

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy)

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Biolchi L.G., Unguendoli S., Bressan L., Giambastiani B.M.S., Valentini A. (2022). Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy). COASTAL ENGINEERING, 173, 1-19 [10.1016/j.coastaleng.2022.104081].

Availability:

This version is available at: <https://hdl.handle.net/11585/859978> since: 2024-05-14

Published:

DOI: <http://doi.org/10.1016/j.coastaleng.2022.104081>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy)

Luis Germano Biolchi^{a,b,c,d,*}, Silvia Unguendoli^d, Lidia Bressan^d,
Beatrice Maria Sole Giambastiani^a, Andrea Valentini^{a,d}

^a University of Bologna, Department of Biological, Geological, and Environmental Sciences (BiGeA), Ravenna, Italy

^b University of Algarve, Portugal

^c University of Cádiz, Spain

^d Regional Agency for Protection, Environment and Energy of Emilia-Romagna, Hydro-Meteo-Climate Service (Arpae-SIMC), Bologna, Italy

ARTICLE INFO

Keywords:

Early Warning System

XBeach

Morphodynamic model

Coastal-marine ensemble forecasting

Beach forecasting

Emilia-Romagna coast

Transnational multi-model ensemble (TMES)

ABSTRACT

During the last three decades, ensemble modelling has switched the focus from deterministic to probabilistic outcomes after its successful application in meteorological forecasting. This work involves the application of Ensemble Prediction System (EPS)-based results as forcing for a coastal EWS employing the morphodynamic model XBeach in a so-called (semi-)probabilistic way. First, calibration following the GLUE approach is performed for a profile in Cesenatico (Emilia-Romagna coast, Italy), while the (semi-)probabilistic system is implemented subsequently for two nearby locations. Ensemble mean and standard deviation from the Transnational Multi-Model Ensemble (TMES) forecasting system are combined in varied ways and used to force XBeach. A testing period of two months is analyzed (March and April 2020) together with the already operational deterministic implementation with one specific day of high sea conditions being used to assess the performance of the system. The deterministic results present higher outcome variability compared to the usage of the TMES mean and mean plus/minus one standard deviation (SD). Adding two SDs to the TMES mean results in higher variability than the deterministic approach. The (semi-)probabilistic system shows high potential as it provides more information on possible outcomes. However, its implementation has to be carefully designed as the application of the TMES mean plus SDs might result in false threshold exceedance and unproportionate responses.

1. Introduction

Among the 37 megacities (with 10 million inhabitants or more) distributed over the globe, 24 are located in coastal zones (Blumberg and Bruno, 2018) adequately symbolizing the worldwide increase of seaside communities in the last decades. Estimations indicate the trend will continue (e.g. Neumann et al., 2015) accompanied by rising sea levels and higher frequency and magnitude of coastal related hazards (Oppenheimer et al., 2019). Among the most recurrent perils, storm surges figure on the top of the list and can be defined as “oscillations of the water level in a coastal or inland body of water in the time range of a few minutes to a few days, resulting from forcing from atmospheric weather systems” (WMO, 2011). Loss and damage of infrastructure as well as casualties under more extreme conditions are regularly associated with these abnormally high-water levels.

On this basis, the development and implementation of Early Warning Systems (EWSs) are decisive as they allow for timely measures prior to the arrival of the flooding waters. According to Basher (2006), EWSs can be divided in four main elements: risk knowledge; monitoring and warning service; dissemination and communication; and response capability. Operational forecasting, substantiated by consolidated knowledge of the involved processes, is normally conducted by environmental and meteorological agencies globally mainly through numerical modelling. Forecasting outcomes supply decision-makers at local, regional, or national levels with relevant information and comprise an essential part of monitoring and warning procedures.

Still nowadays, many short to medium range operational forecasting systems follow a deterministic approach, meaning a single simulation used to produce a short-term prediction. For instance, in the northwest Adriatic sea, several environmental and civil protection agencies have

* Corresponding author. Viale Silvani, 6D, Bologna, Italy.

E-mail address: lgbiolchi@arpae.it (L.G. Biolchi).

developed and maintain storm surge operational forecasting systems based on hydrodynamic modelling mostly deterministically (e.g. Bajo and Umgiesser, 2010; Ferrarin et al., 2013; Mariani et al., 2015; Russo et al., 2013).

However, due to the highly variable, complex behavior of the atmosphere and the oceans (mostly nearshore processes), several sources of uncertainty arise which are not addressed by an individual run. Hence, in the early 1990s, probabilistic frameworks have been developed and firstly implemented for meteorological applications (Buizza, 2019) through Ensemble Prediction Systems (EPSs). They can be applied differently depending on the available structure, covered spatial scale, type of prediction, sources of uncertainties, and on the reliability of available observations (Cloke and Pappenberger, 2009). The most common ensemble generation mechanisms are the single system (based on the perturbation of initial conditions), the multi-model (based on the usage of results from different models), and the lagged averaged ensembles (results from the same model but from different runs) (Dietrich et al., 2008).

After meteorology, several other natural/Earth science fields began to implement EPS-based operational forecasting systems around the planet. Hydrology (e.g. Davolio et al., 2008; Jasper et al., 2002) and oceanography (e.g. Lenartz et al., 2010) are among the sciences that currently benefit from EPSs with the Northwest Adriatic Sea recently being the target of an operational multi-model ensemble system that combines the outputs of several regional hydrodynamic models (Ferrarin et al., 2020).

Even though calibrated and validated hydrodynamic models provide consistent information on total water levels and can be checked with predefined alert thresholds, these models do not consider morphological-hydrodynamical interactions. Thus, for sandy beaches where the geological/geomorphological characteristics strongly interact with hydrodynamic components, the application of morphodynamic models (e.g. XBeach - Roelvink et al. (2009)) better represent nearshore processes and tend to result in more reliable predictions.

For the Emilia-Romagna coastline (Northwest Adriatic Sea), the current deterministic beach forecast produced by the Hydro-Meteorology Service of the Regional Agency for Protection, Environment and Energy (Arpa-SIMC) uses outputs of hydrodynamic and wave forecasting systems to run XBeach and calculate maximum water levels for each time step. The results are then checked against predefined thresholds, that vary from natural and urbanized shores, and thoroughly used as a Decision Supporting System (DSS) by regional authorities on the decision-making process. The methodology applied is described in Harley et al. (2016).

Before the implementation of any numerical model, calibration is essential to correctly address the relevant processes strongly impacting the accuracy and skill of the outputs. Several methodologies are prone to be followed including a manual one-at-a-time testing of parameters to which the model is known to be sensitive for. However, in some cases the models have many parameterizations and this task turns into a time consuming practice. Moreover, in the aftermath it is difficult to understand how the interrelation of the parameters physically affect the final product and an implementation with parameter values coming from different runs might end up uncertain in many aspects. Recent applications of generalized uncertainty estimations through hundreds or even thousands of parameter combinations provide a final parameter set that has been tested and resulted in the best model performance (e.g. Simons et al., 2017a).

The objective of the present paper is to investigate the application of the recent multi-model ensemble (Ferrarin et al., 2020) results as forcing to an XBeach-based coastal EWS system. As the work of Ferrarin et al. (2020) does not include the further morphodynamic model application, the difference between deterministic and (semi-)probabilistic approaches at the very end of the modelling chain are yet to be fully understood. Additionally, this innovative operational suite that results in a morphodynamic model providing (semi-)probabilistic forecasts has not

been applied in the region (Adriatic basin) and possibly in the world (as far as the authors are aware).

Investigating the aforementioned uncertainties strongly involves identifying the best model setting for local applications. Hence, the GLUE methodology (explored in section 3) is followed using model parameters known to strongly affect local XBeach applications as means to achieve a trustworthy combination.

The structure of the paper comprehends the presentation of the study area in Section 2 followed by the methodology in Section 3. Results and discussion are shown in Section 4 while section 5 presents the conclusions.

2. Study area

The Emilia-Romagna coastline extends for 130 km between the Po Delta (Northern boundary) and the Cattolica town (Southern) (Harley et al., 2016). Tourism activities in the area are economically important mostly during summertime and enhance the necessity of structural and non-structural measures to mitigate storm surge damaging effects. Anthropogenic structures are found along the coast, protecting more than 50% of the shoreline in specific locations (Perini et al., 2008). Moreover, winter dunes are built parallel to the coast (around September–October) to protect beach establishments and nearby houses during the stormy season after which the dunes are destroyed and the sand restored to its “normal” placement. The study area is shown in Fig. 1.

Two wind patterns over the Adriatic sea during colder periods are the most recurrent in the region: *Bora* and *Sirocco* (Pandžić and Likso, 2005). *Sirocco* conditions are observed when southerly winds originating from low-pressure synoptic systems in Northern Africa advance towards the boreal Adriatic. As they blow over the longer axis of the Adriatic Sea, *Sirocco* produces long waves and piles up water in the north of the basin. Most devastating incoming waves tend to arrive from the Northeast, caused by *Bora* storms, as strong winds funnel and increase their velocity downhill in the mountainous Croatian coast forming high, steep, relatively short period waves over the shorter Adriatic axis. More energetic waves are related to the latter, while higher surges are associated with *Sirocco* (Harley et al., 2016).

Local hydrodynamic conditions involve a microtidal regime (varying between 30 and 80 cm between neap and spring tides, respectively) with strong diurnal and semidiurnal components. Waves normally arrive from the East with 91% of the cases below 1.25m as a consequence of a restricted fetch (Armaroli et al., 2012).

In the region, profiles in three coastal localities covered by the aforementioned regional coastal EWS have been used in the present work: Cesenatico for calibration, and Marina Romea and Marina di Ravenna for the (semi-)probabilistic implementation (Fig. 1).

2.1. The operational Early Warning System in Emilia-Romagna (Italy)

Hierarchically coming first in the operational chain is the Numerical Weather Prediction (NWP) model which follows one of the Consortium for Small-Scale Modeling (COSMO) implementations in Italy (Stappeler et al., 2003) – referred as COSMO-I. Two domains are currently addressed by COSMO-I: one covering the whole Mediterranean with a horizontal resolution of 5 km (COSMO-5M) and one covering only the Italian territory with a higher resolution of 2.2 km (COSMO-2I). COSMO outputs are used as atmospheric forcing for the sea-state and for the oceanographic forecasting systems operationally implemented and maintained by Arpa-SIMC.

Simulating Waves Nearshore (SWAN) is the third-generation wave action model, which focuses on wave propagation in coastal waters (Booij et al., 1999; Ris et al., 1999), used for sea-state forecasting. A coarser grid over the whole Mediterranean (horizontal resolution of about 25 km–1/4° on Earth spherical coordinates) with two nested, finer ones around the Italian territory (8 km) and in specific regions (0.8 km), like Emilia-Romagna, comprise this part of the chain. More information

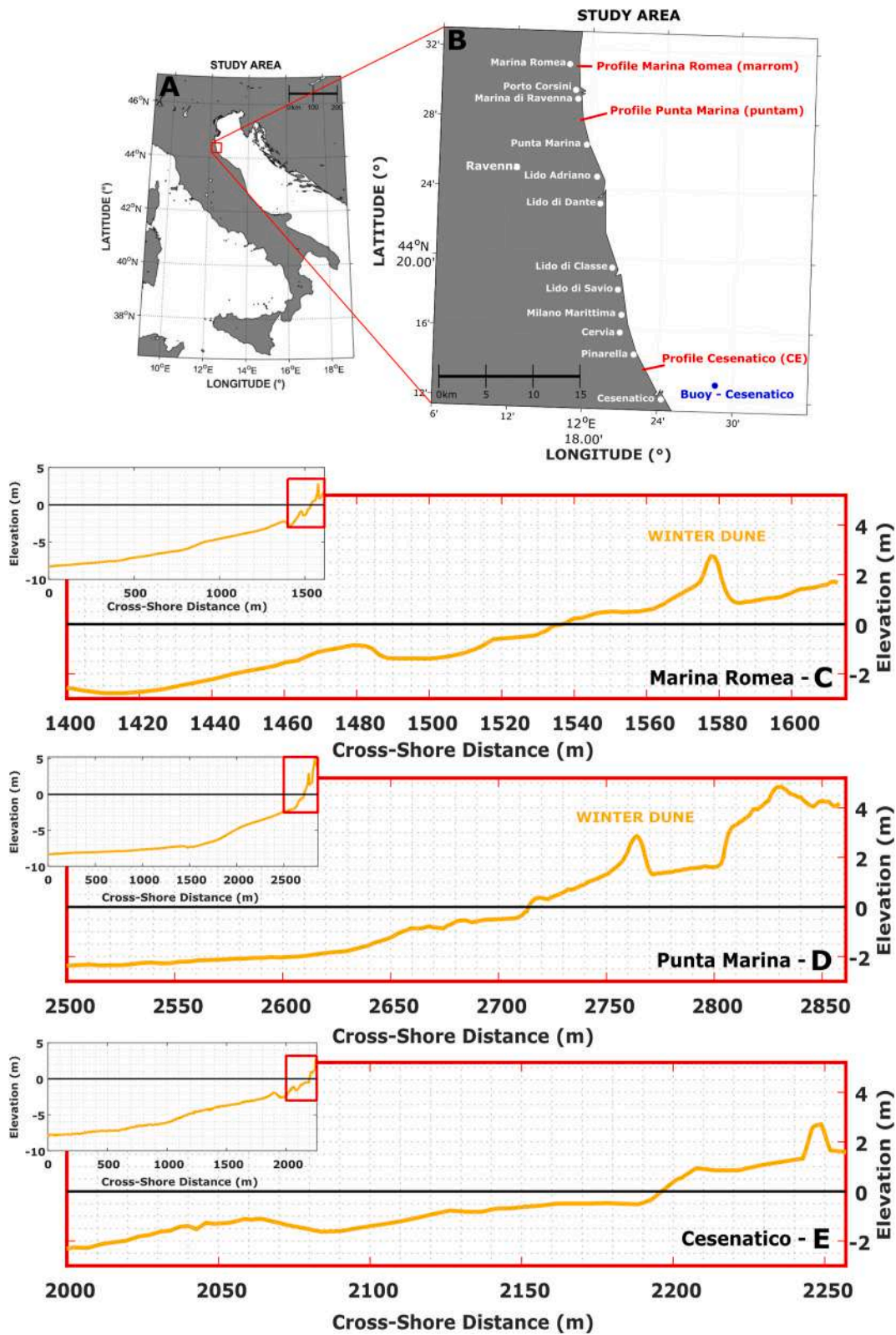


Fig. 1. A) Emilia-Romagna coastline location. B) *in-situ* measured cross-shore transects and Cesenatico wave buoy location. C, D, and E) Gridded profiles of Marina Romea, Punta Marina and Cesenatico, respectively (entire cross-shore profile on the left, and submerged and emerged beach profile zoomed in). Cesenatico topobathymetric data was collected in 2016 as described in Unguendoli (2018). Bathymetric datasets for Punta Marina and Marina Romea are from surveys conducted in July 2019 while the subaerial measurements were conducted in early 2020.

about the sea-state forecasting system (called MEDITARE) can be found in the work of Valentini et al. (2007).

