Storm Surge Ensemble Prediction System for Lagoons and Transitional Environments

JACOPO ALESSANDRI,[a,b] NADIA PINARDI,[b] IVAN FEDERICO,[c] AND ANDREA VALENTINI[d]

[a] *Interdepartmental Research Centre for Environmental Sciences (CIRSA), University of Bologna, Ravenna, Italy*
[b] *Department of Physics and Astronomy (DIFA), University of Bologna, Bologna, Italy*
[c] *Ocean Prediction and Applications (OPA), Euro-Mediterranean Center on Climate Change (CMCC), Lecce, Italy*
[d] *Hydro-Meteo-Climate Service of the Agency for Prevention, Environment and Energy of Emilia-Romagna, Arpae-SIMC, Bologna, Italy*

(Manuscript received 7 March 2023, in final form 13 June 2023, accepted 17 July 2023)

ABSTRACT: We developed a storm surge ensemble prediction system (EPS) for lagoons and transitional environments. Lagoons are often threatened by storm surge events with consequent risks for human life and economic losses. The uncertainties connected with a classic deterministic forecast are many, thus, an ensemble forecast system is required to properly consider them and inform the end-user community accordingly. The technological resources now available allow us to investigate the possibility of operational ensemble forecasting systems that will become increasingly essential for coastal management. We show the advantages and limitations of an EPS applied to a lagoon, using a very high-resolution unstructured grid finite element model and 45 EPS members. For five recent storm surge events, the EPS generally improves the forecast skill on the third forecast day compared to just one deterministic forecast, while they are similar in the first two days. A weighting system is implemented to compute an improved ensemble mean. The uncertainties regarding sea level due to meteorological forcing, river runoff, initial boundaries, and lateral boundaries are evaluated for a special case in the northern Adriatic Sea, and the different forecasts are used to compose the EPS members. We conclude that the largest uncertainty is in the initial and lateral boundary fields at different time and space scales, including the tidal components.

SIGNIFICANCE STATEMENT: Storm surges are extreme sea level events that may threaten densely populated coastal areas. The purpose of this work is to improve the extreme sea level forecast for transitional areas with the understanding of what are the most important forcing generating uncertainties and find a technique to reach a reliable sea level forecast. This is achieved by implementing an ensemble prediction system running 45 members for each event considered. Results show that initial and lateral boundary conditions provide most of the uncertainty, including the tidal components. The weighting system applied to find the ensemble mean improves the forecast skill on the third forecast day while it is comparable with the deterministic forecast in the first two days.

KEYWORDS: Ocean; Coastal flows; Sea level; Storm surges; Ensembles

1. Introduction

The operational sea level (SL) forecasting is a widely developed service used to inform about storm surge hazards that represent a potential threat for human life and activities (Chaumillon et al. 2017; Forzieri et al. 2016). The need for operational storm surge forecasting systems is well known and addressed by many authors (Li and Nie 2017; Pinardi et al. 2017; Umgiesser et al. 2021). However, the forecast lead time is still limited because is affected by the uncertainties regarding the initial conditions, as well as lateral boundary forcing specifications and the limited predictability of the atmospheric forcing.

Addressing the problem of storm surge forecast reliability involves recasting the single, deterministic forecast in probabilistic terms (Gneiting and Katzfuss 2014). Atmospheric seasonal and subseasonal forecasting systems feature several ensemble prediction systems (EPSs), while EPSs are still in their infancy for short term ocean forecasting and coastal forecasting.

The notion of ensemble forecasting was first proposed by Lorenz (1963), who demonstrated the sensitivity to initial conditions for a simple nonlinear system. In the 1960s it was already clear that there was a "limit to deterministic predictability" in weather forecasting (Palmer 2018). During the 1980s and 1990s ensemble weather forecasting systems were developed leading to an operational system at the European Centre for Medium-Range Weather Forecasts (ECMWF; Palmer et al. 1992; Molteni et al. 1996; Buizza 2019) and at the National Meteorological Center (NMC; Toth and Kalnay 1993). A further development was the concept of multimodel ensemble and superensemble (Krishnamurti et al. 2000). Considering outputs from different models to create an ensemble system limits the systematic errors that can affect each individual model. Another application of a multimodel uncertainty estimation is used for IPCC climate projections (IPCC 2023).

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WAF-D-23-0040.s1.

Corresponding author: Jacopo Alessandri, jacopo.alessandri2@unibo.it

DOI: 10.1175/WAF-D-23-0040.1

In oceanography, the multimodel superensemble concept was applied for the Mediterranean Sea SST forecasting by Pistoia et al. (2016) using a multiple linear regression technique applied to a multiphysics and multimodel dataset. The Mediterranean Sea was the subject of early ensemble systems for the determination of the ocean response to surface wind uncertainty (Pinardi et al. 2008, 2011; Milliff et al. 2011).

The first storm surge EPS for coastal systems was made operational by Flowerdew et al. (2009) with further verification of the results in Flowerdew et al. (2010) where the sensitivity of the surge forecast to meteorological forcing and initial conditions was investigated. A recent study by Biolchi et al. (2022) applies EPS techniques for a coastal early warning system (EWS) employing a morphodynamic model. A multimodel storm surge EPS is operational in the North Sea, combining a series of storm surge forecasting systems using a Bayesian model average (BMA) to weight each individual forecast (Beckers et al. 2008). The same methodology was applied in the western Mediterranean (Pérez et al. 2012). A multimodel EPS approach was also developed for the Adriatic Sea by Ferrarin et al. (2020), where several operational forecasting systems were used to generate a SL ensemble mean (EM) and to assess its uncertainty. Outside Europe, a multimodel EPS has been applied in the New York coastal area (Di Liberto et al. 2011). However, most of the previous studies focus mainly on systems forced only by meteorological ensembles without considering the lateral boundary condition uncertainty which we demonstrate instead to be crucial for limited area model, together with initial conditions. Therefore, we believe that a dedicated ensemble methodology should be developed for lagoons and transitional environments, as the relative importance of each forcing and forcing uncertainty could be different in small and large ocean domains. In this study we focus on the development of an EPS methodology for a lagoon in an operational fashion, using a numerical model at very high resolution, ranging from 2 km in the offshore open boundary to a maximum resolution of 10 m inside the Goro Lagoon, that fully resolves the coastline and bathymetry. The novelty is that we explore a methodology to create a set of ensemble members to tackle the specific lagoon/transitional waters modeling uncertainties. In such areas the initial and lateral boundary conditions (Chu 1999), the tidal signal from the lateral boundaries as well as the atmospheric forcing and the river runoff may all represent a source of uncertainty. The physical parameterizations and the specific numerical schemes used in the ocean model add further uncertainty to the system but they will not be explored in this initial work. The results obtained could in principle be easily generalized for other lagoons and transitional environments that share similar characteristics and are subjected to the same forcings (Niedda and Greppi 2007; Umgiesser et al. 2014), and hence have comparable uncertainties (Kjerfve and Magill 1989).

We tested the EPS methodology considering five recent extreme events in a lagoon, using forcing from different meteorological and ocean operational models for initial and lateral boundary conditions, together with river/inlet discharge perturbations. We assessed the sources of uncertainties along with the performance of the ensemble considering specific ensemble mean algorithms.

In section 2 the storm surge, high SL events, and the circulation model are described. Section 3 defines the ensemble methodology. The results are shown in section 4. Section 5 contains summary and conclusions.

2. Data and methods

a. Storm surge events

The area of study is represented in Fig. 1. It contains the whole Goro Lagoon together with the Po River branches of Goro and Volano. The bathymetry of the lagoon is shallow, about 1.50 m on average and it gently slopes to the continental shelf of the northern Adriatic Sea. The lagoon has a wide opening of about 4 km and is thus exposed to sea level anomalies entering from the open Adriatic Sea.

All the extreme events used in this work were collected from Perini et al. (2020, 2019), which details the extreme events and their impacts on the Emilia-Romagna (ER) coast. We chose all the events that had an impact on the coast, except for event 2 that was chosen because it showed an interesting overestimation of the predicted sea level by a few members of the ensemble. Sea level observations collected at Faro and Porto Garibaldi (Fig. 1) are available every 10 min, thus giving a very good determination of extremes that are listed in Table 1.

Figure 2 shows the sea level observations at Faro (from late 2019 to early 2021), together with the surge residual (SR, orange line) and the fundamental and second seiche mode contribution to SL (green line).

SR is calculated from SL using a detiding procedure described in appendix A. The residual signals only contain the surge contribution due to local wind, atmospheric pressure, and large-scale nontidal remote forcing. Likewise, the tidal filter is used to compute the contribution of the seiches. The filter is inverted and the frequencies of the seiches are the only ones retained in the computation.

In Fig. 2 event peak sea level timing is indicated by the vertical dashed lines. The horizontal dashed lines indicate the 99th percentile of the sea level, in black for the SL data (0.64 m) and in red for the SR time series (0.49 m) considering all the data available starting from April 2016. It is quite common to consider percentiles between 95 and 99 in extreme value analyses of the sea level (Wahl et al. 2017; Kirezci et al. 2020). In this work the 99th percentile is just taken as a reference threshold for the considered events. The same percentile computation was also done for the Porto Garibaldi station, where data are available starting from July 2009. The values found were 0.7 and 0.53 m for the SL and SR, respectively. It is worth noting that all the peak events considered for this work exceed both the observed SL and SR 99th percentile.

b. The circulation model

The circulation model used is the implementation of the System of Hydrodynamic Finite Element Modules (SHYFEM) model (Umgiesser et al. 2004; Bellafiore and Umgiesser 2010; Micaletto et al. 2022) in the Goro Lagoon, the so-called

