Ensemble Storm Surge Modelling of the Goro Lagoon (Adriatic Sea, Northern Italy)

Jacopo Alessandri[1,2], Nadia Pinardi[2], Ivan Federico[3] and Andrea Valentini[4]

[1] Interdepartmental Research Centre for Environmental Sciences (CIRSA), University of Bologna, Ravenna, Italy

[2] Department of Physics and Astronomy (DIFA), University of Bologna, Bologna, Italy

[3] Ocean Prediction and Applications (OPA), Euro-Mediterranean Center on Climate change (CMCC), Lecce, Italy

[4] Regional Agency for Prevention, Environment and Energy of Emilia-Romagna (Arpae), Bologna, Italy

Corresponding author: Jacopo Alessandri (jacopo.alessandri2@unibo.it)

Key Points:

- Storm surge forecast skill depends mainly on initial and lateral open boundary condition uncertainties, including the tidal components.
- EPS members derived from different large scale ocean models, atmospheric forcing forecasts and an ad-hoc perturbations of the river runoff.
- Development of an ensemble weighting method provides an improved ensemble mean forecast skill suitable for operational systems.

Abstract

We developed a storm surge Ensemble Prediction System (EPS) for the Goro lagoon (GOLFEM-EPS) in the Northern Adriatic Sea. The lagoon is threatened every year by storm surge events with consequent risks for human life and economic losses. We show the advantages and limitations of an EPS with 45 members, using a very high-resolution unstructured grid finite element model. 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 run-off, initial and lateral boundaries are evaluated, 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.

Plain Languages Summary

Storm surges are extreme events generated by strong winds, low atmospheric pressure, tides and possibly other non-local phenomena. The Goro Lagoon, on the Italian side of the northern Adriatic Sea, is subject to storm surge, especially when a south-easterly wind (locally named Scirocco) occurs. We developed an ensemble system, called GOLFEM-EPS, to forecast with a numerical model the

sea level in the Goro lagoon and the adjacent areas. The "ensemble" approach consists in running a high number of forecasts (the "members", 45 in this work) with different conditions in order to obtain different estimates and account for errors and inaccuracies in the model set-up and in the forcing data. GOLFEM-EPS generally improves the forecast skill on the third day of forecast while it is similar in the first two days to the reference forecast. The uncertainty connected to river-run off, meteorological forcing, initial and boundary conditions is evaluated. Initial, lateral boundary conditions and tidal components provide the largest uncertainty to the system.

1 Introduction

The operational forecast of sea level (SL) is a widely developed service used to prevent flooding and storm surge hazards that represent a potential threat for human life and activities (Chaumillon et al., 2017; Forzieri et al., 2016). However, the forecast lead time is affected by the limited predictability of the atmospheric forcing and the uncertainties regarding the initial conditions, as well as lateral boundary forcing specifications.

Addressing the problem of storm surge forecast reliability involves recasting the single, deterministic forecast in probabilistic terms (Gneiting and Katzfuss, 2014). Atmospheric seasonal and sub-seasonal 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 non-linear system. In the 1960s it was already clear that there was a "limit of deterministic predictability" in weather forecasting (Palmer, 2018). During the 1980s and 1990s ensemble weather forecasting systems were developed leading to an operational system at ECMWF (European Medium Range Weather Forecast, Palmer et al., 1992; Molteni et al., 1996; Buizza, 2019) and at NMC (National Meteorological Center; 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, 2021).

In oceanography, the multimodel super-ensemble concept was applied for the Mediterranean Sea SST forecasting by Pistoia et al., (2016) using a multiple linear regression technique applied to a multi-physics and multi-model 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 studied. 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 mulitimodel EPS has been applied in the New York coastal area (Liberto et al., 2011).

In this study we focus on the development of an EPS methodology for a limited area, coastal model at very high resolution. The novelty is that we explore a methodology to create a set of ensemble members to tackle the specific coastal modelling uncertainties. In coastal domains, the dimension of the domain, the bathymetry, the type of 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 parametrizations and the specific numerical schemes used in the ocean model add further uncertainty to the system.

In the Adriatic Sea, storm surges are more common when Scirocco winds (south easterly) blow for several days (Orlić, Kuzmić, and Pasarić 1994). However, Bora winds (north easterly), which usually have the largest amplitude, are the most threatening, especially if concomitant with high tides and preceded by Scirocco winds (Lionello et al., 2021).

Seiches can also be generated in response to the winds. The Scirocco blowing over the Adriatic Sea can trigger the fundamental mode of seiches with a period of 21.2 *hr* and a decay time of 3.2 *days* (Cerovečki et al.,1997) and amplitude rising up to 50 cm (Godin and Trotti 1975). Hence, seiches can contribute to high SL values if they overlap with a storm surge (Vilibić, 2006; Bajo et al., 2019). The second and third modes of Adriatic Seiches have a 10.7 *hr* and 6.7 *hr* period respectively (Raicich, 1999) and a smaller amplitude.

This study focuses on the Goro Lagoon (Fig. 1), which is exposed to storm surges that threaten the towns around the lagoon (Bondesan et al., 1995). In November 1966, a dramatic flood (Garnier et al., 2018) caused by a persistent Scirocco wind, caused severe damage to buildings, roads and cultivated lands. In Goro town almost 80 families were moved away from their homes. In October 2018, another Scirocco event caused flooding in the Goro harbour. Thus, we tried to define an EPS that could

potentially be beneficial to the local population if made operational and converted into an early warning decision support system.

The forecasting model is a baroclinic very high-resolution model (GOLFEM) described in Maicu et al., (2021). We tested the EPS methodology considering five recent extreme events, using forcing from different meteorological and ocean operational models for initial and lateral boundary conditions, together with river 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

2.1 Sea Level Extreme Events in the Goro Lagoon

The GOLFEM-EPS area of interest 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.

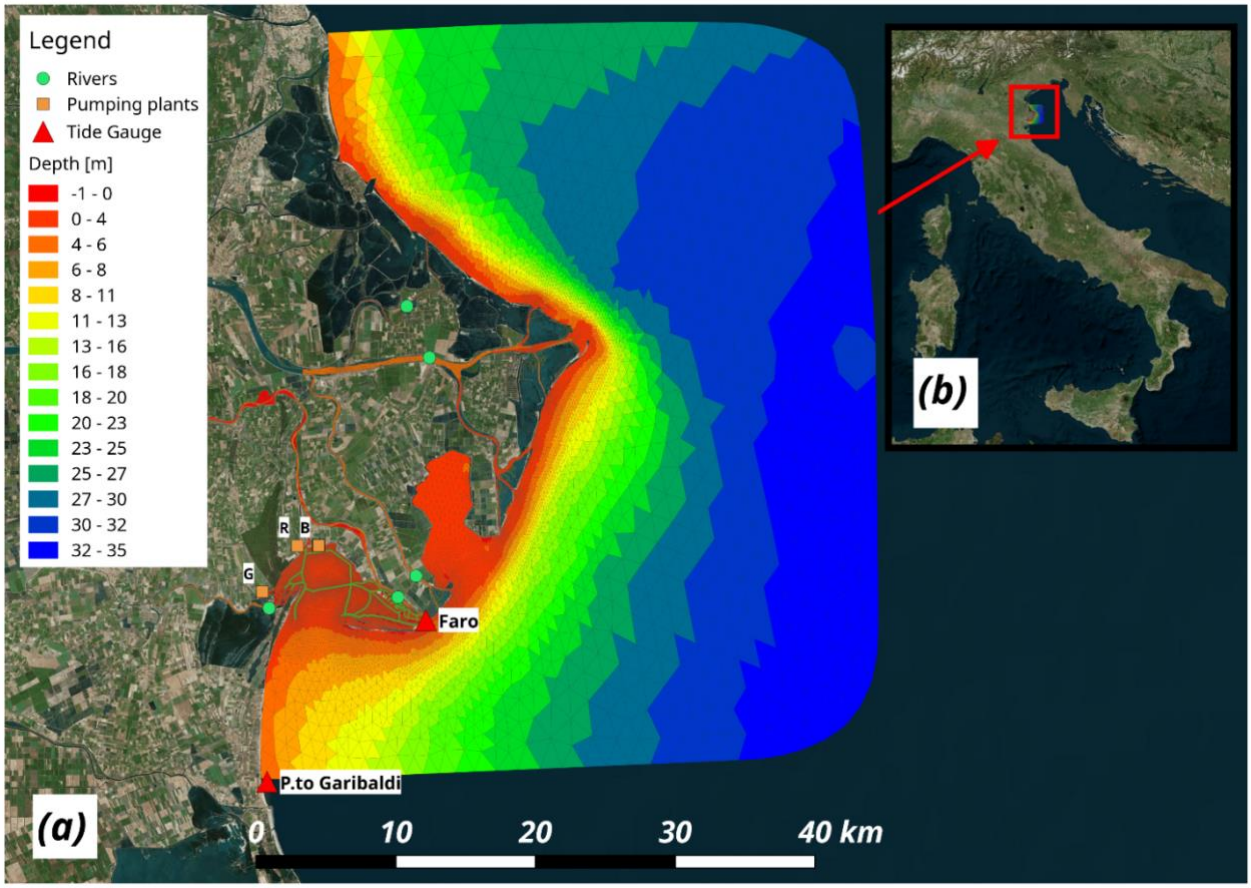

Figure 1: (a) Bathymetry and coastlines of the Goro Lagoon. The red triangles indicate the position of Faro and Porto Garibaldi tide gauges. The green circles indicate the position of rivers, starting from the south-west, the Po di Volano, the Po di Goro, the Po della Donzella, the Po di Pila and the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

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 ER coast. Sea level observations collected at Faro and Porto Garibaldi (Fig. 1) are available every 10 minutes, thus giving a very good determination of extremes. Table 1 describe the five tide events from both Porto Garibaldi and Faro tide gauges.