Adriatic Sea hydrodynamic forecasting follows an Italian implementation of the Regional Ocean Modeling System (ROMS) (Shchepetkin and McWilliams, 2005) referred to as *AdriaROMS* (Chiggiato and Oddo, 2008). The latest *AdriaROMS* version has been implemented and developed by Arpaè in collaboration with the Marche Polytechnic University (Russo et al., 2013). *AdriaROMS* runs in a curvilinear orthogonal grid regularly spaced in 2 km, with 20 vertical terrain-following levels. The variables used as atmospheric forcing are from the COSMO-5M. Open sea boundary conditions (temperature, salinity, and current velocity) are obtained from the Copernicus CMEMS Mediterranean Ocean Model (Clementi et al., 2019). Additionally, four tidal components are derived from the Oregon State University (OSU) model (Egbert and Erofeeva, 2002) and used in the domain.

Dedicated sub-models use the hydrodynamic and wave outputs as boundary and/or forcing conditions. This is the case of *XBeach*, which has been operationally implemented together with the University of Ferrara and the Geological, Seismic and Soil Survey of Emilia-Romagna during the EU-FP7 MICORE project and is maintained by Arpaè. By the time this work was developed, the system covered eight sites for 22 cross-shore profiles, providing 72-h deterministic forecasts displayed in a WebGIS platform.

3. Methodology

Topo-bathymetric data collection, the *XBeach* morphodynamic model, calibration, and the usage of the multi-model ensemble variables were part of the implementation procedures and are presented in the next subsections.

3.1. Topographic and bathymetric data

Topo-bathymetric data for the Cesenatico transect was collected in 2016, as reported by Unguendoli (2018), and it has been used in the present work for calibration purposes only. For Marina Romea and Marina di Ravenna, where the (semi-)probabilistic approach was implemented, topographic surveys were carried out in February 2020 using a real-time kinematic differential global positioning system (RTK DGPS), while bathymetric data was collected in July 2019 with a single-beam bathymetric device. The elevation measurements were converted into orthometric height based on the geoid undulation reported in the monographs of Arpaè Coastal Geodesic Network for topographic and bathymetric monitoring (benchmark PCG0100 for Marina Romea profile, and benchmark SAPC0700 for Punta Marina profile).

The decision to combine a topographic profile from a different period with respect to the bathymetric measurements arose from the need to have an updated subaerial representation. As the bathymetric profile had not been highly altered and it could be combined adequately with the new topographic measurements, it has been decided to keep the most up to date values. Furthermore, the application of the EWS approach presented here has in the backshore and foreshore zones fundamental components in terms of maximum vertical excursion of water and how critical the situation might be for nearby infrastructure/natural areas.

The *in-situ* measured data was used to generate mono-dimensional computational grids with a resolution of about 20m (offshore) decreasing to a minimum of 1m (onshore). Varying the grid resolution allows for a decrease in computational effort, optimizing the simulations. Figure 1(c-e) shows the final grids used in the simulations.

It is important to emphasize that *XBeach* has its abilities limited to short-term responses to specific events. One of the contributing factors is that most morphodynamic models misrepresent post-storm recovery both in timescale and processes themselves (van Rijn et al., 2003). For this reason, the (semi-)probabilistic implementation follows the

deterministic as the forecasts were conducted using always the same initial topo-bathymetric conditions.

Even though this approach has its limitations, the obtention of daily topo-bathymetric profiles in order to feed the system is still extremely difficult to accomplish. Presently, the implemented deterministic EWS runs with summer and winter profiles following the building up of the winter dunes as previously explained. Besides waiting for the winter dunes construction/destruction (for summer activities), the measurement campaigns also depend on funding availability, sea-state conditions, authorization from the responsible authorities, and several logistic and staff constraints.

3.2. *XBeach*

XBeach is based on the approach proposed by Sallenger (2000) and it was first developed to model the impacts of hurricanes on sandy barrier islands. In order to understand the behavior of sandy shorelines hit by storms, the model attempts to accurately reproduce the morphodynamic processes during intense events combining subaerial and subaqueous processes (Roelvink et al., 2015).

Many improvements in the model have been observed since its first version and applications have increased substantially as the model capabilities allow its use in different coastal settings. As examples of different applications, it is possible to include: hurricane impacts on barrier islands (e.g. Lindemer et al., 2010; van der Lugt et al., 2019; McCall et al., 2010; Smallegan et al., 2016; van Verseveld et al., 2015); storms on sandy beaches (e.g. Armaroli et al., 2013; Harley et al., 2011, 2016; Harley and Ciavola, 2013; Pender and Karunarathna, 2013; Unguendoli, 2018); mega cusp formation (e.g. Orzech et al., 2011); tidal inlet evolution (e.g. Pacheco et al., 2011); macro-tidal areas with large dune fields (e.g. Dissanayake et al., 2014, 2015); synthetic storms on sandy beaches (e.g. Schambach et al., 2018); reef environments (e.g. Lashley et al., 2018; Ortiz and Ashton, 2019; Osorio-Cano et al., 2019); and wave-vegetation interaction (e.g. Phan et al., 2014; van Rooijen et al., 2016).

XBeachX, implemented in surfbeat mode (non-stationary), was the version used in this work. This mode is indicated when focus on swash zone processes is needed as it solves the short-wave envelope on the scale of wave groups by resolving wave- and wind-driven (when applied) currents, swash runup and rundown, and infragravity waves using non-linear shallow water equations. For details regarding the model equations and general technical aspects, the reader can refer to Roelvink et al. (2009), Roelvink et al. (2015) as well as the *XBeachX* manual (Deltares, 2018) and references thereafter.

3.3. *XBeach* calibration

Model calibration followed the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) as applied to *XBeach* by Simmons et al. (2017a). This approach is divided in four main steps: (1) selection of the parameters (as well as their ranges) to be tested; (2) generation of a large number of combinations (e.g. 10,000) using the chosen parameters, with (3) the model being run for each of the combinations. Finally, (4) the likelihood score for each resulting simulation was calculated.

Data availability and similar geomorphological characteristics relative to Marina Romea and Marina di Ravenna drove the selection of Cesenatico to be the calibration site. A storm that hit the region's coastline in 2016 has been picked to calibrate the profile with the incoming wave and sea level conditions presented in the following subsection. After, each of the calibration steps are described.

3.3.1. Storm data

On the 28th of February 2016, strong weather resulted in higher wave height and period conditions, reaching the peak in the evening between the 28th and the 29th (Fig. 2). During the storm, the maximum

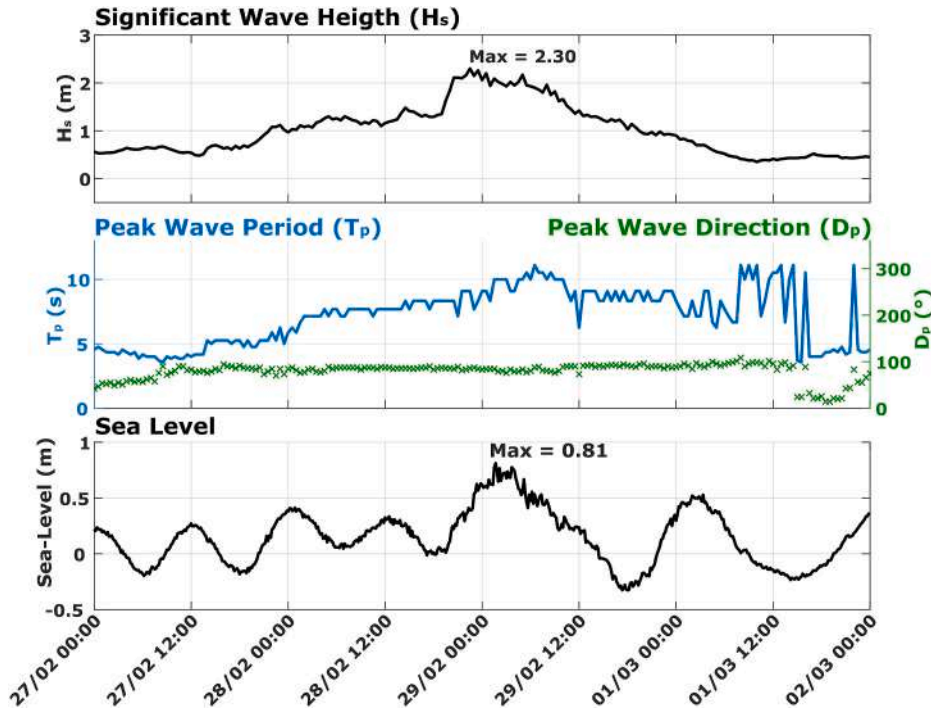


Fig. 2. Measured H_s , T_p , D_p , and sea-level during the storm that hit the Emilia-Romagna coast at the end of February/beginning of March of 2016. Wave data was measured every 30 min by a buoy deployed just offshore of Cesenatico (buoy location shown in Fig. 1B). Sea level data was obtained every 10 min by a tide gauge located in the Municipality of Rimini.

observed wave height was 2.30m and the maximum sea-level reached 0.81m. Between 8.30pm of the 28th until around 5.00am of the 29th, waves higher than 2 m were recorded concurrently to surge elevations higher than 0.5m.

Peak wave direction (D_p) remained from the Eastern quadrant varying between 78.8° and 88.6° (with a mean value of 83.6°). Peak wave period (T_p) varied between 7.14s and 10.0s (mean value of 9.0s). Wave and sea level measurements presented in this subsection were collected by the Cesenatico wave buoy (Fig. 1) and by a tide gauge installed in Rimini, respectively, and subsequently used in the calibration.

3.3.2. Parameter selection and generation of combinations

As an initial step, previous XBeach applications for the Emilia-Romagna coastal areas were assessed to verify the parameters to which the model had shown a higher sensitivity (Table 1). Values associated with better performances specifically for Cesenatico were given preference, resulting in the selection of *facua*, *gamma*, *fw*, *bedfriccoef*, and *smax* for the first GLUE attempt (Table 2).

Facua influences sediment transport varying wave asymmetry and skewness; the higher the value, the more sediment is transported onshore. *Gamma* is the breaker parameter in the wave formulations (Baldock or Roelvink) with lower values inducing earlier breaking. *Fw* and *bedfriccoef* refer to the short-wave and bed-friction coefficients, respectively, with lower values inducing more friction and less erosion for the latter. *Smax* refers to the maximum Shields parameter for equilibrium sediment concentration (Deltares, 2018).

Based on the first GLUE results (10,000 simulations) using the Tested Range (1st GLUE), a second set of 10,000 simulations was conducted with narrower parameter ranges. Additionally, *smax* was replaced by *dryslp*, which refers to the critical bed slope for dry areas, in the Tested Range (2nd GLUE) as shown in Table 2. The replacement occurred as XBeach showed a very low sensitivity to *smax* during the first GLUE application (Fig. 5E).

3.3.3. Likelihood score attribution

The Brier Skill Score (BSS), as proposed by van Rijn et al. (2003), was used to quantify the model skill for each run. The BSS is given by:

$$BSS = 1 - \frac{\sum(|z_0 - z_m|)^2}{\sum(|z_0 - z_b|)^2}, \quad (1)$$

where z_0 is the post-storm measured elevation, z_m is the final modeled elevation, and z_b is the pre-storm measured elevation. BSS values reflect the model skill on reproducing morphological evolution and can be divided into five categories as shown in Table 3.

Two BSS values were calculated for each run: one for the whole profile and one for the subaerial portion only ($z_b > 0$). The subaerial BSS calculation followed the results obtained by Armaroli et al. (2013) in which the authors showed differences between the BSS for the whole profile and for the subaerial beach alone associated to varying model performances on the subaerial and subaqueous domains.

Following the GLUE steps, the definition of a BSS threshold is needed to divide the simulations in behavioral and non-behavioral. The latter refers to the simulations in which the parameter set shows "no skill when used in the model" (Simmons et al., 2017a). Hence, the modeler has to decide on a threshold above which the simulations better agree with the *in-situ* topo-bathymetric measurements considering the purpose of the intended modelling application.

Meaningful classes are already provided by the BSS scale (Table 3) and can, therefore, be used to delimit the threshold. For instance, if one wants simulations with a minimum of reasonable qualification, the threshold can be set as 0.3. Values higher than the threshold imply that a given simulation is behavioral, while values lower than the threshold determine a non-behavioral run and the likelihood value is set to zero.

As seen later in the results, two thresholds were considered in the present work: 0.3 for the whole profile and 0.7 for the subaerial portion alone. This reflects general XBeach performances differing between the underwater and emerged areas as well as our own preliminary results showing a much better skill for the subaerial beach alone.

After the skill assessment and the definition of behavioral and non-

Table 1

XBeach applications assessed for the Emilia-Romagna coastline. The table also depicts the specific location of the application, type of calibration performed, parameters tested, as well as their ranges, and the total number of runs. Refer to the text for parameter descriptions.

