Ocean Modelling 133 (2019) 129–144

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Towards robust pan-European storm surge forecasting

T. Fernández-Montblanc[a], M.I. Vousdoukas[b,c,*], P. Ciavola[a], E. Voukouvalas[d], L Mentaschi[b], G. Breyiannis[b], L. Feyen[b], P. Salamon[b]

[a] Universitá degli Studi di Ferrara, Department of Physics and Earth Sciences, Via Saragat, 1, 44122 Ferrara, Italy
[b] European Commission, Joint European Research Centre (JRC), Via Enrico Fermi 2749, I-21027 Ispra, Italy
[c] Department of Marine Sciences, University of the Aegean, Mitilene, Greece
[d] Engineering Ingegneria Informatica S.p.A., 56 Via S. Martino Della Battaglia, 00185 Rome, Italy

ARTICLE INFO

Keywords:
Storm surge
Operational forecasting
Coastal flooding
Marine storms
Natural hazards
Numerical modelling

ABSTRACT

Operational forecasting systems are important for disaster risk reduction. In this work we implement a coupled storm surge and tidal model on an unstructured grid over Europe towards the development of a pan-European Storm Surge Forecasting System (EU-SSF). The skill to predict tidal, surge and total water levels was evaluated based on measurements from 208 tidal gauge stations. Results show satisfactory performance for the two atmospheric forcing datasets tested, a High Resolution Forecast and ERA-INTERIM reanalysis, both provided by the European Center for Medium range Weather Forecast. For tidal predictions, the total RSS is equal to 0.197 m, lower than the values estimated by the global tidal model FES2004, and outperformed only by FES2012 (RSS = 0.05 m), which however is a product of data assimilation. Storm surge validation results show good predictive skill, with 0.04 m < RMSE < 0.21 m and %RMSE within 4%–22%. Coupling with tides results in improved storm surge level predictions, with RMSE reducing by up to 0.033 m. The areas benefiting most from the coupling are the North Sea and the English Channel, resulting in up to 2% reduction of the %RMSE. Increasing the resolution of atmospheric forcing also improves the predictive skill, leading to a reduction of RMSE up to 0.06 m in terms of the extremes, especially in shallow areas where wind is the main driver for surge production. We propose a setup for operational pan-European storm surge forecasting combining tidal levels from the FES2012 model and storm surge residuals from the EU-SSF setup which couples meteorological and astronomic tides.

1. Introduction

Coastal flooding is a major natural hazard that can inflict large social, economic and environmental impacts (Forzieri et al., 2016; Hinkel et al., 2014; Vousdoukas et al., 2018a), as exemplified by storm Xynthia that hit the Atlantic coast in 2010 (Bertin et al., 2014). A main driver of coastal hazard is storm surge, an anomalous rise of the sea level induced by low pressure and its associated wind. When storm surges coincide with spring tide, especially low lying areas are prone to coastal flooding. Several such events happened along the European coastline (Ciavola et al., 2017; Garnier et al., 2017). The most notable example is the big storm in 1953 (Wolf and Flather, 2005) that struck the coasts of the UK, Netherlands, Belgium and Germany, producing surges over 3 m and leading to a major disaster. With more than 2000 casualties in the Netherlands it marked the launch of the Dutch national water management plans and huge investments in coastal protection. In 2010 the storm Xynthia hit the Atlantic French coast in the Bay of Biscay. For this event the surge itself was not exceptional, but due to the co-occurrence with spring tide (Bertin et al., 2014) it resulted in extensive coastal flooding, 47 casualties and monetary damage of around 1.5 billion Euros (DDTM-17, 2011). Another example is the storm Xavier that took place in December 2013. It also coincided with high astronomical tides and affected extensive areas in northern Europe (Wadey et al., 2015), causing environmental damage, infrastructure disruptions, flooding and/or damage to buildings and private property.

The above events have occurred during the past and present century, when increasing population and development in coastal zones has amplified coastal risk (Bouwer, 2011; Visser et al., 2014). Global warming and corresponding sea level rise is expected to result in a further intensification of coastal hazards (Vitousek et al., 2017; Vousdoukas et al., 2018b), highlighting the need for disaster risk management measures. Overall, disaster risk reduction has been shown to reduce losses and mortality from extreme floods (Bouwer and Jonkman, 2018; Kreibich et al., 2017). Short-term forecasting can be

* Corresponding author at: European Commission, Via Enrico Fermi 2749, 21027 Ispra, Italy.
E-mail addresses: frntms@unife.it (T. Fernández-Montblanc), Michail.VOUSDOUKAS@ec.europa.eu (M.I. Vousdoukas).

https://doi.org/10.1016/j.ocemod.2018.12.001
Received 21 May 2018; Received in revised form 20 October 2018; Accepted 3 December 2018
Available online 04 December 2018

complementary to adaptation actions aimed at medium and long term flood risk reduction. For example, Early Warning Systems (Harley et al., 2011; Harley et al., 2016; Vousdoukas et al., 2012b), can assist emergency response operations by providing useful information about the timing and spatial distribution of hazardous conditions. This is crucial to support decisions on resource deployment and evacuation during an event (Cheung et al., 2003).

In Europe there are several regional operational storm-surge forecasting systems that resolve tidal and atmospheric forcing (Sembiring et al., 2015; Zampato et al., 2016). Several of these systems contribute to the Copernicus Marine Environment Monitoring Service (CMEMS). The latter combines information from different regional domains (Iberia-Biscay-Ireland Regional Seas, Black Sea, Baltic Sea, European North-West Shelf Seas and Mediterranean Sea) into one unified platform (http://marine.copernicus.eu). CMEMS offers also a global sea level forecast service (Lellouche et al., 2018; Lellouche et al., 2013) but that only considers wind forcing and omits pressure effects. On the contrary, the Dynamic atmospheric Corrections (DAC) are produced by CLS using the Mog2D model from Legos and distributed by Aviso+, with support from Cnes (https://www.aviso.altimetry.fr/) and takes into account only the inverse barometer effect, omitting winds.

Apart from these regional systems and a few exceptions (Muis et al., 2016), there is a lack of systems that cover continental-and-beyond scales and resolve all major physical factors contributing to extreme sea levels. Such systems can complement regional ones by providing homogenous information over extensive spatial domains, including areas not served by national/regional systems.

Against the foregoing, the present contribution reports on the implementation and validation of a pan-European Storm Surge Forecasting System (EU-SSF), aspiring to become a future candidate for inclusion in the European Flood Awareness System (www.efas.eu). The manuscript is structured as follows: In Section 2 the model set-up, calibration and validation are outlined. In Section 3 results from the validation are presented along with the outcomes of a sensitivity analysis on the resolution of the atmospheric forcing and the effects of tide/storm surge interactions. Section 4 provides a discussion of the results and a foresight related to model improvement and optimization of the operational setup. Finally, Section 5 summarizes the main conclusions.

2. Methods

2.1. Model setup

The SCHISM model (http://ccrm.vims.edu/schismweb) was used to simulate the tide-surge generation and propagation. SCHISM (Semi-implicit Cross-scale Hydroscience Integrated System Model) is based on the original SELFE code (Zhang and Baptista, 2008), with many enhancements and upgrades, including an extension to large-scale eddying regime and a seamless cross-scale capability from creek to ocean (Zhang et al., 2016). The model solves the full Navier–Stokes equation over unstructured grids. In the present study, SCHISM was configured in its 2D barotropic mode, which can be considered as a reasonably accurate approximation for the given application. It further increases the model efficiency, hence allowing simulations at continental scales with affordable computational and storage requirements. The model configuration accounts for the combined effects of wind, atmospheric pressure gradients and tides.

SCHISM was forced by sea level pressure and 10 m wind speed atmospheric fields from ERA-Interim as well as by a ECMWF high resolution forecast (see Section 2.2). The surface stress in SCHISM is computed using a bulk formula with the drag coefficient computed according to Pond and Pickard (2013). Tidal forcing includes the astronomic potential over the whole domain and boundary conditions along the open boundaries (see Fig. 1a). The main tidal constituents in the area are considered (MM, MF, O1, K1, P1, Q1, M2, S2, N2 and K2) and are available from the Finite Element Solution 2012 (FES2012) (Carrère et al., 2012).

The computational domain extends from 25W^{o} to 42E^{o} and from 25N^{o} to 76N^{o} (Fig. 1a). The computational grid is an unstructured mesh with triangular elements and varying resolution ranging from 78 km in the western boundary of the domain to a higher resolution (10 km) along the European shallow coastal areas. The simulations are based on bathymetric data from the European Marine Observation and Data Network (EMODNET), originally provided in angular coordinates at a resolution of 1/8′-minute (0.0021°) (http://www.emodnet.eu/bathymetry).

The model calibration and validation was based on model simulations spanning from 1 February 2010 to 1 March 2016, a period characterized by a high marine storm activity, including extreme storm surge events that caused severe impacts along the European coastline (Bertin et al., 2014; Spencer et al., 2015; Vousdoukas et al., 2012a). The validation was performed considering the whole time series as well as the upper tail of the data distribution in order to evaluate the model's skill to reproduce extreme events. To ensure model stability and optimize computational times, the validation period was divided in 1-year simulations, including an additional month as spin up time, and then the resulting water level time series were merged. The model output provides water level every 3 h, at 11,366 locations distributed every 10 km along the European coastline. Numerical simulations were run in parallel on a high performance computing facility using 32 processors (2.6 GHz), subdividing the computational domain into 32 subdomains.

2.2. Numerical experiments

A set of model runs were conducted in order to assess the model's performance, as well as its sensitivity to factors like the resolution of atmospheric forcing and non-linear interactions between tidal and atmospheric forcing. Hereinafter the storm surge (η_{SSL}) refers to the anomalous sea level rise driven by wind and pressure, whereas the non-tidal residual (η_{RE}) refers to the anomalous sea level rise including as components both the atmospheric effect and the tide-surge interaction.

The set of performed numerical experiments is listed in Table 1. The experiment E0 neglects meteorological factors in order to validate the model's skill to reproduce tidal elevations only. The experiment E1 considers only the atmospheric forcing from ERA-INTERIM (6 h temporal resolution and 0.75° × 0.75° spatial resolution) without tidal forcing to assess performance in terms of storm surge forecasting in a fictional non-tidal setup. The fully forced run E2 combines tidal and atmospheric forcing (ERA-INTERIM). Once the non-tidal residual was subtracted from E2 results, it allowed the evaluation of the effect of the interaction between the storm surge and the tidal component by comparison with the E1 results, as well as to evaluate the model's performance in reproducing the non-tidal residual.

An additional group of runs was performed with the purpose of assessing the influence of the atmospheric forcing resolution in the model's performance. To that end, a second, higher-resolution dataset based on the global ECMWF forecast was used (ECMWF-HRES; https://www.ecmwf.int/en/forecasts/datasets/set-i). The dataset is available after 2010 and has a spatial resolution of 0.1° × 0.14°. Since the forecast provides atmospheric forcing parameters for a 72 h time window on a daily basis, the latest forecast available for each day is considered creating a dataset which covers the whole testing period. Two additional runs are forced by ECMWF-HRES, one similar to E1 (E3; only atmospheric forcing), and one similar to E2 (E4; tidal and atmospheric effects). Although the original temporal resolution of ECMWF-HRES is 1 h, here the dataset is subsampled at 6 h intervals to match the ERA-INTERIM dataset. Keeping the same time signatures to eliminate temporal resolution bias is avoided when comparing the runs forced from the two datasets.

In addition to the tidal performance of EU-SSF, we assess the performance of available global tidal models such as the TPXO7.2 ocean tide model (Egbert and Erofeeva, 2002), Finite Element Solution 2004

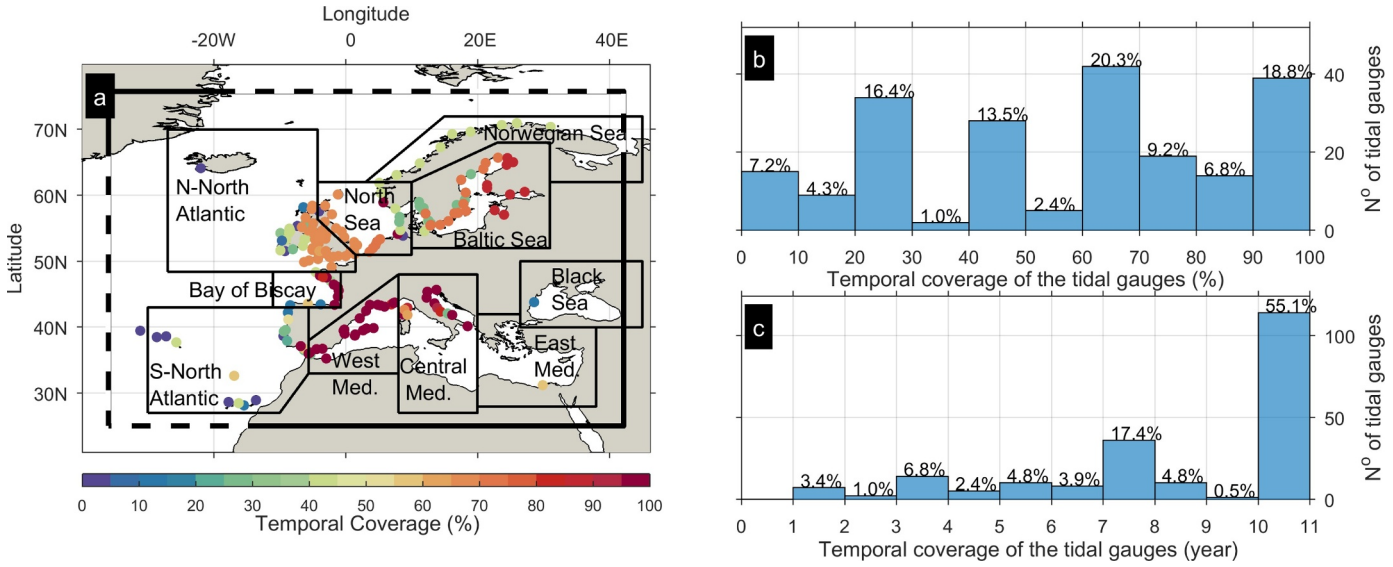

Fig. 1. Map of Europe showing the model computational domain (thick black line), open boundaries (thick black dashed line), the 10 different regions defined for the analysis of the model results (thin continuous black lines), and the location of tide gauge stations (colour scale indicates the temporal coverage of the tidal gauges for the validation period from 2010 to 2016) (a). Histogram of the tide gauge data used for the storm surge validation (temporal coverage as percentage of the storm surge validation period from 2010 to 2016) (b) and as number of years covered by the data used for tidal analysis and tidal validation (c).