Fig. 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 de-tiding procedure described in the Appendix A. The residual signals only conserve the surge contribution due to local wind, atmospheric pressure, and large scale non-tidal 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. The events listed in Table 1 all have a non-negligible surge component.

The time of the peak sea level recorded for each event 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. Considering percentiles between 95 and 99 is quite common 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 m 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.

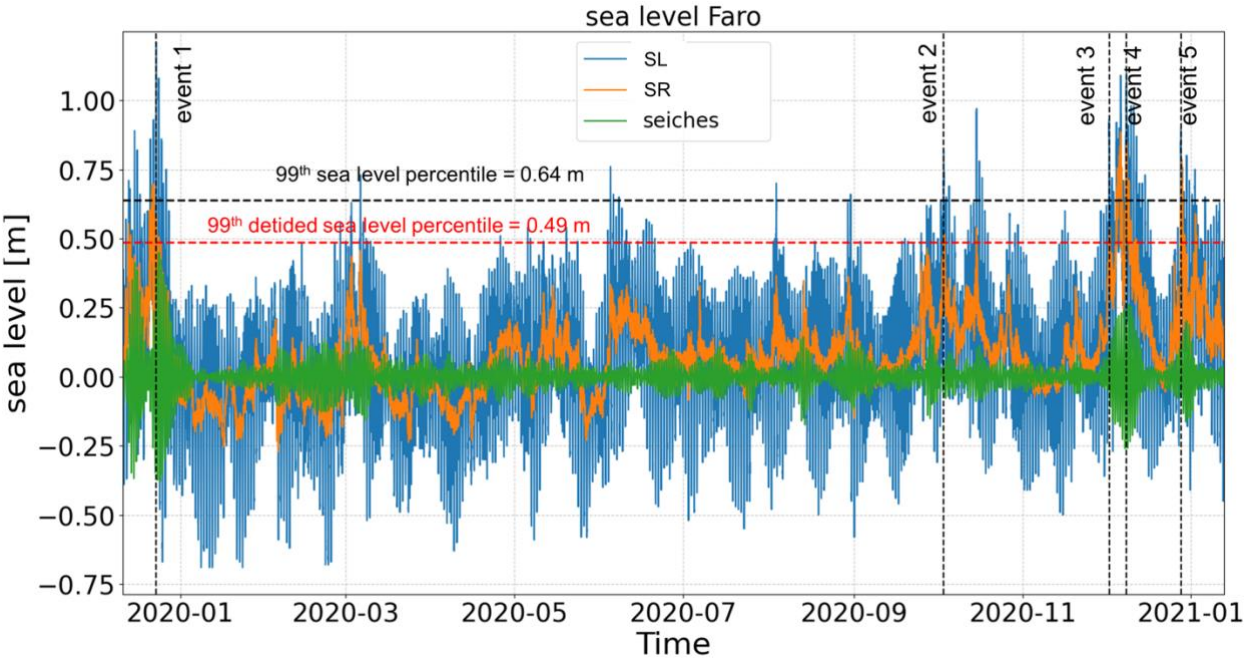

Figure 2: SL observations (blue), SR (orange) and seiches (green) at Faro (see Fig. 1 for reference position). The vertical black dashed lines indicate the 5 extreme events considered in this paper. The horizontal dashed black and red lines indicate the 99th percentile of the SL and SR, respectively.

EVENT NO.	DATE	PORTO GARIBALDI TIME(UTC), SL(M), SR(M)	FARO TIME(UTC), SL(M), SR(M)	IMPACTS	PREVAILING WINDS
1	2019-12-23	07:40 -> **1.17, 0.56**	07:50 -> **1.21, 0.60**	Uniform large impacts	NW
2	2020-10-03	10:50 -> **0.86, 0.51**	10:30 -> **0.82, 0.51**	No impacts	SE
3	2020-12-02	08:20 -> **1.02, 0.61**	08:30 -> **0.93, 0.54**	Uniform large impacts	N-NE
4	2020-12-08	15:30 -> **1.06, 0.96**	15:10 -> **1.10, 0.85**	Ferrara province	E-SE
5	2020-12-28	08:30 -> **0.90, 0.72**	09:00 -> **0.90, 0.78**	Minor impacts Volano	SE

Table 1: The five events analyzed in this study, with date, hour, max observed sea level (SL) and max observed surge residual (SR) at P.to Garibaldi and Faro (Fig.1). Impacts and prevailing winds for each event are also indicated (NW-North Westerly; SE-South Easterly; N-NE-North-North Easterly; E-SE- East-South Easterly; SE- South Easterly).

2.2 The circulation model

GOLFEM is the implementation of the SHYFEM model (Umgiesser et al., 2004; Bellafiore et al., 2010, Micaletto et al., 2022) in the Goro Lagoon. SHYFEM is an open-source (https://github.com/SHYFEM-model/shyfem) unstructured grid baroclinic ocean model that solves

the primitive equations for the ocean 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). GOLFEM has been extensively calibrated, validated and described in Maicu et al. (2021).

GOLFEM is nested in large scale ocean models and is forced by atmospheric numerical weather prediction fields. The river runoff is also imposed as a lateral open boundary condition. All the ocean and meteorological products characteristics used to initialize and force the different members are described in the next section and are summarized in Table S1 of the Supporting information (SI).

2.3 Nesting in large scale circulation models

GOLFEM-EPS 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; 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 (Clementi et al., 2017, 2021, hereafter MED-MFC) is a coupled current-wave model of the Mediterranean Sea and the adjacent Atlantic area. It is composed of the general circulation model NEMO (Nucleus for European Modelling of the Ocean-NEMO; Madec, 2008) and coupled 2-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 correct the model output. 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-Cimate 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 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-newsletter, 2004). ADRIAC provides three day forecasts per day, with a one-day spin-up, 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 three-day hourly forecast with a one-day spin-up as for ADRIAC.

2.4 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 three days and 6 hours 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 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 00 and 12 UTC providing a 72-hour 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 $\sim$ 2.2 km and 65 vertical layers. It provides a *48 hr* forecast (hourly output) and two forecasts per day at 00 and 12 UTC.

2.5 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 repartition equation that fits data from Arpav (2012). At the western side of the lagoon, the Po of Volano runoff is provided as the sum of all the pumping plant discharges which constitute the final flow of the Po of Volano. The runoff is imposed as a lateral boundary

condition. 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.

3 The Ensemble Prediction System

3.1 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).

		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

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

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 00:00 UTC, forecast lateral boundary conditions and surface meteorological fields are used for a three days forecast, up to +72 *hr*. ADRIAC and ADRIAROMS models do not have an operational data assimilation scheme. For ADRIAC and ADRIAROMS the analyses are thus simulations forced by meteorological and open ocean lateral boundary analyses. Since COSMO-2I provide only a 48 *hr* forecast, from +48 *hr* to +72 *hr*, the meteo fields from COSMO-5M are used.

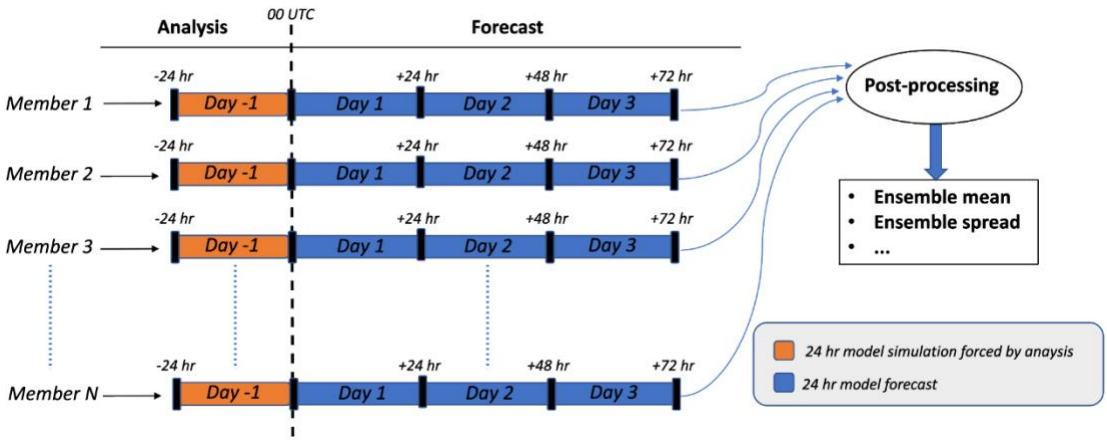

Figure 3: The GOLFEM-EPS modelling scheme. The post-processing creates ensemble mean, ensemble spread and the weighted ensemble mean.

Fig. 4 shows the GOLFEM-EPS basin average kinetic energy (KE) for one member simulation (exp-1) 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 hours the simulations were found to reach a similar KE. This explains the choice of only one day as a spin-up time for the forecast, which is a reasonable time for the ocean limited-area domains. It is also well known that the spin-up time decreases with the scale of the implemented computational domain, and one day of spin-up has also been used by authors dealing with very limited coastal models (Gaeta et al., 2016).