Publication	Location	Calibration Procedure	Parameter Tested	Parameter Range	Number of Runs
Harley et al. (2011)	Lido di Dante/Lido di Classe	One at a time	<i>dryslp</i> <i>wetslp</i> <i>gammax</i>	0.5 and 2.0 0.1 and 0.5 0.5 and 5.0	6
Armaroli et al. (2013)	Lido di Classe/Bevano Area	One at a time	<i>dryslp</i> <i>wetslp</i> <i>gammax</i>	0.5 and 2.0 0.1 and 0.5 0.5 and 5.0	56
Simmons et al. (2015)	Lido di Classe	GLUE	<i>eps</i> <i>facua</i> <i>gamma</i> <i>gammax</i> <i>smax</i> <i>wetslp</i>	0.001 to 0.1 0 to 1.0 0.4 to 0.9 0.4 to 5.0 -1.0 to 3.0 0.1 to 1.0	15,000
Harley et al. (2016)	Lido di Classe	One at a time	<i>smax</i> <i>gamma</i> <i>gammax</i> <i>eps</i> <i>wetslp</i> <i>facua</i>	0.8 0.42 1.5 0.01 and 0.1 0.5 0.15	352 (32 model runs for 11 profiles)
Simmons et al. (2017a)	Lido di Classe	GLUE	<i>eps</i> <i>facua</i> <i>gamma</i> <i>gammax</i> <i>smax</i> <i>wetslp</i> <i>cmax</i> <i>smax</i> <i>Lws</i> <i>facua</i> <i>break</i>	0.001 to 0.1 0 to 1.0 0.4 to 0.9 0.4 to 5.0 -1.0 to 3.0 0.1 to 1.0 0 to 1.0 - 1.0 to 3.0 0 to 1.0 0 to 1.0 roelvinck1, Baldock, roelvinck2, roelvinck_daly, janssen	330,000 (twice 15,000 for each of 11 profiles)
Unguendoli (2018)	Cesenatico	One at a time	<i>gamma</i> <i>turb</i> <i>fw</i> <i>delta</i> <i>eps</i> <i>umin</i> <i>bedfriccoef</i> <i>wetslp</i> <i>dryslp</i>	0.4 to 0.9 none, wave_average, bore_averaged 0 to 1.0 0 to 1.0 0.001 to 0.1 0 to 0.2 3.5E-5 to 0.9 0.1 to 1.0 0.1 to 2.0	not specified in the work

Table 2

Parameters tested with the total range covered in the first and second GLUE applications. The table also shows the default XBeach values for each tested parameter.

Parameter Tested	Tested Range (1st GLUE)	Tested Range (2nd GLUE)	Default Value (Default Range)
<i>facua</i>	0.1 to 0.5	0.1 to 0.3	0.1 (0.0–1.0)
<i>gamma</i>	0.2 to 0.5	0.3 to 0.5	0.55 (0.4–0.9)
<i>fw</i>	0.1 to 0.7	0.1 to 0.5	0 (0–1.0)
<i>bedfriccoef</i>	30 to 50	40 to 55	55 (-) ^a
<i>smax</i>	-0.5 to 1.0	-	-1.0 (- 1.0 to 3.0)
<i>dryslp</i>	-	0.8 to 1.5	1.0 (0.1–2.0)

^aNo minimum or maximum values for the Chezy coefficient (*bedfriccoef*) have been found.

Table 3

BSS scores and model quality.

Qualification	BSS Value
Excellent	1.0–0.8
Good	0.8–0.6
Reasonable/Fair	0.6–0.3
Poor	0.3–0.0
Bad	<0.0

behavioral simulations, the BSS for each run was divided by the total sum of the BSSs, providing a likelihood measure as follows:

$$L_{BSS} = \frac{BSS_i}{\sum_{i=1}^n BSS_i} \quad (2)$$

where n is the total number of simulations. With the likelihood values and the distinction between behavioral and non-behavioral runs, it was possible to proceed with parameter optimization, sensitivity, and uncertainty analyses as shown next.

3.3.4. Sensitivity analysis, parameter optimization, and uncertainty estimation

For parameter optimization and sensitivity analysis, the range of values for a given parameter was divided in bins. The L_{BSS} of every simulation was added to the nearest bin. Each resulting bin contained the sum of the weighted likelihood values around it, with the total sum for all the bins being equal to the total likelihood sum (equals to one). Both parameter optimization and sensitivity analyses were performed for each tested parameter.

For the sensitivity analysis, the weighted density was transformed into a cumulative likelihood curve for both the behavioral and non-behavioral simulations. The Kolmogorov-Smirnov (K-S) D statistic (D_{stat}) (Simmons et al., 2017a; Thorndahl et al., 2008) was then calculated based on the maximum vertical distance between the behavioral

and non-behavioral cumulative curves, providing a value between zero and one. The larger the K-S D statistic, the higher the model sensitivity to that parameter.

As a final step and to provide a second model performance evaluation, the subaerial eroded volume (SEV) was calculated as the difference between the subaerial profiles before and after the storm (or simulation). Assuming a 1m profile width allowed for the extrapolation from m^2 to m^3/m .

3.4. Coastal storm impact index (SII)

Definitions for the Safe Corridor Width (SCW) and Building-Waterline Distance (BWC) must be introduced as they are the SIIs used by Arpae-SIMC. As defined by Harley et al. (2016) the SCW is “a measure of the amount of dry beach available between the dune foot and waterline for safe passage by users”, while the BWD is “the amount of dry beach available between the seaward edge of a building and the model-derived waterline”. SCW is used in natural areas with frontal dunes, while the BWC implementation covers anthropized beaches (with man made structures such as beach bars). The SIIs are calculated as

following:

$$SCW(t) = X_{df} - X_{wl}(t), \quad (3)$$

$$BWD(t) = X_b - X_{wl}(t), \quad (4)$$

where X_{df} is the surveyed cross-shore position of the dune foot, X_{wl} is the modeled cross-shore position of the water line, and X_b is the cross-shore position of the seaward edge of the building. For each time step, the respective SII is calculated based on the maximum water level forecasted value and the reference building/dune foot of the profile.

3.5. Coastal EWS-EPS implementation

As part of the EU ADRION I-STORMS (Integrated Sea Storm Management Strategies) project (<https://istorms.adrioninterreg.eu/>), the outputs of the Transnational Multi-Model Ensemble System (TMES) were made available as one of the six Integrated Web System (IWS) components (Ferrarin et al., 2020). Currently, the TMES combines five wave- and six sea level-forecasting systems interpolated to a common regular latitude-longitude grid with a resolution of 0.02° for the Adriatic

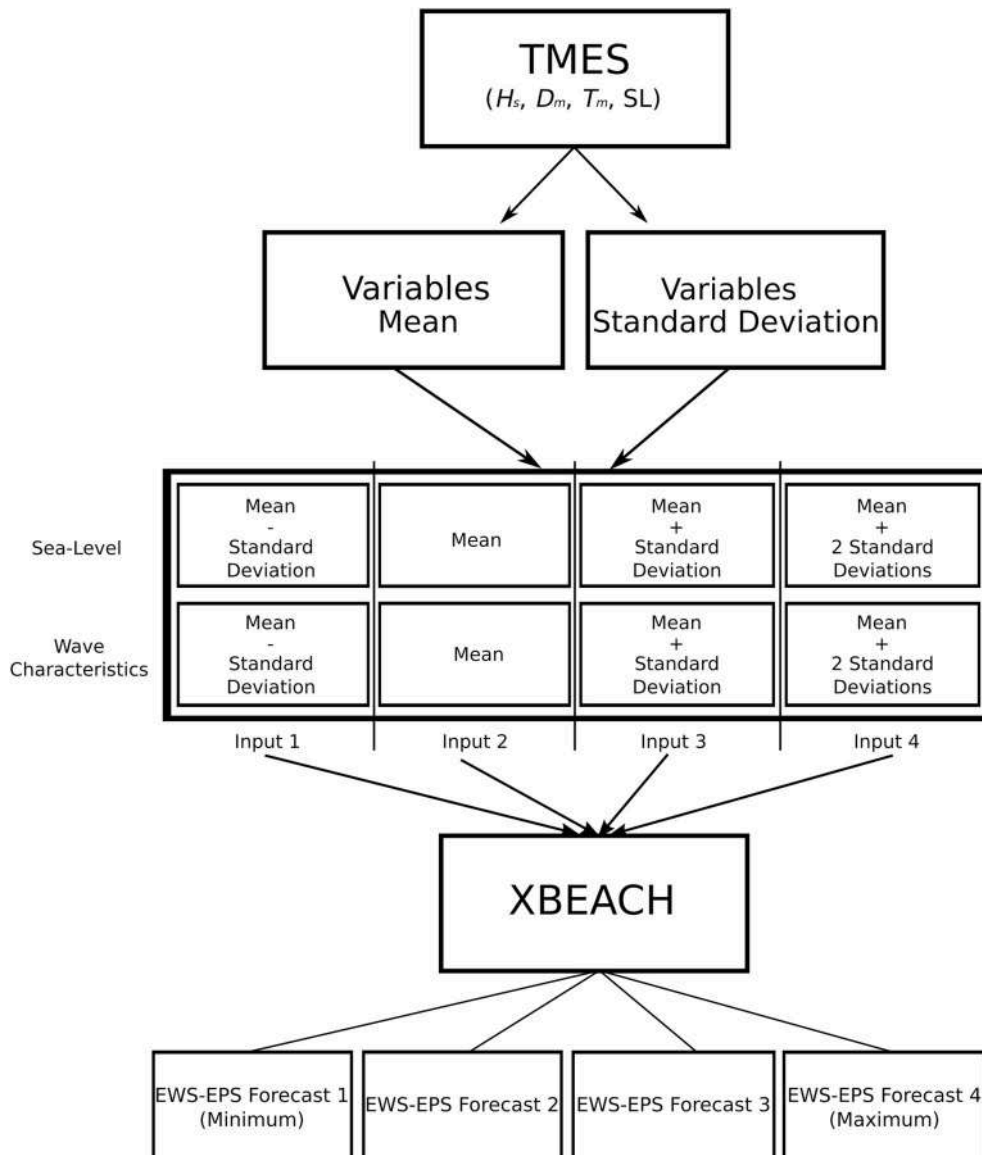


Fig. 3. Scheme of TMES forecast retrieval converted to files that are subsequently used as boundary conditions to run XBeach. Each output consists of two files: one for the wave parameters and one for sea level.

Sea. Based on the different forecasts, +48 h ensemble mean and standard deviation (SD) are calculated with the output variation estimated by adding/subtracting the SD to/from the ensemble mean.

In the present work, four TMES combinations were used as boundary conditions to the operational XBeach as schematized in Fig. 3. The (semi-)probabilistic implementation comprehended the TMES mean, the TMES mean minus one SD, plus one SD and plus two SDs. Whenever mean plus or minus a SD is mentioned hereinafter, it implies that the addition or subtraction of SDs has been performed to the sea level and the three wave parameters used: significant wave height (H_s), mean wave period (T_m), and mean wave direction (D_m).

With the XBeach outputs forced by TMES it was possible to assess minimum and maximum forecasted water levels for each time step. For both profiles where this approach was implemented (Marina di Ravenna and Marina Romea), the BWD estimated the distance between the maximum water level and the local reference building following the already operational deterministic system.

The already implemented deterministic approach (that follows the hierarchical framework previously presented and provides +72 h of forecast) was analyzed together with the (semi-)probabilistic results as means to have a better picture on the improvements given by the new framework. Two months of daily forecasts were evaluated and are presented in this study: March and April 2020.

4. Results and discussion

In this section the results are divided in two main branches: the calibration outputs and the results of the (semi-)probabilistic implementation using TMES as forcing.

4.1. XBeach calibration

An initial analysis of the first GLUE application results indicated a better model performance for the subaerial portion of the profile relative to when the whole profile was analyzed. Hence and as forecasting applications require high skill modelling practices, BSSs values of 0.3 and 0.7 were chosen to be the behavioral/non-behavioral thresholds for the whole profile and for the subaerial part only, respectively. The former (0.3) determines that a minimum of reasonable qualification is necessary for a run to be considered behavioral, while the latter (0.7) lies within the good qualification range (Table 3).

In terms of general results, for the first GLUE application most of the runs (64.73%) for the whole profile remained under the 0.3 BSS threshold indicating a poor model performance (Fig. 4 - top left panel). When only the subaerial portion was considered (0.7 BSS threshold), the non-behavioral runs represented 62.23% of the total (Fig. 4 - bottom left panel). Still in the first GLUE application, the mean BSS was 0.25 with a SD of 0.10, with the best model performance reaching 0.44, while the worst reached 0.03. The scenario improved when only the subaerial profile was considered, with a mean BSS of 0.58 and values fluctuating between 0.28 and 0.88. The mean value indicated the model performance of being reasonable but a higher SD (0.22) was observed when compared to the whole profile.

For the second GLUE application, with narrower parameter ranges and *dryslp* instead of *smax*, the whole profile mean BSS increased to 0.37 (Fig. 4 - top right panel). Still for the whole profile, a decrease in the SD was observed (0.04) with the best model performance score being 0.44 and the worst performance valuing 0.03. Similar to the first GLUE application, the mean BSS value for the subaerial beach overcame the BSS for the whole profile, reaching 0.75 (Fig. 4 - bottom right panel). The SD valued 0.09 with the best and worst performances scoring 0.87 and 0.24, respectively.

For the second GLUE application and keeping the thresholds, a substantial decrease in the number of non-behavioral simulations was observed, representing 5.12% for the whole profile and 15.54% for the subaerial portion. Accordingly, the number of behavioral simulations

drastically increased, reaching 94.88% for the whole profile and 84.46% for the emerged beach.

4.1.1. Sensitivity analysis

D_{stat} results calculated as the maximum vertical distance between behavioral and non-behavioral curves were used to assess model sensitivity to each tested parameter. The results are presented in Fig. 5. Even though the main objective of the cumulative likelihood analysis was not parameter optimization, higher behavioral curve slopes provided a first idea around which value better model performances were observed.