Table 1
List of numerical experiments.

NAME	Type of forcing Atmospheric	Tidal	Atmospheric resolution Era-interim	ECMWF-HRES
E0		X		
E1	X		X	
E2	X	X	X	
E3	X			X
E4	X	X		X

(FES2004) (Lyard et al., 2006) and Finite Element Solution 2012 (FES2012) (Carrère et al., 2012). The tidal predictive skill of EU-SSF is compared with the one of the above tidal models.

2.3. Model validation

The model validation is based on a ground-truth dataset from 208 tidal gauges available from the JRC Sea Level Database (http://webcritech.jrc.ec.europa.eu/SeaLevelsDb) and EMODNET web site (http://www.emodnet-physics.eu/Map/), from which 192 are finally used after a quality control to remove spikes and other erroneous observations (Fig. 1a). The tidal gauge stations provide satisfactory spatial representation of the European seas (Fig. 1a), with the exception of the East Mediterranean and Black Sea, where only one station in each region provides usable measurements. To filter out long-term sea level variations the water-level time series, each station was de-trended by removing the one-year moving average. Hereafter, a tidal harmonic analysis was performed using the t-tide package (Pawlowicz et al., 2002), separating the astronomical tide (η_{tide}) and the non-tidal residual (η_{RE}) components. Note that η_{RE} not only contains the meteorological contributions, but also a non-linear tidal-surge interaction component (Haigh et al., 2016) and the combination of both components results in the total water level (TWL):

$$\text{TWL}=\eta_{\text{tide}} + \eta_{\text{RE}} \tag{1}$$

The temporal resolution of the water level measurements normally varies from few minutes to 1 h; while few stations recorded data during the entire validation period (Fig. 1a), 88.4% and 57.5% of the tidal stations respectively cover more than 20% and 50% of the validation period (Fig. 1b). Additionally, 55.1% of the 192 tidal gauges used for validation provide data for more than 10 years and only 13.6% of the tidal gauges cover less than 5 years (Fig. 1c).

The model validation was performed along the European coastline considering each tide gauge location and its corresponding nearest model grid point. The model's ability to reproduce the astronomic tide is evaluated in terms of the BIAS (ε_A), relative BIAS (δ_A) and BIAS for the tidal phase (ε_P), for each tidal constituent:

$$\varepsilon_A^i = a_o^i - a_p^i \tag{2}$$

$$\delta_A^i = \frac{a_o^i - a_p^i}{a_p^i} \tag{3}$$

$$\varepsilon_P^i = p_o^i - p_p^i \tag{4}$$

In addition, the model accuracy to reproduce the tide is evaluated by calculating the vectorial difference of the main diurnal and semi-diurnal constituents: the root mean square deviation of the amplitude (d):

$$d_j^i = \sqrt{\left(a_{j,o}^i cosp_{j,o}^i - a_{j,p}^i cosp_{j,p}^i\right)^2 + \left(a_{j,o}^i sinp_{j,o}^i - a_{j,p}^i sinp_{j,p}^i\right)^2} \tag{5}$$

where a is the amplitude (m), p is the phase (degrees), subscript j refers to the tidal gauge station, subscript i refers to the tidal constituent, subscript o is the observed data from the tide gauge and subscript p is the model's predicted result.

The root mean square deviation of the amplitude (RMS) for each constituent is defined as follows (Tsimplis, 1995):

$$RMS^i = \sqrt{\frac{1}{2N}\sum_{j=1}^{N}(d_j^i)^2} \tag{6}$$

where N is the number of the tide gauges considered and d_j^i is the vectorial difference defined for each location and each tidal constituent.

To evaluate the total effect of the main tidal constituents the root sum of square (RSS) (Arabelos et al., 2011; Ferrarin et al., 2013) is computed as:

$$RSS = \sqrt{\sum_{i=1}^{n} RMS_i^2} \tag{7}$$

The ability of the model to reproduce the storm surge, non-tidal residual and total water level was based on the root mean square error (RMSE) and the corresponding relative root mean squared error (% RMSE):

$$RMSE = \sqrt{\frac{\Sigma_k^n \left(\eta_o^k - \eta_p^k\right)^2}{n}} \tag{8}$$

$$\%RMSE = \frac{\sqrt{\frac{\Sigma_k^n \left(\eta_o^k - \eta_p^k\right)^2}{n}}}{\max(\eta_o)} \cdot 100 \tag{9}$$

where n is the number of measurements in the time series at a given location, η_o is the water level from measurements (observed non-tidal residual or TWL), and η_p is the predicted water level, storm surge or non-tidal residual.

The effect in model performance of changing the model setup was evaluated in terms of RMSE difference computed as:

$$\Delta RMSE = RMSE_{adv} - RMSE_{base} \tag{10}$$

where $RMSE_{adv}$ is the advanced configuration (i.e. including tidal process or ECMWF-HRES atmospheric forcing) and $RMSE_{base}$ is the base configuration (i.e. omitting tidal process or ERA-INTERIM atmospheric forcing).

The effect of the tidal-surge interaction and atmospheric forcing accuracy/resolution was evaluated in terms of $BIAS_{adv}$ for the whole time series computed as:

$$BIAS_{adv} = \frac{1}{N} \sum (\eta_{adv} - \eta_{base}) \tag{11}$$

Where η_{adv} correspond to the storm surge, non-tidal residual or total water level achieved using the advanced configuration and η_{base} correspond to the base configuration. In the case of the extremes values the effect will be evaluated in terms of the differences in percentile P99 ($\Delta P99 = P99_{\eta_{adv}} - P99_{\eta_{base}}$).

3. Results

3.1. Tidal prediction validation

The present section discusses validation results related to the principal tidal constituent semi-diurnal (M2, S2) and diurnal (O1 and K1) obtained from the E0 numerical experiment, as well as of the TPXO7.2, FES 2004 and FES 2012 tidal models. The main semidiurnal tidal constituent M2 is reproduced satisfactorily in the model with relative BIAS $\delta_a < 15\%$ for most of the tidal gauge stations (Fig. 2). The highest δ_a is found along the Mediterranean and Baltic Sea (circa 50%), but the performance in these micro-tidal areas is less important and BIAS remains below $\varepsilon_a < 0.06$ m. This is also the case for >56% of the tidal gauge stations (Fig. A1b); while the maximum BIAS in M2 prediction is found at the eastern edge of the English Channel ($0.2 < \varepsilon_a < 0.4$ m). The M2 phase is also well reproduced, with $\varepsilon_p < 15^\circ$ for 74% of the tidal stations (Fig. A1c) and the worst performance observed along the Baltic Sea and southern North Sea (Fig. 2c).

The spatial pattern of the S2 constituent validation results is similar to the M2, with $\varepsilon_a < 0.015$ m and $\varepsilon_a < 0.09$ m for 44.7% and more than 15% of the tidal stations considered, respectively (Fig. 2e). The phase is also satisfactorily simulated (Fig. 2f), with $\varepsilon_p < 15\%$ for >76.9% of the tidal gauges (Fig. A1f). However, a poorer performance is observed along the Baltic coast, and the North Sea. The diurnal constituent K1 (Fig. 2g), shows the worst performance; however it contributes very little to the TWL in macro- and meso-tidal areas. Amplitude bias $\varepsilon_a < 0.01$ m is found for almost 55% of the tidal gauges considered (Fig. A1h), and $\varepsilon_a < 0.03$ m for > 85% of the tidal gauges. Similarly, the validation for the K1 tidal phase shows $0^\circ < \varepsilon_p < 15^\circ$ for 44.2% of the tidal gauge stations, and $15^\circ < \varepsilon_p < 30^\circ$ for 34.6% of the stations (Fig. A1i). The model's performance is better for the O1 diurnal tidal constituent with $\varepsilon_a < 0.01$ m for 77% of the tidal stations (Fig. 2k and **Fig. A1**k). The O1 phase error varies between $0^\circ < \varepsilon_p < 15^\circ$ for 50.5% of the tidal stations and from 15° to 30° for 30.3% of the tidal gauge stations (Fig. A1l). The absolute error in O1 phase is larger in the Irish Sea, west Mediterranean and Baltic Sea (Fig. 2l).

To better interpret the EU-SSF tidal prediction performance we have also performed a similar validation analysis for the tidal levels provided by the global tide models TPXO7.2, FES 2004 and FES2012. The total RSS for EU-SSF is equal to 0.198 m (Table 2), lower than the values estimated for FE2004 and TPXO7.2, which have respectively RSS values of 0.42 and 0.39 m. However, the FES2012 model outperforms EU-SSF with RSS = 0.148 m. Similarly, RMS values for the M2 and S2 components for EU-SSF are respectively 0.181 m and 0.071 m, markedly higher than the ones of FES2012 (0.14 m and 0.046 m).

3.2. Non-tidal residual water level validation

The non-tidal residual validation is performed after subtracting the astronomical tide signal estimated by harmonic analysis from the water levels generated from the E2 experiment. The model shows a good predictive skill, with 0.04 m < RMSE < 0.21 m (Fig. 3a, Table 3). More than 91% of the tidal gauges show RMSE values lower than 0.125 m (Fig. 3b), and only 2% of the stations result in RMSE > 0.15 m. In general, RMSE values are below 0.1 m, with the exception of the North Sea (0.10 m < RMSE < 0.14 m) and the East coast of the Baltic Sea (Gulfs of Bothnia, Finland and Riga; 0.1 m < RMSE < 0.15 m).

In areas characterized by lower non-tidal residual values the % RMSE shows the opposite trend to RMSE; e.g. see the highest values found in the Mediterranean Sea and S-North Atlantic. %RMSE ranges from 4% to 22% (Fig. 3c), with 70% of the tidal gauges showing values of %RMSE < 12.5%. Only 6.2% of the stations show %RMSE > 17.5% (Fig. 3d).

The model shows good performance in the reproduction of non-tidal residual extremes and the upper tail (>99th percentile) RMSE is below 0.25 for 79% of tidal gauge stations (Fig. A2b). The higher RMSE values are observed along the North Sea, west coast of Bay of Biscay, Gulf of Bothnia in the Baltic Sea, as well as the north Adriatic Sea (0.3 m < RMSE < 0.4 m; Fig. A2a). The %RMSE for the extreme events remains below 25% for most of the tidal gauge stations (64%) (Fig. A2d). The largest %RMSE values (30%–40%) are located in the Mediterranean and in the S-North Atlantic.

3.3. Benefits of coupling storm surge and tide

The inclusion of tidal processes in the simulations (case E2; see Table 1) allows to resolve the non-linear interactions between meteorological and astronomic tides. This clearly leads to an improvement in predictive skill compared to when tidal forces are omitted (E1). The RMSE reduces up to 0.033 m (14%; Fig. 4a). The improvement in performance is more pronounced along the North Sea and the English Channel, resulting in up to 2% reduction of the %RMSE (Fig. 4b).