The GLOBAL and MED-MFC products are provided without tides. In the case of non-tidal models, the tidal sea level is extracted from the TPXO model and added to the GOLFEM open boundaries. Despite the importance of baroclinic pressure gradients during storm surge events (Staneva et al., 2016), preliminary simulations have compared the sea level of baroclinic and barotropic runs, highlighting only a negligible difference, probably due to the very small and shallow areas considered. We argue that for SL forecasts, the baroclinicity may be site-by-site dependent. However, building a multi-hazard system means that all processes need to be considered from the start. Incremental improvements will thus be made, also considering the progress in SL forecasting.

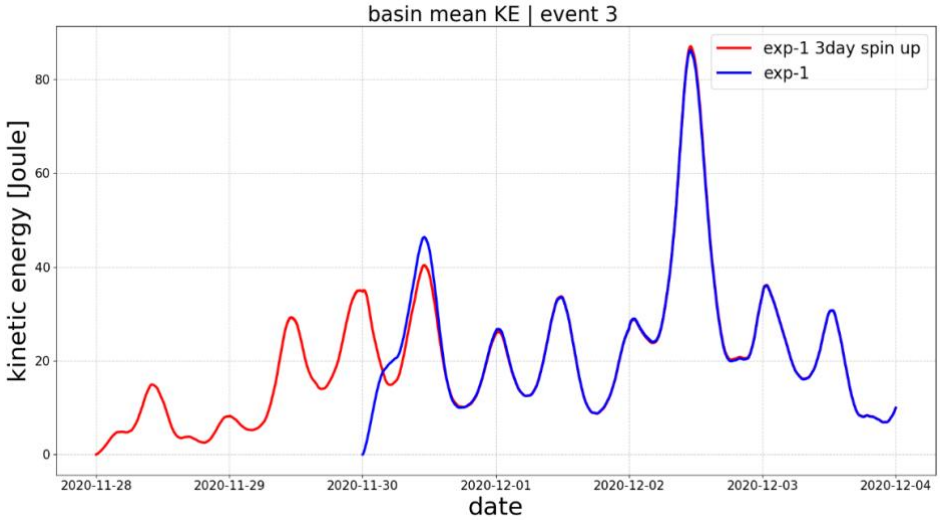

Figure 4: Mean basin kinetic energy of exp-1 considering one (blue) and three (red) days of spin-up.

3.2 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 $X = \{x_i^j\}_{(i,j)\in\{1,\dots,T\}\times\{1,\dots,m\}}$ 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 F_i of all times i. The EM is a simple average of the member forecasts defined as:

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

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. The specific WEM methodology is detailed in Appendix B.

4 Results

4.1 EPS Skill

The SL and SR are shown for events 2, 3 in Fig. 5 and events 4 and 5 in Fig. 6. We will discuss event 1 in the last part of this section. The thick green line (exp-1) is the member forced by the highest resolution models (Adriac and COSMO-2I; see 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 four events.

The ensemble spread grows and peaks at the time of the extreme event. This is another benefit of the EPS which alerts the forecaster of the possibility of a large event even if not capturing the correct amplitude. 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.

The performances 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), for each group there is a smaller dispersion due to the different meteorological forcings and to a lesser extent, due to 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.

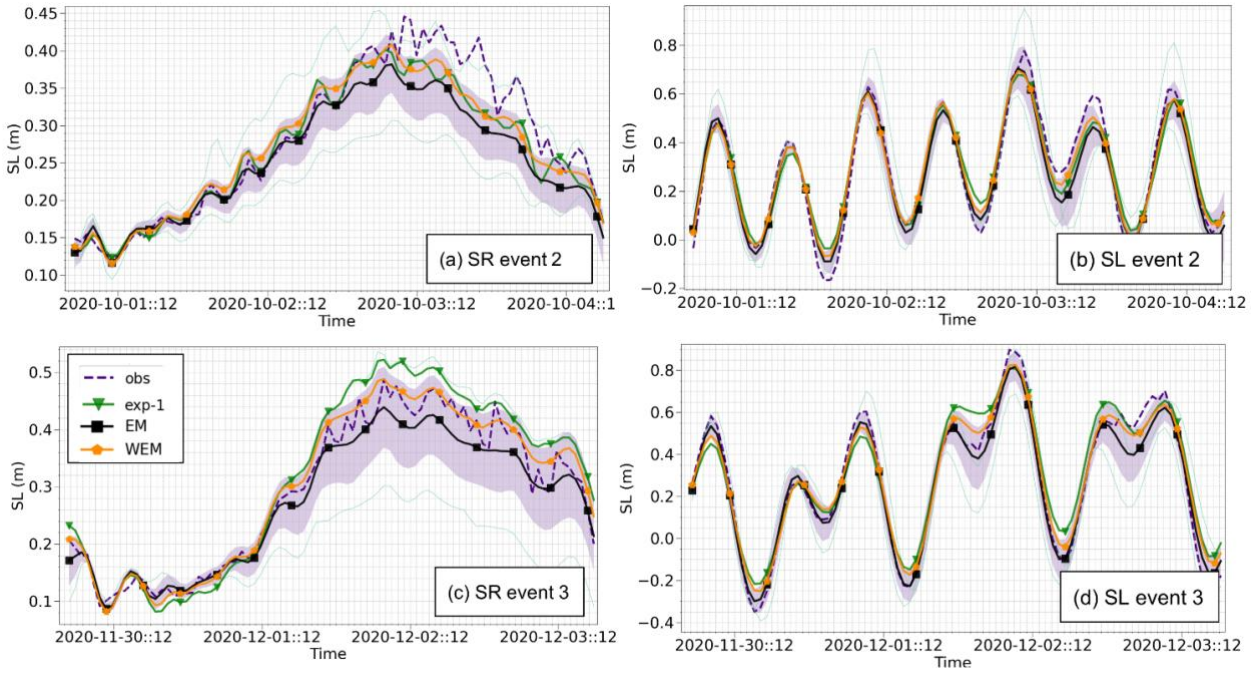

Figure 5: SR (left column) and total SL (right column) forecast comparison at Faro for event 2 (a,b) and event 3 (c,d). The green line shows exp-1 considering the reference deterministic forecast as described in Table 2. EM and WEM indicate the Ensemble Mean and the Weighted Ensemble Mean. The light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

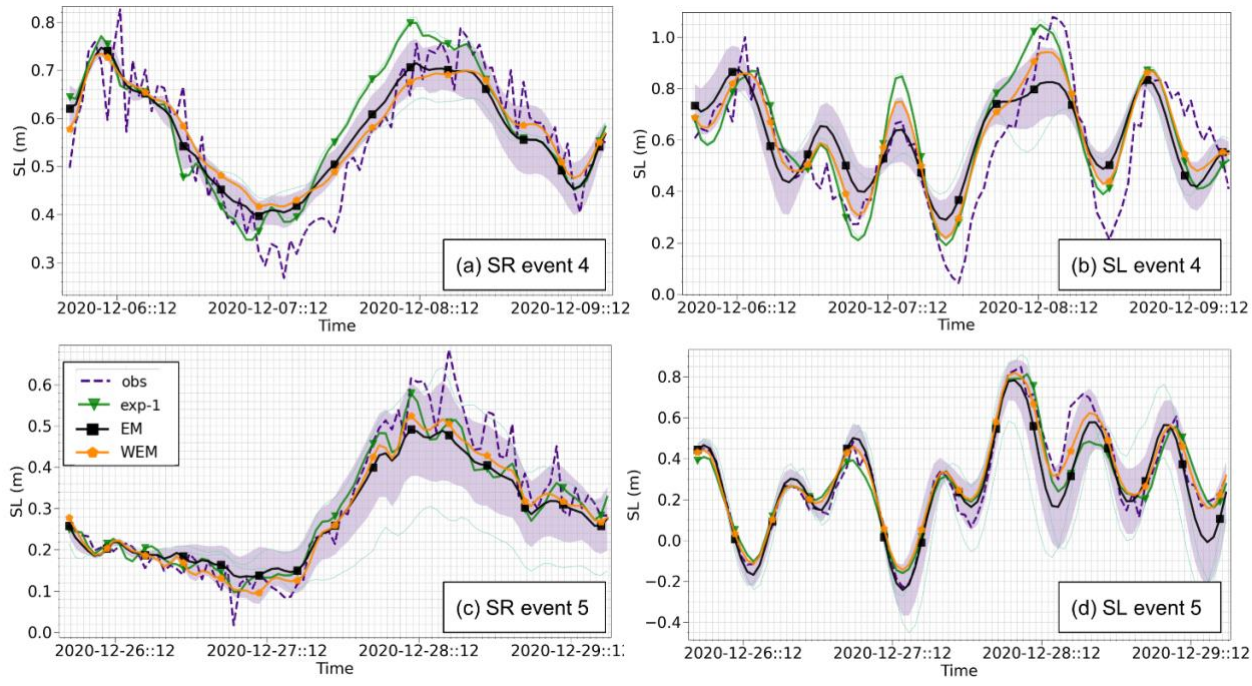

Figure 6: As in Fig. 5 but for event 4 (a,b) and event 5 (c,d).