The *facua* parameter presented higher D_{stat} results in the first GLUE application valuing 0.44 and 0.43 for the whole profile and the emerged beach, respectively (Fig. 5A). Still in the first GLUE, higher slopes for *facua* values between 0.1 and 0.3 indicated better model performances. The non-behavioral runs presented higher slopes for values between 0.3 and 0.5 (poorer model performance within this range). A very low model sensitivity was observed in the second GLUE application with *facua* ranging between 0.1 and 0.3 (Fig. 5B). Table 4 summarizes the sensitivity analysis results.

Γ D_{stat} results reached 0.34 (whole profile) and 0.32 (subaerial portion) for the first GLUE (Fig. 5C). Higher behavioral curve slopes between 0.35 and 0.5 indicated better model performances, while poorer performances occurred when Γ values ranged between 0.2 and 0.35. Similarly to *facua*, model sensitivity decreased considerably for the second GLUE application (Fig. 5D), reaching 0.22 for the sub-aerial beach and 0.18 for the whole profile.

Model sensitivity to *fw* resulting from the first GLUE (Fig. 5E) reached 0.17 for the whole profile, while the subaerial beach resulted in a D_{stat} 0.04. A considerable increase in D_{stat} results was observed in the second GLUE application (Fig. 5F), reaching 0.63 and 0.61 for the whole and subaerial profiles, respectively.

For the whole profile, *bedfriccoef* D_{stat} values of 0.33 and 0.07 resulted from the first (Fig. 5G) and second GLUE approaches (Fig. 5H), respectively, while the subaerial beach resulted in D_{stat} values of 0.34 and 0.17. In general, a higher variation of D_{stat} values was observed in the whole profile analysis. As in the first GLUE application the model sensitivity to *smax* was very low (Fig. 5I), *dryslp* was chosen to replace the previous parameter (Fig. 5J). However, the overall sensitivity to both parameters was similar, with D_{stat} results below 0.05.

4.1.2. Parameter optimization

The results presented in Fig. 6 show the parameter values associated with the best subaerial (red dashed vertical lines) and the best whole profile (blue dashed vertical lines) model performances. Also, the higher weighted densities represent the parameter range with which the model has achieved better results.

For *facua*, better model performances were achieved with values between 0.1 and 0.3 in the first GLUE application (Fig. 6A), while a more evenly distributed histogram was seen for the second GLUE (Fig. 6B) with a slight density increment around 0.2. The best whole domain model performance occurred with a *facua* value of 0.13, while the best subaerial result happened with *facua* equals 0.11 for the first GLUE. The second GLUE application showed *facua* values of 0.11 and 0.20 related to the best simulations for the whole and subaerial profiles, respectively.

Γ weighted histograms for the first GLUE application show a better model performance for values between 0.3 and 0.5 (with a peak between 0.44 and 0.47), with the best whole profile and subaerial beach BSS performances being associated with Γ values of 0.48 and 0.41, respectively (Fig. 6C). The second GLUE application resulted in a more evenly distributed histogram similarly to *facua*. Γ values of 0.37 and 0.38 (Fig. 6D) were associated with the best BSS model performances for the whole and subaerial profiles, respectively.

The parameter *fw* in the first GLUE application had higher densities distributed to lower values for the whole profile with the best model performance being associated with a value of 0.16 (Fig. 6E). Considering only the subaerial beach, the best model performance was with an *fw* of

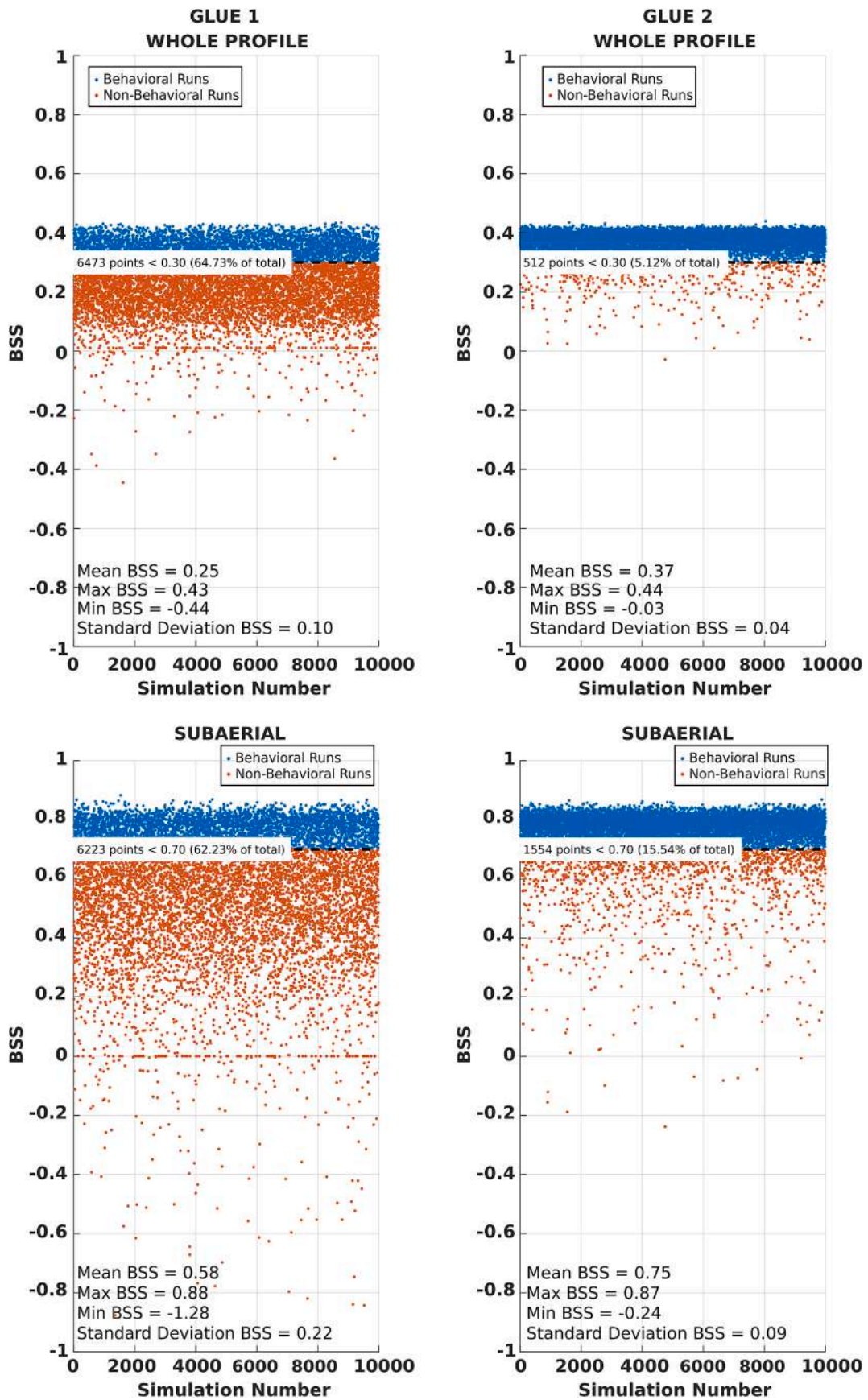


Fig. 4. Calculated BSS for every simulation. The panels on the left side show the results obtained during the first GLUE application, while the panels on the right present the second application results.

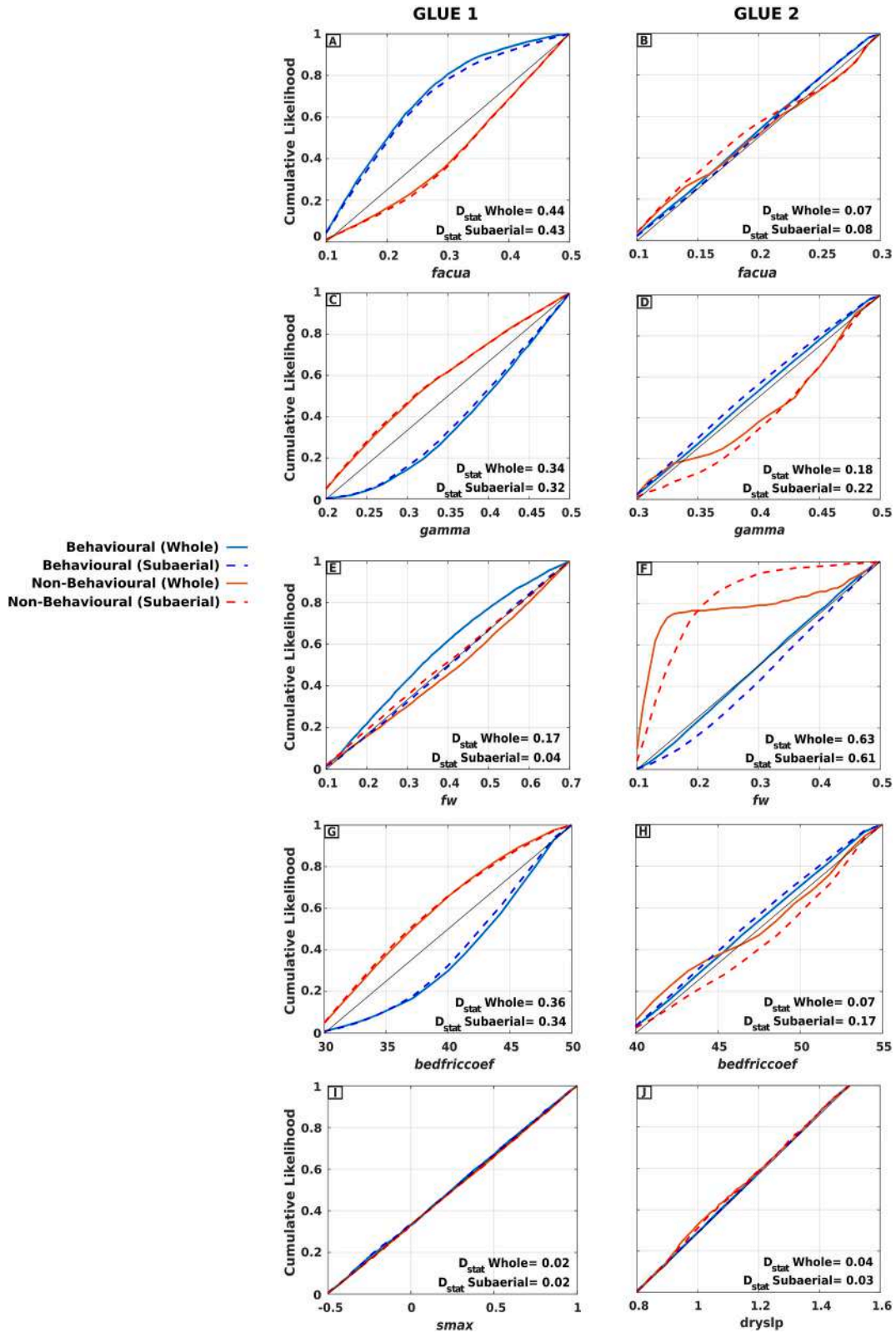


Fig. 5. Cumulative Likelihood for each parameter tested during the first (left) and second (right) GLUE applications. The indicated D_{stat} refers to the Kolmogorov-Smirnov D statistic used to quantify the model sensitivity to that parameter. The full and dashed lines represent the results for the whole profile and only for the subaerial section, respectively. Subfigures A) and B) refer to *facua*, C) and D) to *gamma*, E) and F) to *fw*, G) and H) to *bedfriccoef*, and I) and J) to *smax* and *dryslp*, respectively.

Table 4

Summary of model sensitivity based on D_{stat} for both GLUE applications subdivided in beach compartments.

	facua	gamma	fw	bedfriccoef	smax	dryslp
GLUE 1						
Whole Profile	0.44	0.34	0.17	0.33	0.02	–
Subaerial	0.43	0.32	0.04	0.34	0.02	–
GLUE 2						
Whole Profile	0.07	0.18	0.63	0.07	–	0.04
Subaerial	0.08	0.22	0.61	0.17	–	0.03

0.20 and a small shift in general behavior: higher density values associated to slightly higher fw values. For the second GLUE application (Fig. 6F), higher densities were observed with fw ranging between 0.26 and 0.5, while the best subaerial performance was related to a fw of 0.29. A value of 0.18 was associated with the best whole profile performance and a higher density concentration was observed between 0.26 and 0.34.

Bedfriccoef results for the first GLUE approach (Fig. 6G) indicated a trend of higher densities associated with increasing values, reaching the peak between 48 and 50. Although the general trend indicated higher densities increasing with higher parameter values, the best whole profile simulation was associated with a *bedfriccoef* value of 31.79 and the best subaerial run happened with a *bedfriccoef* of 31.37. For the second GLUE application (Fig. 6H), *bedfriccoef* values of 40.13 and 50.03 were associated with the best run for the whole and subaerial profiles, respectively. The results of the second approach also show a more evenly distributed histogram with higher densities concentrated between 43 and 47.

Smax and *dryslp* histograms show an evenly distribution of densities across the whole range of tested values. In the first approach, the best performance for the whole profile happened with a *smax* value of 0.18 (Fig. 6I), while the best behavioral emerged beach evaluation happened with a value of 0.03. After *smax* was replaced by *dryslp* for the second GLUE application, the best whole and emerged profile simulations occurred with values of 0.97 and 1.15 (Fig. 6J), respectively. A summary of the results is shown in Table 5.

4.1.3. Uncertainty analysis

In order to assess the uncertainties related to the implemented calibration approach, the pre- and post-storm observed profiles were plotted together with the best-modeled results. Additionally, weighted density histograms were built using the behavioral runs to analyze the model performance relative to the SEVs and are all shown in Fig. 7.

The final observed SEV, based on the pre- and post-storm measured topo-bathymetric data, was $2.08\text{m}^3/\text{m}$. In terms of profile alterations caused by the storm, the berm erosion characterized the main morphological change in the subaerial profile (Fig. 7A, 7B, 7C, and 7D). Still in the emerged beach, it was possible to observe sediment accumulation near the dune toe, represented by an increase in elevation on the post-storm measured profile (black dashed lines). Underwater offshore bar migration occurred in the subaqueous profile (Fig. 7B and 7D between 2000 and 2150m cross-shore distance).