Considering only the changes in the extreme events (>99th percentile), the differences in RMSE vary within -0.086 m < ΔRMSE < 0.037 m (Fig. A3a). Coupling with tides improves the forecasting skill of the extreme η_{RE} at the English Channel (0.02 m < ΔRMSE < 0.04 m), Bristol Channel (ΔRMSE < 0.086 m), South Irish Sea, Celtic Sea, and the N–North Atlantic region (0.01 < ΔRMSE < 0.03 m). Similarly, %RMSE appears to reduce by up to 4% in these areas (Fig. A3b). The opposite effect is observed along the central and eastern North Sea, western coast of the Jutland Peninsula, northern Irish Sea, and North Channel, where the %RMSE

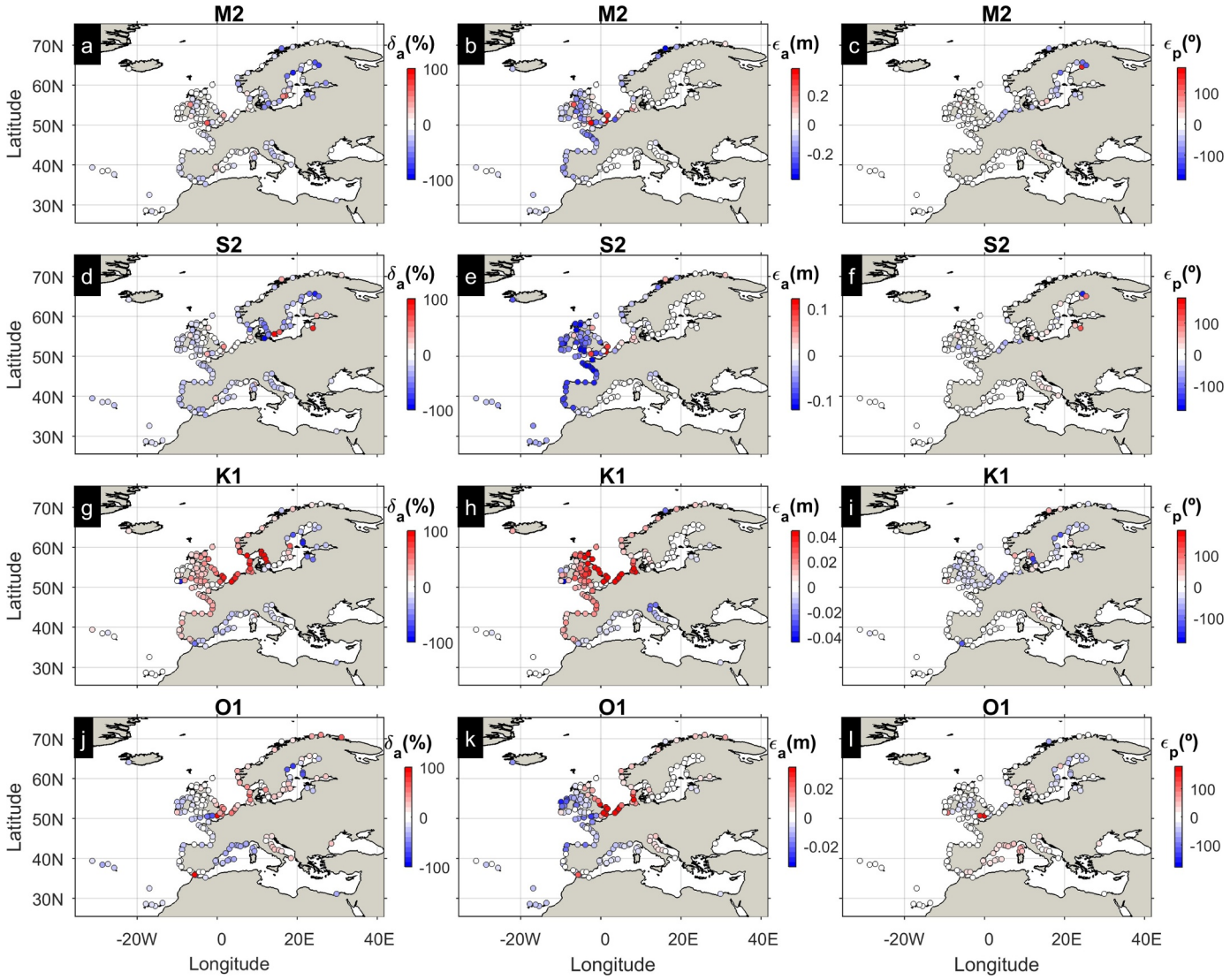

Fig. 2. Tidal validation results for the principal tidal constituents (M2, S2, O1, K1). Scatter plots of the relative amplitude bias (δ_a(%); a,d,g,j); absolute amplitude bias (ε_a(m); b,e,h,k); and absolute phase bias (ε_p(m); c,f,I,l). Points indicate the locations of the tidal gauges and warm/cold colours express over/under-estimation.

Table 2
Summary of tidal validation results against tidal gauge data and comparisons with TPXO7.2, FES 2004 and FES 2012. The table includes RMS and RSS for the main tidal components. Values are in metres.

	RMS semi-diurnal		RMS diurnal		RSS
MODEL	M2	S2	K1	O1	
EU-SSF	0.181	0.071	0.03	0.016	0.198
FES2004	0.397	0.145	0.019	0.014	0.423
FES2012	0.14	0.046	0.012	0.008	0.148
TPXO7.2	0.365	0.129	0.013	0.01	0.386

increases around 0.01–0.02 m (1%–2%).

3.4. Sensitivity to atmospheric forcing accuracy/resolution

The validation considering the whole time series allowed us to find that the higher resolution atmospheric forcing (case E3; see Table 1) allows a reduction of the η_{SSL} RMSE by 0.02 m along the Norwegian and Adriatic Sea, as well as most of the west Mediterranean in comparison to the ERA-INTERIM (E1) run. In addition, a RMSE reduction of 0.03 m is observed in the western coast of the Jutland Peninsula (Fig. 5a). The %RMSE is reduced up to 2% in the same areas. The opposite effect is observed in the Baltic Sea, Bay of Kattegat, Bay of Biscay, N-North Atlantic, and west-southwest coast of the North Sea. In those areas using the high resolution forcing the RMSE increases by 0.01 m–0.04 m. The increase in %RMSE is around $2 < \%RMSE < 4$ (Fig. 5b).

A similar analysis considering only the extreme storm surges (upper tail distribution above 99th percentile showed in Fig. A4) shows that the higher resolution forcing (ECMWF-HRES) improves the predictive skill in enclosed basins (Fig. A4a). For example, RMSE reduces by 0.05 m in the Gulf of Venice (Adriatic Sea), Gulf of Bothnia (Baltic Sea) and Black Sea. The %RMSE is reduced by 10% in the Gulf of Venice and Gulf of Bothnia, and by 15% in the Black Sea. A smaller (5%) %RMSE reduction is observed for the Mediterranean, Norwegian and Baltic Sea. Nevertheless, the higher resolution forcing appears to produce an increase (4%) in %RMSE along the N-North Atlantic and the English Channel (Fig. A4b).

4. Discussion

4.1. Tidal predictive skill

Our EU-SSF satisfactorily describes tidal dynamics along European coastlines, but its performance may still be poorer than that of regional

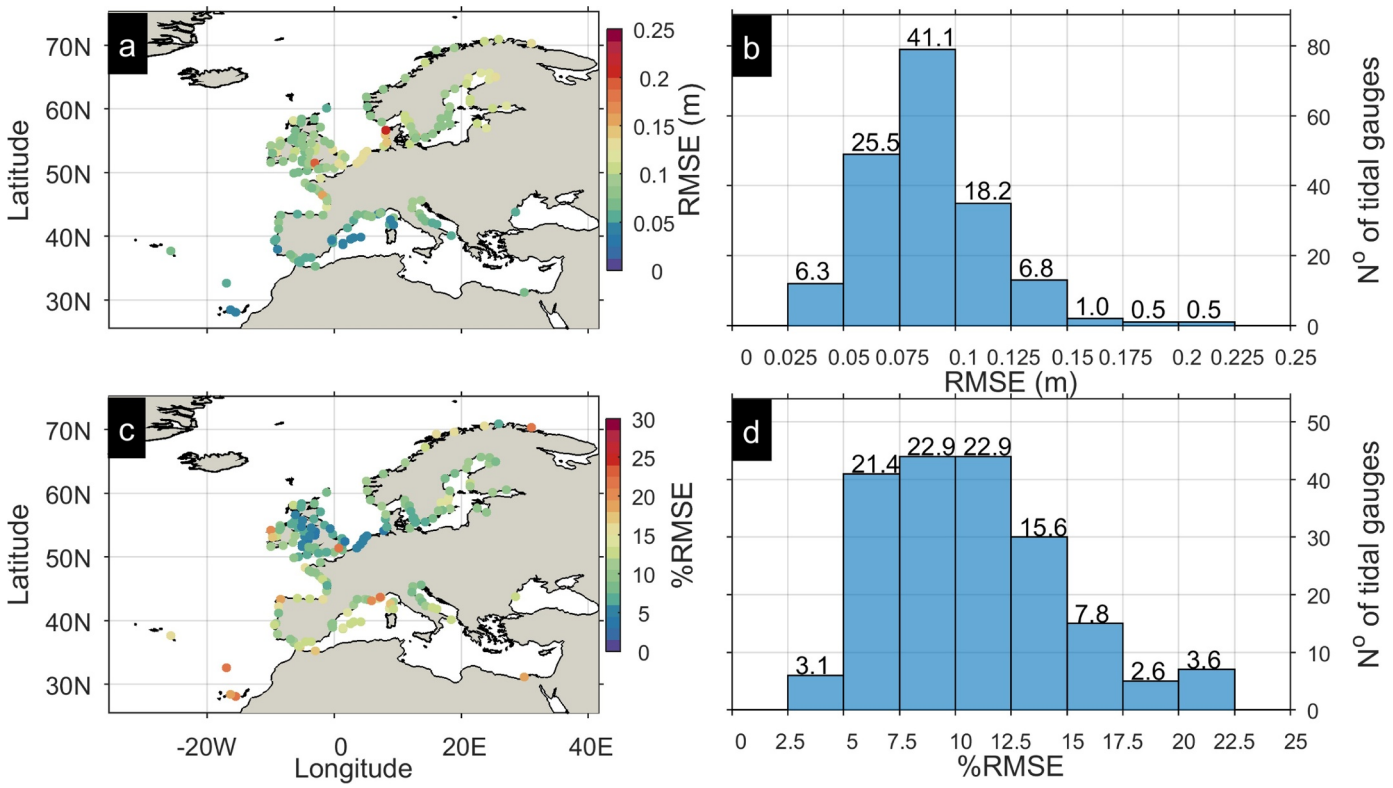

Fig. 3. Non-tidal residual (η_{RE}) validation results at the tidal gauge stations considered for the coupled tidal and storm surge simulation (E2). (a,c) Map scatter plots of RMSE, %RMSE respectively; (b,d) RMSE and %RMSE histograms for all tidal gauges with the vertical axis showing the count and the text labels above the bars the percentage of all tidal gauges belonging in the specific bin.

models. For example, in the central Mediterranean Sea, Ferrarin et al. (2013) reported a RMSE < 0.01 m for the main tidal constituents (M2,S2,K1 and O1) and a RSS of 0.015 m. These values are lower than the 0.009 m < RMSE < 0.026 m and RSS = 0.04 m reported herein. Tsimplis et al. (1995) reported for the whole Mediterranean Sea a 0.008 m < RMSE < 0.014 m and a RSS = 0.023 m, compared to the 0.008 m < RMSE < 0.027 m and RSS = 0.036 estimated by our EU-SSF.

For the western Iberian peninsula, Quaresma et al. (2013) reported a BIAS between $0.006 < \varepsilon_a < 0.04$ m for the amplitude of M2 and S2, and for the phase between $0.4 < \varepsilon_p < 7^{\circ}$. In comparison, our EU-SSF shows an amplitude and phase bias of respectively $0.003 < \varepsilon_a < 0.18$ m and $0.3 < \varepsilon_p < 37^{\circ}$ for S-North Atlantic. In the Bay of Biscay, Bertin et al. (2012) report an amplitude BIAS $0.01 < \varepsilon_a < 0.04$ m and a phase BIAS below 11° for M2, whereas our EU-SSF M2 amplitude BIAS is between $0.07 < \varepsilon_a < 0.35$ m and the phase BIAS is between $1.4 < \varepsilon_p < 13^{\circ}$. Idier et al. (2017) report an amplitude relative BIAS for the North Sea and N-North Atlantic, for constituents M2, S2, N2, K2, and M4 below 5% in most of the tidal stations, values which are lower than the presently reported mean relative amplitude BIAS ($4\% < \delta_a < 14\%$) calculated for M2 and S2. However, our EU-SSF outperforms most global tidal models (Table 2, Fig. A5), with the exception of FES2012. The latter does not come as a surprise since FES2012's skill is enhanced by data assimilation of long-term altimetry data (Topex/Poseidon, Jason-1, Jason-2, ERS-1, ERS-2 and ENVISAT (Carrère et al., 2012).

As in previous studies, the accuracy in tidal solution is reduced in shallow water areas with complex bathymetry and geometry (Zijl et al., 2013). Tidal dynamics are strongly affected by the bathymetry in those areas and corresponding transformation processes are not fully resolved for various reasons. These include the limit in the amount of detail in the bathymetric data, the resolution of the computational grid (10 km along the coast), and the simplifications imposed by a 2D barotropic model. These aspects are particularly important in areas with complex

Table 3
Statistics of the model performance to reproduce the non-tidal residual (η_{RE}) along the 10 defined European regions for the coupled tidal and storm surge simulation (E2).

