https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


# Modelling the barotropic sea level in the Mediterranean using data assimilation

Marco Bajo[1], Christian Ferrarin[1], Georg Umgiesser[1,2], Andrea Bonometto[3], and Elisa Coraci[3]

[1]Institute of Marine Sciences, National Research Council, Castello 2737/F, 30122 Venice, Italy
[2]Klaipėda University, Coastal Research and Planning Institute, H.Manto 84, 92294 Klaipėda, Lituania
[3]Institute for Environmental Protection and Research, S. Marco, 4665, 30122 Venice, Italy

**Correspondence:** Marco Bajo (marco.bajo@ve.ismar.cnr.it)

**Abstract.** This paper analyses the variability of the sea level in the Mediterranean Sea and its reproduction with a barotropic model, with and without applying data assimilation. The impact of data assimilation is considered in hindcast and forecast simulations, considering its usefulness for both reanalysis studies and short-term forecasts. We used a two-dimensional finite element barotropic model with an ensemble Kalman filter, assimilating hourly sea-level observations from 50 stations along all the Mediterranean coasts. The results show a great improvement given by data assimilation in hindcast simulations for the reproduction of astronomical tide, surge and total sea level, even in coastal areas far from the assimilated stations (e.g. the Eastern Mediterranean Sea). The improvement is consistent also in forecasts, especially for the first day ( 37% average error reduction), and in case of storm surge events with a strong presence of seiche oscillations. Since these oscillations depend only on the initial state and not on the boundary conditions, they are corrected very effectively by data assimilation. Finally, based on observations, this article estimates the periods of the normal barotropic modes (seiches) in the Adriatic Sea, where they have been extensively studied, and in the Mediterranean Sea, where the present documentation is scarce.

## 1 Introduction

The Mediterranean Sea and the Adriatic Sea (Fig. 1), which is part of it, are among the most studied basins in the world due to their historical and geopolitical importance. However, the barotropic oscillations of the sea level (Pugh, 1996), which are mainly the astronomical tide, the surge and the seiches (i.e., free oscillations following the normal barotropic modes of the basin), are mainly studied in the Adriatic where they are more energetic and can cause flooding in coastal areas and in particular in Venice (Cavaleri et al., 2019; Ferrarin et al., 2021).

In the Adriatic Sea, the shallow water of the continental shelf, present in the central part and especially in the north, favours the growth of barotropic sea-level anomalies, whether they are of astronomical tides, surges or seiches. Indeed, the Northern Adriatic Sea is one of the Mediterranean regions (together with the Gulf of Gabes) experiencing the highest tidal oscillation (about 1 m at spring tide; Tsimplis et al., 1995). The presence, especially in autumn, of strong winds from the southeast (Scirocco), which blow along the main axis of the basin, favour storm surge events in the north; events that in turn can trigger seiche oscillations of considerable intensity (Međugorac et al., 2016). Floods on the northern coasts of the Adriatic may therefore depend on the superimposition of astronomical tides, surges and seiches resulting from previous storm surge events.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


In densely populated cities with important cultural heritage, such as Venice, Dubrovnik and Alexandria, it is essential to have a correct forecast of the sea level from nowcasting up to about five days of lead time. While the astronomical tide, given its periodic nature, is easily predictable (at least where observations are available), the surge, which depends mainly on the wind, is often affected by errors, mostly due to the incorrect representation of the meteorological forcing (Barbariol et al., 2022). Consequently, the seiche oscillations, resulting from an initial storm surge event, are also incorrect and the error propagates in the following days with the same period and damping as those of the triggered barotropic mode. The forecast error can be reduced by improving the meteorological forcings, as regards the surge component, and improving the initial state regarding the seiche oscillations (Bajo et al., 2019). The use of data assimilation (DA) techniques (Kalnay, 2002; Evensen, 2009; Carrassi et al., 2018), which aims to reduce the error of the initial state from which to start for a forecast simulation, should be able to reduce both the error on the seiche oscillations and, in the short term, errors in the surge, due to underestimates of the wind, in the case of events ongoing at the time of the forecast (Bajo et al., 2017, 2019).

Regarding the situation outside the Adriatic Sea, the barotropic components of the sea level are much weaker. However, the western Mediterranean basin is subject to strong Mistral events (north-west wind) and, in the southern part of the Mediterranean, small but intense cyclones with tropical dynamics (called medicanes) can sometimes form. These extreme weather events have already caused flooding in the past even in areas traditionally not affected by these events (Scicchitano et al., 2021).

Considering the seiche oscillations, although they are less energetic than in the Adriatic Sea, their correct reproduction improves the reproduction of the total sea level. These oscillations, in the Mediterranean Sea, are little studied and, to our knowledge, there is only one scientific work with the application of a simple barotropic model, which predicts the form and period of the most important modes (Schwab and Rao, 1983). Therefore, their study in the Mediterranean Sea has scientific significance, but their correct reproduction is still necessary to reproduce the sea level accurately. Finally, it should be remembered that these modes of oscillation can also be triggered by extremely violent events, such as tsunamis.