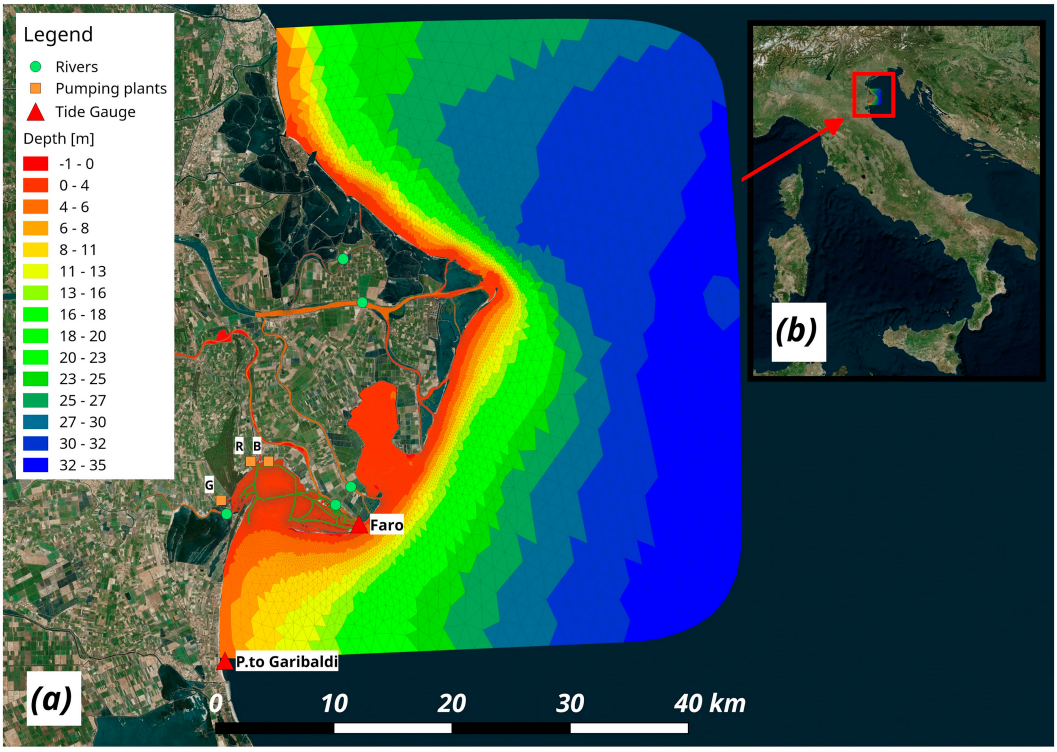

FIG. 1. (a) Bathymetry and coastlines of the Goro Lagoon. The red triangles indicate the position of the Faro and Porto Garibaldi tide gauges. The green circles indicate the position of rivers, starting from the southwest, the Po di Volano, the Po di Goro, the Po della Donzella, the Po di Pila, and the Po di Maistra. The orange squares represent the three pumping plants discharging in the Goro lagoon: Giralda (G), Romanina (R), and Bonello (B). (b) The domain position in the Adriatic Sea (source map: OpenStreetMap).

GOLFEM. SHYFEM is an open-source (https://github.com/SHYFEM-model/shyfem) unstructured grid baroclinic ocean model that solves the oceanic primitive equations under Boussinesq and hydrostatic approximations. It has been already applied in operational (Federico et al. 2017) and relocatable (Trotta et al. 2021) forecasting systems, and for storm surge events (Park et al. 2022). The river runoff is imposed as a lateral open boundary condition, and GOLFEM has been extensively calibrated, validated by Maicu et al. (2021). The model resolution goes up to 10 m along the coastlines and along the channels of the lagoon where high to very high-resolution bathymetry was used. EMODnet data at 250 m were used for the offshore part of the domain and merged with data from a coastal multibeam survey (about 1-m resolution for each transect and up to 10-m depth) operated by Arpae in 2012, while sparse points of single beam measurements with a variable resolution from a few meters to tens of meters were used inside the Goro Lagoon. This allows us to define an ensemble prediction system that deals with irreducible uncertainties due to initial conditions and forcings.

c. *Nesting in large-scale circulation models*

GOLFEM is nested within five ocean larger scale circulation models. Starting with the largest, the Copernicus Marine Service global model (Le Traon et al. 2019), offers daily global forecasts (hereafter referred to as GLOBAL;

TABLE 1. The five events analyzed in this study, with date, hour, max observed sea level (SL), and max observed surge residual (SR) at Porto Garibaldi and Faro (Fig. 1). Impacts and prevailing winds for each event are also indicated [northwesterly (NW), southeasterly (SE), north-northeasterly (N-NE), east-southeasterly (E-SE), and southeasterly (SE)]. SL and SR values are highlighted with bolded values.

		Time, SL (m), SR (m)			
Event No.	Date	Porto Garibaldi	Faro	Impacts	Prevailing winds
1	23 Dec 2019	0740 UTC, **1.17, 0.56**	0750 UTC, **1.21, 0.60**	Uniform large impacts	NW
2	3 Oct 2020	1050 UTC, **0.86, 0.51**	1030 UTC, **0.82, 0.51**	No impacts	SE
3	2 Dec 2020	0820 UTC, **1.02, 0.61**	0830 UTC, **0.93, 0.54**	Uniform large impacts	N-NE
4	8 Dec 2020	1530 UTC, **1.06, 0.96**	1510 UTC, **1.10, 0.85**	Ferrara province	E-SE
5	28 Dec 2020	0830 UTC, **0.90, 0.72**	0900 UTC, **0.90, 0.78**	Minor impacts Volano	SE

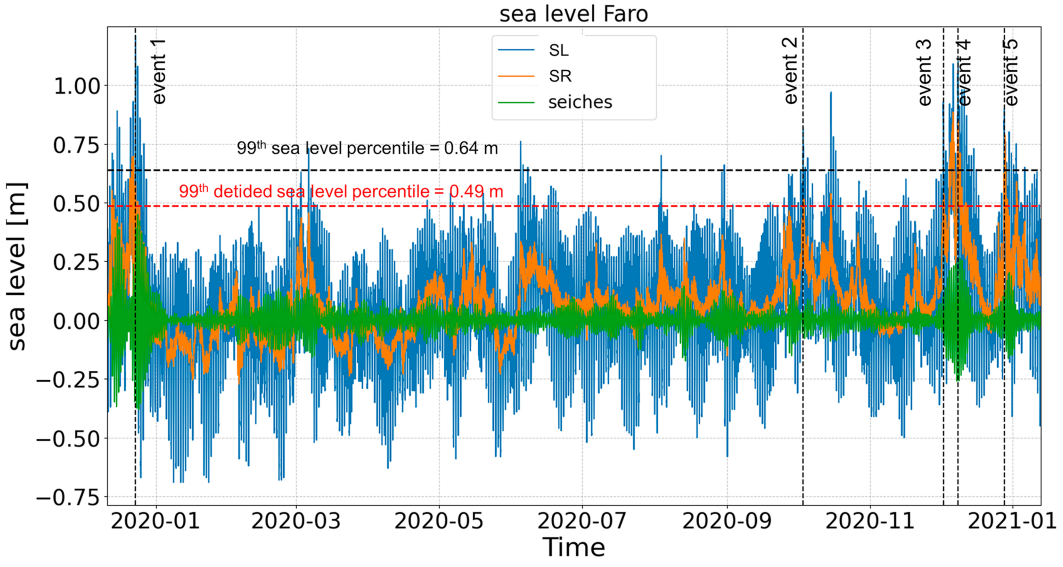

FIG. 2. SL observations (blue), SR (orange), and seiches (green) at Faro (see Fig. 1 for reference position). The vertical black dashed lines indicate the five extreme events considered in this paper. The horizontal dashed black and red lines indicate the 99th percentile of the SL and SR, respectively.

Lellouche et al. 2018) with a horizontal grid resolution of 1/12° (~8 km) and 50 vertical levels. It uses the SAM-2 data assimilation scheme, based on a reduced-order Kalman filter. It is forced at the surface with ECMWF meteorological fields. On a daily basis, the GLOBAL operational system provides a 10-day forecast (daily output), and every week it computes the best analysis for the previous 14 days.

The Copernicus Marine Service Mediterranean Sea model (MED-MFC; Clementi et al. 2017, 2021) is a coupled current-wave model of the Mediterranean Sea and the adjacent Atlantic area. It is composed of the general circulation model Nucleus for European Modeling of the Ocean (NEMO; Madec 2016) and coupled two-way with the third-generation spectral wave model WaveWatchIII (Tolman 2009). The horizontal grid is at 1/24° resolution (~4 km), with 141 unevenly vertical z levels. The analyses and forecasts are forced by ECMWF atmospheric fields and use climatological discharges from 36 rivers, including the Po. The model uses a 3D variational data assimilation scheme to optimally estimate the forecast initial condition. The eight major tidal constituents were recently added providing a different forecasting and analysis current model, hereafter called MED-MFC-T.

ADRIAC is an Adriatic Sea scale forecasting model operational at the Hydro-Meteo-Climate Service of the Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Arpae-SIMC (Bressan et al. 2017). It is based on the COAWST model (Warner et al. 2010), which is a coupled ocean–atmosphere–wave–sediment transport model. The ocean part is simulated with the ROMS model (Shchepetkin and McWilliams 2005) at a resolution of about 1 km and 30 σ layers. Hourly discharge data from the Po are used, while a climatology is used for the other 48 Adriatic rivers. Initial and lateral boundary conditions are provided by the MED-MFC model for currents, salinity, temperature, and sea level. Tides (eight components: K1, O1, P1, S1, K2, S2, M2, and N2) are given at the Otranto strait computed by the TPXO model (Egbert and Erofeeva 2002). Meteorological forcing is provided by COSMO-2I and COSMO-5M models (Gastaldo et al. 2021; Steppeler et al. 2003; COSMO 2004). ADRIAC provides 3-day forecasts per day, with a 1-day spinup, using the analyses of the meteorological forcing and boundary conditions.

ADRIAROMS is another operational model based on ROMS model and implemented by Arpae-SIMC (Russo et al. 2013). It covers the entire Adriatic basin with a horizontal resolution of about 2 km and 20 σ layers. Initial and boundary conditions are provided by CMEMS MED-currents. Tides (four components: K1, O1, S2, M2) are computed and provided by TPXO to the Otranto strait. Meteorological forcing is provided by the COSMO-5M model. Rivers are the same as in ADRIAC. ADRIAROMS provides a 3-day hourly forecast with a 1-day spinup as for ADRIAC.

All the ocean model products characteristics are summarized in Table S1 in the online supplemental material.

d. *Meteorological forcing fields*

Three operational meteorological products are used as input fields by the circulation model to compute momentum, heat and water fluxes at the air–sea interface via bulk formulas (Maicu et al. 2021). The ECMWF weather forecast fields are based on the deterministic high-resolution global model at 12.5 km of nominal resolution (ECMWF-IFS; Owens and Hewson 2018). It provides 10-day forecast fields every three hours for the first 3 days and 6 h for the subsequent days.

COSMO-5M is a regional operational meteorological model (Garbero and Milelli 2020). It is based on the COSMO model (Steppeler et al. 2003) and covers the Mediterranean region. It has a horizontal resolution of about 5 km and 45 vertical layers. It is initialized by the deterministic analysis of

TABLE 2. EPS members with the relative meteorological and ocean forcing and river perturbations.