Region	RMSE (m) Mean	Std	Median	Min	Max	%RMSE (%) Mean	Std	Median	Min	Max
Black Sea	0.05	–	–	–	–	13	–	–	–	–
East Med.	0.06	–	–	–	–	19	–	–	–	–
Central Med.	0.07	0.02	0.07	0.04	0.09	11	3	11	8	17
West Med.	0.06	0.01	0.06	0.04	0.09	14	3	13	10	21
S-North Atlantic	0.06	0.01	0.06	0.05	0.08	14	5	12	8	22
Bay of Biscay	0.09	0.02	0.09	0.06	0.15	12	3	12	7	19
N-North Atlantic	0.09	0.02	0.09	0.07	0.19	9	4	8	4	21
North Sea	0.14	0.03	0.10	0.06	0.21	8	3	7	4	19
Baltic Sea	0.10	0.02	0.10	0.08	0.15	10	2	10	7	16
Norwegian Sea	0.10	0.02	0.10	0.08	0.13	14	5	14	7	21

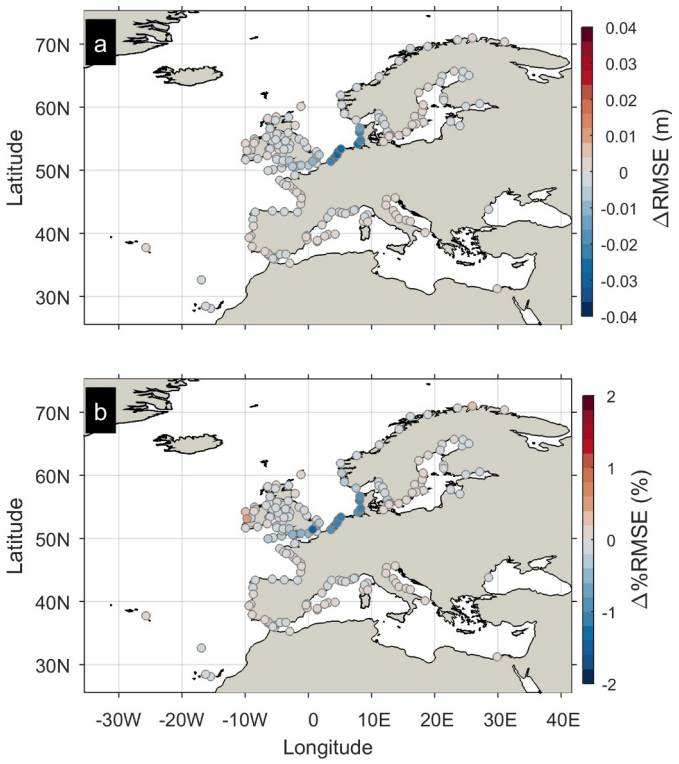

Fig. 4. Difference between storm surge (η_{SSL}) and non-tidal residual (η_{RE}) predictive skill obtained from simulations which respectively omit or include tide-surge coupling. Scatter plots show differences in absolute (a) and %RMSE (b) for all the considered tidal gauges. Warm colours imply that the run which omits tidal processes has a higher predictive skill or lower error.

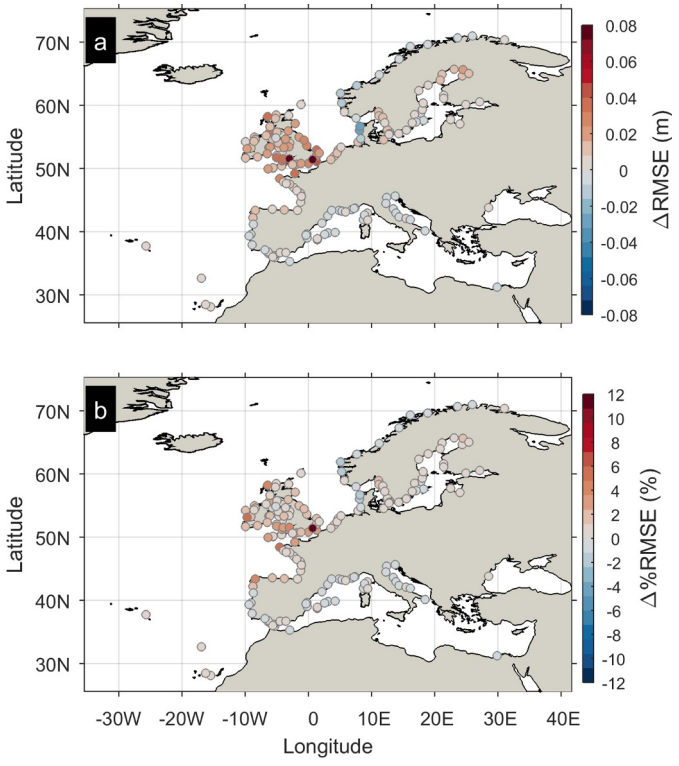

Fig. 5. Difference in storm surge (η_{SSL}) predictive skill between simulations forced by ERA-INTERIM and the higher resolution forecast (ECMWF-HRES). Scatter plots show differences in absolute (RMSE; a) and relative RMSE (% RMSE; b) for all the considered tidal gauges. Warm colours imply that the ERA-INTERIM run has higher predictive skill.

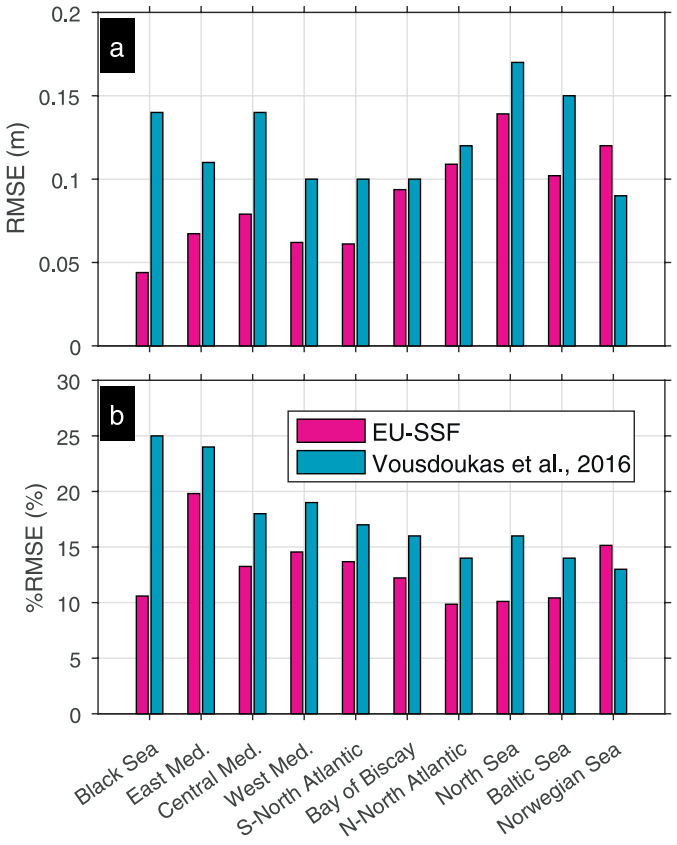

Fig. 6. Model performance comparison of the storm surge component (η_{SSL}) from the E1 simulation with the European storm surge model presented by Vousdoukas et al. (2016): Comparisons of RMSE (a) and %RMSE (b), averaged for each of the different regions defined in Fig. 1.

tidal dynamics such as the German Bight and/or exchange areas between two ocean basins characterized by different seawater properties, e.g. the Gibraltar Strait or the Baltic Sea.

4.2. Non-tidal residual and storm surge water level prediction

Validation results demonstrate that the EU-SSF satisfactorily reproduces the non-tidal residual water levels. The storm surge calculated from E1 improves substantially the model presented by Vousdoukas et al. (2016a), as (i) the use of an unstructured grid allows a higher grid resolution nearshore without increasing the computational cost (it takes 0.95 CPU hours per month simulated running in parallel using 32 processors with 2 CPUs of 2.6 GHz); (ii) EU-SSF uses a higher resolution bathymetry (EMODnet) compared to the GEBCO used previously. As a consequence, the EU-SSF results have a lower RMSE in all the oceanographic basins compared to Vousdoukas et al. (2016a), with the exception of the Norwegian Sea, where RMSE increases by 0.03 m (Fig. 6a). The average RMSE is reduced by 0.03 m, ranging from 0.01 m in the Bay of Biscay, to 0.1 m in the Black Sea. Similarly, the mean %RMSE reduces by 5%, ranging from 3% in the S-North Atlantic, to 14% in the Black sea; with the exception of the Norwegian Sea where the %RMSE increases by 2% (Fig. 6b).

Direct comparison with validation results presented in previous studies is not straightforward, due to inconsistencies in the validation periods and the ground-truth sea level data. Even though most previous studies cover smaller domains, which allows modelling at higher resolution, the performance is found to be comparable. For example, the non-tidal residual RMSE along the Dutch coast presently ranges from 0.12 to 0.14 m, while Sembiring et al. (2015) report RMSE values with annual variation within 0.1 m and 0.2 m. O'Neil et al. (2016) report

Table 4
Statistics of model performance along the 10 defined European regions for the EU-SSF, GLOBAL and NWS (North-West shelf) from CMEMS (http://marine.copernicus.eu).

Region	RMSE(m) EU-SSF Mean	Std	RMSE (m) CMEMS-GLOBAL Mean	Std	RMSE (m) CMEMS-NWS Mean	Std
Black Sea	0.05	–	0.13	–	–	–
East Med.	0.06	–	–	–	–	–
Central Med.	0.07	0.02	0.08	0.03	–	–
West Med.	0.06	0.01	0.08	0,02	–	–
S-North Atlantic	0.06	0.01	0.08	0.003	0.04	–
Bay of Biscay	0.09	0.02	0.11	0.02	0.09	0.03
N-North Atlantic	0.09	0.02	–	–	0.09	0.02
North Sea	0.14	0.03	0.17	0.08	0.08	0.03
Baltic Sea	0.10	0.02	0.16	–	0.10	–
Norwegian Sea	0.10	0.02	0.10	0.02	0.04	–
Regional average	0.08		0.11		0.07	

average RMSE around 0.1 m along the Scottish and English coast, similar to 0.09 m presented. In the French Atlantic and English Channel coast, Muller et al. (2014) report RMSE values around 0.08 m, while present values are 0.10 m. For the non-tidal residual at the Bay of Biscay during the Xynthia storm, Bertin et al. (2012) report RMSE values ranging from 0.12 to 0.17 m compared to 0.11 < RMSE < 0.35 m presently reported for the upper tail values (99th percentile). However, the model in Bertin et al. (2012) uses higher spatial and temporal resolution than ERA-INTERIM and is specifically calibrated for the area affected by the Xynthia storm. Moreover, the presently estimated RMSE = 0.1 m for the Baltic Sea is similar to the one reported for a finer resolution model (~5.5 km resolution) (Kowalewski and Kowalewska-Kalkowska, 2017). In southern Europe and central west Mediterranean Cid et al. (2014) report an average RMSE of 0.08 m, slightly higher than the RMSE of 0.06 m in our EU-SSF.

Table 4 shows a comparison of the EU-SSF (non-tidal residual from E2) predictive skill with respect to the global (1/12° resolution) and North-West shelf sea surface height analysis provided by CMEMS (1/15° latitudinal and 1/12° longitudinal resolution). The global model does not account for tidal and barometric effects. Therefore, the DAC correction has been linearly added to the original dataset before the comparison with the non-tidal residual extracted from the tidal gauge database employed to validate EU-SSF. In the case of the North West shelf (NWS), which accounts for the tide, the non-tidal residual is extracted before the comparison with the ground truth dataset.

The EU-SSF shows a regional average RMSE of 0.08 m, compared to the 0.11 m calculated for the CMEMS-GLOBAL (Table 4). Significant improvement of the predictive skill is identified in all the regions with the exception of the Norwegian Sea. Additionally, a wide range of RMSE was reported for the CMEMS-GLOBAL as indicated by the larger standard deviation of the RMSE. The EU-SSF only shows a slightly poorer performance in comparison with the higher resolution model CMEMS-NWS (regional average RMSE 0.08 m and 0.07 m, respectively). CMEMS-NWS performs substantially better than EU-SSF along the North Sea and the Norwegian Sea (RMSE reduction 50%). This is not surprising since the finer resolution of the CMEMS-NWS allows to capture the tidal dynamics and surge propagation in the North Sea and German Bight. Part of the superior performance of CMEMS-NWS could be attributed to the data assimilation which is included in the system's workflow.