Using the default model configuration (green dashed lines in Fig. 7A and 7C), part of the frontal dune was eroded following a general rectilinearization trend for the entire profile with a resulting SEV of $27.45\text{m}^3/\text{m}$. Hence, it is important to emphasize the necessity of thorough calibration to avoid erosion overestimation.

The simulation that resulted in the best subaerial BSS for the first GLUE application (blue dashed line in Fig. 7A - 0.88 excellent performance) accurately represented the berm erosion. However, a SEV overestimation occurred, reaching $3.70\text{m}^3/\text{m}$ against the observed $2.08\text{m}^3/\text{m}$. The simulation that better reproduced the SEV (red dashed line in Fig. 7A), reaching the exact $2.08\text{m}^3/\text{m}$, scored 0.71 in the BSS evaluation. It is possible to see in Fig. 7A that the best SEV simulation did not fully represent the berm erosion, counter-balancing the misfit relative to sediment accumulation near the dune toe. In the underwater

domain, the bar migration was not accurately reproduced with the modeled results following the pre-storm measurements.

For the second GLUE application (Fig. 7C and 7D), a similar scenario was observed in what refers to the poor representation of the sediment accumulation near the dune toe. Once again, the berm was fully eroded in the best subaerial BSS simulation (blue dashed line in Fig. 7C - reaching 0.87) and partially eroded in the best SEV run (reaching 0.72 in the BSS evaluation). Like what happened during the first GLUE application, the underwater bar migration was also misrepresented. The SEV for the best subaerial BSS run was of $3.69\text{m}^3/\text{m}$, and the best eroded value simulation reached exactly the observed amount of $2.08\text{m}^3/\text{m}$.

Minimum and maximum values for each grid point were also extracted and are shown as the upper and lower limits of the shaded blue areas in Fig. 7B and 7D. Higher variations were observed in the berm area and towards the upper beach/dune toe. A higher variation was observed in the first GLUE compared to the second, but the pattern was very much alike: higher variation around the berm area and its surroundings.

The density weighted histograms in Fig. 7E and 7F present the SEV distribution of the emerged beach behavioral runs. For the first GLUE application, most of the behavioral runs had SEVs between 1.5 and $3.5\text{m}^3/\text{m}$, with the highest density concentrated between 2.0 and $2.5\text{m}^3/\text{m}$. A more normal-shaped distribution centered around $3.5\text{m}^3/\text{m}$ resulted from the second GLUE application, with the highest density ranging between 3.0 and $3.5\text{m}^3/\text{m}$. Table 5 summarizes both GLUE applications and the values obtained by XBeach default configuration run.

The parameter values associated with the best subaerial BSS simulation (shadowed blue row in Table 5) of the second GLUE were used for the operational (semi-)probabilistic testing on the profiles of Marina Romea and Punta Marina. The selection of this specific model configuration followed the correct representation of the berm erosion that made possible for the water level to be accurately reproduced as the storm surge strengthened.

4.1.4. Calibration discussion

The first and second GLUE applications present differences in the BSS values (Fig. 4). A wider distribution of results was observed for the first GLUE, with a larger difference between minimum and maximum values as well as larger SDs. As for the second GLUE application the parameter ranges were narrower, a lower number of physical interactions was reproduced resulting in a more limited output distribution.

Furthermore, Fig. 4 strongly evidences a disparity between BSS values for the whole profile relative to the subaerial portion only. As seen in the calibration profile (Cesenatico – Fig. 7), two submerged bars were observed before and after the storm, with the latter presenting an offshore migration relative to the original position. Higher energetic conditions during the storm were responsible for the offshore migration which was poorly represented by XBeach. Still in Fig. 4 but considering the subaerial profile only, the berm erosion represents the most significant variation observed in the post-storm profile. This process was also misrepresented by the model as shown by the blue shaded area in Fig. 7. Misfits on both underwater and the emerged beach account for the lower skill observed in the whole profile relative to the subaerial beach. Armadori et al. (2013) also showed divergences in model skill when subdividing the profile.

Parameter calibration plays a major role in representing (or misrepresenting) the aforementioned processes. Thus, finding an optimized parameter set through calibration is fundamental. For instance, the simulations with the default configuration led to subaerial erosion overestimation and a very low model skill (Table 5). A rectilinearization of the profile was observed with full erosion of both the berm and submerged bars (Fig. 7 - Default Config green profile). Other authors (e.g. Elsayed and Oumeraci, 2017; Simmons et al., 2017a; Splinter and Palmsten, 2012) have also found that running XBeach with the default configuration tends to overestimate erosion and straighten the active

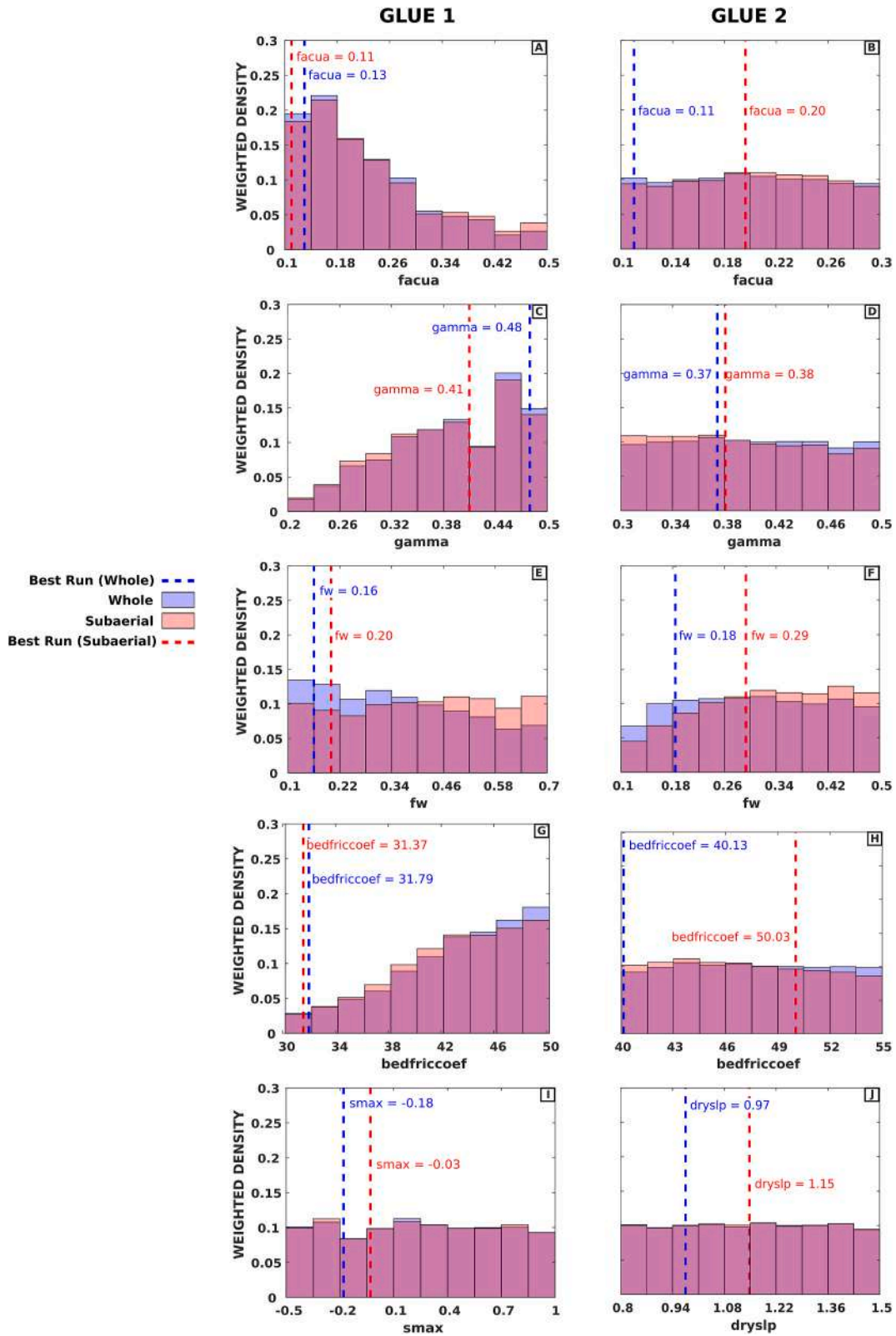


Fig. 6. Weighted density histograms for each parameter tested during the first (left) and second (right) GLUE applications. The blue and red bars represent the values for the whole profile and for the subaerial section only, respectively. Blue dashed vertical lines represent the parameter value associated with the best whole profile BSS evaluation, while the red dashed vertical lines refer to the best subaerial BSS simulation. Subfigures A) to J) follow the same pattern as in Fig. 5. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 5

Results of both GLUE applications showing the best subaerial BSS simulation, best SEV simulation, best whole profile BSS simulation and the default run. For each simulation, the associated value of each parameter is also shown.

	FIRST GLUE APPLICATION						BSS Value	SEV (m ³ /m)
	<i>facua</i>	<i>gamma</i>	<i>fw</i>	<i>bedfriccoef</i>	<i>smax</i>			
Best Subaerial BSS simulation	0.11	0.41	0.20	31.37	-0.03	0.88	3.70	
Best SEV Simulation	0.25	0.42	0.19	32.82	-0.46	0.71	2.08	
Best whole profile BSS simulation	0.13	0.48	0.16	31.79	-0.18	0.43	4.69	
Default	0.10	0.55	0.00	55.00	-1.00	-1.99	27.45	
						Observed SEV	2.08m ³ /m	
	SECOND GLUE APPLICATION						BSS Value	SEV (m ³ /m)
	<i>facua</i>	<i>gamma</i>	<i>fw</i>	<i>bedfriccoef</i>	<i>dryslp</i>			
Best Subaerial BSS simulation	0.20	0.38	0.29	50.03	1.15	0.87	3.69	
Best SEV Simulation	0.26	0.31	0.32	43.41	0.82	0.72	2.08	
Best whole profile BSS simulation	0.11	0.37	0.18	40.13	0.97	0.44	4.37	
Default	0.10	0.55	0.00	55.00	1.00	-1.99	27.45	
						Observed SEV	2.08m ³ /m	

profile.

It is important to stress that the BSS evaluation is not fully correlated to the subaerial volume erosion (Fig. 7E and 7F). The best SEV simulations were the ones balancing the misrepresentation of the dune toe sediment accumulation and the berm erosion. In contrast, the best subaerial BSS simulations fully eroded the berm but also poorly represented the dune toe sediment accumulation, resulting in erosion overestimation. Hence, using more than one method to estimate model skill and accuracy is a valid approach to better address uncertainties and understand how the physical processes were represented in the simulations.

During calibration, the role played by the selection of behavioral threshold values strongly influenced the sensitivity evaluation. Higher threshold values tend to narrow the amount of results to be used in the weighted density analysis (Fig. 6 and subfigures). For instance, if for the second GLUE application the subaerial BSS threshold was set as 0.8, there would be fewer behavioral simulations. In that scenario, one of the possible outcomes would be weighted density histograms with narrower ranges of parameter values associated with higher densities. As for the purpose of the present work the results obtained with the second GLUE application already provided a good agreement with the *in-situ* measurements, it was decided not to further explore possible outcomes by altering the threshold values.

Through parameter optimization, the GLUE approach helped in decreasing parameter-related uncertainty. Hence, the selection of parameters to be tested has to be carefully conducted (e.g. based on previous studies that have addressed model sensitivity in the area of interest). However, through testing a given parameter set generated by randomly combining the values, it becomes difficult (and time consuming) to individualize the effects of each parameter in the simulations. Following a one parameter at a time approach (Armaroli et al., 2013; Dissanayake et al., 2014; Harley et al., 2016; Unguendoli, 2018; Vousdoukas et al., 2012) makes it simpler to distinguish how a given parameter influences the simulations and it might be used as an important initial step to choose parameters prone to be tested in a “GLUE-like” procedure.

If XBeach shows a high sensitivity to more than one parameter in a given region, the one-at-the-time approach might not be the best option. In this case, applying a “GLUE-like” procedure in which several important parameters are combined differently by means of a randomized scheme tends to provide the modeler a solid final set to be used. Even though this procedure does not consider parameter interrelation, many combinations help on reaching substantial conclusions.

Among the cons of a generalized estimation following thousands of combinations are the time consumption and storage space. The former can be minimized through automating the procedure as much as

possible and running the simulations in parallel among different processors. When such a facility is not available, the modeler might try other calibration strategies that require less computer power. Furthermore, the storage space to absorb the amount of data generated might not be at one’s disposal in regular processing computers creating management difficulties. However, as shown in the present work, a GLUE-like calibration indeed provides a more extensive outlook when it comes to testing several parameters that the model has already shown sensitivity to for a given area.

Also for brevity, the specific details on each parameter behavior during the calibration steps are not discussed here.

4.2. EWS-EPS implementation

The (semi-)probabilistic implementation results from March 1 until April 30 of 2020 are presented in this subsection. XBeach-based forecasts were run using four combinations of TMES outputs as previously described. Results from the already implemented deterministic forecast were also assessed for a total of five +48 h h forecasts (Fig. 8).

In Fig. 8, the minimum BWD reached by each of the five forecasts was plotted with the starting day of the simulation. Marina Romea (Marrom) presented BWD values lower than Punta Marina (Puntam). The average BWD, considering all days and forecasts, was 63.82m for Marina Romea and 83.22 for Punta Marina. During the analyzed time span, none of the forecasts reached neither the medium (orange line) nor the high (red orange) hazard thresholds.

For both profiles, the maximum BWD values were associated with the forecasts using the ensemble mean minus one SD (green circles) followed by the ensemble mean forecasts (blue cross markers). The ensemble mean plus one SD (magenta circles) stayed between the previous two combinations and either the deterministic (brown cross markers) or the mean plus two SDs (red circles). Lower BWD values were related to the latter two combinations, which represented higher water levels arriving closer to the reference building of each profile.

Throughout the investigated period, higher energetic conditions were observed between the 21st and the 27th of March (blue shaded area in Fig. 8). Among these dates, a slight convergence of the minimum BWD results towards lower values can be seen. In the next subsection, the forecast results of the 22nd are shown to exemplify the outcomes of the new (semi-)probabilistic implementation. Only the results for Marina Romea are presented both for brevity and as the outputs were very similar under non-stormy conditions.