		Ocean models				
Meteorological models	River runoff	ADRIAC	ADRIAROMS	MED-MFC	GLOBAL	MED-MFC-T
COSMO 2I	Data	Exp-1	Exp 10	Exp 19	Exp 28	Exp 37
	−30%	Exp 2	Exp 11	Exp 20	Exp 29	Exp 38
	+30%	Exp 3	Exp 12	Exp 21	Exp 30	Exp 39
COSMO 5M	Data	Exp 4	Exp 13	Exp 22	Exp 31	Exp 40
	−30%	Exp 5	Exp 14	Exp 23	Exp 32	Exp 41
	+30%	Exp 6	Exp 15	Exp 24	Exp 33	Exp 42
ECMWF	Data	Exp 7	Exp 16	Exp 25	Exp 34	Exp 43
	−30%	Exp 8	Exp 17	Exp 26	Exp 35	Exp 44
	+30%	Exp 9	Exp 18	Exp 27	Exp 36	Exp 45

COMet-LETKF (the model used by the Italian Air Force) and takes the ECMWF-IFS fields as boundary conditions. It runs twice per day, at 0000 and 1200 UTC providing a 72-h forecast (hourly output). The COSMO-5M input fields are given at hourly frequency for 3 forecast days.

COSMO-2I is the highest resolution operational weather forecast model (Gastaldo et al. 2021) covering the Italian domain. It is nested in COSMO-5M. The initial state is computed from the KENDA-LETFK system (Schraff et al. 2016; Gastaldo et al. 2021), and has a resolution of ~2.2 km and 65 vertical layers. It provides a 48-h forecast (hourly output) and two forecasts per day at 0000 and 1200 UTC. All the atmospheric model products characteristics are summarized in Table S1.

e. River runoff conditions

The hourly discharge from the Po River is measured by Arpae at Pontelagoscuro. At the eastern side of the Goro Lagoon (Fig. 1), the Po of Goro runoff is computed as a percentage of the Pontelagoscuro discharge values, based on a rule that fits data from ARPAV (2012). At the western side of the lagoon, the Po of Volano runoff is completely regulated by several pumping plants that discharge along the sides of the river. In addition to the two Po branches, three pumping stations were considered for the freshwater discharge in the lagoon. Three simulations were performed: one with the nominal discharge of the Po of Volano, Po of Goro, and the three pumping plants, and two others were defined by adding and subtracting 30% of the discharges. The 30% value was extracted from the analysis of the climatological daily Po runoff time series between 2010 and 2021. The mean of the runoff is 1490 $m^3 s^{-1}$ while its standard deviation is 450 $m^3 s^{-1}$, which approximately correspond to 30% of the mean. We considered this value to be a good estimate of the river's variability. Moreover, considering higher perturbation values would lead to unrealistic discharges.

3. The ensemble prediction system

a. The ensemble methodology

GOLFEM-EPS has 45 members. The GOLFEM-EPS members are composed of a suitable combination of atmospheric and ocean models and perturbations to the river flow (Table 2).

The EPS methodology is shown in Fig. 3: one day before the start of the forecast, GOLFEM-EPS is initialized and forced with surface analysis fields and lateral analysis boundary conditions from the circulation model outputs. Starting from the initial time of the forecast, always set at 0000 UTC, forecast lateral boundary conditions and surface meteorological fields are used for a 3-day forecast, up to +72 h. ADRIAC and ADRIAROMS models do not have an operational data assimilation scheme. For ADRIAC and ADRIAROMS the

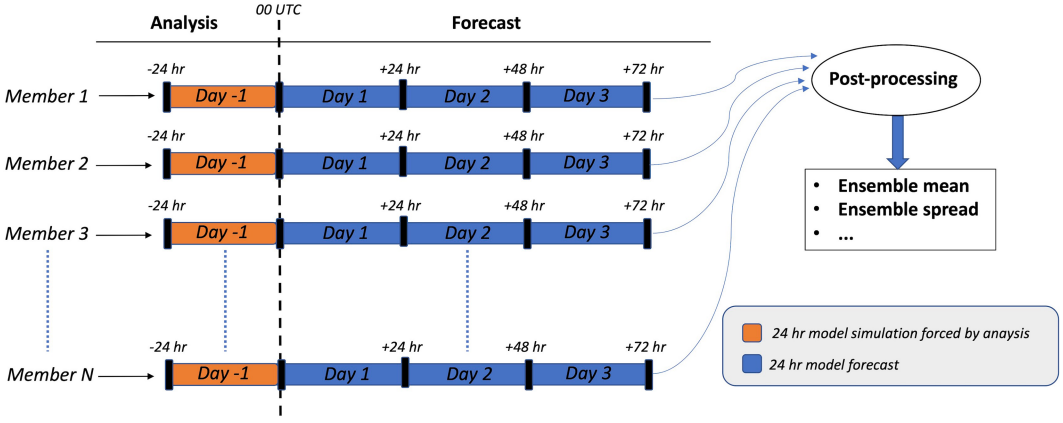

FIG. 3. The GOLFEM-EPS modeling scheme. The postprocessing creates ensemble mean, ensemble spread, and the weighted ensemble mean.

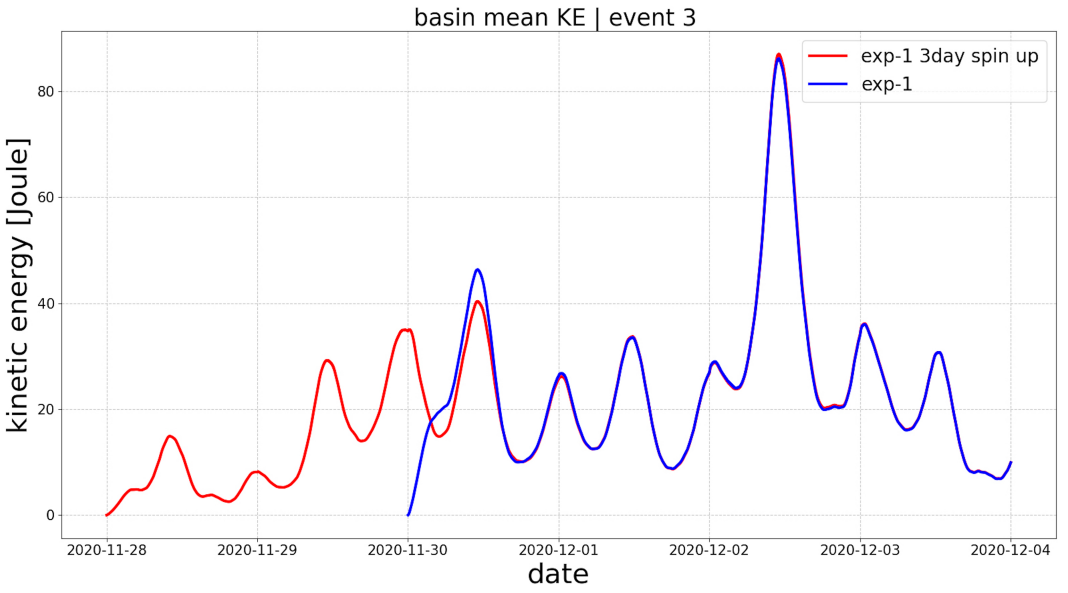

FIG. 4. Mean basin kinetic energy of exp-1 considering 1 (blue) and 3 (red) days of spinup.

analyses are thus simulations forced by meteorological and open ocean lateral boundary analyses. Since COSMO-2I provides only a 48-h forecast, from +48 to +72 h, the meteorological fields from COSMO-5M are used.

Figure 4 shows the GOLFEM-EPS basin average kinetic energy (KE) for one member simulation (exp-1 in Table 2) initialized three days and one day before the nominal start of the forecast. The KE is initially zero for both simulations (all simulations are initialized with a zero-velocity field). After about 14 h the simulations reach a similar KE. This explains the choice of only one day as a spinup time for the forecast, which is a reasonable time for the lagoon/transitional water domains. It is also well known that the spinup time decreases with the scale of the implemented computational domain, and one day of spinup has also been used by authors dealing with much more limited coastal model domains (Gaeta et al. 2016).

The GLOBAL and MED-MFC products are provided without tides. In the case of nontidal models, the tidal sea level is extracted from the TPXO model and added to the GOLFEM open boundaries.

b. *The ensemble mean*

The ensemble forecast is usually considered to be a better estimate of the forecast since the members sample the probability distribution function of the forecasts. However, interpreting multiple forecasts can be complex, so very often the ensemble mean (EM) is used for the sake of simplicity.

Using the notation of Salighehdar et al. (2017), a forecast matrix is defined as $\mathbf{X} = \{x_i^j\}_{(i,j)\in\{1,\ldots,T\}\times\{1,\ldots,m\}}$, where T and m are the forecast lead time and the total number of forecasts available, respectively. The point x_i^j thus represents the sea level or surge at time i predicted by member j. The EM produced by an EPS is denoted by a vector $\mathbf{F} = F\{_i\}_i$ of all times i. The EM is a simple average of the member forecasts defined as

$$F_i = \overline{x_i} = \frac{1}{m}\sum_{j=1}^{m} x_i^j. \quad (1)$$

In this case each member has the same weight. However, during extreme events, when usually the forecast uncertainty is greatest, this may not be the best solution. Another method is to evaluate the weights of each member based on the performance achieved during a training period. Here the correlation method is tested to compute a weighted EM (WEM). In addition, only a subset k of the m members can be considered to compute a better average. The first k forecasts are chosen based on the performance during the training period. However, selecting the value of k is subjective and should be chosen according to the performance of the ensemble model at each specific location (Salighehdar et al. 2017). In our case the best performance was obtained with $k = 21$ which was used consistently in all our cases. Since the performance of each member can be different from case to case, due to different forcing conditions, new weights are computed for each event, considering the simulation range from −24 h to 0 as the weights training period (Fig. 3). The uncertainty of the ensemble system (spread) is computed as the standard deviation of the ensemble member distribution. The specific WEM methodology is detailed in appendix B.