4.3. Importance of resolving tide-surge interactions

Considering tidal forces is important for storm surge forecasting as

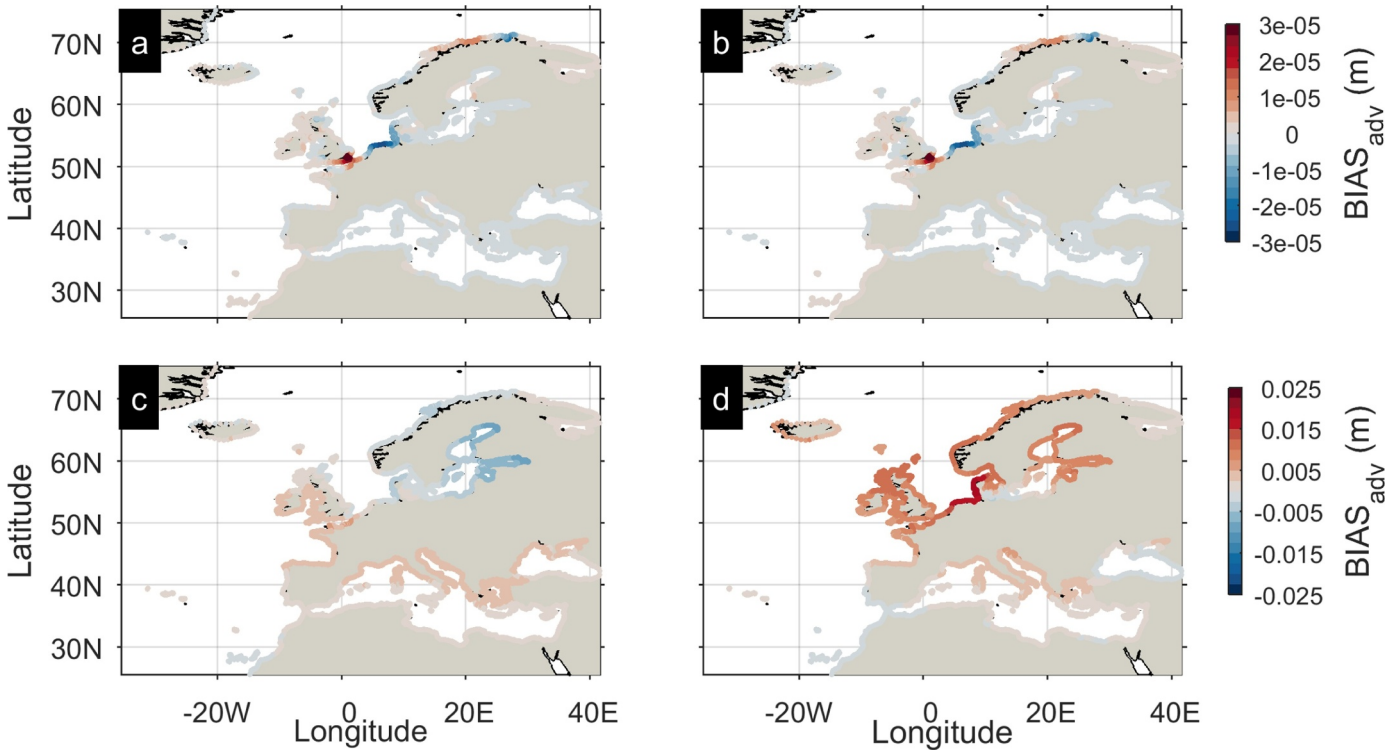

Fig. 7. Effect of tidal-surge interaction and atmospheric forcing accuracy/resolution in the water level components along European coastline. (a) BIAS comparing the non-tidal residual vs the storm surge water level; the former is obtained by subtracting the water levels of the only tidally-forced simulation E0, from the coupled run E2; the latter is obtained from the simulation E1 which omits tidal forcing. (b) BIAS comparing the total water level (TWL) from E2 with that obtained by the linear addition of E0 and E1. (c) BIAS comparing the storm surge level obtained from the high resolution atmospheric forcing (no tidal forcing; E3) and E1. (d) BIAS of the TWL from the high resolution atmospheric forcing coupled run (E4) and E2. Warm colours indicate that most advanced configuration (coupled, or higher resolution) produces higher water level values.

they allow resolving (i) the interaction between astronomic and meteorological flows; and (ii) the tidal water depth modulations which drive changes in bottom friction and on the overall depth-dependent circulation patterns (Arns et al., 2015). Thus coupled tidal-storm surge models usually have higher skill (Brown et al., 2010; Zijl et al., 2013), especially along meso- and macro-tidal areas that are strongly affected by storm surge phenomena (North Sea and N-North Atlantic; Fig. 4). For that reason on the coast of the Jutland Peninsula and in the southern coast of the North Sea the non-tidal residual is lower than the storm surge estimated without resolving tide-surge interactions (see negative $BIAS_{adv}$; Fig. 7a,b). An inverse pattern is observed in the western English Channel as indicated by the scarcely positive $BIAS_{adv}$. Larger differences exist in the case of the extreme events (upper tail event above the 99th percentile). The non-tidal residual is lower than the storm surge in the western coast of the Jutland Peninsula and along the southern coast of the North Sea, while the opposite effect is observed along the west coast of Ireland, the Celtic and Irish Seas, and the Bristol Channel (Fig. A6a). Such findings are in agreement with previous studies (Horsburgh and Wilson, 2007).

The spatial pattern of the observed storm surge amplification/reduction is related to the shape of the tidal wave and the depth dependence of the surge production and propagation processes. As the water level gradient is proportional to the wind stress divided by the water depth (Pugh, 1987), then the storm surge is higher in shallow depths. Storm surges and tides can be interpreted as shallow water waves with propagation velocity $(g{\cdot}h)^{1/2}$, where h is the water depth and g is the gravitational acceleration. Based on the above, the depth decrease results in a reduction of the storm surge propagation velocity and increase of the amplitude (Idier et al., 2017). As a result asymmetries in the tidal distribution can enhance or reduce non-tidal residuals, depending on whether they result in higher or lower frequency of water levels below MSL.

For example in the case of West-Terschelling (North Sea) the tidal distribution is bimodal and negative skewed, favouring positive tides (Fig. 8a); while the distribution of the non-tidal residual water level is bimodal with the two peaks correlated to low and high tidal levels, respectively (Fig. 8b). The result is that while the mean non-tidal residual is higher during negative tides, the mean η_{RE} for the extremes (above 99th) is higher during positive tides (in both cases statistical significance of the differences of the means is tested using two-sample *t*-tests for *p*-value < 0.05). Therefore omitting tides results in over-estimation of the non-tidal residual (uncoupled vs coupled run; see also Fig. A6a).

The contrary is expected to happen (storm surge under-prediction) in areas where tides follow an asymmetrical bimodal distribution with water levels below MSL being more frequent; like the station of Hinkley Point (Bristol Channel; Fig. 8a). The result is a unimodal negatively skewed joint distribution of η_{RE} and tidal levels, which shows that extremes occur mostly during negative tides (Fig. 8c). This fact has been confirmed by the two-sample *t*-test that reveals a statistically significant higher mean (p-value < 0.05) for the extreme surges (above 99th) occurred during negative tides at Hinkley Point. The underestimation of the non-tidal residual when tides are omitted is shown is shown in the uncoupled vs coupled run comparison (Fig. A6a). We have also estimated the conditional probability of η_{RE} exceeding the 99th percentile during negative and positive tides and we see that in both sites it is higher during low tide than during high tide (1.02% vs 0.88% for West–Tershelling and 1.22% vs 0.74% for Hinkley Point). The above confirms that extreme η_{RE} values are favored by negative tides.

Large-scale models have a tendency to under-estimate extreme storm surges (Calafat et al., 2014). The shortcoming above can paradoxically result in an improved model performance in areas where omitting tide-surge interactions results in an over-prediction of storm surges. This is the case of the Bristol Channel, when only extreme events are considered (Fig. A3). On the other hand, in areas like the western coast of the Jutland Peninsula and the southern coast of the North Sea (Fig. A6b), the storm surges are higher than the non-tidal residual, resulting in further under-estimation, even though the setup is improved. Of course all the above are not model setup recommendations, but insights on potential interactions between epistemic error factors which sometimes can cancel each other.

4.4. Importance of atmospheric forcing accuracy/resolution

If the whole time series is considered, the atmospheric resolution results in minor changes in storm surge ($\boldsymbol{BIAS_{adv}} < = 0.005$ m)

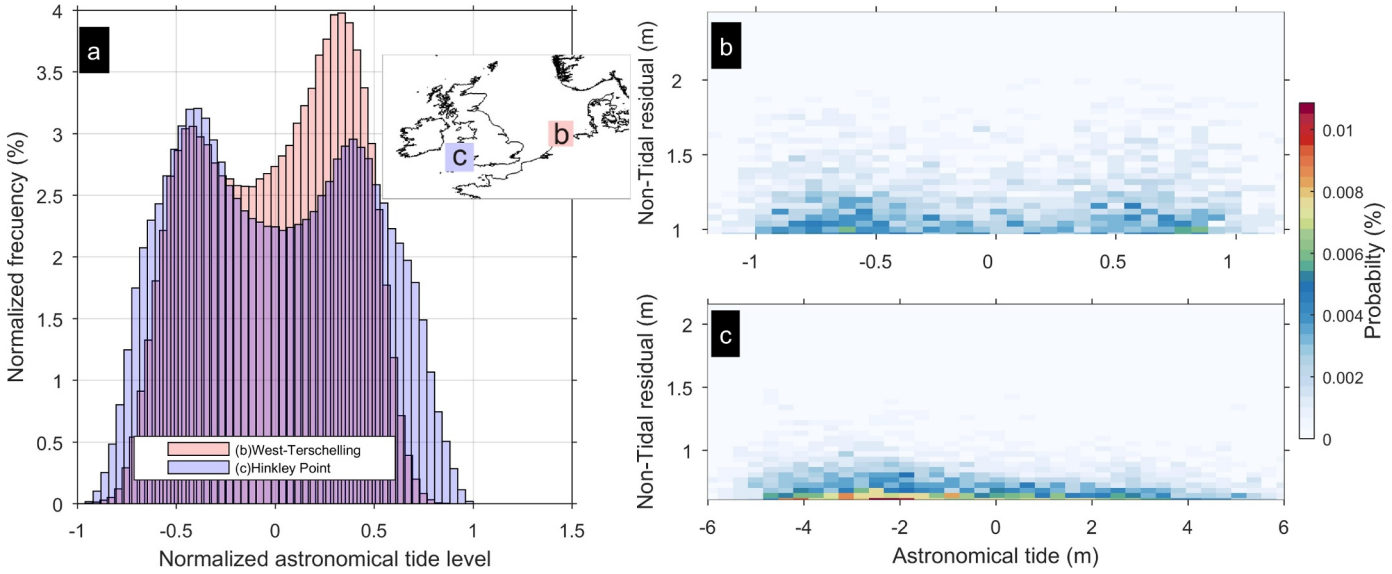

Fig. 8. (a) Histogram of normalized astronomical tide level. Probability distribution of the astronomical tide (η_{tide}) and the non-tidal residual (η_{RE}) above percentile 99th at West-Terschelling (b) and Hinkley Point (c).

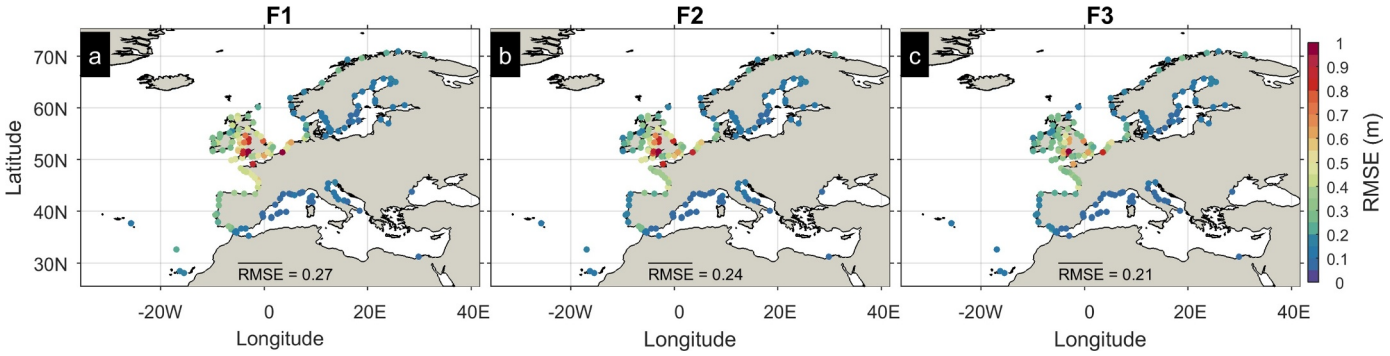

Fig. 9. Inter-comparison of RMSE for the different configurations tested for the TWL forecast. (a) F1: Direct calculation from a coupled run (E2; tides and storm surge). (b) F2: Computation from linear addition of the FES2012 tidal model solution to the storm surge estimated after omitting the tidal flows (E1). (c) F3: Linear addition of the FES2012 tidal model solution to the non-tidal residual extracted from E1.

(Fig. 7c), although the relative importance becomes higher when the tidal effects are included. The larger values of $\boldsymbol{BIAS_{adv}}$ (up to 0.02 m in Jutland Peninsula and the southern coast of the North Sea) observed in the case of TWL (Fig. 7d) reveal the relevance of the atmospheric resolution for the computation of tide-surge interaction.