The use of the DA is very useful not only in the forecast, but also in hindcast, allowing obtaining reanalysis of the total sea level, or some of its components, much more accurate than those obtainable without. In particular, the calculation of the harmonic components of the astronomical tide is usually possible in areas where there is a tide gauge. In areas where there are no measurements, one is forced to calculate the components based on model data which can be inaccurate. The DA can therefore have a positive impact both in the reproduction of the astronomical tide, locally, and in the study of its spatial structure. Similarly, from the reanalysis of the surge component, it is possible to extract spatial information on its propagation and the shape of the seiche oscillations.

This paper analyses the model reproduction of the various components of the marine barotropic level described up to now, using a finite element hydrodynamic model. The impact of the assimilation of 50 sea-level coastal stations, done through an ensemble Kalman filter, is also analysed. The results are analysed and discussed both in forecast and hindcast (reanalysis in the case of the use of the DA). A two-month simulation period was chosen, November and December 2019, in which one of the most extreme storm surge events ever recorded in Venice was recorded and, after this in December, a very energetic event

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


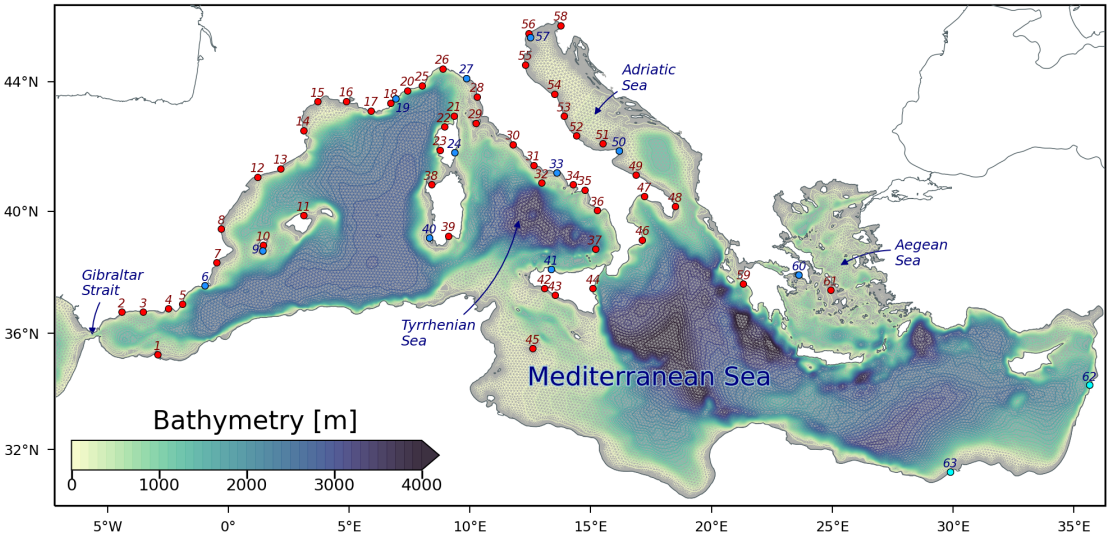


**Figure 1.** Bathymetry of the Mediterranean Sea with the unstructured grid superimposed. The red and blue dots mark the location of assimilated and validation tide gauges, respectively.

of seiche oscillations followed. This period is therefore very particular from a meteorological-marine point of view, not only in the Adriatic Sea but in the entire Mediterranean basin.

Below we report first the methodology followed, with a description of the hydrodynamic model used (Section 2.1), of the observations and their processing (Section 2.2) and the DA (Section 2.3), ending with a brief description of the simulations performed (Section 2.4). Then follows the calibration of the DA (Section 3.1), the model results in hindcast mode (Section 3.2) and in forecast mode (Section 3.3), and a part dedicated to the seiche oscillations (Section 3.4). Finally, the discussion (Section 4) and conclusions (Section 5) are reported.

## 2 Methods

### 2.1 The hydrodynamic model

The hydrodynamic model we use is called SHYFEM (System of HydrodYnamic Finite Element Module) and was created at the CNR in Venice (Umgiesser and Bergamasco, 1993), where it is largely developed continuously. Its code is available under an open-source license and freely downloadable from the Web (https://github.com/SHYFEM-model/shyfem). SHYFEM is composed of a hydrodynamic core that solves the shallow water equations with the finite element technique and with a semi-implicit time-stepping algorithm, which allows a remarkable speed of execution. Various terms in the equations can be turned on or off, such as momentum advection terms, Coriolis terms, baroclinic terms of density gradients, and tidal potential. The

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


model can be used in two or three-dimensional modes and various formulations of bottom stress and wind stress are available. Finally, the model can be coupled to various modules or models (e.g., waves, Lagrangian, ecological).