4.2.1. Forecast results for 22/03/2020

Fig. 9 presents the deterministic run (green solid line), the TMES mean run (green dotted line) and the limits of the TMES based forecasts

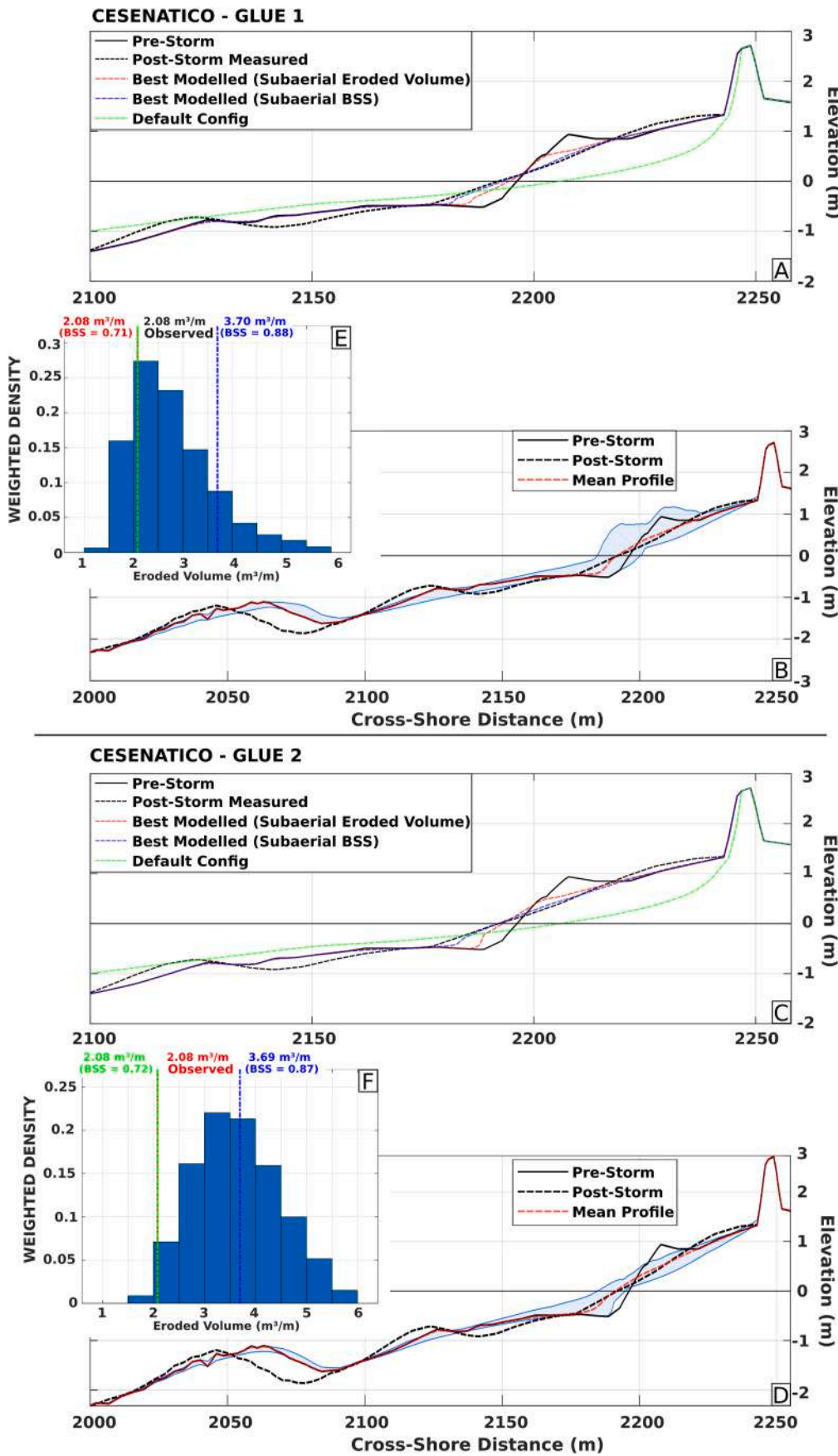


Fig. 7. Subfigures A and C show the pre- (full black line) and post-storm (black dashed line) measured profiles, best BSS (blue dashed line) and SEV (red dashed line) modeled profiles, as well as the final modeled profile using the default XBeach configuration (green dashed line). Subfigures B and D show the pre- (full black line) and post-storm (black dashed line) measured profiles with the mean profile of all the behavioral runs (red dashed line); the blue shaded area represents the final minimum and maximum elevation values reached by the behavioral simulations. Subfigures E and F represent the weighted densities for the behavioral runs relative to SEV. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

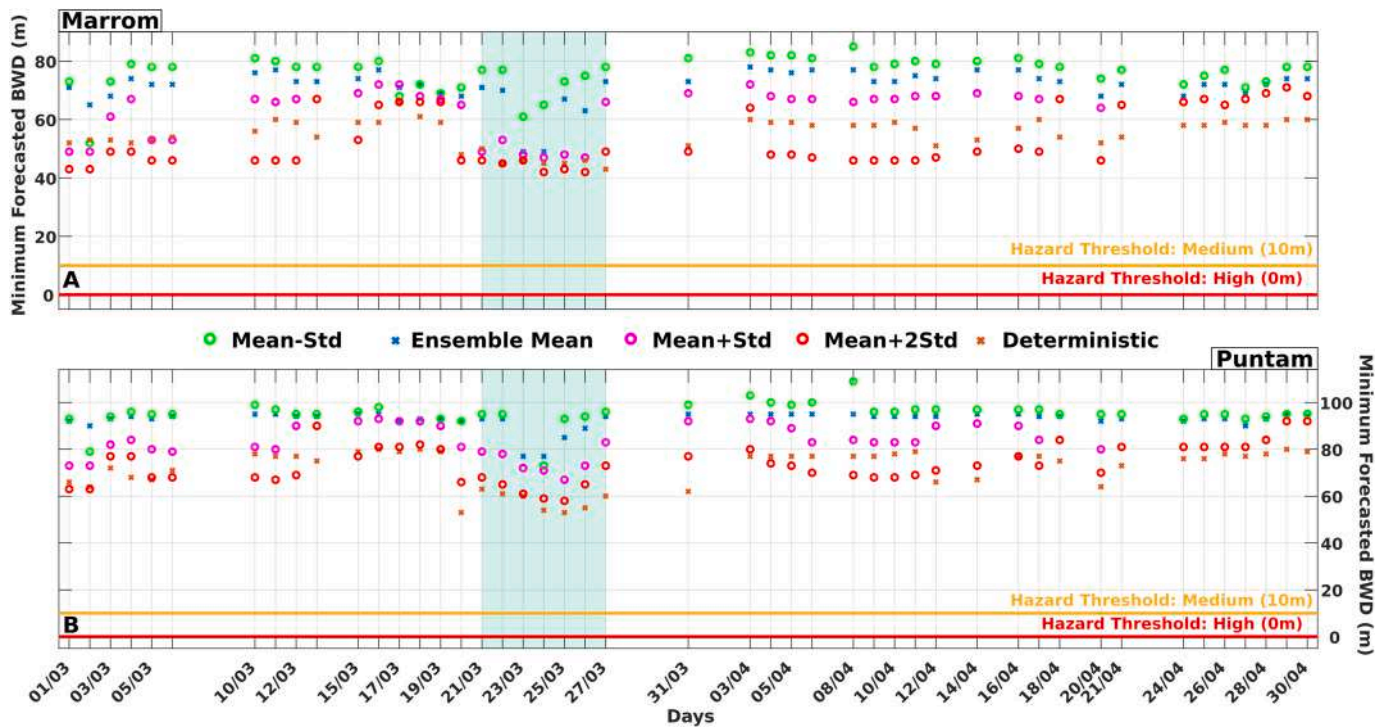


Fig. 8. Minimum BWD reached by each 48h forecast plotted based on the starting day for Marina Romea (A - Marrom) and Punta Marina (B - Puntam). Blue and brown cross markers represent the simulations using the TMES mean and the deterministic approach, respectively. The green, magenta and red circles refer to the forecasts using the TMES mean minus one SD, plus one SD and plus two SDs, respectively. More intense sea conditions reached the study area between the 21st and the 27th (blue shaded area). High and medium hazard thresholds are shown as orange and red lines, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

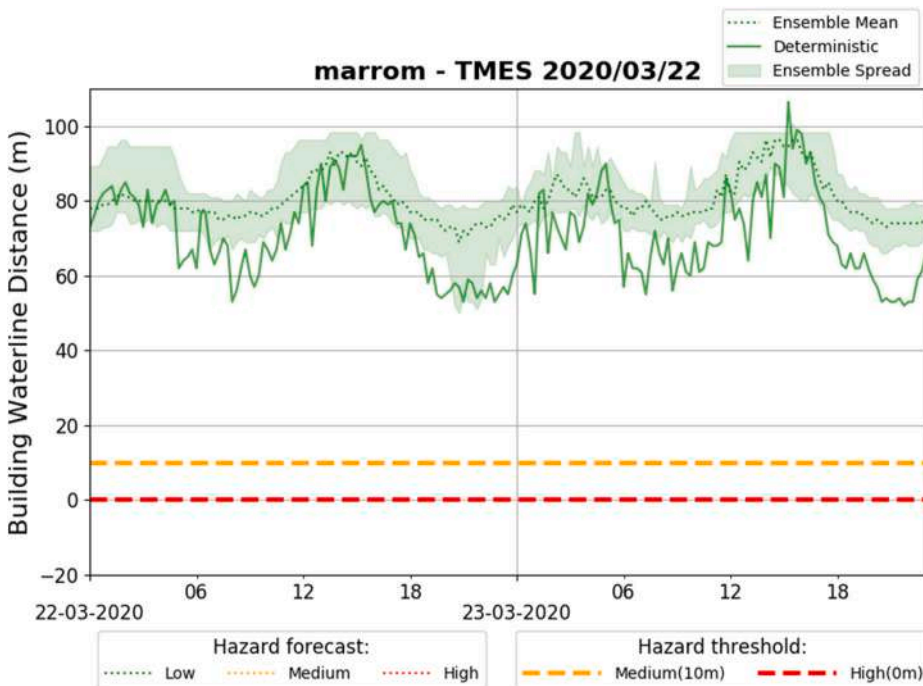


Fig. 9. Marina Romea (marrom) deterministic and (semi-)probabilistic forecast results for the 22/03/2020. The lower green shaded limit is represented by the simulation using the TMES results plus one SD. As during the forecasted period the BWD has never overcome the medium and high thresholds, no continuous orange or red lines are seen in the graph (as they would only occur below 10m and 0m, respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

using the mean minus one (upper limit of the green shaded area) and plus one (lower limit) SDs, respectively. Fig. 10 is different from Fig. 9 as the simulation using the TMES mean plus two SDs (instead of one) represents the lower limit of the green shaded area.

In Figs. 9 and 10 it is possible to see an oscillation in the BWD values

with an approximate period of 24 h with two peaks and two troughs related to the semi-diurnal tidal characteristic of the region. Within the low frequency oscillation, higher frequency fluctuations can be observed.

In the troughs of Fig. 9, the deterministic run presents values below

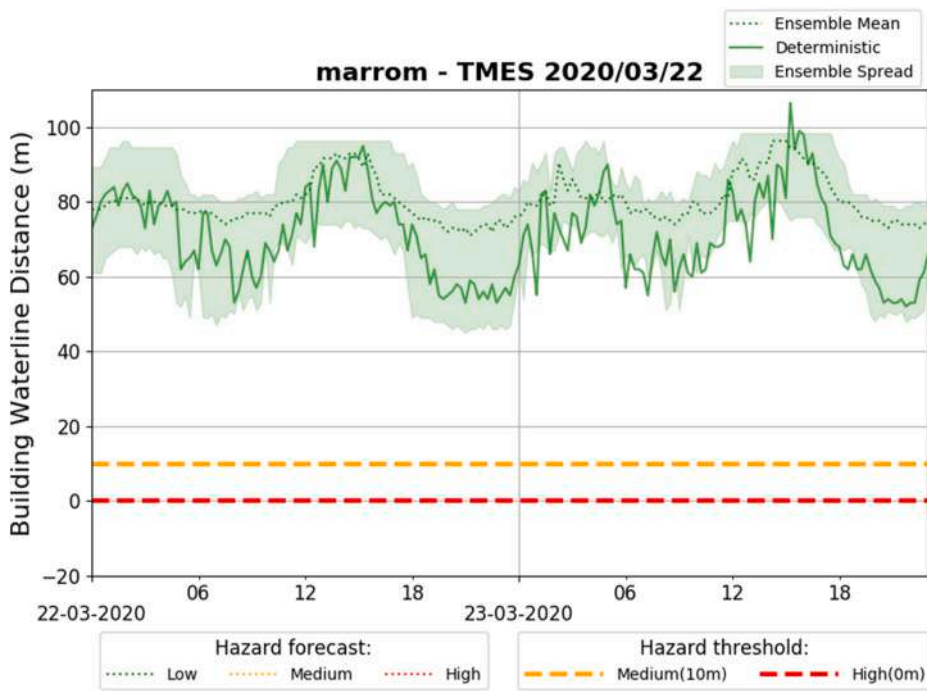


Fig. 10. Marina Romea (marrom) deterministic and (semi-)probabilistic forecast results for the 22/03/2020. The lower green shaded limit is represented by the simulation using the TMES results plus two SDs. As during the forecasted period the BWD has never overcome the medium and high thresholds, no continuous orange or red lines are seen in the graph (as they would only occur below 10m and 0m, respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the shaded area inferior limit (indicating higher water levels). In the peaks, the deterministic forecast remained mostly between the shaded area limits with an exception between 12pm and 18pm of the 23/03. Higher BWD values indicated lower water levels and the water line being further away from the reference building.