In both events 2 and 3, the maximum surge occurred with the maximum tidal amplitude, with some impacts at the level of the coast. 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. The coincidence of remotely forced

coastally trapped sea level maxima with tidal and local wind forcing has also been examined in detail for the US East Coast (Park et al., 2022). A correlation between the arrival time of remote signals was found to be the most important predictor of sea level extremes.

The Taylor diagrams provide an indication of the members that most contribute to build 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 (exp-1 to exp-9; ρ between 0.85 and 0.95; ~ 0.5 for event 1 SR) and MED-MFC (exp-19 to exp-27; ρ between 0.85 and 0.99; ~ 0.7 for event 1 SR). Members initialized/forced by ADRIAROMS (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 (exp-28 to exp-36; ρ between 0.4 to 0.9; ~ -0.4 for event 1 SR).

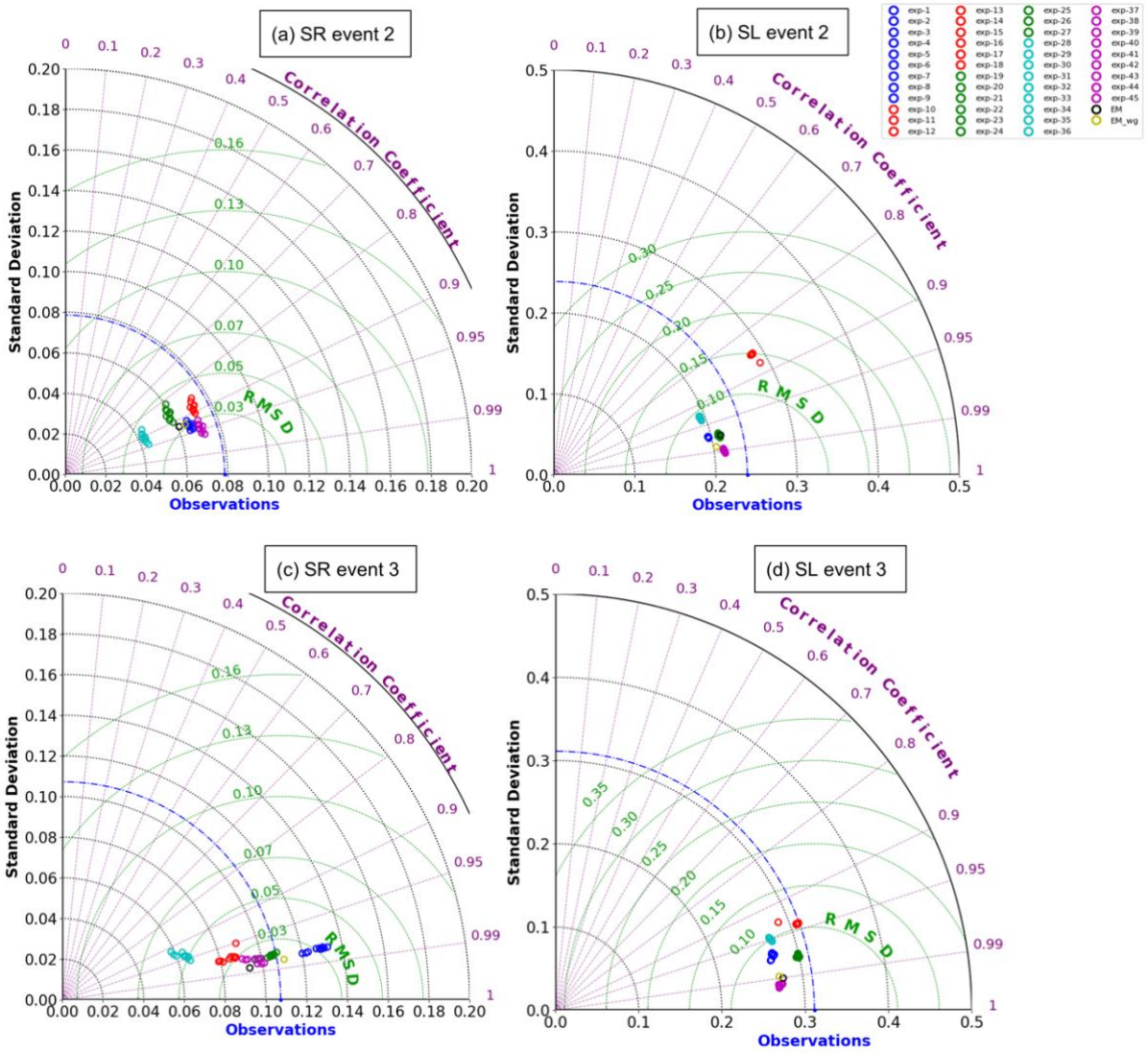

Figure 7: Taylor diagrams for SR and total SL for event 2 (a,b) and event 3 (c,d).

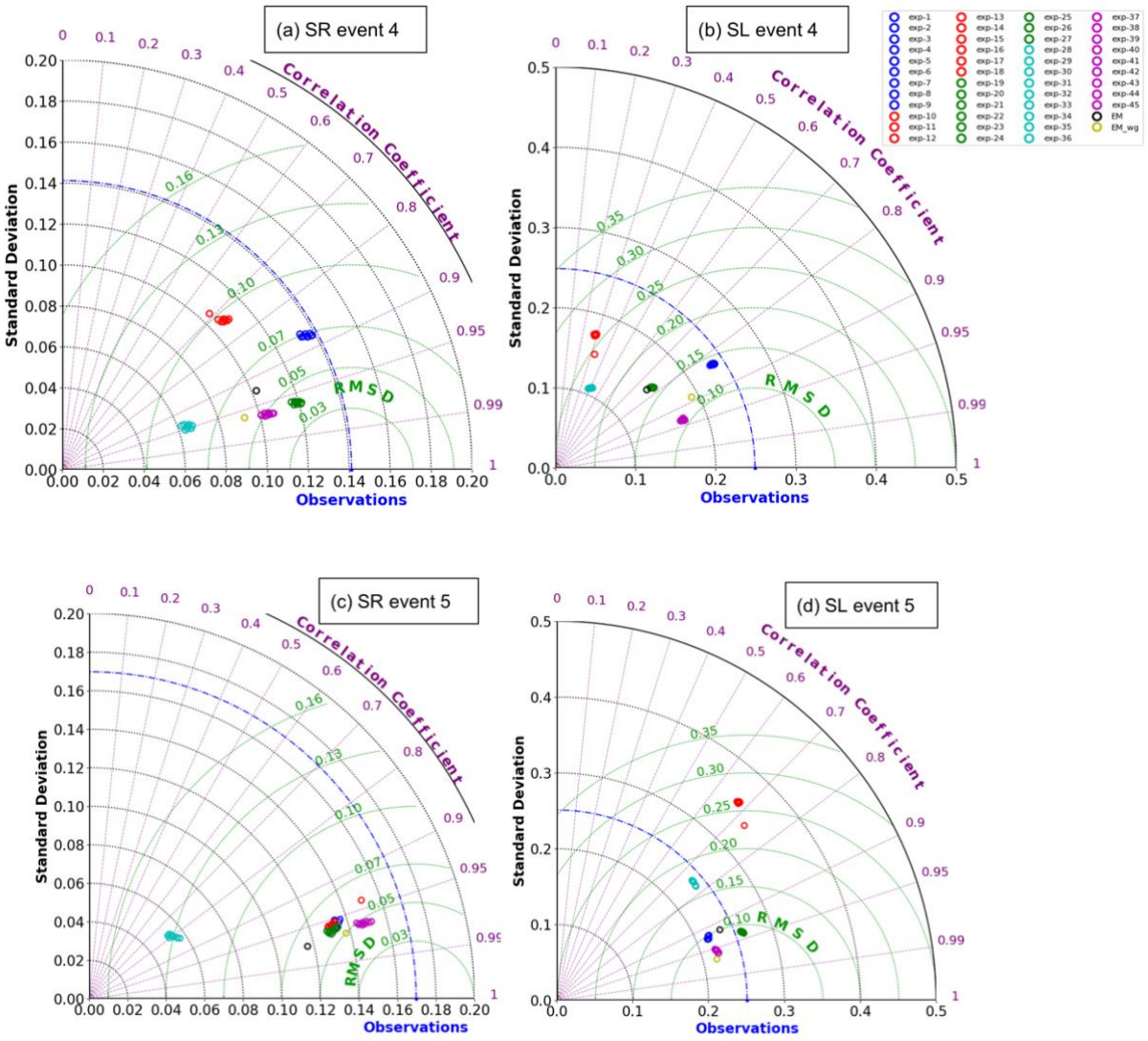

Figure 8: Taylor diagrams for SR and total SL for event 4 (a,b) and event 5 (c,d).

If SL Taylor diagrams are considered, we note again that members forced by MED-MFC-T have the best performance. This suggests that a Mediterranean scale ocean model may be more effective in representing the non-linear interactions that occur between tides, bathymetry and Adriatic basin modes compared to Adriatic scale models (ADRIAC or ADRIAROMS) that imposes TPXO tides at the southern boundary.

Event 1 is a special case because the contribution of the Adriatic Sea seiches reached values close to 40 cm (Fig. 2). Fig. 9a shows that the SL peak coincides with a decreasing SR (Fig. 9a and 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 (Fig. 9c and d). However, EM and WEM benefit from error compensation and show satisfactory results in terms of SR, although none of the members reproduced the peak of SL occurring on 23 December 2019 at 07:50. 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.