In this application, we use a two-dimensional barotropic formulation given by the following equations:

$$\begin{aligned}
&\frac{dU}{dt} - fV = -H\left(g\frac{\partial \zeta}{\partial x} + \frac{1}{\rho_w}\frac{\partial p_a}{\partial x}\right) + A_H\nabla^2 U + \frac{1}{\rho_w}\left(\tau_{wx} - \tau_{bx}\right)\\
&\frac{dV}{dt} + fU = -H\left(g\frac{\partial \zeta}{\partial y} + \frac{1}{\rho_w}\frac{\partial p_a}{\partial y}\right) + A_H\nabla^2 V + \frac{1}{\rho_w}\left(\tau_{wy} - \tau_{by}\right)\\
&\frac{\partial \zeta}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0,
\end{aligned} \tag{1}$$

where the independent variables are the time, $t$, and the spatial variables $x$ and $y$. $U(x,y,t)$ and $V(x,y,t)$ are the transports along $x$ and $y$, $f(y)$ is the Coriolis coefficient, $H(x,y,t)$ is the sum of the sea depth with $\zeta(x,y,t)$, which is the variable level with respect to the resting state; $g$ is the gravitational acceleration, $\rho_w$ is the average density of seawater, $p_a(x,y,t)$ is the atmospheric pressure at sea level and $A_H$ is the horizontal coefficient of turbulent viscosity, formulated with Smagorinsky (1963), using a dimensionless coefficient equal to 0.2; while $\nabla^2[\cdot]$ is the two-dimensional Laplacian operator. $\tau_{bx}(x,y,t)$ and $\tau_{bx}(x,y,t)$ are the components of the stress at the bottom, expressed with a linear-quadratic formulation with coefficient 0.0025 (Bajo et al., 2019); $\tau_{wx}(x,y,t)$ and $\tau_{wx}(x,y,t)$ are the components of wind stress, expressed with the formulation proposed by Hersbach (2011) and with a Charnock coefficient equal to 0.02.

Furthermore, for the simulations that calculate the tidal level or the total sea level, the terms of tidal potential are also active and four semi-diurnal components ($M_2$, $S_2$, $N_2$ and $K_2$) and four diurnal components ($K_1$, $O_1$, $Q_1$ and $P_1$) are calculated. This formulation or very similar formulations for SHYFEM have been used in the past with success in many works on the storm surge (Bajo et al., 2017, 2019; Cavaleri et al., 2019; Ferrarin et al., 2021) or the total sea level or the tide (Ferrarin et al., 2013, 2018).

The model is applied on a mesh of the Mediterranean Sea, which extends into the Atlantic Ocean up to about 7° W and has about 163,000 triangular elements. The size of the elements is variable, with a gradually greater resolution from the open sea to the coasts (Fig. 1). The bathymetry derives from the 2020 dataset of the European Marine Observation and Data Network (https://www.emodnet-bathymetry.eu/), which was bilinearly interpolated on the mesh.

The simulations use, as forcing at the surface, 10-m wind and mean sea level pressure fields provided by the BOLAM model (Mariani et al., 2015), which is a hydrostatic model at 8 km, with hourly fields, nested in the IFS ECMWF operational model (https://www.ecmwf.int/en/publications/ifs-documentation). For hindcast simulations, the first forecast days chained together are used, while each daily forecast simulation uses the entire forecast up to five days ahead. Lateral boundary conditions are closed everywhere except in Gibraltar, where Dirichlet conditions are imposed, with the sea level imposed and water transports being free. For the open boundary sea level, we used the E.U. Copernicus Marine Service Information (https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1). The total sea level is used for total sea level simulations, while the de-tided sea level, called Non-Tidal Residual (NTR), is used in the surge simulations. The simulations with the astronomical tide are forced at the boundary of Gibraltar with the tidal signal derived from the same Copernicus model.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


Hereafter we will refer with NTR to the residual level obtained by removing the tide from the total water level. The NTR is mainly made up of a meteorological component, forced by wind and pressure, called surge, and by a further possible spurious signal due to seasonal and non-seasonal baroclinic forcing, which is generally at low frequency (in DA part of this signal is removed with the MDT correction explained later). The NTR represents our best estimate of the surge component, calculated from the model.

The simulations with the DA use the same forcings and boundary conditions as the deterministic simulations (without the DA) but these are perturbed to obtain an ensemble of simulations. This procedure is described in Section 2.3.

### 2.2 In-situ observations

Sea-level observations were retrieved from the European Joint Research Center database (https://data.jrc.ec.europa.eu/). As shown in Fig. 1, tide gauges are concentrated in the western and central Mediterranean Sea, mostly along the Spanish, French and Italian coasts, while on the northern African coast there is only one station (Melilla) and few stations are present in the eastern Mediterranean Sea. The Adriatic Sea has stations only along the Italian coast and not on the eastern coast, but they are still quite numerous. The stations in the Mediterranean Sea were divided into 50 stations to be assimilated and 13 for validation (Tab. A1). Data is recorded every 10 minutes in the period of October-December 2019. We processed it with the SELENE quality check software (https://puertos-del-estado-medio-fisico.github.io/SELENE/; Pérez et al., 2013) for spikes and outliers detection, stability test, date and time control, flagging and interpolation of short gaps. Subsequently, the quality-checked data were elaborated with the Python binding of UTide (https://github.com/wesleybowman/UTide; http://www.po.gso.uri.edu/~codiga/utide/utide.htm), based on the least squares fitting, to separate the tidal periodic part from the non-periodic part (NTR) in the total sea level. We kept the eight most energetic tidal constituents in the harmonic analysis ($M_2$, $S_2$, $N_2$, $K_2$, $K_1$, $O_1$, $P_1$, $Q_1$), which are the most important in the Mediterranean Sea (CITE). The NTR was further processed by applying a 2-hour moving average, to remove high-frequency signals. The harmonic analysis was not possible for stations 62 and 63, due to a lack of continuous data. Therefore, these stations were used only for the validation of the total sea level.