When the forecast using the mean plus two SDs (Fig. 10) is considered, the inferior limit of the shaded area expanded towards lower values. The deterministic forecast remained within the shaded area limits during almost the entire forecast, with one exception between 12am and 6m of the 22nd and one between 6am and 6pm of the 23rd. Using the TMES plus two SDs increased the amplitude of the shaded area limits and encompassed the deterministic forecast within its lower and upper boundaries.

Table 6 summarizes the results of each forecast conducted for Marina Romea on the 22/03. The minimum forecasted BWD and the lower mean value throughout the 48-h forecast were reached by the simulation using the TMES mean plus two SDs. In terms of hazard levels, the hazard thresholds of 10m (orange solid line) and 0m (red solid line) were never exceeded.

4.2.2. EWS-EPS discussion

In order to analyze the (semi-)probabilistic results, the TMES and the applied methodology were examined. As the boundary conditions are from wave and sea level ensembles, the approach presented in this work differs from other probabilistic storm surge EWSs implemented worldwide. For instance, in the UK and in the Netherlands there have been probabilistic approaches using hydrodynamic models forced by atmospheric ensemble members (e.g. Flowerdew et al., 2010, 2013; de Vries, 2009). Unlike those applications, the present work accounts for the

morphologic-hydrodynamic interactions in shallow waters calculated by a local implementation of XBeach. Model optimization by means of calibration plays an important role to better represent simulations on sandy beaches presenting a complex morphodynamic behavior. Furthermore, here the results are not given in probabilistic terms, hence being called “(semi-)probabilistic” as its foundations come from probabilistic outputs but the final answer is not yet fully probabilistic (e.g. probability of threshold exceedance).

With the usage of SDs combined with the ensemble mean, the results provided a better indication of forecast uncertainty when compared to the deterministic approach alone (Figs. 9 and 10). However, in order to calculate the probability of threshold exceedance, a larger number of members would be necessary. This task can be accomplished by running the system merging the wave and sea level ensemble components in different ways.

Nevertheless, using two SDs from the mean in every wave and sea level parameter should encompass the possible variations within this range. As an example, the results of TMES mean plus/minus one SD showed a very limited spread compared to the total variation of the deterministic forecast alone (Fig. 9). Using only one SD, the results of the spread rarely reached BWD values as low as the minimum achieved by the deterministic forecasts. By contrast, the forecasts using the TMES mean plus two SDs (Fig. 10) presented a higher spread with the lower limit being more similar to the deterministic simulations and encompassing a larger spectrum of scenarios that included the results of the simulations with one SD.

As a measure of dispersion, the more SDs, the higher the spread from the mean. If a normal distribution is considered, 68% of the data is contained around one SD from the mean, while a distance of two SDs

Table 6

Minimum, maximum, and mean BWD values calculated for each forcing condition. The results represent the forecast of 22/03 for the Marina Romea profile.

	TMES mean - 1SD	TMES mean	TMES mean + 1SD	TMES mean + 2SD	Deterministic
min BWD (m)	78.00	69.00	50.00	45.00	52.00
max BWD (m)	100.66	96.35	89.33	80.00	106.56
mean BWD (m)	88.32	80.51	73.07	63.05	70.76

includes approximately 95% of the data. Thus, the combination of mean and more SDs could allow for more information, as well as an increased possibility to forecast worst-case scenarios. Notwithstanding, a problem arises as the higher the number of SDs, the higher the chances of false threshold exceedance. In this sense, analyzing a longer series of the EWS-EPS forecasting results is fundamental to define which combination would perform better under flooding conditions.

With a larger number of members it would be also possible to determine whether the distribution of the TMES results in terms of *Hs* and sea level fully follow a normal distribution. If the distribution is not normal, there is a partial devaluing of the whole meaningfulness behind the application proposed in this work and other implementation considering the real distribution could be developed. Future TMES- XBeach usage could evaluate if a larger number of combinations would provide additional information that is not covered by the TMES mean plus two SDs, as well as determine the most appropriate way to merge the wave and sea level boundary conditions.

Heretofore, it seems conceivable an initial full-operational implementation involving two standard deviations. This approach could be useful to fully test the application of the EWS-EPS until more energetic sea storms hit the coast of Emilia-Romagna and more data is gathered. As up to now the results of the (semi-)probabilistic implementation have been compared to the deterministic approach, the system validation using *in-situ* water levels (maximum vertical excursion) should be a priority. As a means to overcome this difficulty, the implementation of georeferenced camera imagery (e.g. Dusek et al., 2019) could be used to extract water levels during storms allowing for a performance evaluation of the deterministic and the (semi-)probabilistic forecasts.

The results discussed in this and in the previous sections only involved combination and parameter related uncertainties, but they are not the only sources associated with the implemented procedure. Initial topo-bathymetric differences and the model setup also account for divergences between the deterministic and the (semi-)probabilistic implementation. In the former, the optimal parameter set currently implemented (deterministic framework) followed a one parameter at-a-time calibration (Unguendoli, 2018) being different from the values found in the present work through the GLUE. Moreover, the final gridded profiles used in the forecasts were slightly different mostly on their subaerial portion but built based on the same bathymetry.

Another source of uncertainty prone to be tackled in future works refer to the parameter values applied to Marina Romea and Punta Marina. The optimized parameter set was obtained through the GLUE approach tested in Cesenatico and based on one storm only. Transferability of calibrated parameters has already been associated with poorer model performances, even for short distances (Bugajny et al., 2013; Splinter and Palmsten, 2012), when compared to locally adjusted values. Also, using only one storm might not be sufficient to cover the total spectrum of storms and conditions reaching the area. Simmons et al. (2017b) have already shown the increase in model skill when more than one storm is used in the calibration phase and a strong recommendation is made in order to continuously update the parameter set.

In what refers to the TMES, a distance-weighted average remapping of the nearest neighbors (Ferrarin et al., 2020) is used to combine the different models and generate the mean and standard deviation, so far not accounting for model weighting. For instance, if one of the models used in the ensemble tends to provide a better representation of sea level and wave conditions along the Adriatic, giving a higher weight to that model could imply a better representation by the ensemble of the real conditions. Another option could involve a geographical model weighting: models that better represent the conditions in the Southern or Northern Adriatic would be given a higher weight in the geographic location in which they tend to perform better.

The (semi-)probabilistic implementation attempted here could be improved by the aforementioned adjustments both on TMES, as well as on the XBeach chain (e.g. calibration and topo-bathymetric initial conditions). Nonetheless, the usage of the TMES ensemble as forcing might

be “too soft” even with different combinations of the mean and SD values. Hence, testing an XBeach-based forecast employing directly the outputs of the models used to build the TMES could provide even more information to the forecaster. As EWSs have to alert mainly about incoming extreme conditions, using the mean and SDs could smoothen possible hazardous conditions.

Considering the aforementioned limitations, it should be emphasized that the XBeach based EWS is the last step of an operational chain and its results alone are not the only sources of information to decision makers. Currently, two operational oceanographic models and nested SWAN implementations as well as a series of *in situ* measuring devices are also available during the decision-making step. Hence, all the recommendations and findings presented previously are valuable in improving similar systems.

5. Conclusions

In a context of increasing magnitude and frequency of storms and considering the recent increment in coastal communities, the present work enhances the importance of development and usage of storm surge EWSs. Based on current developments of probabilistic modelling approaches, it is possible to develop more reliable uncertainty measures associated with the forecasting procedures.

By using a morphodynamic model instead of a hydrodynamic model alone, it is possible to address the nearshore morphodynamic interactions that play a major role on sandy beaches presenting a complex bathymetry. In this sense, the application of XBeach, whether deterministically or probabilistically, tends to provide a better representation of flooding waters combining hydrodynamics and morphology.

The importance of appropriately calibrating a model with as many parameters as XBeach is also stressed in this work. Following what has been widely reported in the literature, XBeach usage with the default configuration tends to overestimate erosion. The GLUE approach application proves to be a very useful tool to address parameter related uncertainty if the tested parameters are carefully chosen. However, its implementation demands powerful resources in the form of computer processing capacity, as well as hardware space. Hence, the GLUE approach might not be feasible for short term projects to be developed using “regular processing” computers.

Still on the importance of calibration, the present work shows that using only one method of evaluation of the modeled outcomes might not fully address the complexities associated with morphodynamic modelling. This conclusion is possible as the simulation with the higher BSS is not the same as the best simulation considering the SEV. Whenever possible, more than one skill and/or accuracy measures should be used to evaluate model performance.

In terms of the (semi-)probabilistic implementation, the performance of this approach is very satisfactory even with all the differences relative to the already implemented deterministic system. By using the TMES outputs, more information is available to the modeler/forecaster in terms of water level vertical excursion ranges that might or might not be hazardous. The results enhance the multitude of possible operational applications associated with such a complex morphodynamic model being forced by wave and sea-level parameters.

A longer TMES series should be gathered to provide a higher range of results reproducing distinct sea conditions. In this way, it will be possible to better evaluate an appropriate combination of the ensemble mean and SDs to be applied as means to provide reliable forecasts also under conditions that exceed the BWD flooding thresholds. Furthermore, using a higher number of mean and SD combinations would allow for the generation of probabilistic outcomes in terms of threshold exceedance. The latter proposition has to be thoroughly tested as there may not be an increment in the amount of information if compared to what is provided by the forecasts using the mean plus two SDs.

As a final remark, collecting pre- and post-storm topo-bathymetric data for the profiles of Marina Romea and Marina di Ravenna would

allow for a site specific GLUE application. This would be important to evaluate the transferability of the calibrated parameters between them and Cesenatico.

Author credit statement

LGB carried out part of the topo-bathymetric measuring campaigns, designed the model calibration steps in the internal system, and operational implementation of the EWS-EPS system. SU assisted with XBeach calibration and implementation inside the internal system. LB assisted the calibration step, the operational implementation of the operational system and the connection with the multi-model ensemble. BMSG carried out part of the topo-bathymetric data collection, pre-processing and data analysis. AV conceived the EWS-EPS implementation, coordinated the work group and assisted during the XBeach calibration and the operational system implementation.

Funding

This work was supported by the European Union under the 2014–2020 Interreg V-A Italy-Croatia CBC Programme within the STREAM - Strategic Development of Flood Management – strategic project, ID 10249186.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The first author is thankful to the Erasmus Mundus Joint Master Degree (EMJMD) in Water and Coastal Management (WACOMA) Programme (University of Bologna - Italy, University of Cadiz – Spain, and University of Algarve - Portugal) for the scholarship provided as part of the 2018–2020 cohort. The authors would like to acknowledge all the help provided by CNR-ISMAR, specially to Dr. Amedeo Fadini and Dr. Christian Ferrarin, on interacting with the TMES operational system, extracting outputs from its database to be used for the coastal EWS-EPS implementation.

References

Armaroli, C., Ciavola, P., Perini, L., Calabrese, L., Lorito, S., Valentini, A., Masina, M., 2012. Critical storm thresholds for significant morphological changes and damage along the Emilia-Romagna coastline, Italy. *Geomorphology* 143–144, 34–51. <https://doi.org/10.1016/j.geomorph.2011.09.006>.

Armaroli, C., Grotoli, E., Harley, M.D., Ciavola, P., 2013. Beach morphodynamics and types of foredune erosion generated by storms along the Emilia-Romagna coastline, Italy. *Geomorphology* 199, 22–35. <https://doi.org/10.1016/j.geomorph.2013.04.034>.

Bajo, M., Umgiesser, G., 2010. Storm surge forecast through a combination of dynamic and neural network models. *Ocean Model.* 33, 1–9. <https://doi.org/10.1016/j.ocemod.2009.12.007>.

Basher, R., 2006. Global early warning systems for natural hazards: systematic and people-centred. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 364, 2167–2182. <https://doi.org/10.1098/rsta.2006.1819>.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* 6, 279–298.

Blumberg, A.F., Bruno, M.S., 2018. *Overview*. In: *The Urban Ocean*. Cambridge University Press, pp. 1–11. <https://doi.org/10.1017/9781108123839.002>.

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions: 1. Model description and validation. *J. Geophys. Res.* 104, 7649–7666. <https://doi.org/10.1029/98JC02622>.

Bugajny, N., Furmańczyk, K., Dudzińska-Nowak, J., Papińska-Swerpel, B., 2013. Modelling morphological changes of beach and dune induced by storm on the Southern Baltic coast using XBeach (case study: Dziwnow Spit). *J. Coast Res.* 65, 672–677. <https://doi.org/10.2112/si65-114.1>.

Buizza, R., 2019. Introduction to the special issue on “25 years of ensemble forecasting.”. *Q. J. R. Meteorol. Soc.* 145, 1–11. <https://doi.org/10.1002/qj.3370>.

Chiggiato, J., Oddo, P., 2008. Operational ocean models in the Adriatic Sea: a skill assessment. *Ocean Sci.* 4, 61–71. <https://doi.org/10.5194/os-4-61-2008>.

Clementi, E., Pistoia, J., Escudier, R., Delrosso, D., Drudi, M., Grandi, A., Lecci, R., Creti, S., Ciliberti, S., Coppini, G., Masina, S., Pinardi, N., 2019. Mediterranean sea analysis and forecast (CMEMS MED-currents, EAS5 system). https://doi.org/10.25423/CMCC/MEDSEA_ANALYSIS_FORECAST_PHY_006_013_EAS5.

Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: a review. *J. Hydrol.* 375, 613–626. <https://doi.org/10.1016/j.jhydrol.2009.06.005>.

Davolio, S., Miglietta, M.M., Diomede, T., Marsigli, C., Morgillo, A., Moscatello, A., 2008. A meteorological prediction system based on a multi-model approach for precipitation forecasting. *Nat. Hazards Earth Syst. Sci.* 8, 143–159. <https://doi.org/10.5194/nhess-8-143-2008>.

de Vries, H., 2009. Probability forecasts for water levels at the coast of The Netherlands. *Mar. Geodes.* 32, 100–107. <https://doi.org/10.1080/01490410902869185>.

Deltares, 2018. XBeach Documentation: Release XBeach v1.23.5527 XBeachX FINAL.

