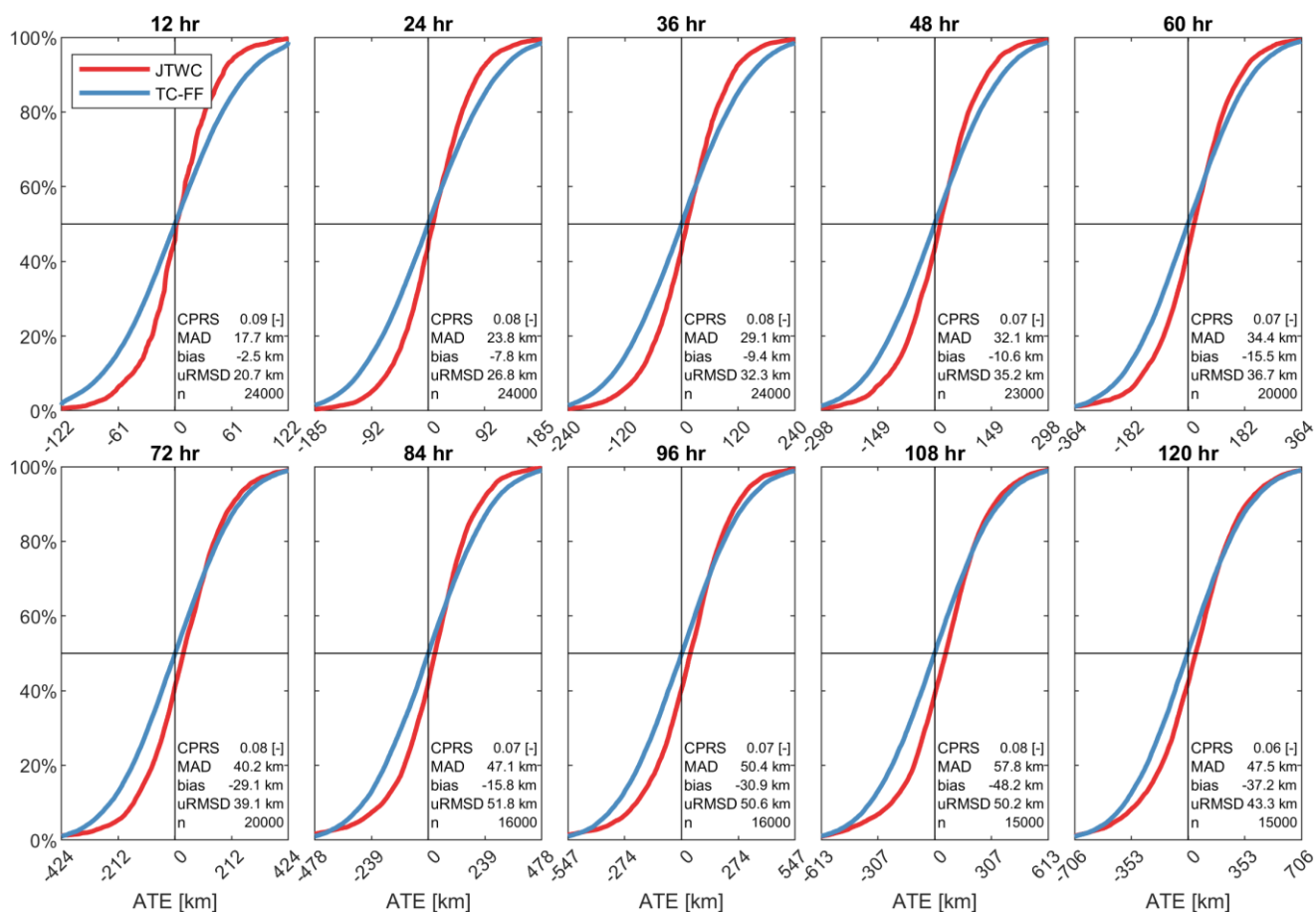
Name	MAE [m]	RMSE [m]	Δ M2 A [m]	Δ M2 ϕ [°]	Δ S2 A [m]	Δ S2 ϕ [°]
Bazaruto	0.13	0.15	-0.10	-7	0.01	-2
Bartolomeu Dias	0.12	0.15	-0.14	1	-0.11	-1
Chiloane	0.30	0.41	0.20	-10	0.08	-15
Beira	0.17	0.20	0.00	0	0.00	0
Chinde	0.21	0.25	-0.08	-13	-0.01	-12
Quelimane	0.26	0.32	-0.14	-15	-0.09	-21
Pebane	0.21	0.25	-0.14	-11	-0.09	-15
Median	0.21	0.25	-0.10	-10	-0.01	-12