4. Results

a. *EPS skill*

The SL and SR at Faro (Fig. 1) are shown for events 2 and 3 in Fig. 5 and events 4 and 5 in Fig. 6. We will discuss event 1

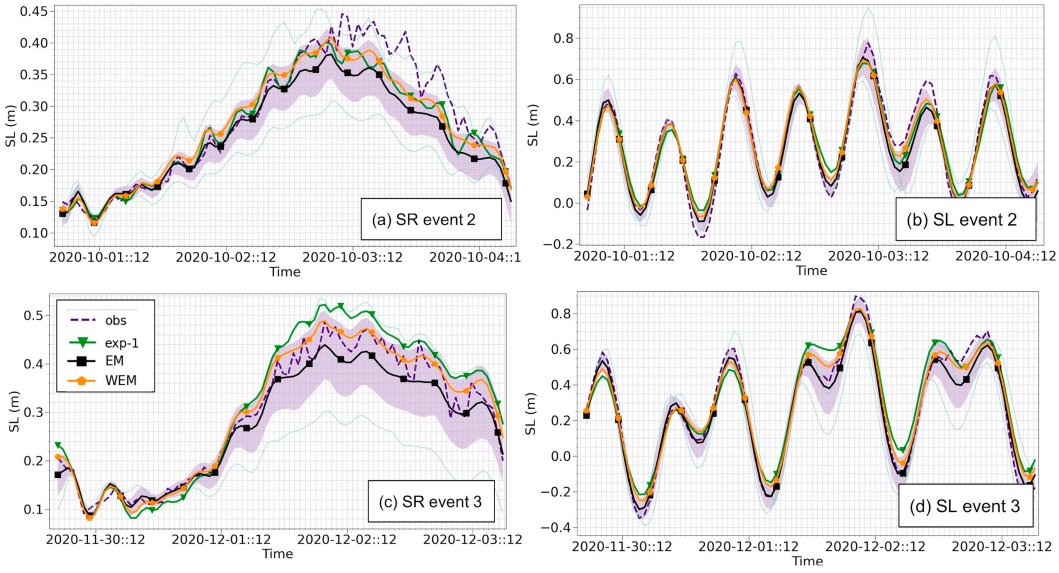

FIG. 5. (left) SR and (right) total SL forecast comparison at Faro for (a),(b) event 2 and (c),(d) event 3. The green line shows exp-1 considering the reference deterministic forecast as described in Table 2. EM and WEM indicate the ensemble mean and the weighted ensemble mean. The light green lines represent the members with maximum and minimum sea level values during the storm surge peak. The shaded areas are the ensemble spreads.

in the last part of this section because of the special contribution of seiches. The thick green line (Figs. 5, 6, and 9) is the member forced by the highest resolution models (ADRIAC and COSMO-2I; exp-1 in Table 2). This is the deterministic reference model. The thick black and orange lines are the EM and WEM, respectively. Qualitatively the deterministic forecast overestimates the SR in two out of the five events.

The ensemble spread peaks at the time of the extreme event thus denoting a growth of the uncertainty. This is another benefit of the EPS, the spread of which alerts the forecaster of the possibility that an extreme event may occur. For SR, the maximum spread is at the peak of the event with values of 4.5, 7, 7, and 12 cm for event 2, 3, 4, and 5, respectively. Interestingly, in event 2 one of the members forced by ADRIAROMS overestimates the SL at the peak time (see the upper light green line in Fig. 5b), as do other members with the same nesting model (not shown). Since no overestimation is observed in the SR (Fig. 5a) the error may be attributed to a wrong tidal signal provided by ADRIAROMS.

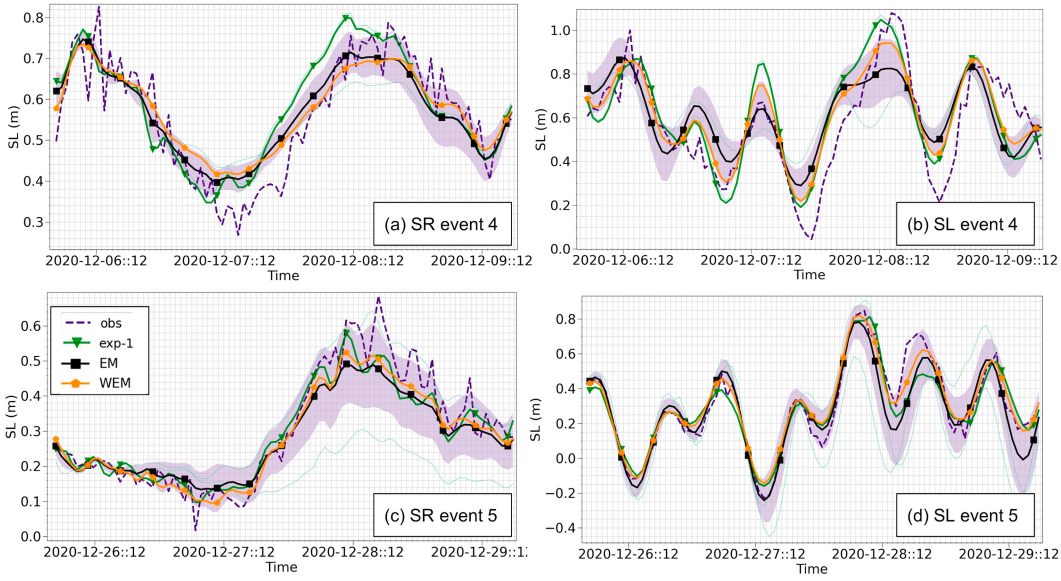

FIG. 6. As in Fig. 5, but for (a),(b) event 4 and (c),(d) event 5.

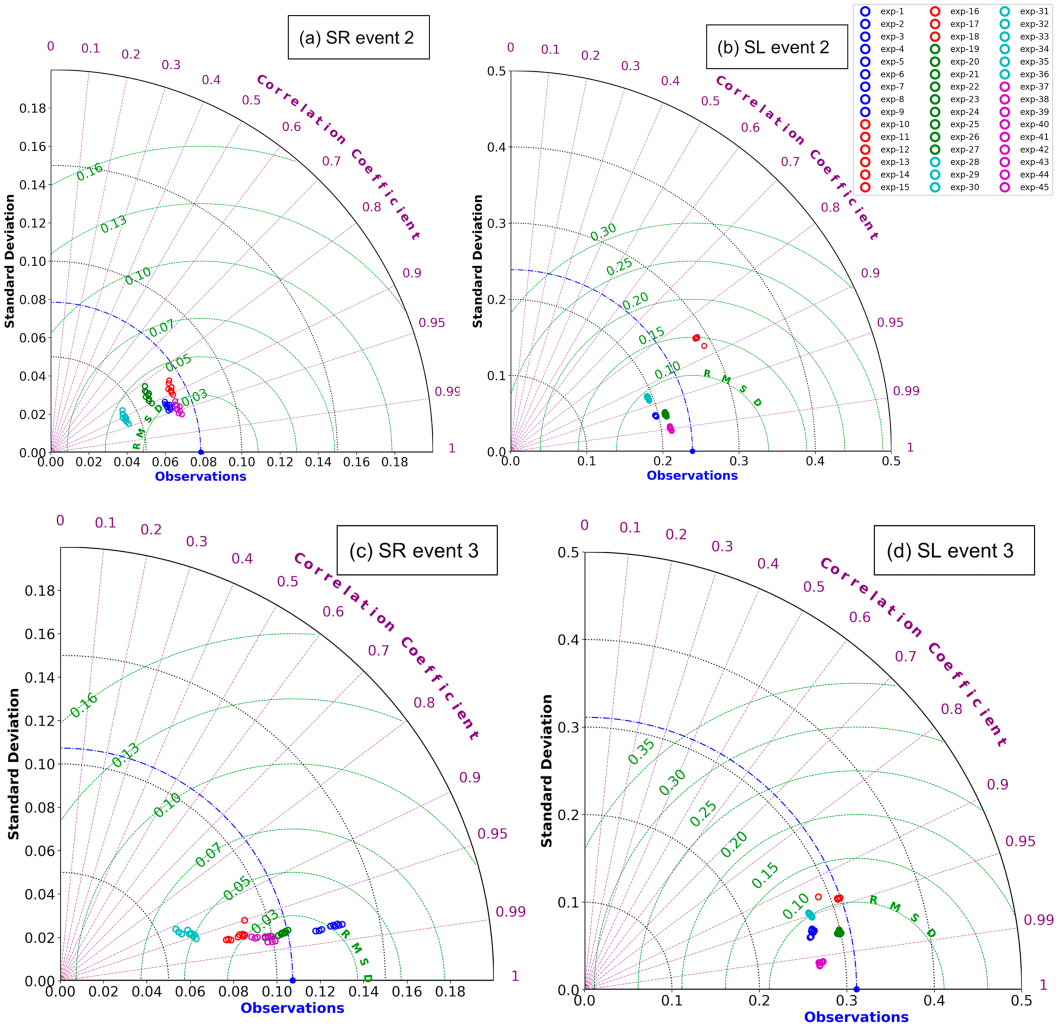

FIG. 7. Taylor diagrams for SR and total SL for (a),(b) event 2 and (c),(d) event 3.

The performance of individual members during the forecast time are summarized for each event with Taylor diagrams (Figs. 7 and 8). An evident feature is the grouping of members with the same ocean initial and lateral boundary conditions. This is even clearer when the SL is considered (Figs. 7b,d and 8b,d). If SR is considered (Figs. 7a,c and 8a,c there is a small dispersion due to the different meteorological forcing and the different river forcing. The initial/lateral boundary conditions (including the tidal forcing) therefore seem to be the greatest source of uncertainty (both for SR and SL), followed by the meteorological forcing and river forcing.

In both events 2 and 3, the maximum surge occurred with the maximum tidal amplitude. Conversely, event 4 reached the peak during a tidal minimum. A spectral analysis of the observed sea level revealed that there was a high contribution from the fundamental Adriatic seiche (~0.25 m), as shown in Fig. 2. We argue that a correlation between the arrival time of remote signals (seiches) and local forcings might be in some cases an important predictor of sea level extremes as demonstrated also for transitional waters along the U.S. East Coast (Park et al. 2022).

The Taylor diagrams provide an indication of the members that most contribute to building the final WEM for each event. In most cases members initialized/forced with MED-MFC-T (from exp-37 to exp-45) show the best scores (ρ between 0.95 and 0.99; ~0.7 for event 1 SR) followed by members initialized/forced with ADRIAC (from exp-1 to exp-9; ρ between 0.85 and 0.95; ~0.5 for event 1 SR) and MED-MFC (from exp-19 to exp-27; ρ between 0.85 and 0.99; ~0.7 for event 1 SR). Members initialized/forced by ADRIAROMS (from exp-10 to exp-18) usually show the worst performance (ρ between 0.3 and 0.95; ~0.1–0.2 for event 1 SR) together with the members initialized/forced by the GLOBAL product (from exp-28 to exp-36; ρ between 0.4 and 0.9; ~−0.4 for event 1 SR).