The observations may have different reference datum according to the monitoring network to which they belong. Therefore, in order to have no bias errors in DA, the measurements were referred to the mean sea level of the deterministic simulation of the model, which can be defined as the mean dynamic topography (MDT) of the model (Vidard et al., 2009; Byrne et al., 2021). The MDT of the model is given only by the barotropic terms present in the equations and lacks the baroclinic part due to the temperature and salinity gradients.

### 2.3 The data assimilation system

In this section and the following ones, we will use some terminologies and concepts typical of the DA, for an introduction to these concepts and the various techniques we recommend reading Carrassi et al. (2018).

The code used for the DA is based on the routines developed and described in Evensen (2003, 2004) and available at https://github.com/geirev/EnKF_analysis. The routines have been adapted and extended to be used by the SHYFEM model, allowing the use of different DA techniques, such as the Ensemble Kalman Filter (EnKF) and the Ensemble Square Root

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


Filter (EnSRF), with different numerical schemes (https://github.com/marcobj/shyfem). Furthermore, various routines have been created to perturb the forcing and boundary conditions obtaining ensembles of arbitrary size. In this work, we used an EnKF technique, with the correction described in Evensen (2004) to avoid the loss of rank in the observation covariance matrix (Kepert, 2004).

As described in Section 2.4, each set-up was decided after numerous test simulations. It was therefore decided to use an ensemble of 81 members (80 perturbed + 1 control) and not to use local analysis (Carrassi et al., 2018). The system uses adaptive inflation (Evensen, 2009), to avoid narrowing of the ensemble spread, and the observations are considered independent (in fact they come from different stations) and have an error set to 2 cm. In order to discard innovations that are too high, a simple technique, which checks the values of the variances of the background matrices and observations, is used (Järvinen and Undén, 1997; Storto, 2016).

Finally, to dampen shocks in the analysis solution near the lateral boundary conditions, in Gibraltar, the analysis solution is relaxed to the background one, approaching the boundary. The two states, background and analysis, are weighted through a Gaspari-Cohn function (Gaspari and Cohn, 1999), with a radius of about 250 km starting from the open boundary nodes. In each node we have the following solution:

$$A_a(x,y) = A_b(x,y)f(x,y) + (1 - f(x,y))A_a^*(x,y), \tag{2}$$

where $x$ and $y$ define the position in the grid, $A_b$ is the matrix of the background states, $A_a^*$ is that of the analysis states not corrected, $A_a$ is the one corrected and $f$ is the function of Gaspari-Cohn, equal to 1 in the open boundary nodes.

The ensemble of 81 members in all the DA simulations performed was created by perturbing the initial state, forcing and boundary conditions and a few model parameters. The perturbation of the initial state is performed only for the levels, with a technique similar to that for atmospheric pressure (described later), while the water transport is not perturbed. However, for the forecast, this is used only for the first simulation, while the following ones start from an ensemble of states saved from the simulations of the previous day. Even for hindcast/reanalysis simulations, the initial state perturbation is of little importance, as the simulation lasts two months and forcing and boundary conditions are far more important. The perturbations of the forcing and boundary conditions are maintained for the entire duration of the DA. The forecast simulations perform a day of assimilation and then start from the average analysis state and use the unperturbed forcings and boundary conditions.

The perturbations are calculated so that in each spatial point the average is approximately equal to the non-perturbed value (no bias) and the standard deviation is equal to the estimated error; furthermore, the perturbations must have a Gaussian distribution. Perturbing the conditions at the lateral boundary is not very complicated, we have applied the same perturbations to all the nodes of the open boundary. As for the atmospheric fields, their perturbation is more complex because, in addition to satisfying the previous condition, there must be a spatial correlation and a physical coherence between the variables (wind and pressure). We therefore first perturbed the atmospheric pressure field, through a technique to generate two-dimensional pseudo-random fields (Evensen, 1994, 2003), imposing a decorrelation length of about 400km and a standard deviation of 35 hPa. From these two-dimensional waves of pressure perturbations, we calculated the corresponding perturbations for the velocity components. If the pressure perturbation in one point is $\delta P$, the perturbations for the wind components, in geostrophic

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


equilibrium, are:

$$\delta u = -\frac{\delta P}{\delta y}\frac{1}{\rho_a f}$$
$$\delta v = \frac{\delta P}{\delta x}\frac{1}{\rho_a f}. \tag{3}$$

Using these perturbation fields to be applied to the unperturbed fields of wind and pressure at an instant $t$, we obtain perturbed fields with a physical coherence.