693

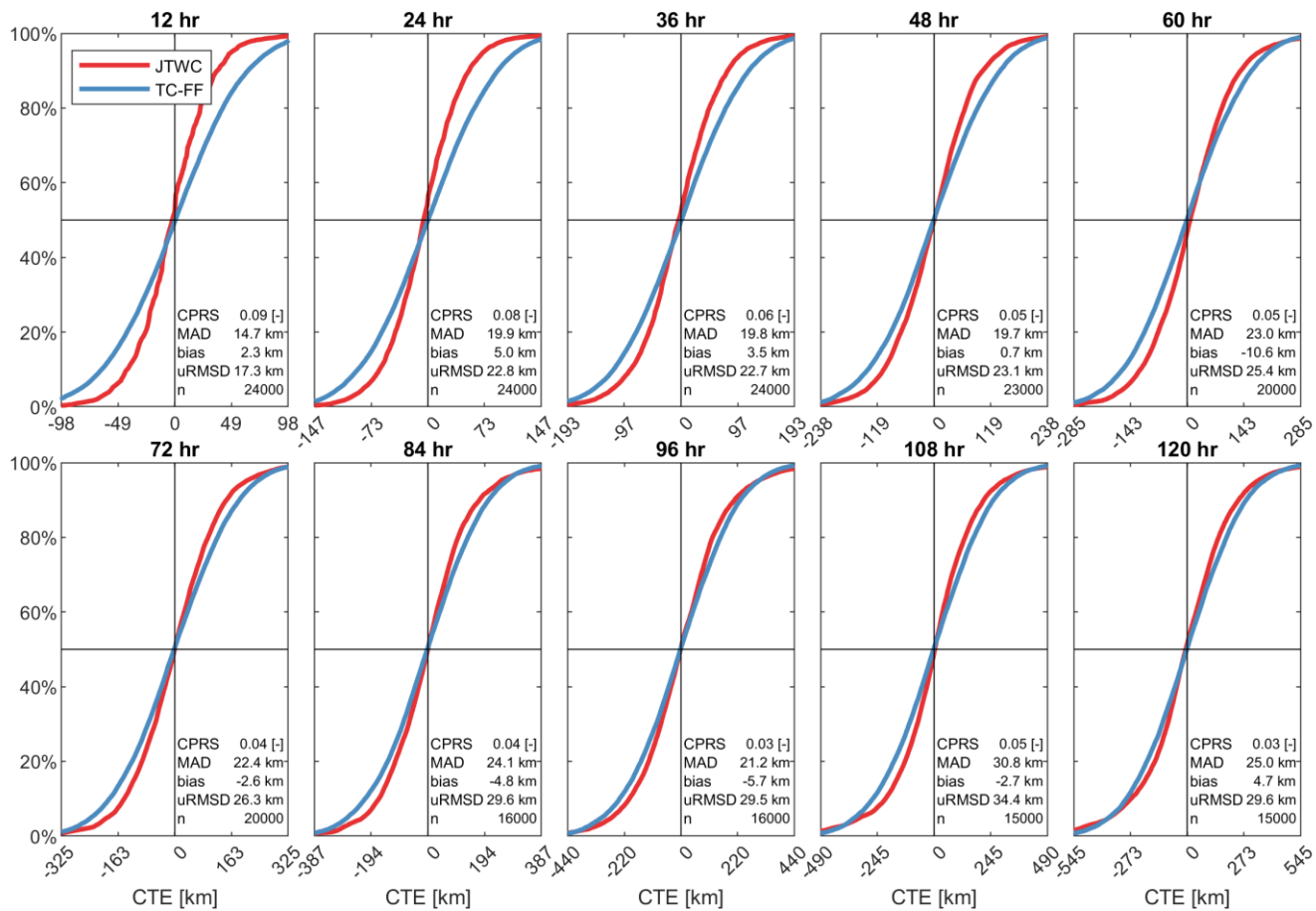


694 **8.2 Additional figures for Section 5.2.2. 'Influence of simplification in TC-FF'**

695 Figure 12, Figure 13 and Figure 14 provide additional information for Section 5.2.2. 'Influence of simplification in TC-FF'.