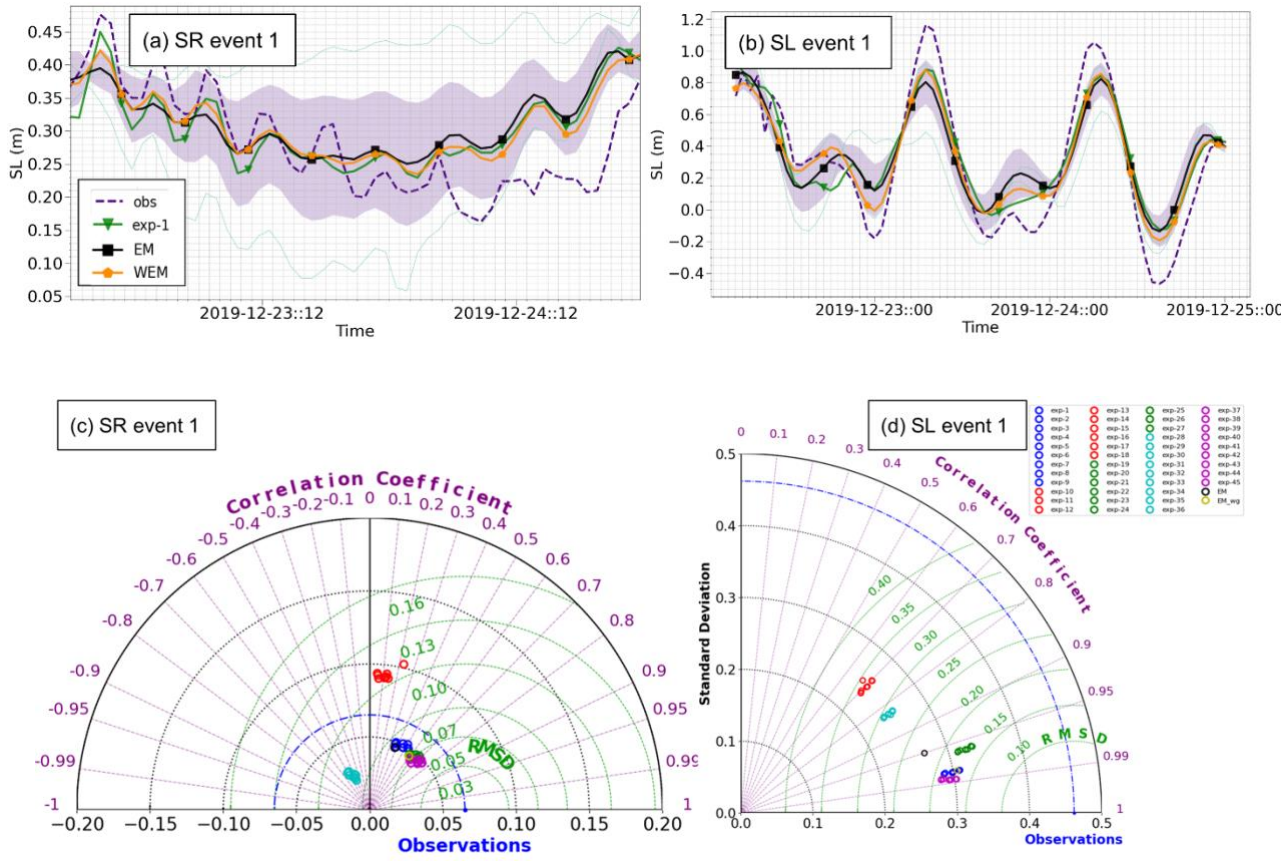

Figure 9: SR (a) and SL (b) forecast comparison for event 1. The light green lines represent the members with maximum and minimum sea level values during the peak. The shaded area is the ensemble spread. Taylor diagrams for SR (c) and SL (d) are also shown.

4.2 Ensemble Member Analysis

Any analysis of an EPS should consider its degree of the over or under-dispersiveness of the members. To do this, we analyzed the root mean square error (RMSE) as a function of the ensemble spread (Figs. 10a and 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.

In an ideal case the RMSE should be linearly proportional to the ensemble spread. Fig. 10a shows that there is a slight over-dispersion 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 over-dispersion of the ensemble is due to the very different initial and lateral boundary conditions used. If the SL is considered, over-dispersion appears for spread values greater than 0.13

cm, while a slight under-dispersion 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 *hr*, while after 48 *hr*, both EM and WEM performances are slightly better. The SL RMSE shows almost a 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.

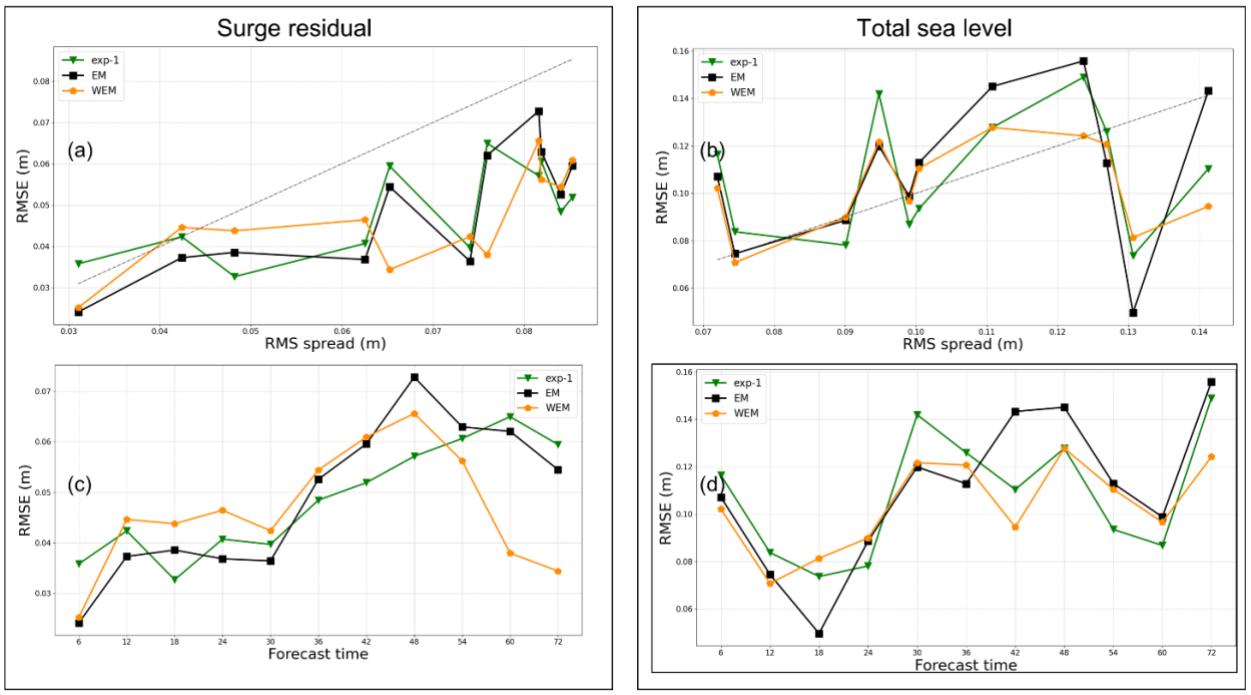

Figure 10: RMSE against ensemble spread and forecast time for SR (a,c) and total sea level (b,d). Data are computed aggregating all the events at Porto Garibaldi and Faro stations.

We conclude this section by analyzing the spatial distribution of the ensemble spread at the time of each peak event, taking into account only the SR. Considering first the river and meteorological forcing induced ensemble spread, the river run-off 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 small size of the domain, and the fact that the storm surge component due to remote wind effects is important, as discussed earlier.