Again for the atmospheric fields, in addition to this kind of perturbation, a temporal perturbation has also been introduced in which, from a field at time $t$, an ensemble of equal fields is generated but with reference time $t + dt_n$, where $dt_n$ are perturbations belonging to a Gaussian distribution as well. Other perturbations, again with a Gaussian distribution, have been introduced for the bottom stress and, in the simulations with the tidal potential, for the calibration coefficient of the loading tide.

Finally, as regards the forcing and boundary conditions that vary over time, the errors at a given instant $t_1$ must be related to the errors at the instant of the next field, at time $t_2$. This type of error, or noise, which is defined as "red noise", is applied by calculating a weight dependent on the time interval between the two fields and on a decay time:

$$\alpha = 1 - \frac{t_2 - t_1}{\tau}, \tag{4}$$

where $\tau$ is the decay time. The perturbation $\xi_2$, at time $t_2$, becomes a linear combination of the perturbation $\xi_1$, at time $t_1$, and the newly calculated perturbation $\xi_2^*$:

$$\xi_2 = \alpha\xi_1 + \sqrt{1-\alpha^2}\xi_2^*. \tag{5}$$

### 2.4 Results' production and post-processing

In this study, we performed simulations with and without DA, in hindcast and forecast modes. The hindcast simulations extend from the beginning of November to the end of December 2019 and were performed considering the total sea level, the surge component and the astronomical component. The hindcast simulations with DA, which can be defined reanalysis simulations, assimilate the data from the 50 stations every hour, throughout the two months of simulation. From the ensemble of reanalysis states, the average is calculated, as the best estimate of the real state of the physical system, and is used in the analysis of the results.

The forecast simulations were performed by calculating the total sea level and the meteorological component, the surge. The simulations of the astronomical tide have not been carried out, as they are not very useful since there is always the possibility, even for future periods, to calculate pseudo-observations in stations where the harmonic constants are known and use them to perform a simulation with DA.

The forecast simulations were performed identically to those that would be performed by an operational model. The period is the same as considered in the hindcast, November and December 2019. However, the simulations are performed daily and each is composed of a hindcast day and five forecast days of which, for the sake of brevity, we will show the results for the first

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


three. The simulations with DA assimilate the data from the 50 stations, every hour, in the 24 hours of the hindcast period. From the final average analysis state, saved at 00 UTC each day, the forecast simulation is started, which lasts 5 days. In addition, the final analysis states of the entire ensemble are saved to be used by the simulation of the next day, as initial states. In this way, the DA always restarts from a state of analysis and is equivalent to the cycle performed in hindcast, except for the perturbation of the forcing and boundary conditions, which is redone every day.

To analyse the results, each daily forecast simulation was divided into five parts and each part was chained with the corresponding one of the previous and following days. In this way, continuous results are obtained for each forecast day, which can be directly compared with the measures in the two months considered.

To compare the results with the observations, we calculated the standard deviations of the model and observed data, the correlation between them and the Centered Root Mean Squared Error (CRMSE). The standard deviations and CRMSEs were normalised to the standard deviation of the observations at each station and represented by Taylor diagrams (Taylor, 2001). Bias error plots were also calculated, in which bias is calculated as the mean of the differences between the modelled and observed values; while the CRMSE represented in the same plots is not normalised.

Regarding the spectral analysis, the NTR and the surge calculated by the model for December 2019, in which there is a strong presence of seiche signals, were used. The power spectral density was estimated with the Welch method (Welch, 1967), dividing the period into 8-day windows with 50% overlap. The fast Fourier transform length is rounded up to the nearest integer power of 2 by zero padding.

## 3 Results

### 3.1 Calibration of the data assimilation

Before running the definitive simulations, which were used to produce the results, we carried out numerous experiments to determine the best parameters of the DA. The results of these tests (partially shown in this paper), showed how, in this study case, the EnKF gives better results than the EnSRF and that it is not convenient to use data localisation techniques (local analysis). Although in fact, the localisation in most cases brings advantages, in our case the observations available are mainly arranged only on the northern side of the computational domain. This implies that to obtain a spatially uniform analysis correction, a very wide localisation radius should be used, reaching the other end of the basin.

Furthermore, the correlation radius of a variable (the sea level in our case) between a point and its neighbours increases with the speed of propagation of the perturbations of this variable which, in the case of barotropic perturbations, is high (in the western basin with an average depth of about 2000 m there is a speed of 140 m/s). To avoid spurious correlations, without the use of local analysis, we have increased the number of members of the ensemble. However, the system remains quite fast, considering that these are two-dimensional barotropic simulations. We then performed, in hindcast mode, some DA tests by varying the number of members of the ensemble. In Fig. 2 we report the Centered Root Mean Squared Error (CRMSE) of the total sea level averaged in the validation stations, using a different number of members of the ensemble. The error is reduced