If SL Taylor diagrams are considered, we note again that members forced by MED-MFC-T have the best performance. This suggests that a large-scale ocean model (MED-MFC-T)

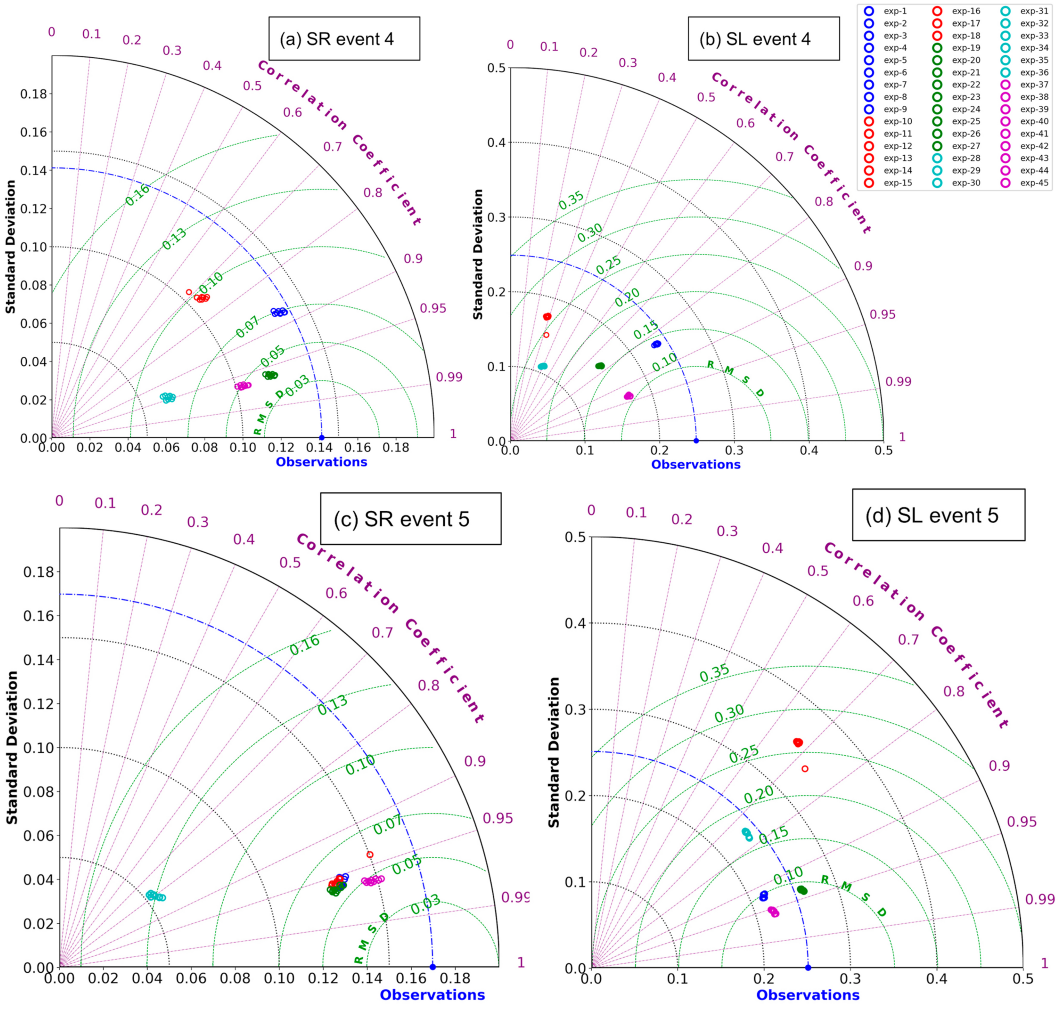

FIG. 8. Taylor diagrams for SR and total SL for (a),(b) event 4 and (c),(d) event 5.

may be more effective in representing the nonlinear interactions that occur between tides, bathymetry and basin modes compared to regional scale models (ADRIAC or ADRIAROMS) that imposes TPXO tides at their open boundary.

Event 1 is a special case because the contribution of the Adriatic Sea seiches reached values close to 40 cm (Fig. 2). Figure 9a shows that the SL peak coincides with a decreasing SR (Figs. 9a,b). The storm surge is thus generated mainly by the seiches and the tides.

Of the events analyzed, this case shows the worst performance for most of the ensemble members, with a negative correlation for the one forced by GLOBAL (Figs. 9c,d). However, EM and WEM benefit from error compensation and show satisfactory results in terms of SR, although none of the members reproduced the SL peak amplitude occurring at 0750 UTC 23 December 2019. This highlights the fundamental importance of tide-seiche resonant phenomena without the contribution from local meteorological forcing and underline some limitation of regional and large-scale ocean models in simulating such interaction.

The members initialized/forced by the GLOBAL model have usually the worst results. This is not surprising since the GLOBAL product has the lower resolution and its daily output is not able to catch appropriately the surge components due to wind during extreme events. Indeed, members forced by GLOBAL are most of the times excluded by the weighting procedure.

b. *Ensemble member analysis*

Any analysis of an EPS should consider its degree of the over- or underdispersiveness of the members. To do this, we analyzed the root-mean-square error (RMSE) as a function of the ensemble spread (Figs. 10a,b), aggregating all the events in Porto Garibaldi and Faro, considering 6-h intervals. The RMSE spread plot provides information on the dispersion of the ensemble members: if the growth of error is slower than the growth in spread, the ensemble is called overdispersive; the contrary is underdispersive.

In an ideal case the RMSE should be linearly proportional to the ensemble spread. Figure 10a shows that there is a slight

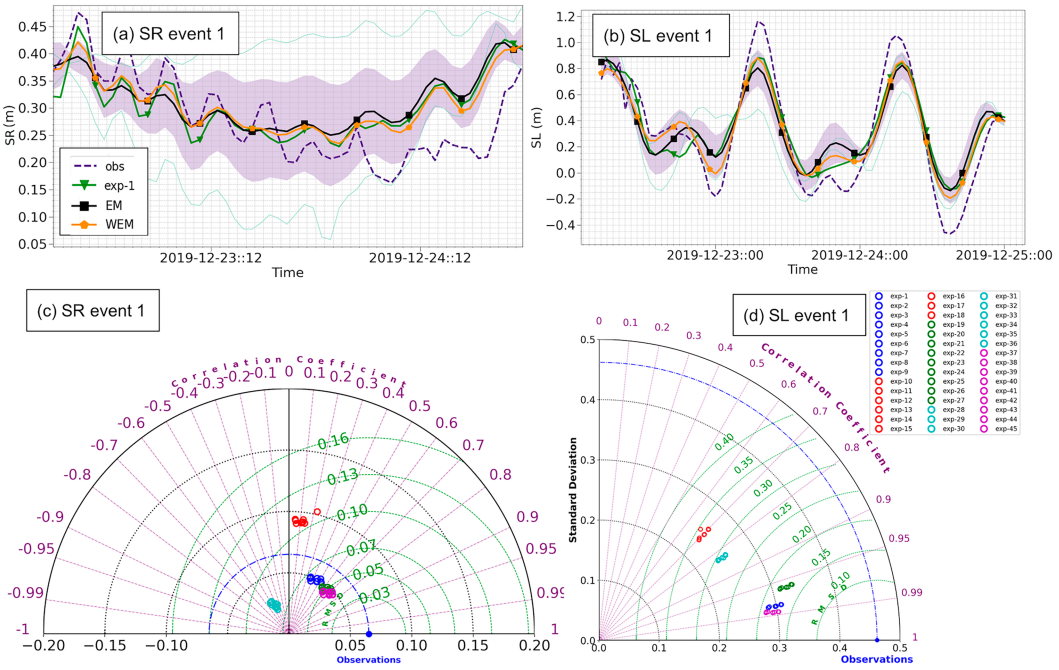

FIG. 9. (a) SR and (b) SL forecast comparison for event 1. The light green lines represent the members with maximum and minimum sea level values during the peak. The shaded area is the ensemble spread. Taylor diagrams for (c) SR and (d) SL are also shown.

overdispersion of the members for spreads greater than 0.05 cm. This implies that during peak events, there is more dispersion of the ensemble members than expected. We argue that this overdispersion of the ensemble is due to the very different initial and lateral boundary conditions used. If the SL is considered, overdispersion appears for spread values greater than 0.13 cm, while a slight underdispersion appears for smaller spread values, probably connected to the tidal components.

The SR RMSE error increases with time up to a maximum at 48 h for EM and WEM. The exp-1 seems to perform slightly better between 36 and 48 h, while after 48 h, both EM and WEM performances are slightly better. The SL RMSE shows an almost linear increase with time for all forecasts with slightly better performances for the WEM at the end of the forecast period. The RMSEs considered for all the forecasts and all events are 5.7, 5.5, and 5 cm for EM, exp-1, and WEM, respectively, for SR, while for SL, the RMSEs are 12, 11.5, and 11 cm.

We conclude this section by analyzing the spatial distribution of the ensemble spread at the time of the peak events, considering only the SR. Considering first the river and meteorological forcing induced ensemble spread, the river runoff has a small influence on the SR ensemble spread (Fig. 11a), which remains confined at the river mouths along the coast, reaching low values of 0.3 and 0.4 cm. The meteorological forcing leads to a bigger ensemble spread but with a maximum value of 1–3 cm (Fig. 11b). Although the winds are one of the main drivers of storm surges in the area and may represent a big source of uncertainty, we argue that the low spread values found here are a consequence of the limited area size of the domain, and the fact that the storm surge component due to remote wind effects is important, as discussed earlier.

The ensemble spread due to initial and lateral boundary conditions is clearly a dominant contribution (Fig. 11c), exceeding by one order of magnitude the spread due to local meteorological forcing. The SR spread (Fig. 11c) is almost uniform throughout the entire domain with values between 12 and 13 cm. The tidal signal extracted inverting the filtering procedure reveals that the contribution of tides to the sea level spread is of the same order of magnitude (Fig. 11d), with slightly smaller values between 9 and 10 cm. This suggests that both for computational efficiency and system performance, the storm surge EPS implemented in limited area domains should focus on the initial and lateral boundary condition uncertainties, further substantiating previous results (Chu 1999).

5. Summary and conclusions

The expected increasing number of extreme storm surge events necessitates reliable operational storm surge forecasting systems. The assessment of the forecast uncertainties continues to be a challenge that requires an ensemble approach to be dealt with.