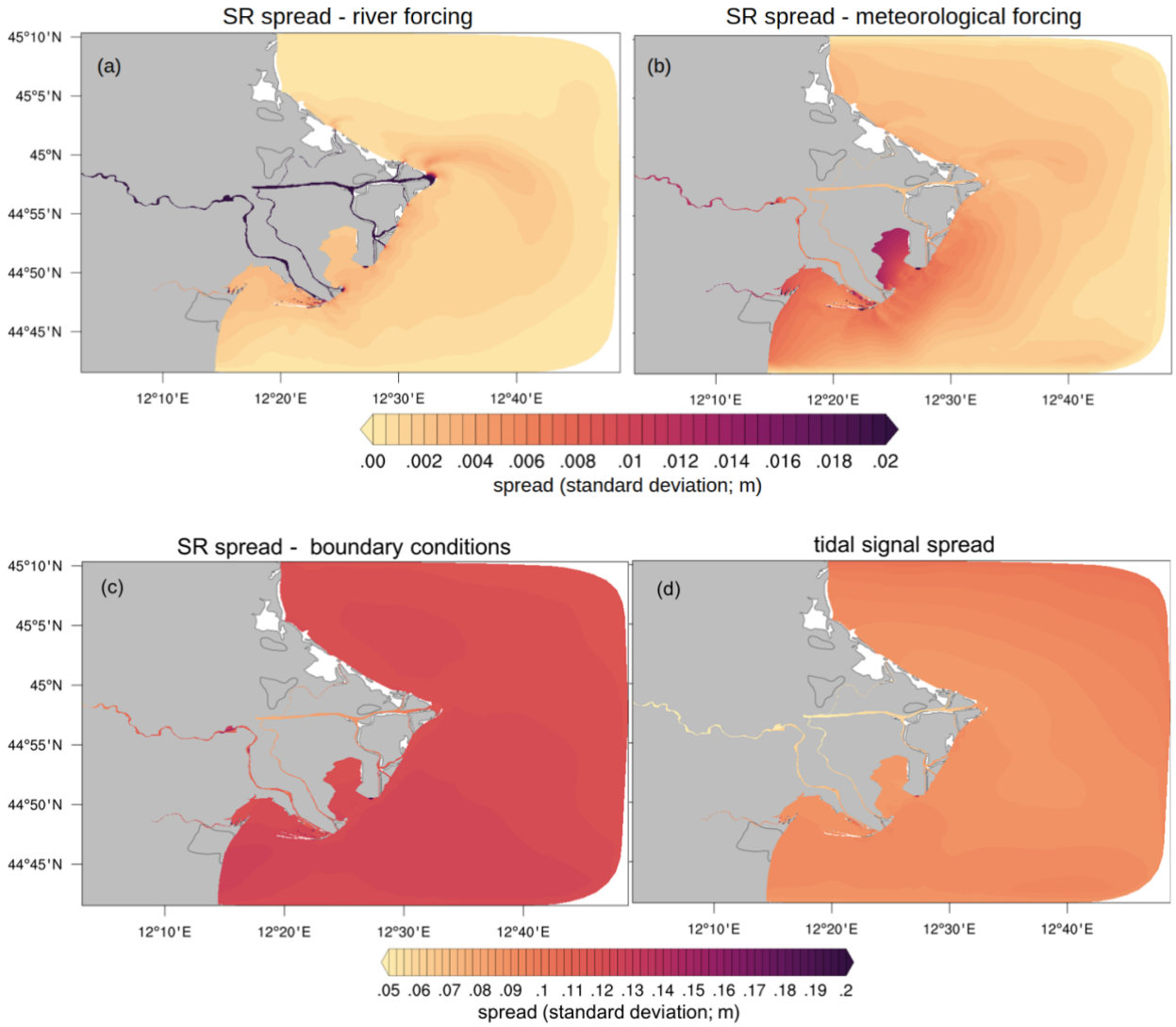

Figure 11: Spread due to river forcing (a), meteorological forcing (b), boundary conditions (c) and tidal signal (d). Data are computed aggregating all the events.

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 small domains should focus on the initial and lateral boundary condition uncertainties, further substantiating previous results (Chu, 1999).

5 Summary and Conclusions

We designed and developed a coastal Ensemble Prediction System for the Goro Lagoon (GOLFEM-EPS) using a baroclinic very high-resolution unstructured grid model. The EPS is composed of 45

393 members produced by model forecasts carried out with different meteorological forcing, initial and
394 lateral boundary conditions from coarser resolution ocean models and perturbations to the river
395 discharge. The forecast lead time is 3 days with a one day spin-up in analysis/simulation mode.
396 The total SL RMSE computed at the two reference stations is 11 and 12 cm respectively for WEM
397 and EM, while it is 11.5 cm for the deterministic reference forecast. One case, event 1, showed large
398 errors, due to the specific importance of seiches and tidal elevation in the process and the limitation
399 of the circulation models to reproduce such basin modes and their non -linear interaction with tides.
400 The correct simulation of the tidal phase is also crucial and a way to account for the shifts in the tidal
401 peak is to consider the skew surge (Williams et al., 2016). Such an approach will be further
402 investigated for the domain considered here.
403 A post-processing ensemble mean (EM) procedure was carried out and a weighted ensemble mean
404 (WEM) methodology was tested based on the performance of the members during the spin-up which
405 was considered to be equivalent to a training period. Of the 45 members used for the ensemble, most
406 of the variability was reached including only members that differed in initial and lateral boundary
407 conditions. An estimate of the uncertainty due to different initial/lateral boundary conditions, river
408 runoff and atmospheric forcings was provided. A perturbation of ± 30 % in the river run off was
409 found to contribute with 3-4 mm in the ensemble sea level spread limited to the river mouths.
410 Meteorological forcing had a greater impact of between 1 and 3 cm, while initial/boundary conditions
411 (including tides) provided most of the sea level uncertainty between 9 and 13 cm.
412 It is found that EM and WEM forecasts have better skill from 48 to 72 *hr* forecast lead time while in
413 the first two days they are equivalent to a single deterministic forecast skill.
414 The conclusions regarding the skill of WEM or EM for 3 days sea level forecast is probably ultimately
415 connected to the specific model domain size and the specific area dynamics. For general ensemble
416 coastal forecasting, several areas should be tried with the same WEM and EM methodologies. For
417 the Goro Lagoon specifically, different model size domains are needed, and the optimal model
418 domain size should be found in order to obtain a reduced SL forecast error. The limitations in
419 validating the model domain size derive from the small number of observations available which
420 makes it difficult to produce a robust probabilistic verification of the EPS. In the future it is possible
421 that coastal scale altimetry will offer a convenient data set for validation/calibration of a storm surge
422 EPS.
423 The lesson learnt from this exercise, most probably a generalizable conclusion, is that limited area
424 storm surge forecasting uncertainty is dominated by initial and lateral boundary conditions and by
425 phase/amplitude errors in tidal components. The former are also influenced by remotely forced
426 propagating signals such as seiches for the specific case of the Adriatic but more generally remotely

forced circulation signals. Given the necessity of having resolutions of ~100 m at the coasts, we conclude that an EPS might be useful also at the scale of the regional models that embed the limited area coastal models.

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 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

$$X'(f) = W(f)X(f) \qquad \textbf{A 1}$$

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 non-zero Fourier coefficients (Forbes, 1988). We tapered the transition bands using a Tukey window (also called cosine-taper), defined as

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

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 of SI 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 of SI. 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, 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 the 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).

$$B^j = \frac{1}{T}\sum_{i=1}^{T}\left(x_i^j - o_i\right) \quad \textbf{B.1}$$

where B^j is the bias for member j. 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

$$\rho_o^j = \frac{\frac{1}{T}\sum_{i=1}^{T}\left(x_i^j - \overline{x^j}\right)(o_i - \overline{o})}{x_{\sigma j} o_\sigma} \quad \textbf{B.2}$$

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

$$w_j = \frac{\rho_o^j}{\sum_{j=1}^{k}\rho_o^j} \quad \textbf{B.3}$$

4. the WEM is computed as the weighted average of the selected forecasts

$$F_i = \sum_{j=1}^{k} w_j\left(x_i^j - B^j\right) \quad \textbf{B.4}$$

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

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i^j - o_i\right)^2} \quad \textbf{B.5}$$

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).

$$RMSE \approx \sqrt{\frac{1}{N}\sum_{t=1}^{N} s_t^2} = \left(\overline{s_t^2}\right)^{\frac{1}{2}} \quad \textbf{B.6}$$

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

$$CRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\left(x_i^j - \overline{x_i^j}\right) - \left(o_i - \overline{o_i}\right)\right)^2} \quad \textbf{B.7}$$

In eqs. B.5, B.6 and B.7, *N* is the maximum number of observations in the forecast period.

Acknowledgments

This work was supported by ARPAE grant to the University of Bologna and the Department of Physics and Astronomy supporting the PhD of Jacopo Alessandri. 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) 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 figures 7, 8 and 9 were made with Skillmetrics python package (Rockford, 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.

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. Technical report 02/2012. Vicenza: Arpa Veneto.

Bajo, M., Meugorac, I., Umgiesser, G., & Orlić M. (2019). Storm Surge and Seiche Modelling in the Adriatic Sea and the Impact of Data Assimilation. *Quarterly Journal of the Royal Meteorological Society* 145 (722): 2070–84. https://doi.org/10.1002/qj.3544.

Beckers, J. V. L., Sprokkereef, E., & Roscoe K. L. (2008). Use of Bayesian Model Averaging to Determine Uncertainties in River Discharge and Water Level Forecasts. In: *proc. 4th International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability*, Toronto, Ontario, Canada.

Bellafiore, D., Umgiesser, G. (2010). Hydrodynamic coastal processes in the north Adriatic investigated with a 3D finite element model. *Ocean Dynamics*, 60(2):255-273. Doi: 10.1007/s10236-009-0254-x.

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

Bloomfield, P. (2000). *Fourier Analysis of Time Series*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471722235.

Bondesan, M., Castiglioni, G. B., Elmis, C., Gabbianellis, G., Marocco R., Pirazzoli, P. A., & Tomasin A. (1995). Coastal Areas at Risk from Storm Surges and Sea-Level Rise in Northeastern Italy. *Journal of Coastal Research* 11 (4): 1354–79.

Bressan, L., Valentini, A., Paccagnella, T., Montani, A., Marsigli, C., & Tesini, M. S.(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.". *Q. J. R. Meteorol. Soc*. 145, 1–11. https://doi.org/10.1002/qj.3370.

Caswell, T. A., Droettboom, M., Lee, A., Andrade, E. S. D., Hunter, J.,Hoffmann, T., Firing, E., Klymak, J., Stansby, D., Varoquaux, N., Nielsen, J. H., Root, B., May, R., Elson, P., Seppänen, J. K., Dale, D., Jae-Joon Lee, McDougall, D., Straw, A., Hobson, P., Gohlke, C., Yu, T. S, Ma, E., Hannah, A., Vincent, A. F., Silvester, S., Moad, C., Kniazev, N., Ernest, E., & Ivanov, P. (2021). Matplotlib v3.4.0. [Software] https://doi.org/10.5281/zenodo.3633844.