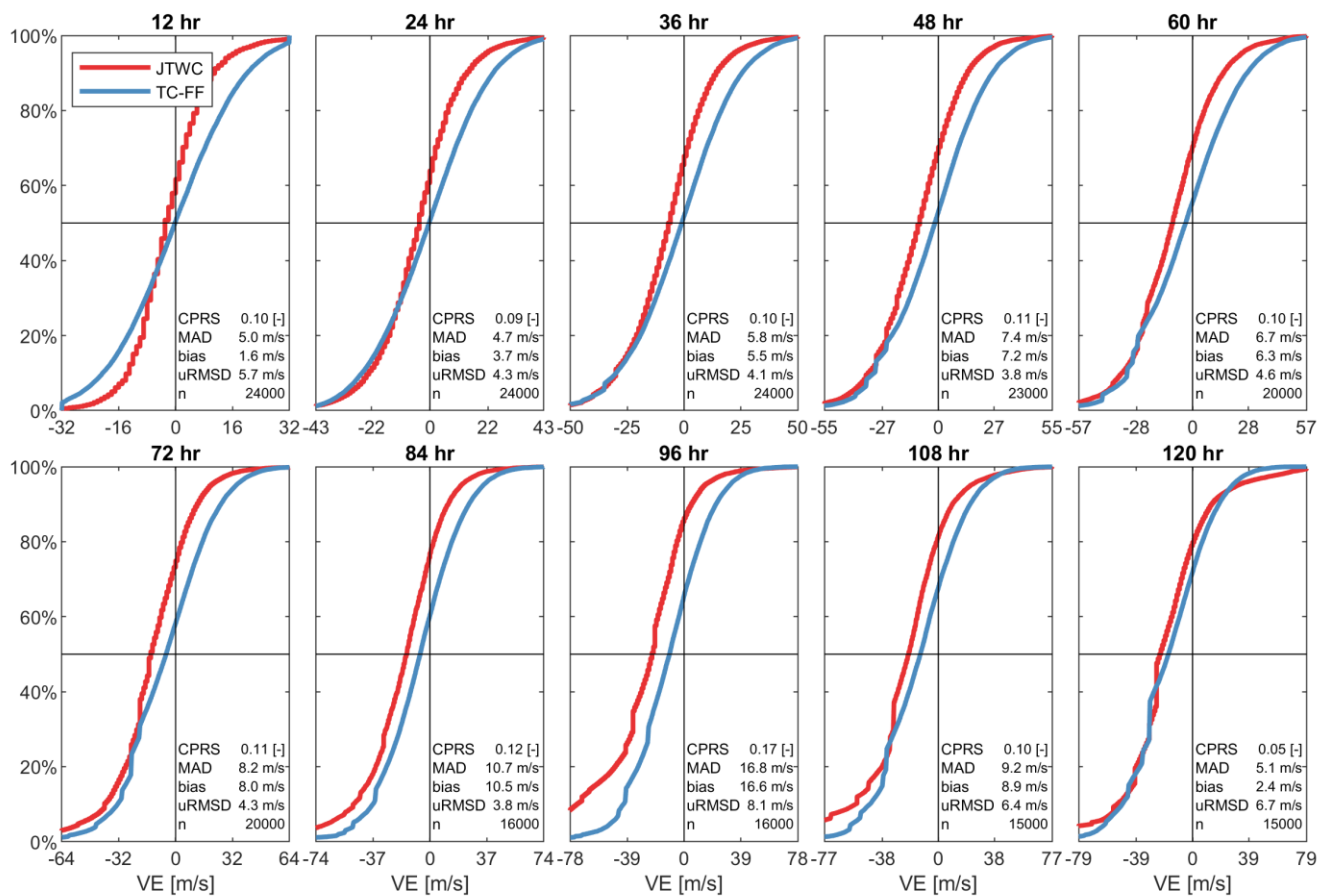


696
 697 **Figure 12. Comparison between the cumulative distribution function (CDF) of the along-track-error (ATE) for JTWC (red;**
 698 **reference) and TC-FF (blue; modeled). The different panels represent different lead times.**