Dietrich, J., Trepte, S., Wang, Y., Schumann, A.H., Voß, F., Hesser, F.B., Denhard, M., 2008. Combination of different types of ensembles for the adaptive simulation of probabilistic flood forecasts: hindcasts for the Mulde 2002 extreme event. *Nonlinear Process Geophys.* 15, 275–286. <https://doi.org/10.5194/npg-15-275-2008>.

Dissanayake, P., Brown, J., Karunaratna, H., 2014. Modelling storm-induced beach/dune evolution: Sefton coast, Liverpool Bay, UK. *Mar. Geol.* 357, 225–242. <https://doi.org/10.1016/j.margeo.2014.07.013>.

Dissanayake, P., Brown, J., Wisse, P., Karunaratna, H., 2015. Effects of storm clustering on beach/dune evolution. *Mar. Geol.* 370, 63–75. <https://doi.org/10.1016/j.margeo.2015.10.010>.

Dusek, G., Hernandez, D., Willis, M., Brown, J.A., Long, J.W., Porter, D.E., Vance, T.C., 2019. WebCAT: piloting the development of a Web camera coastal observing network for diverse applications. *Front. Mar. Sci.* 6 <https://doi.org/10.3389/fmars.2019.00353>.

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

Elsayed, S.M., Oumeraci, H., 2017. Effect of beach slope and grain-stabilization on coastal sediment transport: an attempt to overcome the erosion overestimation by XBeach. *Coast. Eng.* 121, 179–196. <https://doi.org/10.1016/j.coastaleng.2016.12.009>.

Ferrarin, C., Roland, A., Bajo, M., Umgiesser, G., Cucco, A., Davolio, S., Buzzi, A., Malguzzi, P., Drofa, O., 2013. Tide-surge-wave modelling and forecasting in the Mediterranean Sea with focus on the Italian coast. *Ocean Model.* 61, 38–48. <https://doi.org/10.1016/j.ocemod.2012.10.003>.

Ferrarin, C., Valentini, A., Vodopivec, M., Klaric, D., Massaro, G., Bajo, M., De Pascalis, F., Fadini, A., Ghezzi, M., Menegon, S., Bressan, L., Unguendoli, S., Fettich, A., Jerman, J., Ličer, M., Fustar, L., Papa, A., Carraro, E., 2020. Integrated sea storm management strategy: the 29 October 2018 event in the Adriatic Sea. *Nat. Hazards Earth Syst. Sci.* 20, 73–93. <https://doi.org/10.5194/nhess-20-73-2020>.

Flowerdew, J., Horsburgh, K., Wilson, C., Mylne, K., 2010. Development and evaluation of an ensemble forecasting system for coastal storm surges. *Q. J. R. Meteorol. Soc.* 136, 1444–1456. <https://doi.org/10.1002/qj.648>.

Flowerdew, J., Mylne, K., Jones, C., Titley, H., 2013. Extending the forecast range of the UK storm surge ensemble. *Q. J. R. Meteorol. Soc.* 139, 184–197. <https://doi.org/10.1002/qj.1950>.

Harley, M., Armaroli, C., Ciavola, P., 2011. Evaluation of XBeach predictions for a real-time warning system in Emilia-Romagna, Northern Italy. *J. Coast Res.* 1861–1865.

Harley, M.D., Ciavola, P., 2013. Managing local coastal inundation risk using real-time forecasts and artificial dune placements. *Coast. Eng.* 77, 77–90. <https://doi.org/10.1016/j.coastaleng.2013.02.006>.

Harley, M.D., Valentini, A., Armaroli, C., Perini, L., Calabrese, L., Ciavola, P., 2016. Can an early-warning system help minimize the impacts of coastal storms? A case study of the 2012 Halloween storm, northern Italy. *Nat. Hazards Earth Syst. Sci.* 16, 209–222. <https://doi.org/10.5194/nhess-16-209-2016>.

Jasper, K., Gurtz, J., Lang, H., 2002. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. *J. Hydrol.* 267, 40–52. [https://doi.org/10.1016/S0022-1694\(02\)00138-5](https://doi.org/10.1016/S0022-1694(02)00138-5).

Lashley, C.H., Roelvink, D., van Dongeren, A., Buckley, M.L., Lowe, R.J., 2018. Nonhydrostatic and surfbeat model predictions of extreme wave run-up in fringing reef environments. *Coast. Eng.* 137, 11–27. <https://doi.org/10.1016/j.coastaleng.2018.03.007>.

Lenartz, F., Beckers, J.-M., Chiggiato, J., Mourre, B., Troupin, C., Vandembulcke, L., Rixen, M., 2010. Super-ensemble techniques applied to wave forecast: performance and limitations. *Ocean Sci.* 6, 595–604. <https://doi.org/10.5194/os-6-595-2010>.

Lindemer, C.A., Plant, N.G., Puleo, J.A., Thompson, D.M., Wamsley, T.V., 2010. Numerical simulation of a low-lying barrier island's morphological response to Hurricane Katrina. *Coast. Eng.* 57, 985–995. <https://doi.org/10.1016/j.coastaleng.2010.06.004>.

Mariani, S., Casaioli, M., Coraci, E., Malguzzi, P., 2015. New high-resolution BOLAM-MOLOCH suite for the SIMM forecasting system: assessment over two HyMeX intense observation periods. *Nat. Hazards Earth Syst. Sci.* 15, 1–24. <https://doi.org/10.5194/nhess-15-1-2015>.

McCall, R.T., Van Thiel de Vries, J.S.M., Plant, N.G., Van Dongeren, A.R., Roelvink, J.A., Thompson, D.M., Reniers, A.J.H.M., 2010. Two-dimensional time dependent hurricane overwash and erosion modeling at Santa Rosa Island. *Coast. Eng.* 57, 668–683. <https://doi.org/10.1016/j.coastaleng.2010.02.006>.

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

- Oppenheimer, M., Glavovic, B.C., Hinkel, J., van de Wal, R., Magnan, A.K., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M., DeConto, R.M., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., Sebesvari, Z., 2019. sea level rise and implications for low-lying islands, coasts and communities. In: Portner, H.-O., Roberts, D.C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegria, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., Weyer, N.M. (Eds.), IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, pp. 321–445.
- Ortiz, A.C., Ashton, A.D., 2019. Exploring carbonate reef flat hydrodynamics and potential formation and growth mechanisms for motu. *Mar. Geol.* 412, 173–186. <https://doi.org/10.1016/j.margeo.2019.03.005>.
- Orzech, M.D., Reniers, A.J.H.M., Thornton, E.B., MacMahan, J.H., 2011. Megacusps on rip channel bathymetry: observations and modeling. *Coast. Eng.* 58, 890–907. <https://doi.org/10.1016/j.coastaleng.2011.05.001>.
- Osorio-Cano, J.D., Osorio, A.F., Peláez-Zapata, D.S., 2019. Ecosystem management tools to study natural habitats as wave damping structures and coastal protection mechanisms. *Ecol. Eng.* 130, 282–295. <https://doi.org/10.1016/j.ecoleng.2017.07.015>.
- Pacheco, A., Williams, J.J., Ferreira, Ó., Garel, E., Reynolds, S., 2011. Applicability of sediment transport models to evaluate medium term evolution of tidal inlet systems. *Estuar. Coast Shelf Sci.* 95, 119–134. <https://doi.org/10.1016/j.ecss.2011.08.027>.
- Pandžić, K., Likso, T., 2005. Eastern Adriatic typical wind field patterns and large-scale atmospheric conditions. *Int. J. Climatol.* 25, 81–98. <https://doi.org/10.1002/joc.1085>.
- Pender, D., Karunarathna, H., 2013. A statistical-process based approach for modelling beach profile variability. *Coast. Eng.* 81, 19–29. <https://doi.org/10.1016/j.coastaleng.2013.06.006>.
- Perini, L., Lorito, S., Calabrese, L., 2008. Il Catalogo delle opere di difesa costiera della Regione Emilia-Romagna. *Stud. Costieri* 15, 39–56.
- Phan, L.K., van Thiel de Vries, J.S., Stive, M.J., 2014. Coastal mangrove squeeze in the Mekong Delta. *J. Coast Res.* 31 (2), 233–243.
- Ris, R.C., Holthuijsen, L.H., Booij, N., 1999. A third-generation wave model for coastal regions: 2. Verification. *J. Geophys. Res. Ocean.* 104, 7667–7681. <https://doi.org/10.1029/1998JC900123>.
- Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., Lescinski, J., 2009. Modelling storm impacts on beaches, dunes and barrier islands. *Coast. Eng.* 56, 1133–1152. <https://doi.org/10.1016/j.coastaleng.2009.08.006>.
- Roelvink, D., van Dongeren, A., McCall, R., Hoonhout, B., van Rooijen, A., van Geer, P., de Vet, L., Nederhoff, K., Quataert, E., 2015. XBeach Technical Reference. Kingsday Release.
- Russo, A., Coluccelli, A., Carniel, S., Benetazzo, A., Valentini, A., Paccagnella, T., Ravaioli, M., Bortoluzzi, G., 2013. Operational models hierarchy for short term marine predictions: the Adriatic Sea example. In: *Ocean. 2013 MTS/IEEE Bergen Challenges North*. <https://doi.org/10.1109/OCEANS-Bergen.2013.6608139>. Dimens.
- Sallenger, A., 2000. Storm impact scale for barrier islands. *J. Coast Res.* 16, 890–895.
- Schambach, L., Grilli, A.R., Grilli, S.T., Hashemi, M.R., King, J.W., 2018. Assessing the impact of extreme storms on barrier beaches along the Atlantic coastline: application to the southern Rhode Island coast. *Coast. Eng.* 133, 26–42. <https://doi.org/10.1016/j.coastaleng.2017.12.004>.
- Shchepetkin, A.F., McWilliams, J.C., 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* 9, 347–404. <https://doi.org/10.1016/j.ocemod.2004.08.002>.
- Simmons, J.A., Marshall, L.A., Turner, I.L., Splinter, K.D., Cox, R.J., Harley, M.D., Hanslow, D.J., Kinsela, M.A., 2015. A more rigorous approach to calibrating and assessing the uncertainty of coastal numerical models. In: *Australasian Coasts & Ports Conference 2015: 22nd Australasian Coastal and Ocean Engineering Conference and the 15th Australasian Port and Harbour Conference*. Engineers Australia and IPENZ, Auckland, New Zealand, pp. 821–827.
- Simmons, J.A., Harley, M.D., Marshall, L.A., Turner, I.L., Splinter, K.D., Cox, R.J., 2017a. Calibrating and assessing uncertainty in coastal numerical models. *Coast. Eng.* 125, 28–41. <https://doi.org/10.1016/j.coastaleng.2017.04.005>.
- Simmons, J.A., Harley, M.D., Turner, I.L., Splinter, K.D., 2017b. Quantifying calibration data requirements for coastal erosion models: how many storms is enough?. In: *Australasian Coasts & Ports 2017: Working with Nature*. Barton, ACT: Engineers Australia, PIANC Australia and Institute of Professional Engineers New Zealand, pp. 978–984.
- Smallegan, S.M., Irish, J.L., Van Dongeren, A.R., Den Bieman, J.P., 2016. Morphological response of a sandy barrier island with a buried seawall during Hurricane Sandy. *Coast. Eng.* 110, 102–110. <https://doi.org/10.1016/j.coastaleng.2016.01.005>.
- Splinter, K.D., Palmsten, M.L., 2012. Modeling dune response to an East coast low. *Mar. Geol.* 329–331, 46–57. <https://doi.org/10.1016/j.margeo.2012.09.005>.
- Steppeler, J., Doms, G., Schattler, U., Bitzer, H.W., Gassmann, A., Damrath, U., Gregoric, G., 2003. Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorol. Atmos. Phys.* 82, 75–96. <https://doi.org/10.1007/s00703-001-0592-9>.
- Thorndahl, S., Beven, K.J., Jensen, J.B., Schaarup-Jensen, K., 2008. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. *J. Hydrol.* 357, 421–437. <https://doi.org/10.1016/j.jhydrol.2008.05.027>.
- Unguendoli, S., 2018. Propagation of Uncertainty across Modeling Chains to Evaluate Hydraulic Vulnerability in Coastal Areas. University of Bologna.
- Valentini, A., Delli Passeri, L., Paccagnella, T., Patruno, P., Marsigli, C., Cesari, D., Deserti, M., Chiggiato, J., Tibaldi, S., 2007. The sea state forecast system of ARPA-SIM. *Boll. Geofis. Teor. Appl.* 48, 333–349.
- van der Lugt, M.A., Quataert, E., van Dongeren, A., van Ormondt, M., Sherwood, C.R., 2019. Morphodynamic modeling of the response of two barrier islands to Atlantic hurricane forcing. *Estuar. Coast Shelf Sci.* 229, 106404. <https://doi.org/10.1016/j.ecss.2019.106404>.
- van Rijn, L.C., Waslra, D.J.R., Grasmeijer, B., Sutherland, J., Pan, S., Sierra, J.P., 2003. The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models. *Coast. Eng.* 47, 295–327. [https://doi.org/10.1016/S0378-3839\(02\)00120-5](https://doi.org/10.1016/S0378-3839(02)00120-5).
- van Rooijen, A.A., McCall, R.T., van Thiel de Vries, J.S.M., van Dongeren, A.R., Reniers, A.J.H.M., Roelvink, J.A., 2016. Modeling the effect of wave-vegetation interaction on wave setup. *J. Geophys. Res. Ocean.* 121, 4341–4359.
- van Verseveld, H.C.W., van Dongeren, A.R., Plant, N.G., Jager, W.S., den Heijer, C., 2015. Modelling multi-hazard hurricane damages on an urbanized coast with a Bayesian Network approach. *Coast. Eng.* 103, 1–14. <https://doi.org/10.1016/j.coastaleng.2015.05.006>.
- Vousdoukas, M.I., Ferreira, Ó., Almeida, L.P., Pacheco, A., 2012. Toward reliable storm-hazard forecasts: XBeach calibration and its potential application in an operational early-warning system. *Ocean Dynam.* 62, 1001–1015. <https://doi.org/10.1007/s10236-012-0544-6>.
- WMO, 2011. Guide to Storm Surge Forecasting.