Conversely to what was expected, using the high resolution atmospheric forcing leads to lower predictive skill in terms of storm surge levels along most of the European Atlantic ocean coast (Fig. 5). This can be attributed to the fact that the ECMWF-HRES forcing is a forecast product, in comparison to the more accurate ERA-INTERIM dataset that is post-processed through data-assimilation. In addition, using the full temporal resolution of ECMW-HR (1 h) instead of 6 h, used for consistency for comparison with ERA-INTERIM, could improve the final performance of the forecast system. Still, the achieved results have allowed the identification of critical areas in which atmospheric forcing resolution plays a key role in the accuracy of the storm surge prediction; i.e. semi-enclosed basins, like the Black, Baltic, Adriatic, and Mediterranean Sea (Fig. 5). The benefit of increasing the atmospheric forcing resolution for the predictive skill of the extreme storm surges is evident (see RMSE and %RMSE reduction in Fig. A4). This is due to the tendency of large-domain models to underestimate upper-tail events (Muis et al., 2016; Vousdoukas et al., 2016a), particularly the ones characterized by short duration and high energy (Calafat et al., 2014; Conte and Lionello, 2013). As a result of the above, an increased atmospheric resolution allows to ameliorate the underestimation of storm surge and TWL (Fig. A6c and d).

The greater model sensitivity to the atmospheric forcing resolution in semi-enclosed basins is related to the smaller spatial scales and shorter life-cycles of low pressure systems, which demand higher resolution and accuracy in the atmospheric forcing. This is the case for the Mediterranean where pressure-lows are less intense and show smaller spatial and temporal scales (Trigo et al., 1999), and particularly in the Adriatic Sea, where storm surge events are of higher frequency and shorter duration (Cid et al., 2016). In a similar fashion, higher frequency, short duration events have been also reported in the Gulf of Bothnia, Finland and Riga in the Baltic Sea (Wolski et al., 2014). In addition, the aforementioned areas (along with the Azov Sea) are all characterized by shallow waters. Here wind is the main driver of surge production since the water level gradient is proportional to the wind stress divided by the water depth (Pugh, 1987). Consequently, the accuracy of wind representation becomes more relevant.

4.5. Towards an operational pan-European forecast

The following paragraphs discuss potential extensions of the present work, towards an operational pan-European storm surge forecasting system. The appropriate forcing for such application is the high resolution ECMWF forecast; while given the scope, the critical prediction variable is the combined TWL. As a result we assess the model predictive skill considering the following alternatives:

F1: using the output from the coupled setup
F2: combining linearly the tidal forecast from FES2012 with the storm surge prediction from the uncoupled setup (i.e. without tidal flows)
F3: combining linearly the tidal forecast from FES2012 with the non-tidal residual prediction from the coupled setup

Option F3 results in the best performance (RMSE = 0.21 m; Fig. 9c), reducing the average RMSE by 0.06 m compared to option F1 (RMSE = 0.27 m; Fig. 9a) and by 0.03 m with respect to option F2 (RMSE = 0.24 m; Fig. 9b). The improved performance is mainly observed in meso- and macro-tidal areas. A similar spatial pattern, although with larger RMSE reduction (0.04 m from F1 to F2 and 0.07 m from F2 to F3), is observed for the extreme TWLs (> 99th percentile; see Fig. A7). The overall performance of option F3 is considered acceptable for the considered domain, although it is poorer in comparison to the validation results reported by regional models. See for example the mean RMSE = 0.35 m found along the North Atlantic, North Sea, Skagerrak and Kattegat, with respect to RMSE = 0.10 m reported by Zijl (2013). However, in most cases RMSE for TWL remains below 0.25 m, and the TWL prediction error is mainly related to poor skill in reproducing the tidal levels in meso- and macro-tidal areas (Fig. 2).

The above solution combines the benefits from (i) resolving the non-linear interactions between meteorological and astronomic tides in the model in terms of non-tidal residual forecasting; and (ii) having the

most accurate tidal forecast benefiting from the data-assimilation applied in the FES2012 model.

4.6. Foresight

The spatial extent of the model domain imposes some inevitable limitations in the applied setup. Most of these can be addressed with further developments in computational power and code capabilities. Potential model upgrades include increasing the resolution of the computational mesh, the atmospheric forcing and the bathymetry along shallow near-shore areas with complex topography, e.g. bays, ports and estuaries. In these areas the more dynamic circulation processes and complex bathymetry are not fully resolved by the model. The same applies to semi-enclosed basins (e.g. Baltic, Adriatic) and narrow straits with complex circulation (e.g. Gibraltar, English Channel, North Channel in the Irish Sea and Kattegat in Denmark) (Fig. 2). Increasing the resolution would improve model performance in all aspects, and in particular in terms of the representation of the tidal dynamics and the non-linear interactions with storm surge.

The tidal predictive skill in the TWL forecast could be improved after testing and incorporating output from the more recently developed FES2014 (Carrere et al., 2016). A step even further would be to include waves, which contribute to extreme sea levels and thus coastal hazard through wave setup (Vitousek et al., 2017; Vousdoukas et al., 2016b) and runup (Serafin and Ruggiero, 2014; Vousdoukas, 2014). In addition, the inclusion of waves can enhance the model's skill in storm surge estimation (Ferrarin et al., 2013), since (i) wave-current interaction is important in certain locations (Roland et al., 2012); (ii) waves increase the roughness of the sea surface, which controls wind drag and energy transfer from the atmosphere to the ocean, resulting in circulation and wind set-up (Bertin et al., 2015).

Finally, the spatial and temporal coverage of the tidal gauge database is poorer along the East Mediterranean and Black Sea, acting as a bottleneck in model validation and calibration. The use of a complementary dataset from altimetry satellite data could help to assess the model performance in those areas.

5. Conclusions

This contribution presents the development and validation of an unstructured hydrodynamic storm surge and tidal model for Europe: EU-SSF. The tidal surge model accounts for atmospheric pressure, wind and astronomical tide. The model has been validated against sea level measurements from 192 tidal gauge stations in terms of the skill to predict tidal and meteorological water levels, as well as their combination.

Overall, EU-SSF shows satisfactory performance. Regarding tidal prediction, the total RSS is equal to 0.197 m, lower than the values estimated for FE2004, and outperformed only by FES2012 which results in RSS = 0.05 m. Also EU-SSF reproduces satisfactorily the principal tidal constituents M2, S2, O1 and K1. The main semidiurnal tidal constituent M2 is reproduced satisfactorily with average amplitude and phase BIAS found equal to 0.07 m and 14°, respectively.

Non-tidal residual water level validation results show good model predictive skill with RMSE = 0.08 m and %RMSE = 11%, respectively. Including tidal effects results in an improvement in non-tidal residual water level prediction, reducing the RMSE by up to 0.033 m. The areas benefiting more from the coupling are the North Sea and the English Channel, resulting in up to 2% reduction of the %RMSE. The improvement is even more pronounced for the prediction of extreme non-tidal residual water levels (> 99th percentile), for which the differences in RMSE ranges within -0.086 m < ΔRMSE < 0.037 m.

Increasing the resolution of atmospheric forcing also improves the predictive skill, especially in terms of the extremes in shallow areas where wind is the main driver for surge production; like the Azov Sea, as well as the Gulf of Venice (−0.06 m < ΔRMSE < -0.04 m), Bothnia (−0.06 m < ΔRMSE < -0.01 m), the area including the Gulfs of Finland and Riga (−0.01 m < ΔRMSE < 0 m).

The pan-European storm surge model presented in this work and the findings related to the sources of uncertainty can be useful for the design of large-scale storm surge forecasting systems. For the present application we find that the most accurate setup is the one combining tidal levels from the FES2012 model and the non-tidal residual from the EU-SSF setup which couples meteorological and astronomic tides. Such a system can be essential for disaster risk reduction supporting the coordination of real-time emergency response against surge coastal flooding.

Future work will include several developments which are expected to increase the EU-SSF performance: (i) coupling the current model with a wave generation/propagation one; (ii) increasing model resolution in specific areas in order to properly reproduce the tidal dynamics; and (iii) including data assimilation in the forecasting system.

Acknowledgments

The research leading to these results has received funding from the EUh2020 program under EU-H2020 grant agreement no 700099 (ANYWHERE: EnhANcing emergency management and response to extreme WeatHER and climate Events.

Appendix

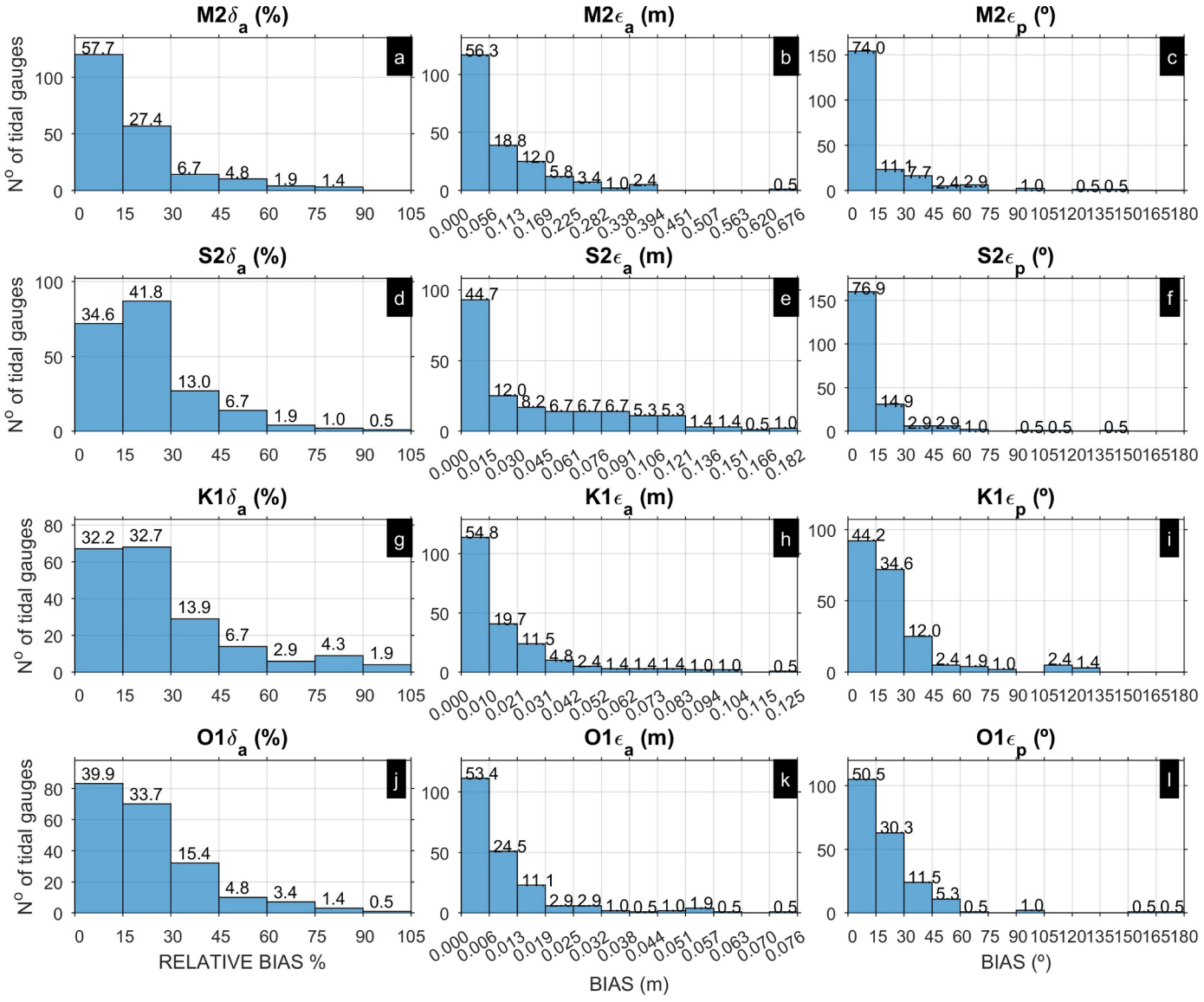

Fig. A1. Tidal validation results of the principal semidiurnal tidal constituents (M2, S2) and diurnal (O1, K1). Histogram of the relative amplitude bias (δ_a(%); a,d,g,j); absolute amplitude bias (ϵ_a(m); b,e,h,k); and absolute phase bias (ϵ_p(m); c,f,i,l).