699

700 **Figure 13. Comparison between the cumulative distribution function (CDF) of the cross-track-error (CTE) for JTWC (red;**
 701 **reference) and TC-FF (blue; modeled). The different panels represent different lead times.**



702

703 **Figure 14. Comparison between the cumulative distribution function (CDF) of the intensity error (VE) for JTWC (red; reference)**
704 **and TC-FF (blue; modeled). The different panels represent different lead times.**

705



706

707 **9 References**

- 708 Agriculture, U. S. D. of. (2009). *National Engineering Handbook Chapter 7 Hydrologic Soil Groups*.
- 709 Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013).
710 GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth*
711 *System Sciences*, 17(3), 1161–1175. <https://doi.org/10.5194/hess-17-1161-2013>
- 712 Ayyad, M., Orton, P. M., El Safty, H., Chen, Z., & Hajj, M. R. (2022). Ensemble forecast for storm tide
713 and resurgence from Tropical Cyclone Isaias. *Weather and Climate Extremes*, 38(May), 100504.
714 <https://doi.org/10.1016/j.wace.2022.100504>
- 715 Bakker, T. M., Antolínez, J. A. A., Leijnse, T. W. B., Pearson, S. G., & Giardino, A. (2022). Estimating
716 tropical cyclone-induced wind, waves, and surge: A general methodology based on representative
717 tracks. *Coastal Engineering*, 176(August 2021). <https://doi.org/10.1016/j.coastaleng.2022.104154>
- 718 Chen, B. F., Kuo, Y. Te, & Huang, T. S. (2023). A deep learning ensemble approach for predicting
719 tropical cyclone rapid intensification. *Atmospheric Science Letters*, 24(5), 1–11.
720 <https://doi.org/10.1002/asl.1151>
- 721 Choi, K.-S., Kim, B.-J., Choi, C.-Y., & Nam, J.-C. (2009). Cluster analysis of Tropical Cyclones
722 making landfall on the Korean Peninsula. *Advances in Atmospheric Sciences*, 26(2), 202–210.
723 <https://doi.org/10.1007/s00376-009-0202-1>
- 724 de Vries, H. (2009). Probability Forecasts for Water Levels at the Coast of The Netherlands. *Marine*
725 *Geodesy*, 32(2), 100–107. <https://doi.org/10.1080/01490410902869185>
- 726 Deltares. (2021). *Beira Coastal Protection Preparation study : flood hazard modelling*.
- 727 Deltares. (2018). *Wind Enhance Scheme for cyclone modelling - User Manual*. 1–110.
- 728 DeMaria, M., Knaff, J. A., Brennan, M. J., Brown, D., Knabb, R. D., DeMaria, R. T., Schumacher, A.,
729 Lauer, C. A., Roberts, D. P., Sampson, C. R., Santos, P., Sharp, D., & Winters, K. A. (2013).
730 Improvements to the operational tropical cyclone wind speed probability model. *Weather and*
731 *Forecasting*, 28(3), 586–602. <https://doi.org/10.1175/WAF-D-12-00116.1>