We designed and developed a coastal EPS for lagoons/transitional waters using a baroclinic very high-resolution unstructured grid model. The model high resolution eliminates the uncertainties due to resolution of the coastal geometry and bathymetry leaving only irreducible uncertainties contained

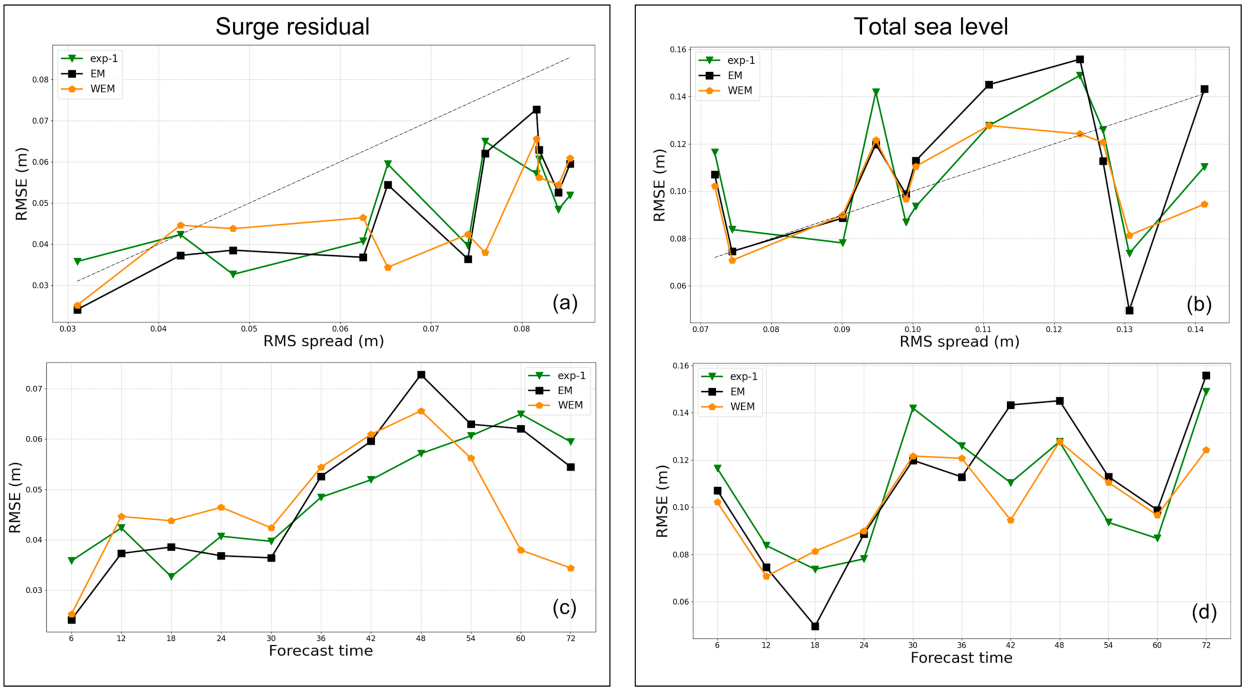

FIG. 10. RMSE against ensemble spread and forecast time for (a),(c) SR and (b),(d) total sea level. Data are computed aggregating all the events at the Porto Garibaldi and Faro stations.

in initial conditions and forcings. Therefore, the EPS is formulated by considering model forecasts carried out with different meteorological forcing, initial, and lateral boundary conditions from coarser resolution ocean models and perturbations to the river/inlet discharges. The forecast lead time is 3 days with a 1-day spinup in analysis/simulation mode.

A postprocessing ensemble mean (EM) procedure was carried out and a weighted ensemble mean (WEM) methodology was tested based on the performance of the members during the spinup, which was considered to be equivalent to a training period. Of the 45 members used for the ensemble, most of the variability was reached including only members that differed in initial and lateral boundary conditions. An estimate of the uncertainty due to different initial/lateral boundary conditions, river runoff and atmospheric forcings was provided. A perturbation of ±30% in the river runoff, corresponding to one standard deviation of the climatological daily Po River discharge, was found to contribute with 3–4 mm to the ensemble sea level spread, limited to the river mouths. Meteorological forcing had a greater impact of between 1 and 3 cm, while initial/boundary conditions (including tides) provided most of the sea level uncertainty between 9 and 13 cm.

It is found that EM and WEM forecasts have better skill from 48- to 72-h forecast lead time while in the first two days they are equivalent to a single deterministic forecast skill. The third day SL RMSE computed at the two reference stations is 11 and 12 cm, respectively, for WEM and EM, while it is 11.5 cm for the deterministic reference forecast. One case, event 1, showed large errors, due to the specific importance of seiches and tidal elevation nonlinear interactions. The correct simulation of the tidal phase is crucial and a way to account for the shifts in the tidal peak is to consider the skew surge (Williams et al. 2016). Such an approach will be further investigated in the future.

In general, we might say that for lagoons/transitional waters with limited connections to the open ocean, the EPS methodology presented here could be taken as a prototype system. For general ensemble coastal forecasting, the domain size might be an added uncertainty and different model domains should be tried with the same WEM and EM methodologies. The limitations to use different model domains derive from the inadequate number of observations in the open ocean, shelf areas which makes it difficult to produce a robust verification of the EPS. In the future it is possible that coastal scale altimetry will offer a convenient dataset for validation/calibration of a storm surge EPS. An improvement of the weight computation could be achieved by using the past n forecasts to build a more robust set of weights for the current forecast. In this work we chose events that are not closely related in time, and so we could not apply this approach; we limited the computation of the weights to the day prior to the forecast start, and even in this case we obtained an improvement of the forecast.

The lesson learnt from this exercise, most probably a generalizable conclusion, is that limited-area storm surge forecasting uncertainty in lagoons is dominated by initial and lateral boundary conditions and by phase/amplitude errors in tidal components. The former are also influenced by remotely forced propagating signals, such as seiches for the specific case of the Adriatic, but more generally remotely forced

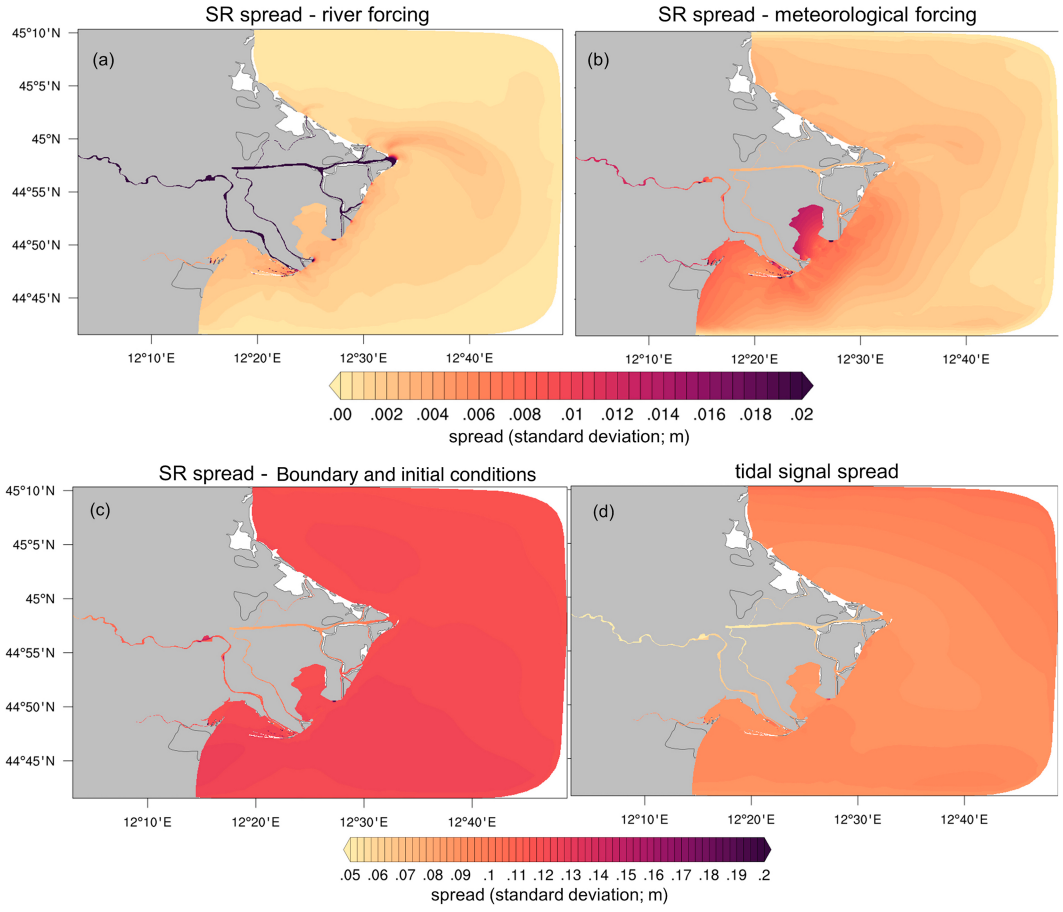

FIG. 11. Spread due to (a) river forcing, (b) meteorological forcing, (c) boundary/initial conditions, and (d) tidal signal. Data are computed aggregating all the events.

circulation signals. Given the necessity of having resolutions of ~10–100 m at the coasts, we conclude that an EPS might be necessary to account for irreducible errors in initial and lateral boundary conditions from coarser resolution models.

Acknowledgments. This work was supported by an ARPAE grant to the University of Bologna and the Department of Physics and Astronomy supporting the PhD of Jacopo Alessandri. Alessandri and Pinardi have received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement 862626, EuroSea Project. All the computations were done on the ZEUS supercomputing facilities, provided by CMCC.

Data availability statement. Model data of COSMO-5M, COSMO-2I, ADRIAC, and ADRIAROMS models were provided by Arpae. Observed sea level and Po river data were provided by Arpae through the dext3r webapp (https://simc.arpae.it/dext3r/). The GLOBAL product was provided by COPERNICUS-CMEMS (https://marine.copernicus.eu/it). ECMWF, MED-MFC, and MED-MFC-T data were provided by CMCC. The open-source model SHYFEM (version 7_5_70; Umgiesser et al. 2020) used in this work is freely available on the Zenodo repository (https://doi.org/10.5281/zenodo.3833857). Figures 2, 4, 5, 6, 9, and 10 were made with Matplotlib version 3.4.0 (Caswell et al. 2021; Hunter 2007), available under the matplotlib license at https://doi.org/10.5281/zenodo.4638398. Taylor diagrams of Figs. 7–9 were made with the Skillmetrics Python package (Rochford 2016; version 1.1.8). Figure 11 was made with PyNGL (version 1.6.1), a graphical Python package developed by the National Center for Atmospheric Research (NCAR). Figure 1 was made with QGIS 3.22 geographic information system.

APPENDIX A

The Tidal Filter Methodology

The surge component (i.e., SR), the seiches and the tides extracted from model results were found using a frequency domain filter procedure based on the Fourier transform. The basic principles of the digital filter used here are described in Thomson and Emery (2014). If we consider a variable $x(t)$ in the time domain t and its transformed $X(f)$ in the frequency domain f found applying the discrete Fourier transform (DFT), the application of the digital filter can be summarized in the following three steps:

1) Take the DFT, $X(f)$ of the original dataset $x(t)$.
2) Multiply $X(f)$ by the appropriate frequency response function $W(f)$ or FRF, of a high-, low-, or bandpass filter.
3) Take the inverse Fourier transform (IFT) of the results to obtain a filtered dataset in the time domain.

The power of this method is its simplicity. Unlike the filters in the time domain that require a convolution (e.g., Doodson filter), in the frequency domain, our solution just entails a simple multiplication:

$$X'(f) = W(f)X(f). \tag{A1}$$

The filtered time series is then simply found by applying the IFT to $X'(f)$. However, the form of $W(f)$ is extremely important in order to have a reliable filter. Ideally the FRF should be near unity in the frequency band to be passed, and zero in the bands to be stopped, with a narrow transition band to prevent contamination by unwanted frequencies.