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


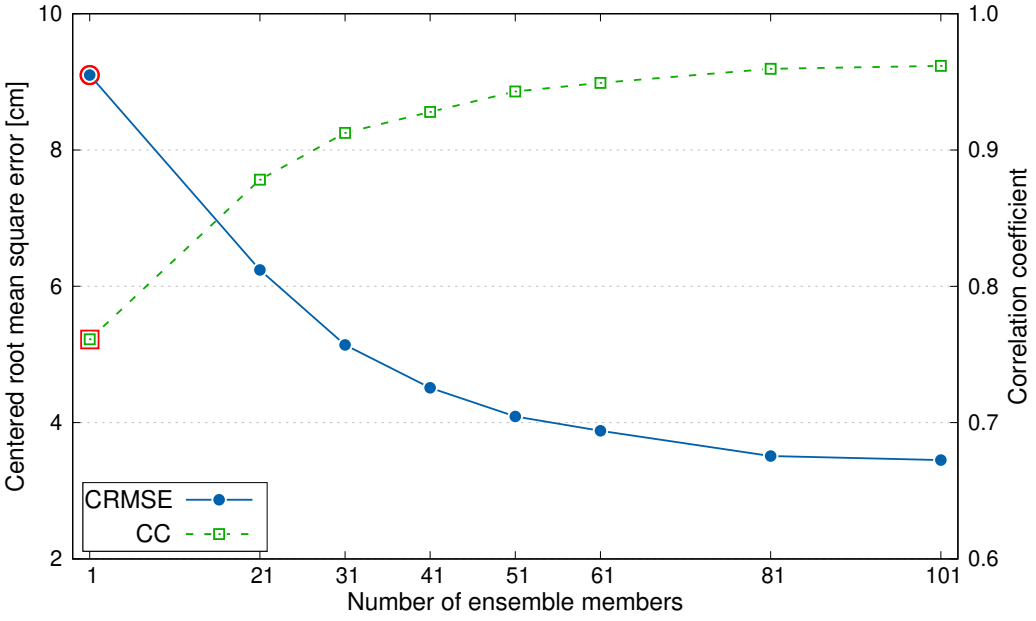


**Figure 2.** Performance of the data assimilation, in terms of CRMSE and correlation coefficient, as a function of the number of ensemble members. The red contour highlights the results of the simulation without data assimilation.

from 9.3 cm, in the case without DA, to 3.6 cm using 101 members, and the correlation increases from 0.75 to 0.95. Since the error pattern is regular and asymptotic, we decided to use 81 members.

### 3.2 Hindcast simulations

We will analyse first the results in hindcast, for the astronomical tide, the surge and the total sea level. In Fig. 3, the first diagram on the left shows the astronomical tide comparison, in which the tide calculated by the model is compared with that calculated by the harmonic constants. The results are quite good even without DA in almost all stations, with a certain tendency to overestimate the signal amplitude (higher standard deviation). Station 60 is an exception, where the results are poor, probably due to its position in the Aegean Sea, in a morphologically complex area. The results with DA are definitely good for all the validation stations, reaching almost perfect agreement (correlation about 0.99), with a small deterioration in station 60, which however improves and still achieves a more than good accuracy (CRMSE from 4 cm to 1 cm).

The central diagram shows the reproduction of the surge signal, compared with the NTR extracted from the observations. In this case, the distribution of the stations in the deterministic simulation is sparse and nothing particular can be deduced, except that also for the surge, station 60 is the worst. The simulation with DA improves considerably all the stations, arriving to have a very concentrated distribution even if not like that of the astronomical tide. For example, in station 60, which is the worst, the CRMSE reduced from 8 cm to 3 cm.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


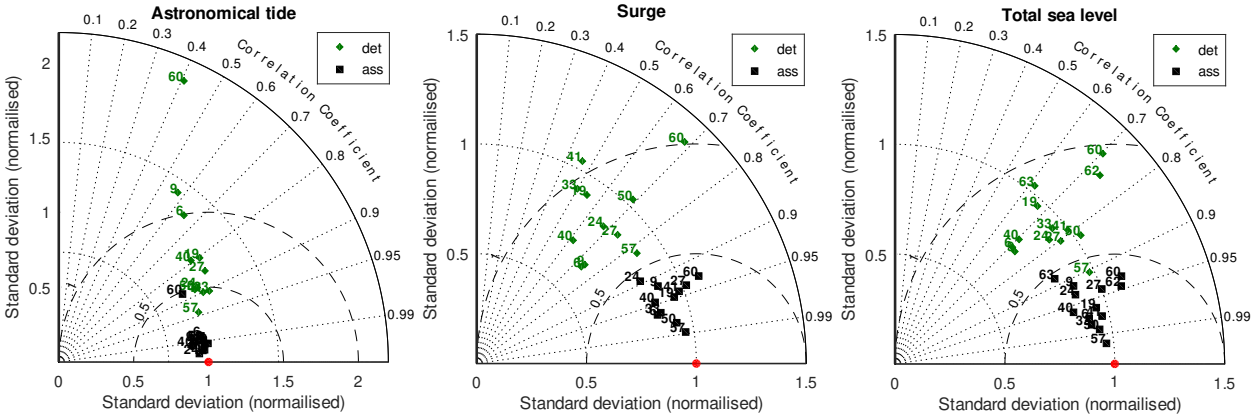


**Figure 3.** Normalised Taylor diagrams of the hindcast simulations. The deterministic simulations (green diamonds) compared to DA simulations (black squares), for the astronomical tide (left), the surge (centre) and the total sea level (right). The red dot indicates the perfect agreement.