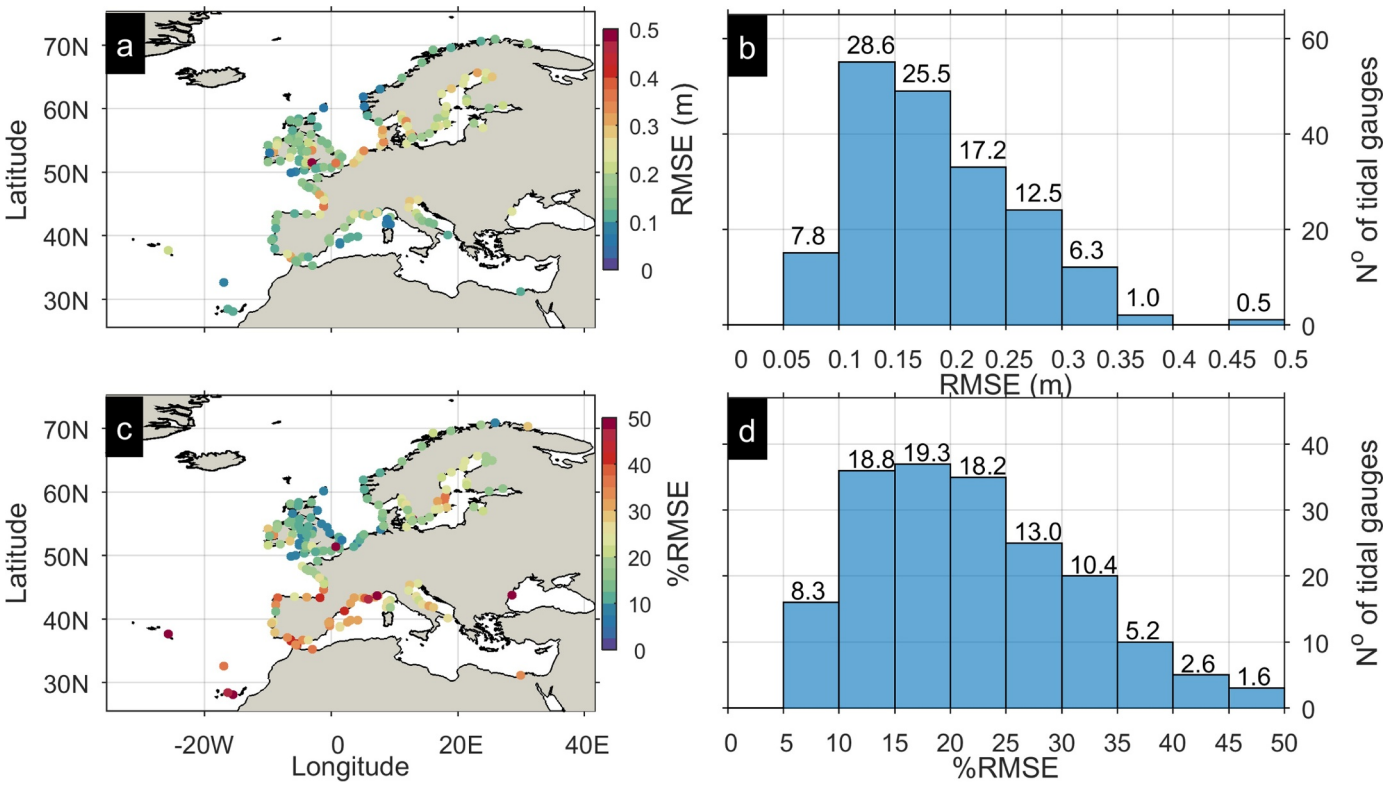

Fig. A2. Extreme non-tidal residual (η_{RE}) water level validation (99th percentile) at the tidal gauge stations considered. The two rows correspond to RMSE, %RMSE, respectively. Left column shows scatter plots of the spatial distribution of the results and right column histograms of the results for all tidal gauges.

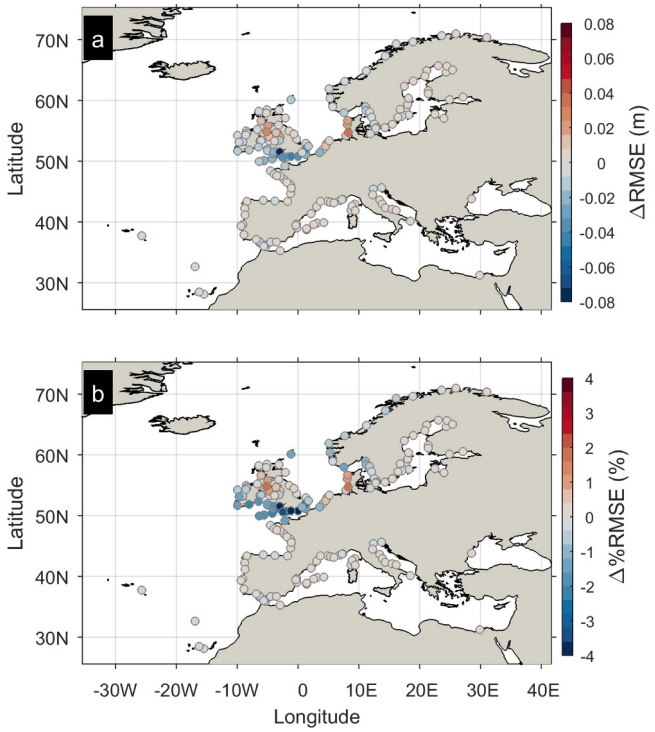

Fig. A3. Improvement in the predictive skill for extreme surge events (above 99th percentile) after including the tide effects. The scatter plot maps show the differences in absolute (a) and %RMSE (b) of the storm surge (η_{SSL}) and non-tidal residual (η_{RE}) against all the considered tidal gauges. Warm colours imply that the run which omits tidal processes has lower predictive skill.

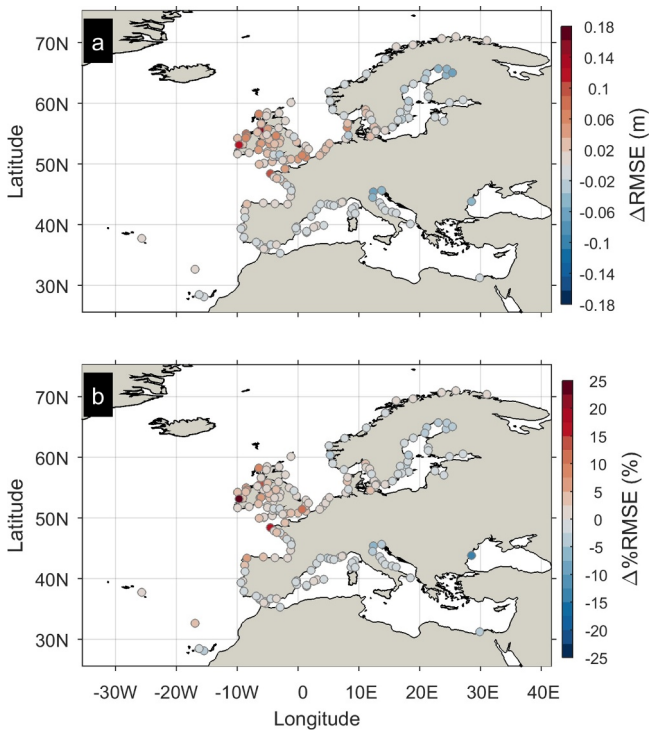

Fig. A4. Difference in storm surge level (η_{SSL}) predictive skill for extreme surge events (above 99th percentile) after using forcing from the higher resolution forecast (ECMWF-HR), instead of ERA-INTERIM. Scatter plots show differences in absolute (a) and %RMSE (b) for all the considered tidal gauges. Warm colours imply that the ERA-INTERIM run has higher predictive skill.

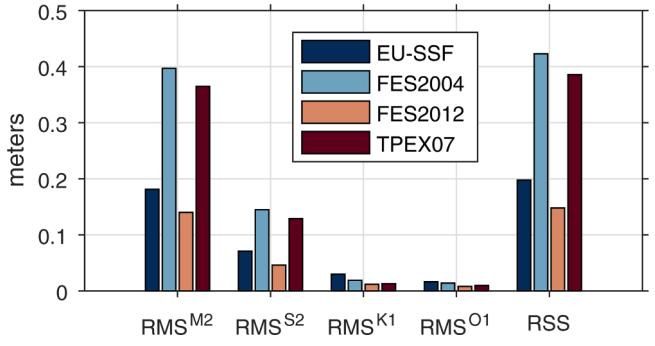

Fig. A5. Inter-comparison of tidal prediction performance for EU-SSF, TPEX07.2, FES2004, and FES2012.

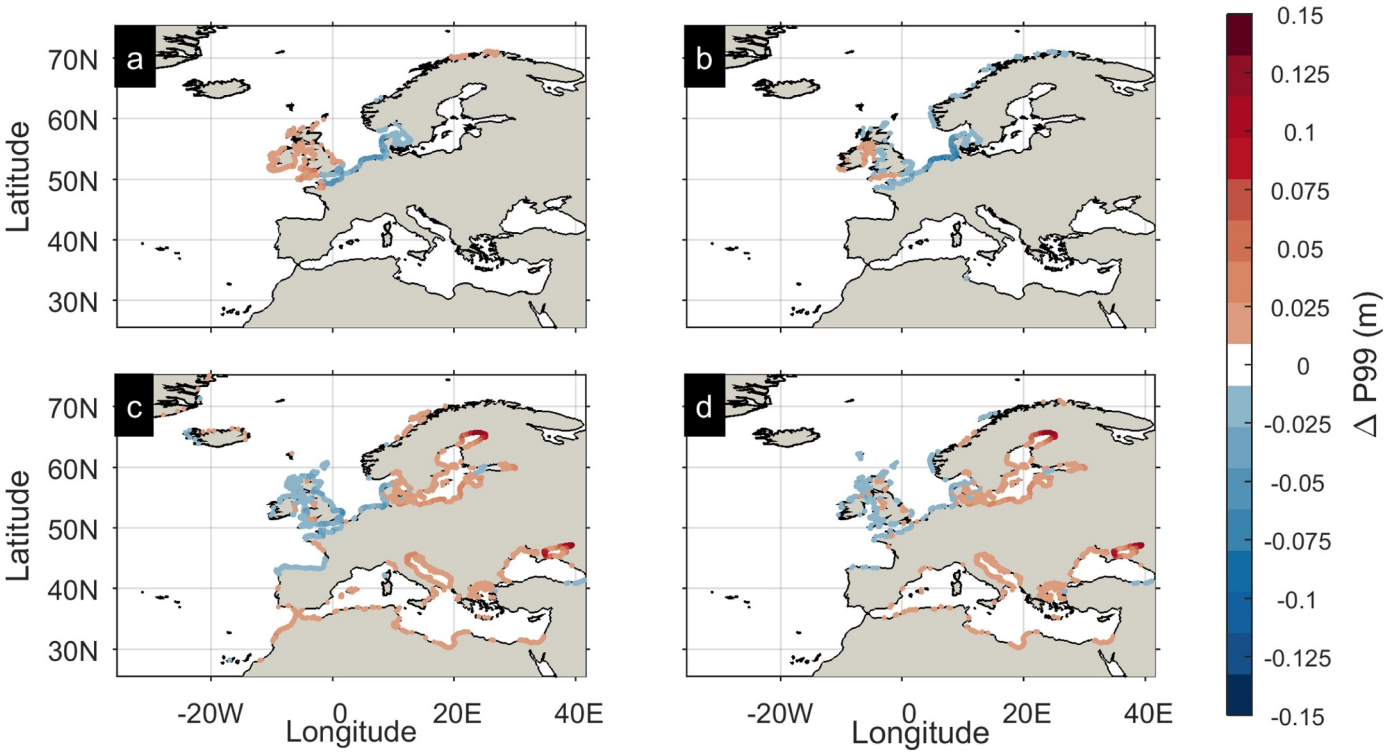

Fig. A6. Effect of tidal-surge interaction and atmospheric forcing accuracy/resolution in the extreme water level components along European coastline (99th percentile (m)). (a) Comparison of the non-tidal residual vs the storm surge water level; the former is obtained by subtracting the water levels of the only tidally forced simulation E0, from the coupled run E2; the latter is obtained from the simulation E1 which omits tidal forcing. (b) Comparison of the total water level (TWL) from E2 with that obtained by the linear addition of E0 and E1. (c) Comparison of the storm surge level obtained from the high resolution atmospheric forcing (no tidal forcing; E3) and E1. (d) Comparison of the TWL from the high resolution atmospheric forcing coupled run (E4) and E2. Warm colours indicate that most advanced configuration (coupled, or higher resolution) produces higher water level values.

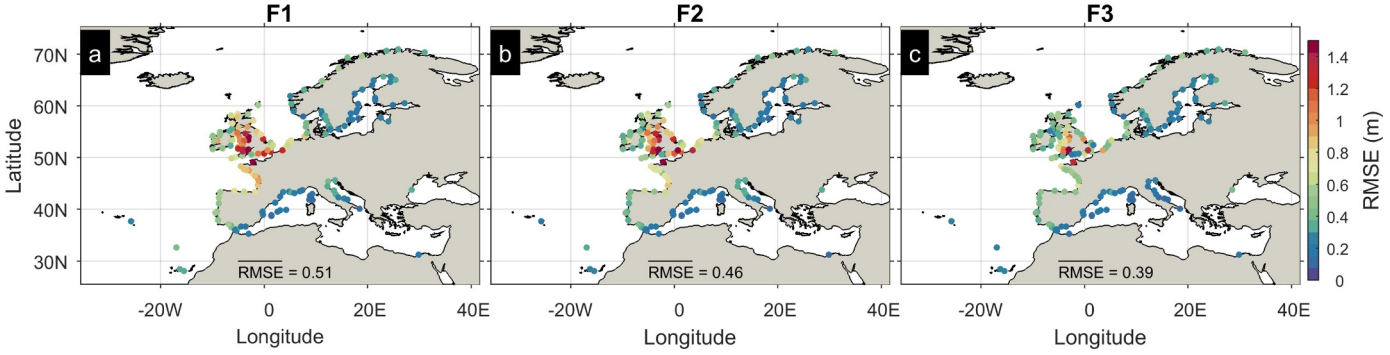

Fig. A7. Inter-comparison of RMSE of the extreme TWL (99th percentile) for the different configurations tested for the TWL forecast. (a) F1: Direct calculation from fully forced run (tides and storm surge). (b) F2: Computation from linear addition of FES2012 tidal model solution and storm surge calculation omitting the tidal flows. (c) F3: Linear addition of FES2012 tidal model solution and non-tidal residual extracted from fully forced run (tides and storm surge).