Unfortunately, a very narrow and steep transition band is the main cause of large Gibbs phenomenon, affecting the time series obtained by IFT of $X'(f)$ which manifests itself as large side lobs in the initial and final part of the time series (ringing). The Gibbs phenomenon can be reduced by "tapering" the filter $[W(f)]$ with a smooth function to ensure a smooth transition to nonzero Fourier coefficients (Forbes 1988). We tapered the transition bands using a Tukey window (also called cosine taper), defined as

$$w(x) = \begin{cases} \frac{1}{2}\left\{1 + \cos\left[\frac{2\pi}{r}(x - r/2)\right]\right\}, & 0 \le x \le \frac{r}{2} \\ 1, & \frac{r}{2} \le x \le 1 - \frac{r}{2}, \\ \frac{1}{2}\left\{1 + \cos\left[\frac{2\pi}{r}(x - 1 + r/2)\right]\right\}, & 1 - \frac{r}{2} \le x < 1 \end{cases} \tag{A2}$$

where x are the L points of the windows, and r is the ratio of the cosine-tapered section length to the entire window length with $0 < r < 1$. A value $r = 0.5$ produces a Tukey window where half of the entire window length consists of segments of a phase shifted cosine with period $2r = 1$. If $r \le 0$ a rectangular window is returned, while for $r \ge 1$ a Von Hann window is generated (Bloomfield 2000).

The energy density spectrum (ESD) of the observed sea level at Faro is shown in Fig. S1 together with the ESD of the filtered signal using the Fourier transform filter. The green line is the FRF which was calibrated to exclude tidal and seiche signals from the time series in order to retain only the surge component. The resulting time series after the filter has been applied is the orange line in Fig. S1. The FRF can be easily inverted and the windows can be moved if there is a need to focus on particular frequencies (e.g., tides or seiches).

On the other hand, the initial and final parts of the time series must be excluded because they are the ones most affected by the Gibbs phenomenon. If in time domain filters, the initial and final part of the time series are automatically excluded, while in frequency domain filters, there is no clear threshold, and which part of the time series should be excluded is subjective. However, Walters and Heston (1982) suggest that in both time domain and frequency domain filters, the same amount of data are lost. The version of he tidal filter used for this work can be found at the Zenodo repository of the Oceanography group of the University of Bologna at the link: https://doi.org/10.5281/zenodo.6478113.

APPENDIX B

The Weighted Ensemble Mean Method

The steps for the WEM are as follows.

1) First the bias between the model output and observations during the training period in a determined station is removed for each member (this step is also done for the simple average):

$$B^j = \frac{1}{T}\sum_{i=1}^{T}(x_i^j - o_i), \tag{B1}$$

where B^j is the bias for member j. The term x_i^j is the member j variable value at time i; o_i is the observation at time i and T is the length of the training period.

2) The standard Pearson correlation coefficient, defined by

$$\rho_o^j = \frac{\frac{1}{T}\sum_{i=1}^{T}(x_i^j - \overline{x^j})(o_i - \overline{o})}{x_{\sigma j} o_\sigma}, \tag{B2}$$

is computed for each member of the ensemble during the training period, where $\overline{x^j}$ is the mean value of member j, $\overline{o}$ is the mean observed value, $x_{\sigma j}$ is the model member j standard deviation, and o_σ is the observed standard deviation.

3) The forecasts are ranked based on the correlation and the first k members are retained and used to compute the weights w_j:

$$w_j = \frac{\rho_o^j}{\sum_{j=1}^{k}\rho_o^j}. \tag{B3}$$

4) The WEM is computed as the weighted average of the selected forecasts:

$$F_i = \sum_{j=1}^{k} w_j(x_i^j - B^j). \tag{B4}$$

The accuracy of the member forecasts and of EM and WEM are evaluated by computing the correlation for the forecast period and RMSE defined by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i^j - o_i)^2} \tag{B5}$$

for both SL and SR.

The RMSE is evaluated against the ensemble spread computed as the root-mean-square of the ensemble variance. For sufficiently large ensemble sizes, the following equation should be approximately verified (Fortin et al. 2014):

$$\mathrm{RMSE} \approx \sqrt{\frac{1}{N}\sum_{t=1}^{N} s_t^2} = (\overline{s_t^2})^{1/2}, \tag{B6}$$

where s^2 indicates the variance of the ensemble.

Taylor diagrams (Taylor 2001) are produced for chosen events, evaluating the model performances in terms of correlation, standard deviation (σ) and CRMSE (centered root-mean-square error) defined as

$$\mathrm{CRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(x_i^j - \overline{x_i^j}) - (o_i - \overline{o_i})]^2}. \tag{B7}$$

In Eqs. (B5)–(B7), N is the maximum number of observations in the forecast period.

REFERENCES

ARPAV, 2012: Sulla ripartizione delle portate del Po tra i vari rami e le bocche a mare del delta: Esperienze storiche e nuove indagini all'anno 2011. Arpa Veneto Tech. Rep. 02/2012, 47 pp.

Beckers, J. V. L., E. Sprokkereef, and K. L. Roscoe, 2008: Use of Bayesian model averaging to determine uncertainties in river discharge and water level forecasts. *Proc. Fourth Int. Symp. on Flood Defence: Managing Flood Risk, Reliability and Vulnerability*, Toronto, ON, Canada, Institute for Catastrophic Loss Reduction, 1–8.

Bellafiore, D., and G. Umgiesser, 2010: Hydrodynamic coastal processes in the north Adriatic investigated with a 3D finite element model. *Ocean Dyn.*, **60**, 255–273, https://doi.org/10.1007/s10236-009-0254-x.

Biolchi, L. G., S. Unguendoli, L. Bressan, M. S. B. Giambastiani, and A. Valentini, 2022: Ensemble technique application to an XBeach-based coastal Early Warning System for the northwest Adriatic Sea (Emilia-Romagna region, Italy). *Coastal Eng.*, **173**, 104081, https://doi.org/10.1016/j.coastaleng.2022.104081.

Bloomfield, P., 2000: *Fourier Analysis of Time Series: An Introduction.* John Wiley and Sons, 269 pp., https://doi.org/10.1002/0471722235.

Bressan, L., A. Valentini, T. Paccagnella, A. Montani, C. Marsigli, and M. S. Tesini, 2017: Sensitivity of sea-level forecasting to the horizontal resolution and sea surface forcing for different configurations of an oceanographic model of the Adriatic Sea. *Adv. Sci. Res.*, **14**, 77–84, https://doi.org/10.5194/asr-14-77-2017.

Buizza, R., 2019: Introduction to the special issue on "25 years of ensemble forecasting." *Quart. J. Roy. Meteor. Soc.*, **145** (Suppl. 1), 1–11, https://doi.org/10.1002/qj.3370.

Caswell, T. A., and Coauthors, 2021: matplotlib/matplotlib v3.1.3. Zenodo, accessed 15 September 2022, https://doi.org/10.5281/zenodo.3633844.

Chaumillon, E., and Coauthors, 2017: Storm-induced marine flooding: Lessons from a multidisciplinary approach. *Earth-Sci. Rev.*, **165**, 151–184, https://doi.org/10.1016/j.earscirev.2016.12.005.

Chu, P. C., 1999: Fundamental problems in coastal ocean prediction. *Proc. Oceanol. Int.*, **99**, 37–46.

Clementi, E., P. Oddo, M. Drudi, N. Pinardi, G. Korres, and A. Grandi, 2017: Coupling hydrodynamic and wave models: First step and sensitivity experiments in the Mediterranean Sea. *Ocean Dyn.*, **67**, 1293–1312, https://doi.org/10.1007/s10236-017-1087-7.

——, and Coauthors, 2021: The new Mediterranean Sea analysis and forecasting system including tides: Description and validation. *EGU General Assembly 2021*, online, EGU, EGU21-13531, https://doi.org/10.5194/egusphere-egu21-13531.

COSMO, 2004: ARPA-SIM (Bologna). *COSMO Newsletter*, No. 4, Consortium for Small-Scale Modeling, Offenbach, Germany, 25–26, http://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter04/chp4.pdf.

Di Liberto, T. D., B. A. Colle, N. Georgas, A. F. Blumberg, and A. A. Taylor, 2011: Verification of a multimodel storm surge ensemble around New York City and Long Island for the cool season. *Wea. Forecasting*, **26**, 922–939, https://doi.org/10.1175/WAF-D-10-05055.1.

Egbert, G. D., and S. Y. Erofeeva, 2002: Efficient inverse modelling of barotropic ocean tides. *J. Atmos. Oceanic Technol.*, **19**, 183–204, https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2.

Federico, I., N. Pinardi, G. Coppini, P. Oddo, R. Lecci, and M. Mossa, 2017: Coastal ocean forecasting with an unstructured grid model in the southern Adriatic and northern Ionian Seas. *Nat. Hazards Earth Syst. Sci.*, **17**, 45–59, https://doi.org/10.5194/nhess-17-45-2017.

Ferrarin, C., and Coauthors, 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., K. Horsburgh, and K. Mylne, 2009: Ensemble forecasting of storm surges. *Mar. Geod.*, **32**, 91–99, https://doi.org/10.1080/01490410902869151.

——, ——, C. Wilson, and K. Mylne, 2010: Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quart. J. Roy. Meteor. Soc.*, **136**, 1444–1456, https://doi.org/10.1002/qj.648.

Forbes, A. M. G., 1988: Fourier transform filtering: A cautionary note. *J. Geophys. Res.*, **93**, 6958, https://doi.org/10.1029/JC093iC06p06958.

Fortin, V., M. Abaza, F. Anctil, and R. Turcotte, 2014: Why should ensemble spread match the RMSE of the ensemble mean? *J. Hydrometeor.*, **15**, 1708–1713, https://doi.org/10.1175/JHM-D-14-0008.1.

Forzieri, G., and Coauthors, 2016: Multi-hazard assessment in Europe under climate change. *Climatic Change*, **137**, 105–119, https://doi.org/10.1007/s10584-016-1661-x.

Gaeta, M. G., A. G. Samaras, I. Federico, R. Archetti, F. Maicu, and G. Lorenzetti, 2016: A coupled wave–3-D hydrodynamics model of the Taranto Sea (Italy): A multiple-nesting approach. *Nat. Hazards Earth Syst. Sci.*, **16**, 2071–2083, https://doi.org/10.5194/nhess-16-2071-2016.

Garbero, V., and M. Milelli, 2020: Reforecast of the November 1994 flood in Piedmont using ERA5 and COSMO model: An operational point of view. *Bull. Atmos. Sci. Technol.*, **1**, 339–354, https://doi.org/10.1007/s42865-020-00027-0.

Gastaldo, T., V. Poli, C. Marsigli, D. Cesari, P. P. Alberoni, and T. Paccagnella, 2021: Assimilation of radar reflectivity

volumes in a pre-operational framework. *Quart. J. Roy. Meteor. Soc.*, **147**, 1031–1054, https://doi.org/10.1002/qj.3957.