- 732 DeMaria, M., Knaff, J., Knabb, R., Lauer, C., Sampson, C., & DeMaria, R. T. (2009). A New Method
733 for Estimating Tropical Cyclone Wind Speed Probabilities. *Weather and Forecasting*, 24(6),
734 1573–1591. <https://doi.org/10.1175/2009WAF2222286.1>
- 735 Doyle, J., Hodur, R., Chen, S., Jin, Y., Msokaitis, J., Wang, S., Hendricks, E., Jin, J., & Smith, T.
736 (2014). Tropical Cyclone Prediction Using COAMPS-TC. *Oceanography*, 27(3), 104–115.
737 <https://doi.org/10.5670/oceanog.2014.72>
- 738 Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal*
739 *of Atmospheric and Oceanic Technology*, 19(2), 183–204. [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2)
740 [0426\(2002\)019<0183:EIMOBO>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2)
- 741 Eilander, D., Couasnon, A., Leijnse, T., Ikeuchi, H., Yamazaki, D., Muis, S., Dullaart, J., Winsemius,
742 H. C., & Ward, P. J. (2022). A globally-applicable framework for compound flood hazard
743 modeling. *EGUsphere*, 2022(April), 1–40. <https://doi.org/10.5194/egusphere-2022-149>
- 744 Emerton, R., Cloke, H., Ficchi, A., Hawker, L., de Wit, S., Speight, L., Prudhomme, C., Rundell, P.,
745 West, R., Neal, J., Cuna, J., Harrigan, S., Titley, H., Magnusson, L., Pappenberger, F., Klingaman,
746 N., & Stephens, E. (2020). Emergency flood bulletins for Cyclones Idai and Kenneth: A critical
747 evaluation of the use of global flood forecasts for international humanitarian preparedness and
748 response. *International Journal of Disaster Risk Reduction*, 50, 101811.
749 <https://doi.org/10.1016/j.ijdr.2020.101811>
- 750 Flowerdew, J., Horsburgh, K., Wilson, C., & Mylne, K. (2010). Development and evaluation of an
751 ensemble forecasting system for coastal storm surges. *Quarterly Journal of the Royal*
752 *Meteorological Society*, 136(651), 1444–1456. <https://doi.org/10.1002/qj.648>
- 753 Goerss, J. S. (2007). Prediction of consensus tropical cyclone track forecast error. *Monthly Weather*
754 *Review*, 135(5), 1985–1993. <https://doi.org/10.1175/MWR3390.1>
- 755 Gonzalez, T., & Taylor, A. (2018). *Development of the NWS' Probabilistic Tropical Storm Surge*
756 *Model*. 2018.
- 757 Hasegawa, H., Kohno, N., & Itoh, M. (2015). Development of Storm Surge Model in Japan
758 Meteorological Agency. *14th International Workshop on Wave Hindcasting and Forecasting &*
759 *5th Coastal Hazard Symposium, May 2010*, 1–6.
- 760 Higaki, M., Hayashibara, H., & Nozaki, F. (2009). Outline of the High Resolution Global Model at the
761 Japan Meteorological Agency. *RSMC Tokyo-Typhoon Center Technical Review*, 11:25–38, 2009.



- 762 Holland, G. J., Belanger, J. ., & Fritz, A. (2010). A Revised Model for Radial Profiles of Hurricane
763 Winds. *American Meteorological Society*, 4393–4401. <https://doi.org/10.1175/2010MWR3317.1>
- 764 Hu, K., Chen, Q., & Fitzpatrick, P. (2012). Assessment of a Parametric Hurricane Surface Wind Model
765 for Tropical Cyclones in the Gulf of Mexico. In *Advances in Hurricane Research - Modelling,*
766 *Meteorology, Preparedness and Impacts*. InTech. <https://doi.org/10.5772/51288>
- 767 IOC IHO and BODC, '. (2003). *Centenary Edition of the GEBCO Digital Atlas, published on CD-ROM*
768 *on behalf of the Intergovernmental Oceanographic Commission and the International*
769 *Hydrographic Organization as part of the General Bathymetric Chart of the Oceans; British*
770 *Oceanog.*
- 771 Jelesnianski, C. P. ., Chen, J., & Shaffer, W. A. . (1992). SLOSH : Sea, Lake, and Overland Surges from
772 Hurricanes. *NOAA Technical Report, April*.
- 773 Joint Typhoon Warning Center. (2022). *Typhoon Idai (2021) Best Track Data [Data File]*. Retrieved
774 from <https://www.metoc.navy.mil/jtwc/jtwc.html>.
- 775 Joint Typhoon Warning Center. (2021). *Annual Tropical Cyclone Report: 2020 [PDF file]*. Retrieved
776 from <https://www.metoc.navy.mil/jtwc/products/atcr/2020atcr.pdf>.
- 777 Lamers, A., Devi S, S., Sharma, M., Berg, R., Gálvez, J. M., Yu, Z., Kriat, T., Cardos, S., Grant, D., &
778 Moron, L. A. (2023). Forecasting Tropical Cyclone Rainfall and Flooding Hazards and Impacts.
779 *Tropical Cyclone Research and Review*. <https://doi.org/10.1016/j.tcr.2023.06.005>
- 780 Lecacheux, S., Rohmer, J., Paris, F., Pedreros, R., Quetelard, H., & Bonnardot, F. (2021). Toward the
781 probabilistic forecasting of cyclone-induced marine flooding by overtopping at Reunion Island
782 aided by a time-varying random-forest classification approach. *Natural Hazards*, 105(1), 227–251.
783 <https://doi.org/10.1007/s11069-020-04307-y>
- 784 Leijnse, T., van Ormondt, M., Nederhoff, K., & van Dongeren, A. (2021). Modeling compound
785 flooding in coastal systems using a computationally efficient reduced-physics solver: Including
786 fluvial, pluvial, tidal, wind- and wave-driven processes. *Coastal Engineering*, 163, 103796.
787 <https://doi.org/https://doi.org/10.1016/j.coastaleng.2020.103796>
- 788 Liang, Q., Du, G., Hall, J. W., & Borthwick, A. G. (2008). Flood Inundation Modeling with an
789 Adaptive Quadtree Grid Shallow Water Equation Solver. *Journal of Hydraulic Engineering*,
790 134(11), 1603–1610. [https://doi.org/10.1061/\(ASCE\)0733-9429\(2008\)134:11\(1603\)](https://doi.org/10.1061/(ASCE)0733-9429(2008)134:11(1603))