References

Arabelos, D.N., Papazachariou, D.Z., Contadakis, M.E., Spatalas, S.D., 2011. A new tide model for the Mediterranean Sea based on altimetry and tide gauge assimilation. Ocean Sci. 7, 429–444.

Arns, A., Wahl, T., Dangendorf, S., Jensen, J., 2015. The impact of sea level rise on storm surge water levels in the northern part of the German Bight. Coastal Eng 96, 118–131.

Bertin, X., Bruneau, N., Breilh, J.-F., Fortunato, A.B., Karpytchev, M., 2012. Importance of wave age and resonance in storm surges: the case Xynthia, Bay of Biscay. Ocean Modell. 42, 16–30.

Bertin, X., Li, K., Roland, A., Bidlot, J.-R., 2015. The contribution of short-waves in storm surges: two case studies in the Bay of Biscay. Cont. Shelf Res. 96, 1–15.

Bertin, X., Li, K., Roland, A., Zhang, Y.J., Breilh, J.F., Chaumillon, E., 2014. A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. Coastal Eng 94, 80–89.

Bouwer, L.M., 2011. Have disaster losses increased due to anthropogenic climate change? Bull. Am. Meteorol. Soc. 92, 39–46.

Bouwer, L.M., Jonkman, S.N., 2018. Global mortality from storm surges is decreasing. Environ. Res. Lett. 13, 014008.

Brown, J.M., Souza, A.J., Wolf, J., 2010. An 11-year validation of wave-surge modelling in the Irish Sea, using a nested POLCOMS–WAM modelling system. Ocean Modell. 33, 118–128.

Calafat, F.M., Avgoustoglou, E., Jordà, G., Flocas, H., Zodiatis, G., Tsimplis, M.N., Kouroutzoglou, J., 2014. The ability of a barotropic model to simulate sea level extremes of meteorological origin in the Mediterranean Sea, including those caused by explosive cyclones. J. Geophys. Res. 119, 7840–7853.

Carrere, L., Lyard, F., Cancet, M., Guillot, A., Picot, N., 2016. FES 2014, a new tidal model – validation results and perspectives for improvements. In: ESA Living Planet Conference. Prague.

Carrère, L., Lyard, F., Cancet, M., Guillot, A., Roblou, L., 2012. FES2012: A new Global Tidal Model Taking Taking Advantage of Nearly 20 Years of Altimetry. 20 Years of

Altimetry, Venice.
Cheung, K.F., Phadke, A.C., Wei, Y., Rojas, R., Douyere, Y.J.M., Martino, C.D., Houston, S.H., Liu, P.L.F., Lynett, P.J., Dodd, N., Liao, S., Nakazaki, E., 2003. Modeling of storm-induced coastal flooding for emergency management. Ocean Eng. 30, 1353–1386.
Ciavola, P., Harley, M.D., den Heijer, C., 2017. The RISC-KIT storm impact database: a new tool in support of DRR. Coastal Eng. 134, 24–32.
Cid, A., Castanedo, S., Abascal, A., Menéndez, M., Medina, R., 2014. A high resolution hindcast of the meteorological sea level component for Southern Europe: the GOS dataset. Clim. Dyn. 43, 2167–2184.
Cid, A., Menéndez, M., Castanedo, S., Abascal, A.J., Méndez, F.J., Medina, R., 2016. Long-term changes in the frequency, intensity and duration of extreme storm surge events in southern Europe. Clim. Dyn. 46, 1503–1516.
Conte, D., Lionello, P., 2013. Characteristics of large positive and negative surges in the Mediterranean Sea and their attenuation in future climate scenarios. Global Planet. Change 111, 159–173.
DDTM-17, 2011. Éléments de mémoire sur la tempête Xynthia du 27 et 28 février 2010.
Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of Barotropic Ocean Tides. J. Atmos. Oceanic Technol. 19, 183–204.
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 Modell. 61, 38–48.
Forzieri, G., Feyen, L., Russo, S., Vousdoukas, M., Alfieri, L., Outten, S., Migliavacca, M., Bianchi, A., Rojas, R., Cid, A., 2016. Multi-hazard assessment in Europe under climate change. Clim. Change 137, 105–119.
Garnier, E., Ciavola, P., Spencer, T., Ferreira, O., Armaroli, C., McIvor, A., 2017. Historical analysis of storm events: case studies in France, England, Portugal and Italy. Coastal Eng 134, 10–23.
Haigh, I.D., Wadey, M.P., Wahl, T., Ozsoy, O., Nicholls, R.J., Brown, J.M., Horsburgh, K., Gouldby, B., 2016. Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. Scientific Data 3, 160107.
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. SI 64, 1861–1865.
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.
Hinkel, J., Lincke, D., Vafeidis, A.T., Perrette, M., Nicholls, R.J., Tol, R.S.J., Marzeion, B., Fettweis, X., Ionescu, C., Levermann, A., 2014. Coastal flood damage and adaptation costs under 21st century sea-level rise. Proc. Natl. Acad. Sci. 111, 3292–3297.
Horsburgh, K.J., Wilson, C., 2007. Tide-surge interaction and its role in the distribution of surge residuals in the North Sea. J. Geophys. Res. 112.
Idier, D., Paris, F., Cozannet, G.L., Boulahya, F., Dumas, F., 2017. Sea-level rise impacts on the tides of the European Shelf. Cont. Shelf Res. 137, 56–71.
Kowalewski, M., Kowalewska-Kalkowska, H., 2017. Sensitivity of the Baltic Sea level prediction to spatial model resolution. J. Mar. Syst. 173, 101–113.
Kreibich, H., Di Baldassarre, G., Vorogushyn, S., Aerts, J.C.J.H., Apel, H., Aronica, G.T., Arnbjerg-Nielsen, K., Bouwer, L.M., Bubeck, P., Caloiero, T., Chinh, D.T., Cortès, M., Gain, A.K., Giampá, V., Kuhlicke, C., Kundzewicz, Z.W., Llasat, M.C., Mård, J., Matczak, P., Mazzoleni, M., Molinari, D., Dung, N.V., Petrucci, O., Schröter, K., Slager, K., Thieken, A.H., Ward, P.J., Merz, B., 2017. Adaptation to flood risk: results of international paired flood event studies. Earth's Future 5, 953–965.
Lellouche, J.M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., Le Traon, P.Y., 2018. Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system. Ocean Sci. 14, 1093–1126.
Lellouche, J.M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., Ferry, N., Desportes, C., Testut, C.E., Bricaud, C., Bourdallé-Badie, R., Tranchant, B., Benkiran, M., Drillet, Y., Daudin, A., De Nicola, C., 2013. Evaluation of global monitoring and forecasting systems at Mercator Océan. Ocean Sci. 9, 57–81.
Lyard, F., Lefevre, F., Letellier, T., Francis, O., 2006. Modelling the global ocean tides: modern insights from FES2004. Ocean Dyn. 56, 394–415.
Muis, S., Verlaan, M., Winsemius, H.C., Aerts, J.C.J.H., Ward, P.J., 2016. A global re-analysis of storm surges and extreme sea levels. Nat. Commun. 7.
Muller, H., Pineau-Guillou, L., Idier, D., Ardhuin, F., 2014. Atmospheric storm surge modeling methodology along the French (Atlantic and English Channel) coast. Ocean Dyn. 64, 1671–1692.
O'Neill, C., Saulter, A., Williams, J., Horsburgh, K., 2016. NEMO-surge: Application of atmospheric forcing and surge evaluation. Met Office Forecasting Research Technical Report 619. Met Office.
Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Comput. Geosci. 28, 929–937.
Pond, S., Pickard, G.L., 2013. Introductory Dynamical Oceanography. Elsevier.
Pugh, D.T., 1987. Tides, surges and mean sea-level: a handbook for engineers and scientists.
Quaresma, L.S., Pichon, A., 2013. Modelling the barotropic tide along the West-Iberian margin. J. Mar. Syst. 109-110, S3–S25.
Roland, A., Zhang, Y.J., Wang, H.V., Meng, Y., Teng, Y.-C., Maderich, V., Brovchenko, I., Dutour-Sikiric, M., Zanke, U., 2012. A fully coupled 3D wave-current interaction model on unstructured grids. J. Geophys. Res. 117, 2156–2202.
Sembiring, L., van Ormondt, M., van Dongeren, A., Roelvink, D., 2015. A validation of an operational wave and surge prediction system for the Dutch coast. Nat. Hazards Earth Syst. Sci. 15, 1231–1242.
Serafin, K.A., Ruggiero, P., 2014. Simulating extreme total water levels using a time-dependent, extreme value approach. J. Geophys. Res. 119, 6305–6329.
Spencer, T., Brooks, S.M., Evans, B.R., Tempest, J.A., Möller, I., 2015. Southern North Sea storm surge event of 5 December 2013: water levels, waves and coastal impacts. Earth Sci. Rev. 146, 120–145.
Trigo, I.F., Davies, T.D., Bigg, G.R., 1999. Objective climatology of cyclones in the Mediterranean Region. J. Clim. 12, 1685–1696.
Tsimplis, M.N., Proctor, R., Flather, R.A., 1995. A two-dimensional tidal model for the Mediterranean Sea. J. Geophys. Res. 100, 16223–16239.
Visser, H., Petersen, A.C., Ligtvoet, W., 2014. On the relation between weather-related disaster impacts, vulnerability and climate change. Clim. Change 125, 461–477.
Vitousek, S., Barnard, P.L., Fletcher, C.H., Frazer, N., Erikson, L., Storlazzi, C.D., 2017. Doubling of coastal flooding frequency within decades due to sea-level rise. Sci. Rep. 7, 1399.
Vousdoukas, M.I., 2014. Observations of wave run-up and groundwater seepage line motions on a reflective-to-intermediate, meso-tidal beach. Mar. Geol. 350, 52–70.
Vousdoukas, M.I., Almeida, L.P., Ferreira, Ó, 2012a. Beach erosion and recovery during consecutive storms at a steep-sloping, meso-tidal beach. Earth Surf. Processes Landforms 37, 583–691.
Vousdoukas, M.I., Ferreira, O., Almeida, L.P., Pacheco, A., 2012b. Toward reliable storm-hazard forecasts: XBeach calibration and its potential application in an operational early-warning system. Ocean Dyn. 62, 1001–1015.
Vousdoukas, M.I., Mentaschi, L., Voukouvalas, E., Bianchi, A., Dottori, F., Feyen, L., 2018a. Climatic and socioeconomic controls of future coastal flood risk in Europe. Nat. Climate Change.
Vousdoukas, M.I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L.P., Feyen, L., 2018b. Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. Nature. Commun. 9, 2360.
Vousdoukas, M.I., Voukouvalas, E., Annunziato, A., Giardino, A., Feyen, L., 2016a. Projections of extreme storm surge levels along Europe. Clim. Dyn. 47, 3171–3190.
Vousdoukas, M.I., Voukouvalas, E., Mentaschi, L., Dottori, F., Giardino, A., Bouziotas, D., Bianchi, A., Salamon, P., Feyen, L., 2016b. Developments in large-scale coastal flood hazard mapping. Nat. Hazards Earth Syst. Sci. 16, 1841–1853.
Wadey, M.P., Haigh, I.D., Nicholls, R.J., Brown, J.M., Horsburgh, K., Carroll, B., Gallop, S.L., Mason, T., Bradshaw, E., 2015. A comparison of the 31 January–1 February 1953 and 5–6 December 2013 coastal flood events around the UK. Front. Marine Sci. 2.
Wolf, J., Flather, R.A., 2005. Modelling waves and surges during the 1953 storm. Philosop. Trans. R. Soc. 363, 1359–1375.
Wolski, T., Wiśniewski, B., Giza, A., Kowalewska-Kalkowska, H., Boman, H., Grabbi-Kaiv, S., Hammarklint, T., Holfort, J., Lydeikaitė, Ž, 2014. Extreme sea levels at selected stations on the Baltic Sea coast. Oceanologia 56, 259–290.
Zampato, L., Bajo, M., Canestrelli, P., Umgiesser, G., 2016. Storm surge modelling in Venice: two years of operational results. J. Oper. Oceanogr. 9, s46–s57.
Zhang, Y., Baptista, A.M., 2008. SELFE: A semi-implicit Eulerian–Lagrangian finite-element model for cross-scale ocean circulation. Ocean Modell. 21, 71–96.
Zhang, Y.J., Ye, F., Stanev, E.V., Grashorn, S., 2016. Seamless cross-scale modeling with SCHISM. Ocean Modell. 102, 64–81.
Zijl, F., Verlaan, M., Gerritsen, H., 2013. Improved water-level forecasting for the Northwest European Shelf and North Sea through direct modelling of tide, surge and non-linear interaction. Ocean Dyn. 63, 823–847.