Gneiting, T., and M. Katzfuss, 2014: Probabilistic forecasting. *Annu. Rev. Stat. Appl.*, **1**, 125–151, https://doi.org/10.1146/annurev-statistics-062713-085831.

Hunter, J. D., 2007: Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95, https://doi.org/10.1109/MCSE.2007.55.

IPCC, 2023: *Climate Change 2021: The Physical Science Basis*. Cambridge University Press, https://doi.org/10.1017/9781009157896, 2392 pp.

Kirezci, E., I. R. Young, R. Ranasinghe, S. Muis, R. J. Nicholls, D. Lincke, and J. Hinkel, 2020: Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st century. *Sci. Rep.*, **10**, 11629, https://doi.org/10.1038/s41598-020-67736-6.

Kjerfve, B., and K. E. Magill, 1989: Geographic and hydrodynamic characteristics of shallow coastal lagoons. *Mar. Geol.*, **88**, 187–199, https://doi.org/10.1016/0025-3227(89)90097-2.

Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216, https://doi.org/10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2.

Lellouche, J. M., and Coauthors, 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, https://doi.org/10.5194/os-14-1093-2018.

Le Traon, P. Y., and Coauthors, 2019: From observation to information and users: The Copernicus Marine Service perspective. *Front. Mar. Sci.*, **6**, 234, https://doi.org/10.3389/fmars.2019.00234.

Li, J., and B. Nie, 2017: Storm surge prediction: Present status and future challenges. *Procedia IUTAM*, **25**, 3–9, https://doi.org/10.1016/j.piutam.2017.09.002.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Madec, G., 2016: NEMO ocean engine. Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace 27, 386 pp., https://www.nemo-ocean.eu/wp-content/uploads/NEMO_book.pdf.

Maicu, F., and Coauthors, 2021: Downscaling with an unstructured coastal-ocean model to the Goro Lagoon and the Po River delta branches. *Front. Mar. Sci.*, **8**, 647781, https://doi.org/10.3389/fmars.2021.647781.

Micaletto, G., and Coauthors, 2022: Parallel implementation of the SHYFEM (System of HydrodYnamic Finite Element Modules) model. *Geosci. Model Dev.*, **15**, 6025–6046, https://doi.org/10.5194/gmd-15-6025-2022.

Milliff, R. F., A. Bonazzi, C. K. Wikle, N. Pinardi, and L. M. Berliner, 2011: Ocean ensemble forecasting. Part I: Ensemble Mediterranean winds from a Bayesian hierarchical model. *Quart. J. Roy. Meteor. Soc.*, **137**, 858–878, https://doi.org/10.1002/qj.767.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, https://doi.org/10.1002/qj.49712252905.

Niedda, M., and M. Greppi, 2007: Tidal, seiche and wind dynamics in a small lagoon in the Mediterranean Sea. *Estuarine Coastal Shelf Sci.*, **74**, 21–30, https://doi.org/10.1016/j.ecss.2007.03.022.

Owens, R., and T. Hewson, 2018: ECMWF forecast user guide. ECMWF, 16559, https://doi.org/10.21957/M1CS7H.

Palmer, T. N., 2018: The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Quart. J. Roy. Meteor. Soc.*, **145**, 12–24, https://doi.org/10.1002/qj.3383.

——, F. Molteni, R. Mureau, R. Buizza, P. Chapelet, and J. Tribbia, 1992: Ensemble prediction. ECMWF Tech. Memo. 188, 43 pp., https://www.ecmwf.int/en/elibrary/75930-ensemble-prediction.

Park, K., I. Federico, E. Di Lorenzo, T. Ezer, K. M. Cobb, N. Pinardi, and G. Coppini, 2022: The contribution of hurricane remote ocean forcing to storm surge along the southeastern U.S. coast. *Coastal Eng.*, **173**, 104098, https://doi.org/10.1016/j.coastaleng.2022.104098.

Pérez, B., and Coauthors, 2012: ENSURF: Multi-model sea level forecast implementation and validation results for the IBIROOS and western Mediterranean regions. *Ocean Sci.*, **8**, 211–226, https://doi.org/10.5194/os-8-211-2012.

Perini, L., L. Calabrese, and P. Luciani, 2019: Mareggiate: Analisi dati del 2019 e aggiornamento della sintesi 1946–2019. Regione Emilia-Romagna, 16 pp., https://ambiente.regione.emilia-romagna.it/it/geologia/geologia/costa/pdf/rapporto_mareggiate_analisi-dati2019.pdf/view.

——, ——, and ——, 2020: Mareggiate e impatti sulla costa: Aggiornamento dei dati al 2020, degli indicatori e analisi delle tendenze. Regione Emilia-Romagna, 32 pp., https://ambiente.regione.emilia-romagna.it/it/geologia/notizie/notizie-2021/mareggiate-e-impatti-aggiornamenti-al-2020-degli-indicatori-e-analisi-delle-tendenze.

Pinardi, N., A. Bonazzi, E. Scoccimarro, S. Dobricic, A. Navarra, A. Ghiselli, and P. Veronesi, 2008: Very large ensemble ocean forecasting experiment using the grid computing infrastructure. *Bull. Amer. Meteor. Soc.*, **89**, 799–804, https://doi.org/10.1175/2008BAMS2511.1.

——, ——, S. Dobricic, R. F. Milliff, C. K. Wikle, and L. M. Berliner, 2011: Ocean ensemble forecasting. Part II: Mediterranean forecast system response. *Quart. J. Roy. Meteor. Soc.*, **137**, 879–893, https://doi.org/10.1002/qj.816.

——, and Coauthors, 2017: From weather to ocean predictions: An historical viewpoint. *J. Mar. Res.*, **75**, 103–159, http://mseas.mit.edu/publications/PDF/Pinardi_etal_from_weather_to_ocean_predictions_TheSea2017.pdf.

Pistoia, J., N. Pinardi, P. Oddo, M. Collins, G. Korres, and Y. Drillet, 2016: Development of super-ensemble techniques for ocean analyses: The Mediterranean Sea case. *Nat. Hazards Earth Syst. Sci.*, **16**, 1807–1819, https://doi.org/10.5194/nhess-16-1807-2016.

Rochford, P. A., 2016: SkillMetrics: A Python package for calculating the skill of model predictions against observations. GitHub, accessed 20 October 2022, http://github.com/PeterRochford/SkillMetrics.

Russo, A., A. Coluccelli, S. Carniel, A. Benetazzo, A. Valentini, T. Paccagnella, M. Ravaioli, and G. Bortoluzzi, 2013: Operational models hierarchy for short term marine predictions: The Adriatic Sea example. *2013 MTS/IEEE OCEANS-Bergen*, Bergen, Norway, Institute of Electrical and Electronics Engineers, 1–6, https://doi.org/10.1109/OCEANS-Bergen.2013.6608139.

Salighehdar, A., Z. Ye, M. Liu, I. Florescu, and A. F. Blumberg, 2017: Ensemble-based storm surge forecasting models. *Wea. Forecasting*, **32**, 1921–1936, https://doi.org/10.1175/WAF-D-17-0017.1.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periáñez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quart.*

J. Roy. Meteor. Soc., **142**, 1453–1472, https://doi.org/10.1002/qj.2748.

Shchepetkin, A. F., and J. C. McWilliams, 2005: The Regional Oceanic Modeling System (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modell.*, **9**, 347–404, https://doi.org/10.1016/j.ocemod.2004.08.002.

Steppeler, J., G. Doms, U. Schättler, H. W. Bitzer, A. Gassmann, U. Damrath, and G. Gregoric, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteor. Atmos. Phys.*, **82**, 75–96, https://doi.org/10.1007/s00703-001-0592-9.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192, https://doi.org/10.1029/2000JD900719.

Thomson, R. E., and W. J. Emery, 2014: Digital filters. *Data Analysis Methods in Physical Oceanography*, 3rd ed. R. E. Thomson and W. J. Emery, Eds., Elsevier, 593–637, https://doi.org/10.1016/B978-0-12-387782-6.00006-5.

Tolman, H. L., 2009: User manual and system documentation of WAVEWATCH III version 3.14. MMAB Rep. 276, 220 pp., http://polar.ncep.noaa.gov/mmab/papers/tn276/MMAB_276.pdf.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

Trotta, F., I. Federico, N. Pinardi, G. Coppini, S. Causio, E. Jansen, D. Iovino, and S. Masina, 2021: A relocatable ocean modeling platform for downscaling to shelf-coastal areas to support disaster risk reduction. *Front. Mar. Sci.*, **8**, 642815, https://doi.org/10.3389/fmars.2021.642815.

Umgiesser, G., D. M. Canu, A. Cucco, and C. Solidoro, 2004: A finite element model for the Venice Lagoon: Development, set-up, calibration and validation. *J. Mar. Syst.*, **51**, 123–145, https://doi.org/10.1016/j.jmarsys.2004.05.009.

——, C. Ferrarin, A. Cucco, F. De Pascalis, D. Bellafiore, M. Ghezzo, and M. Bajo, 2014: Comparative hydrodynamics of 10 Mediterranean lagoons by means of numerical modeling. *J. Geophys. Res. Oceans*, **119**, 2212–2226, https://doi.org/10.1002/2013JC009512.

——, ——, M. Bajo, I. Federico, J. Alessandri, D. Bellafiore, T. Chegini, and A. Pinsky, 2020: SHYFEM-model/shyfem: Covid edition (VERS_7_5_70). Zenodo, accessed 10 November 2021, https://doi.org/10.5281/zenodo.3833857.

——, and Coauthors, 2021: The prediction of floods in Venice: Methods, models and uncertainty (review article). *Nat. Hazards Earth Syst. Sci.*, **21**, 2679–2704, https://doi.org/10.5194/nhess-21-2679-2021.

Wahl, T., I. D. Haigh, R. J. Nicholls, A. Arns, S. Dangendorf, J. Hinkel, and A. B. A. Slangen, 2017: Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nat. Commun.*, **8**, 16075, https://doi.org/10.1038/ncomms16075.

Walters, R. A., and C. Heston, 1982: Removing tidal-period variations from time-series data using low-pass digital filters. *J. Phys. Oceanogr.*, **12**, 112–115, https://doi.org/10.1175/1520-0485(1982)012<0112:RTPVFT>2.0.CO;2.

Warner, J. C., B. Armstrong, R. He, and J. B. Zambon, 2010: Development of a coupled ocean-atmosphere-wave-sediment transport (COAWST) modeling system. *Ocean Modell.*, **35**, 230–244, https://doi.org/10.1016/j.ocemod.2010.07.010.

Williams, J., K. J. Horsburgh, J. A. Williams, and R. N. F. Proctor, 2016: Tide and skew surge independence: New insights for flood risk. *Geophys. Res. Lett.*, **43**, 6410–6417, https://doi.org/10.1002/2016GL069522.