- 791 Lin, N., Emanuel, K., Oppenheimer, M., & Vanmarcke, E. (2012). Physically based assessment of
792 hurricane surge threat under climate change. *Nature Climate Change*, 2(6), 462–467.
793 <https://doi.org/10.1038/nclimate1389>
- 794 Luettich, R. A., Westerink, J. J., & Scheffner, N. W. (1992). ADCIRC: An Advanced Three-
795 Dimensional Circulation Model for Shelves Coasts and Estuaries, Report 1: Theory and
796 Methodology of ADCIRC-2DDI and ADCIRC-3DL, Dredging Research Program Technical
797 Report DRP-92-6. In *Coastal Engineering Research Center (U.S.), Engineer Research and*
798 *Development Center (U.S.).* (Issue 32466, pp. 1–137). [https://erdc-](https://erdc-library.erdc.dren.mil/jspui/handle/11681/4618)
799 [library.erdc.dren.mil/jspui/handle/11681/4618](https://erdc-library.erdc.dren.mil/jspui/handle/11681/4618)
- 800 Matheson, J. E., & Winkler, R. L. (1976). Scoring Rules for Continuous Probability Distributions.
801 *Management Science*, 22(10), 1087–1096. <https://doi.org/10.1287/mnsc.22.10.1087>
- 802 Mori, N., & Shimura, T. (2023). Tropical cyclone-induced coastal sea level projection and the
803 adaptation to a changing climate. *Cambridge Prisms: Coastal Futures*, 1, e4.
804 <https://doi.org/10.1017/cft.2022.6>
- 805 NASA GPM. (2019). *Global Precipitation Measurement (GPM) of Cyclone Idai via*
806 <https://gpm.nasa.gov/tropical-storm-idai-measured-gpm>. Accessed on March 16, 2023.
- 807 National Hurricane Center. (2023). *About NHC Graphics*. Retrieved June 27, 2023, from
808 <https://www.nhc.noaa.gov/aboutnhcgraphics.shtml#WATCHWARN>.
- 809 Nederhoff, K., Giardino, A., van Ormondt, M., & Vatvani, D. (2019). Estimates of tropical cyclone
810 geometry parameters based on best track data. *Natural Hazards and Earth System Sciences*,
811 19(11), 2359–2370. <https://doi.org/10.5194/nhess-19-2359-2019>
- 812 Nederhoff, K., Hoek, J., Leijnse, T., van Ormondt, M., Caires, S., & Giardino, A. (2021). Simulating
813 synthetic tropical cyclone tracks for statistically reliable wind and pressure estimations. *Natural*
814 *Hazards and Earth System Sciences*, 21(3), 861–878. <https://doi.org/10.5194/nhess-21-861-2021>
- 815 Nguyen, D. T., & Chen, S. T. (2020). Real-time probabilistic flood forecasting using multiple machine
816 learning methods. *Water (Switzerland)*, 12(3), 1–13. <https://doi.org/10.3390/w12030787>
- 817 OCHA, U. (n.d.). https://www.unocha.org/sites/unocha/files/CycloneIdai_BusinessGuide_vFinal.pdf.
- 818 Rappaport, E. N. (2014). Fatalities in the united states from atlantic tropical cyclones: New data and
819 interpretation. *Bulletin of the American Meteorological Society*, 95(3), 341–346.
820 <https://doi.org/10.1175/BAMS-D-12-00074.1>