Finally, as regards the simulation with the total sea level, the quality of the simulations with and without DA is very similar to that of the simulations with the surge. Some stations are even better, perhaps thanks to the good accuracy with which the model calculates the astronomical tide (in station 60 the CRMSE goes from 8 cm to 3 cm).

In the case of the total sea level, the comparison was also made for stations 62 and 63 which, as previously mentioned, are the only ones in the eastern basin and are at least a thousand kilometres away from the first assimilated station. It is interesting to note that these stations also have a consistent improvement, the CRMSE goes from 9.6 cm to 4 cm for station 62 and from 10.9 cm to 5.7 cm for station 63. This improvement indicates the good quality of the background covariance matrix, even for model variables that are very distant from each other, obtained thanks to the high number and good independence of the ensemble members.

### 3.3 Forecast simulations

In this section we analyse the forecast results for the surge component and for the total sea level. In Fig. 4 the Taylor diagrams are shown for the first, second and third forecast days, both for the data of the model without DA and for those with DA, in all validation stations. The results are relative to the surge simulations. The effect of the DA on the first forecast day is evident and the distribution is similar, slightly worse, to that obtained in hindcast. Each validation station improves significantly, including 60, which does not have many assimilated stations nearby. Unfortunately, the data in stations 61 and 62 cannot be used in the validation of the surge simulations, as it was not possible to perform the harmonic analysis, due to the few available data.

The improvement is reduced on the second day, while on the third day is almost nil or slightly worse in some stations, although not significantly. The initial state of the system gradually loses its importance as the forecast moves temporally away from it, and so does the decrease in its error. The forecast without DA has a larger initial state error, which mostly counts on the first and second forecast days.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


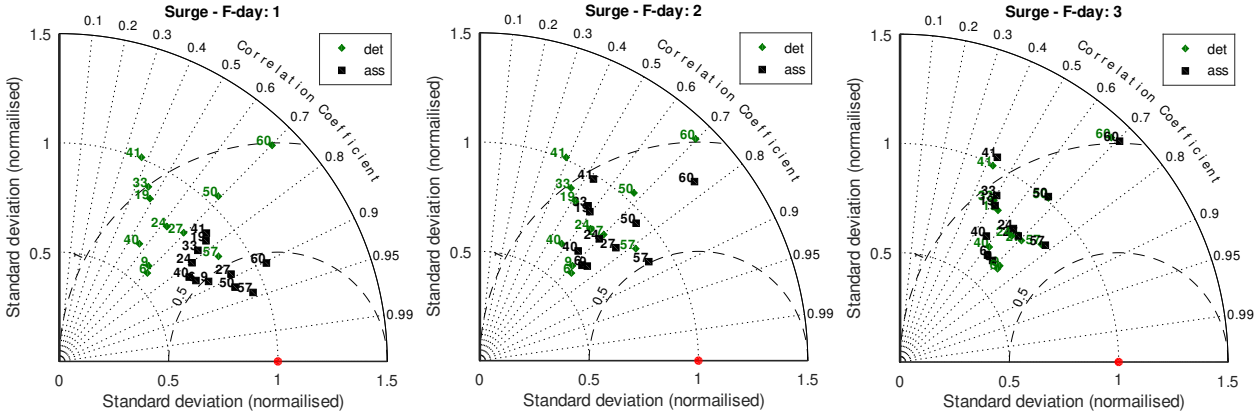


**Figure 4.** Normalised Taylor diagrams of the forecast simulations with the surge. The deterministic simulations (green diamonds) compared to DA simulations (black squares), for the first (left), the second (centre) and the third (right) -day forecast. The red dot indicates the perfect agreement.

In Fig. 5 we can see the bias error for the surge simulations. This graph was not made for the hindcast simulations, as in that case, the bias is almost null. The figure shows that the DA improves the results, especially on the first forecast day, then gives a still positive but weaker correction on the second day, while on the third day the DA slightly worsens the model data, in agreement with what has been seen for the Taylor diagrams. The worsening is in any case contained and relates to the third forecast day, which is of secondary importance compared to the first and second forecast days. Still observing Fig. 5, it can be seen how station 57 deviates from the group, with a much greater bias and CRMSE than the other stations. This is due to the position of this station, in the upper Adriatic, where the surge events and the associated seiche oscillations have much greater intensity than in the rest of the Mediterranean Sea. However, precisely for this reason, and since there are numerous good-quality stations in the Adriatic Sea, the effect of DA is strong, both in the correction of random and systematic errors. The last (i.e., the biases) are almost all positive, denoting a systematic model overestimation of the NTR, even if this is true statistically, while for extreme events the trend is normally the opposite.