Cerovečki, I., Orlić, M., & Hendershott, M. C. (1997). Adriatic Seiche Decay and Energy Loss to the Mediterranean. *Deep Sea Research Part I: Oceanographic Research Papers* 44 (12): 2007–29. https://doi.org/10.1016/s0967-0637(97)00056-3.

Chaumillon, E., Bertin, X., Fortunato, A. B., Bajo, M., Schneider, J. L., Dezileau, L., Patrick Walsh, J., et al. (2017). Storm-Induced Marine Flooding: Lessons from a Multidisciplinary Approach. *Earth-Science Reviews* 165 (February): 151–84. https://doi.org/10.1016/j.earscirev.2016.12.005.

Chu, P. C. (1999). Fundamental Problems in Coastal Ocean Prediction. *Proceedings of Oceanology International 99*, 37–46.

Clementi, E., Goglio, A. C., Aydogdu, A., Pistoia, J., Escudier, R., Drudi, M., Grandi A., Mariani, A., Lyubartsev, V., Lecci, R., Creti, S., Masina, S., Coppini, G., & Pinardi, N. (2021). The New Mediterranean Sea Analysis and Forecasting System Including Tides: Description and Validation. EGU general assembly 2021, online, 19–30 Apr 2021, EGU21-13531 . https://doi.org/10.5194/egusphere-egu21-13531.

Clementi, E., Oddo, P., Drudi, M., Pinardi, N., Korres, G., & Grandi, A. (2017). Coupling Hydrodynamic and Wave Models: First Step and Sensitivity Experiments in the Mediterranean Sea. *Ocean Dynamics* 67 (10): 1293–1312. https://doi.org/10.1007/s10236-017-1087-7.

COSMO-newsletter. (2004). Operational Applications - ARPA-SIM (BOLOGNA). 25-26. Deutsch. WetterDienst (DWD) Offenbach 6.

Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient Inverse Modelling of Barotropic Ocean Tides. *J. Atmosph. Oceanic Technol.* 19 (2): 183–204. https://doi.org/10.1175/1520-0426(2002)019<0183:eimobo>2.0.co;2.

Federico, I., Pinardi, N., Coppini, G., Oddo, P., Lecci, R., & Mossa, M. (2017). Coastal Ocean Forecasting with an Unstructured Grid Model in the Southern Adriatic and Northern Ionian Seas. *Natural Hazards and Earth System Sciences* 17 (1): 45–59. https://doi.org/10.5194/nhess-17-45-2017.

Ferrarin, C., Valentini, A., Vodopivec, M., Klaric, D., Massaro, G., Bajo, M., De Pascalis, F., Fadini, A., Ghezzo, M., Menegon, S., Bressan, L., Unguendoli, S., Fettich, A., Jerman, J., Ličer, M., Fustar, L., Papa, A., and Carraro, E. (2020). Integrated Sea Storm Management Strategy: The 29 October 2018 Event in the Adriatic Sea. *Natural Hazards and Earth System Sciences* 20 (1): 73–93. https://doi.org/10.5194/nhess-20-73-2020.

Flowerdew, J., Horsburgh, K., & Mylne, K. (2009). Ensemble Forecasting of Storm Surges. *Marine Geodesy* 32 (2): 91–99. https://doi.org/10.1080/01490410902869151.

Flowerdew, J., Horsburgh, K., Wilson, C., & Mylne, K. (2010). Development and Evaluation of an Ensemble Forecasting System for Coastal Storm Surges. *Quarterly Journal of the Royal Meteorological Society* 136 (651): 1444–56. https://doi.org/10.1002/qj.648.

Forbes, A. M. G. (1988). Fourier Transform Filtering: A Cautionary Note. *Journal of Geophysical Research* 93 (C6): 6958. https://doi.org/10.1029/jc093ic06p06958.

Fortin, V., Abaza, M., Anctil, F., & Turcotte, R. (2014). Why Should Ensemble Spread Match the RMSE of the Ensemble Mean? *Journal of Hydrometeorology* 15 (4): 1708–13. https://doi.org/10.1175/jhm-d-14-0008.1.

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. *Climatic Change* 137 (1-2): 105–19. https://doi.org/10.1007/s10584-016-1661-x.

Gaeta, M. G., Samaras, A.G., Federico, I., Archetti, R., Maicu, F., & Lorenzetti, G. (2016). A Coupled Wave 3-d Hydrodynamics Model of the Taranto Sea (italy): A Multiple-Nesting Approach. *Natural Hazards and Earth System Sciences* 16 (9): 2071–83. https://doi.org/10.5194/nhess-16-2071-2016.

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

Garnier, E., Ciavola, P., Spencer, T., Ferreira, O., Armaroli, C., & McIvor, A. (2018). Historical Analysis of Storm Events: Case Studies in France, England, Portugal and Italy. *Coastal Engineering* 134 (April): 10–23. https://doi.org/10.1016/j.coastaleng.2017.06.014.

Gastaldo, T., Poli, V., Marsigli, C., Cesari, D., Alberoni, P. P., & Paccagnella, T. (2021). Assimilation of radar reflectivity volumes in a pre-operational framework. *Quarterly Journal of the Royal Meteorological Society. Royal Meteorological Society (Great Britain)*, (qj.3957). doi:10.1002/qj.3957

Gneiting, T. & Katzfuss, M. (2014). Probabilistic Forecasting. *Annual Review of Statistics and Its Application*, 1:1, 125-151

Godin, G., & Trotti, L. (1975). *Trieste, Water Levels 1952-1971: A Study of the Tide, Mean Level and Seiche Activity*. Vol. 28. Miscellaneous Special Publcation.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, May-June 2007, doi: 10.1109/MCSE.2007.55.

IPCC. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Kirezci, E., Young, I.R., Ranasinghe, R., Muis, S., Nicholls, R.J., Lincke, D., & Hinkel, J. (2020). Projections of Global-Scale Extreme Sea Levels and Resulting Episodic Coastal Flooding over the 21st Century. *Scientific Reports* 10 (1). https://doi.org/10.1038/s41598-020-67736-6.

Krishnamurti, T. N., Kishtawal, C.M., Zhang, Z., LaRow, T., Bachiochi, D., Williford, E., Gadgil, S., & Surendran, S. (2000). Multimodel Ensemble Forecasts for Weather and Seasonal Climate. *Journal of Climate* 13 (23): 4196–216. https://doi.org/10.1175/1520-0442(2000)013<4196:meffwa>2.0.co;2.

Lellouche, J.M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., et al. (2018). Recent Updates to the Copernicus Marine Service Global Ocean Monitoring and Forecasting Real-Time 1⁄12 High-Resolution System. *Ocean Science* 14 (5): 1093–1126. https://doi.org/10.5194/os-14-1093-2018.

Le Traon P.Y., Reppucci A., Alvarez Fanjul E., Aouf L., Behrens A., Belmonte M., Bentamy A., Bertino L., Brando V.E., Kreiner M.B., Benkiran M., Carval T., Ciliberti S.A., Claustre H., Clementi E., Coppini G., Cossarini G., De Alfonso Alonso-Muñoyerro M., Delamarche A., Dibarboure G., Dinessen F., Drevillon M., Drillet Y., Faugere Y., Fernández V., Fleming A., Garcia-Hermosa M.I., Sotillo M.G., Garric G., Gasparin F., Giordan C., Gehlen M., Gregoire M.L., Guinehut S., Hamon M., Harris C., Hernandez F., Hinkler J.B., Hoyer J., Karvonen J., Kay S., King R., Lavergne T., Lemieux-Dudon B., Lima L., Mao C., Martin M.J., Masina S., Melet A., Buongiorno Nardelli B., Nolan G., Pascual A., Pistoia J., Palazov A., Piolle J.F., Pujol M.I., Pequignet A.C., Peneva E., Pérez Gómez B., Petit de la Villeon L., Pinardi N., Pisano A., Pouliquen S., Reid R., Remy E., Santoleri R., Siddorn J., She J., Staneva J., Stoffelen A., Tonani M., Vandenbulcke L., von Schuckmann K., Volpe G., Wettre C. & Zacharioudaki A. (2019) From Observation to Information and Users: The Copernicus Marine Service Perspective. *Front. Mar. Sci*. 6:234. Doi: 10.3389/fmars.2019.00234

Liberto, T. D., Colle, B.A., Georgas, N., Blumberg, A.F., & Taylor, A.A. (2011). Verification of a Multimodel Storm Surge Ensemble Around New York City and Long Island for the Cool Season. *Weather and Forecasting* 26 (6): 922–39. https://doi.org/10.1175/waf-d-10-05055.1.

Lionello, P., Barriopedro, D., Ferrarin, C., Nicholls, R., Orlić, M., Raicich, F., Reale M., Umgiesser, G., Vousdoukas, M., & Zanchettin, D. (2021). Extreme Floods of Venice: Characteristics, Dynamics, Past and Future Evolution (review Article). *Natural Hazards and Earth System Sciences* 21 (8): 2705–31. https://doi.org/10.5194/nhess-21-2705-2021.

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* 20 (2): 130–41. https://doi.org/10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2.

Madec, G. 2008. *NEMO Ocean Engine*. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619.