- 821 Resio, D. T., & Irish, J. L. (2015). Tropical Cyclone Storm Surge Risk. *Current Climate Change*
822 *Reports*, 1(2), 74–84. <https://doi.org/10.1007/s40641-015-0011-9>
- 823 Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanni re, B., Mecking, J., Haarsma, R.,
824 Bellucci, A., Scoccimarro, E., Caron, L. P., Chauvin, F., Terray, L., Valcke, S., Moine, M. P.,
825 Putrasahan, D., Roberts, C. D., Senan, R., Zarzycki, C., ... Wu, L. (2020). Projected Future
826 Changes in Tropical Cyclones Using the CMIP6 HighResMIP Multimodel Ensemble. *Geophysical*
827 *Research Letters*, 47(14), 1–12. <https://doi.org/10.1029/2020GL088662>
- 828 Roy, C., & Kovord nyi, R. (2012). Tropical cyclone track forecasting techniques — A review.
829 *Atmospheric Research*, 104–105, 40–69. <https://doi.org/10.1016/j.atmosres.2011.09.012>
- 830 Schwerdt, R. W., Ho, F. P., & Watkins, R. R. (1979). Meteorological criteria for standard project
831 hurricane and probable maximum hurricane wind fields. *Gulf and East Coasts of the United States*.
832 *NOAA Tech. Rep. NWS* 23, 317.
- 833 Suh, S. W., Lee, H. Y., Kim, H. J., & Fleming, J. G. (2015). An efficient early warning system for
834 typhoon storm surge based on time-varying advisories by coupled ADCIRC and SWAN. In *Ocean*
835 *Dynamics* (Vol. 65, Issue 5). <https://doi.org/10.1007/s10236-015-0820-3>
- 836 Taylor, A., & Glahn, B. (2008). Probabilistic guidance for hurricane storm surge. *Proc. 88th AMS*
837 *Annual Meeting*, 8. <https://ams.confex.com/ams/pdfpapers/132793.pdf>
- 838 UN OCHA. (2019). *Business Guide: Cyclone Idai. April*, 1–3.
839 <https://media.ifrc.org/ifrc/document/emergency-appeal-mozambique-cyclone-idai/>
- 840 Van Kampen, N. G., & Reinhardt, W. P. (1983). Stochastic Processes in Physics and Chemistry.
841 *Physics Today*, 36(2), 78–80. <https://doi.org/10.1063/1.2915501>
- 842 van Ormondt, M., Nederhoff, K., & Van Dongeren, A. (2020). Delft Dashboard: a quick setup tool for
843 hydrodynamic models. *Journal of Hydroinformatics*, 22(3), 510–527.
844 <https://doi.org/10.2166/hydro.2020.092>
- 845 Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound
846 flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5(12), 1093–
847 1097. <https://doi.org/10.1038/nclimate2736>
- 848 Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., & Heynert, K. (2013). The
849 Delft-FEWS flow forecasting system. *Environmental Modelling & Software*, 40, 65–77.
850 <https://doi.org/10.1016/j.envsoft.2012.07.010>



- 851 Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017).
852 Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water*
853 *Resources Research*, 53(9), 7968–7986. <https://doi.org/10.1002/2017WR020917>
- 854 Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C.,
855 Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical*
856 *Research Letters*, 44(11), 5844–5853. <https://doi.org/10.1002/2017GL072874>
- 857