Maicu, F., Alessandri, J., Pinardi, N., Verri, G., Umgiesser, G., Lovo, S., Turolla, S., Paccagnella, T., & Valentini, A. (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., Barletta, I., Mocavero, S., Federico, I., Epicoco, I., Verri, G., Coppini, G., Schiano, P., Aloisio, G., & Pinardi, N. (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., Bonazzi, A., Wikle, C.K., Pinardi, N., & Berliner, L.M. (2011). Ocean Ensemble Forecasting. Part I: Ensemble Mediterranean Winds from a Bayesian Hierarchical Model. *Quarterly Journal of the Royal Meteorological Society* 137 (657): 858–78. https://doi.org/10.1002/qj.767.

Molteni, F., Buizza, R., Palmer, T.N., & Petroliagis, T. (1996). The ECMWF Ensemble Prediction System: Methodology and Validation. *Quarterly Journal of the Royal Meteorological Society* 122 (529): 73–119. https://doi.org/10.1002/qj.49712252905.

Orlić, M., Kuzmić, M., & Pasarić, Z. (1994). Response of the Adriatic Sea to the Bora and Sirocco Forcing. *Contin. Shelf Res.* 14 (1): 91–116. https://doi.org/10.1016/0278-4343(94)90007-8.

Owens, R, & Hewson, T. (2018). ECMWF Forecast User Guide. https://doi.org/10.21957/M1CS7H.

Palmer, T. N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P., & Tribbia, J. (1992). Ensemble Prediction, no. 188 (July): 43. https://doi.org/10.21957/igxccor4n.

Palmer, T. (2018). The ECMWF Ensemble Prediction System: Looking Back (more than) 25 years and Projecting Forward 25 years. *Quarterly Journal of the Royal Meteorological Society* 145 (S1): 12–24. https://doi.org/10.1002/qj.3383.

Park, K., Federico, I., Di Lorenzo, E., Ezer, T., Cobb, K.M., Pinardi, N. & Coppini, G. (2022). The Contribution of Hurricane Remote Ocean Forcing to Storm Surge along the Southeastern U.S. coast. *Coastal Engineering*, 173.

Perini, L., Calabrese, L., & Luciani, P. (2019). Mareggiate: Analisi Dati Del 2019 e Aggiornamento Della Sintesi 1946-2019. Bologna: Regione Emilia-Romagna.

Perini, L., Calabrese, L., Luciani P. (2020). Mareggiate e Impatti Sulla Costa: Aggiornamento Dei Dati Al 2020, Degli Indicatori e Analisi Delle Tendenze. Bologna: Regione Emilia-Romagna.

Pérez, B., Brouwer, R., Beckers, J., Paradis, D., Balseiro, C., Lyons, K., Cure, M., Sotillo, M.G., Hackett, B., Verlaan, M. & Fanjul E. A. (2012). ENSURF: Multi-Model Sea Level Forecast Implementation and Validation Results for the IBIROOS and Western Mediterranean Regions. *Ocean Science* 8 (2): 211–26. https://doi.org/10.5194/os-8-211-2012.

Pinardi, N., Bonazzi, A., Dobricic, S., Milliff, R.F., Wikle, C.K. & Berliner, L.M. (2011). Ocean Ensemble Forecasting. Part II: Mediterranean Forecast System Response. *Quarterly Journal of the Royal Meteorological Society* 137 (657): 879–93. https://doi.org/10.1002/qj.816.

Pinardi, N., Bonazzi, A., Scoccimarro, E., Dobricic, S., Navarra, A., Ghiselli, A. & Veronesi, P. (2008). Very Large Ensemble Ocean Forecasting Experiment Using the Grid Computing Infrastructure. *Bulletin of the American Meteorological Society* 89 (6): 799–804. https://doi.org/10.1175/2008bams2511.1.

Pistoia, J., Pinardi, N., Oddo, P., Collins, M., Korres, G. & Drillet, Y. (2016). Development of Super-Ensemble Techniques for Ocean Analyses: The Mediterranean Sea Case. *Natural Hazards and Earth System Sciences* 16 (8): 1807–19. https://doi.org/10.5194/nhess-16-1807-2016.

Raicich, F. F. (1999). A Case Study of the Adriatic Seiches (december 1997). *Nuovo Cimento Della Societa Italiana Di Fisica C* 22 (5): 715–26.

Rochford, P. A. (2016). SkillMetrics: A Python package for calculating the skill of model predictions against observations. [Software]. http://github.com/PeterRochford/SkillMetrics

Russo, A., Coluccelli, A., Carniel, S., Benetazzo, A., Valentini, A., Paccagnella, T., Ravaioli, M., & Bortoluzzi, G. (2013). Operational models hierarchy for short term marine predictions: the Adriatic Sea example, in: OCEANS-Bergen, 2013 MTS/IEEE Jun 10 (pp. 1–6), IEEE.

Salighehdar, A., Ye, Z., Liu, M., Florescu, I., & Blumberg, A.F. (2017). Ensemble-Based Storm Surge Forecasting Models. *Weather and Forecasting* 32 (5): 1921–36. https://doi.org/10.1175/waf-d-17-0017.1.

Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A. & Potthast, R. (2016). Kilometre-Scale Ensemble Data Assimilation for the COSMO Model (KENDA). *Quarterly Journal of the Royal Meteorological Society* 142 (696): 1453–72.

Shchepetkin, A. F. & McWilliams, J.C. (2005). The Regional Oceanic Modeling System (ROMS): A Split-Explicit, Free-Surface, Topography-Following-Coordinate Oceanic Model. *Ocean Modelling* 9 (4): 347–404. https://doi.org/10.1016/j.ocemod.2004.08.002.

Staneva, J., Wahle, K., Koch, W., Behrens, A., Fenoglio-Marc, L. & Stanev, E.V. (2016). Coastal Flooding: Impact of Waves on Storm Surge During Extremes, a Case Study for the German Bight. *Natural Hazards and Earth System Sciences* 16 (11): 2373–89. https://doi.org/10.5194/nhess-16-2373-2016.

Steppeler, J., Doms, G., Schättler, U., Bitzer, H.W., Gassmann, A., Damrath, U. & Gregoric, G. (2003). Meso-Gamma Scale Forecasts Using the Nonhydrostatic Model LM. *Meteorog. Atmos. Phys.* 82 (1-4): 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. *Journal of Geophysical Research: Atmospheres* 106 (D7): 7183–92. https://doi.org/10.1029/2000jd900719.

Thomson, R. E. & Emery, W.J. (2014). Chapter 6 - Digital Filters. In *Data Analysis Methods in Physical Oceanography (third Edition)*, edited by Richard E. Thomson and William J. Emery, Third Edition, 593–637. Boston: Elsevier. https://doi.org/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*. NOAA/NWS/NCEP/MMAB Tech. Note 276, pp 194.

Toth, Z. & Kalnay, E. (1993). Ensemble Forecasting at NMC: The Generation of Perturbations. *Bulletin of the American Meteorological Society* 74 (12): 2317–30. https://doi.org/10.1175/1520-0477(1993)074<2317:efantg>2.0.co;2.

Trotta F., Federico I., Pinardi N., Coppini G., Causio S., Jansen E., Iovino D. & Masina S. (2021). A Relocatable Ocean Modeling Platform for Downscaling to Shelf-Coastal Areas to Support Disaster Risk Reduction. *Front. Mar. Sci.* 8:642815. Doi: 10.3389/fmars.2021.642815

Umgiesser, G., Canu, D.M., Cucco, A., Solidoro, C. (2004). A Finite Element Model for the Venice Lagoon. Development, set-up, calibration and validation. *Journal of Marine Systems*, 51 (1-4):123-145. Doi: 10.1016/j.jmarsys.2004.05.009.

Umgiesser, G., Ferrarin, C., Bajo, M., Federico, I., Alessandri, J., Bellafiore, D., Chegini, T, & Pinsky, A. (2020). SHYFEM-model/shyfem: Covid edition (VERS_7_5_70). [Software] Zenodo. https://doi.org/10.5281/zenodo.3833857.

Vilibić, I. (2006). The Role of the Fundamental Seiche in the Adriatic Coastal Floods. *Continental Shelf Research* 26 (2): 206–16. https://doi.org/10.1016/j.csr.2005.11.001.

Wahl, T., Haigh, I.D., Nicholls, R.J., Arns, A., Dangendorf, S., Hinkel, J. & Slangen, A.B.A. (2017). Understanding Extreme Sea Levels for Broad-Scale Coastal Impact and Adaptation Analysis. *Nature Communications* 8 (1). https://doi.org/10.1038/ncomms16075.

Walters, R. A. & Heston, C. (1982). Removing Tidal-Period Variations from Time-Series Data Using Low-Pass Digital Filters. *Journal of Physical Oceanography* 12 (1): 112–15. https://doi.org/10.1175/1520-0485(1982)012<0112:rtpvft>2.0.co;2.

Warner, J. C., Armstrong, B., He, R. & Zambon, J.B. (2010). Development of a Coupled Ocean-atmosphere-wave-sediment Transport (COAWST) Modeling System. *Ocean Modelling* 35 (3): 230–44. https://doi.org/10.1016/j.ocemod.2010.07.010.1

Williams, J., Horsburgh, K.J., Williams, J.A. & Proctor, R.N.F. (2016). Tide and skew surge indepen-dence: New insights for flood risk. *Geophys. Res. Lett.*,43, 6410–6417, doi:10.1002/2016GL069522.