In Fig. 6 we report the Taylor diagrams relating to the total sea level. In this case, the diagrams are slightly better, both without and with DA, which maintains evident improvements even on the third forecast day. For the total sea level, the comparison is also possible in stations 61 and 62 even if they have a smaller number of data. As seen for hindcast simulations these stations are important given their distance from other similar stations and being the only stations present in the eastern basin. Also for the forecast, as well as for the hindcast, there is a notable improvement.

Fig. 7 shows the bias diagram for the total sea level. In this case, the bias is generally lower than that of the surge, even for the model without DA, even if is still positive in most of the stations, denoting overestimation by the model. The improvements given by the DA are therefore smaller in proportion and present more on the first day of the forecast, as for the simulations with the surge. However, the CRMSE improves on all three forecast days, as already seen in the Taylor diagram, and has values

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


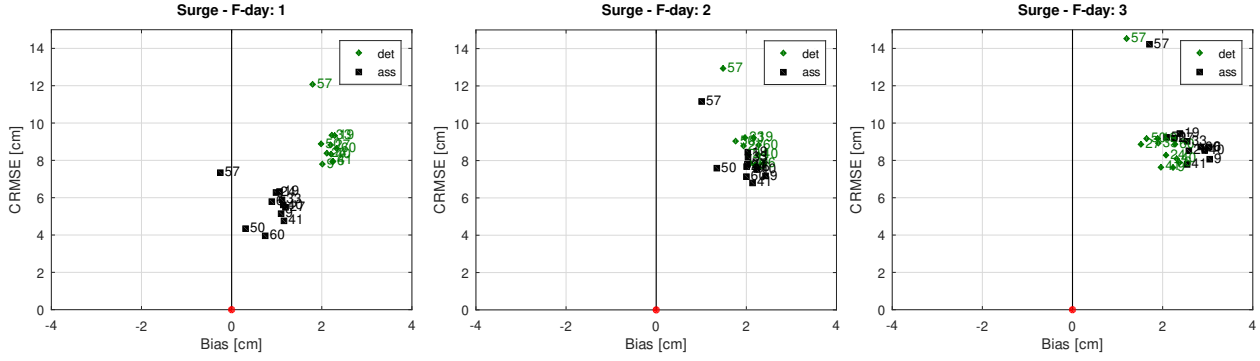


**Figure 5.** Bias diagrams for the first (left), the second (centre) and the third (right) -day forecast of the surge simulations. The deterministic results (green diamonds) are plotted with the assimilation ones (black squares). The red dot is the perfect agreement, while positive bias means an overestimation of the model.

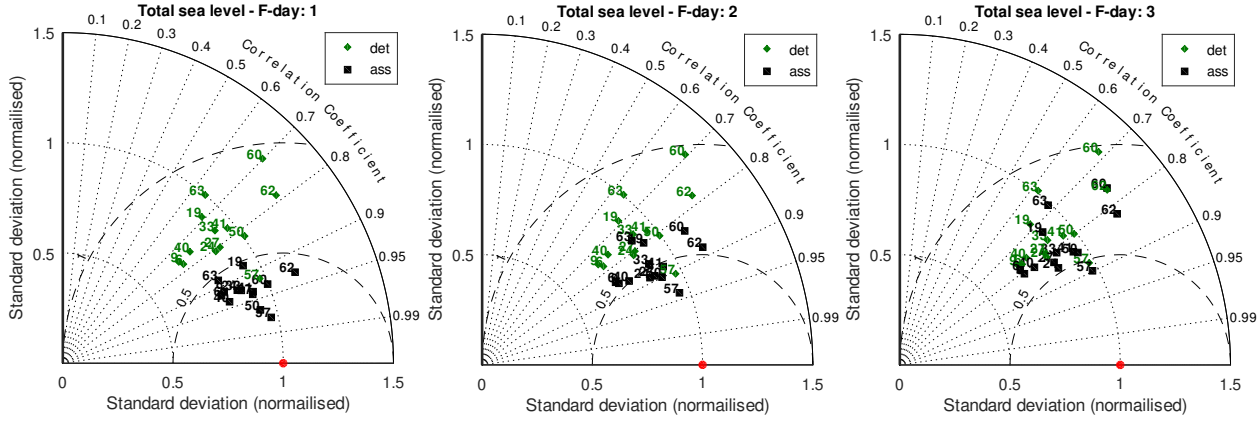


**Figure 6.** Normalised Taylor diagrams of the forecast simulations with the total sea level. The deterministic simulations (green diamonds) compared to DA simulations (black squares), for the first (left), the second (centre) and the third (right) -day forecast. The red dot is the perfect agreement.

similar to those of the surge. However, the error is related to the total sea level, which has an oscillation larger than that of the surge alone.

### 3.4 Seiches

As explained in the introduction, the seiches are free barotropic oscillations of the sea level in a basin, triggered by an initial perturbation. Therefore, their propagation depends solely on a correct initial state and a correct modelling system. Given that DA has the purpose of reducing the error of the initial state, we expect it to have a strong impact on the reproduction of the seiches. Generally, these oscillations in the Mediterranean Sea are not particularly studied, as they are quite small. On the

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


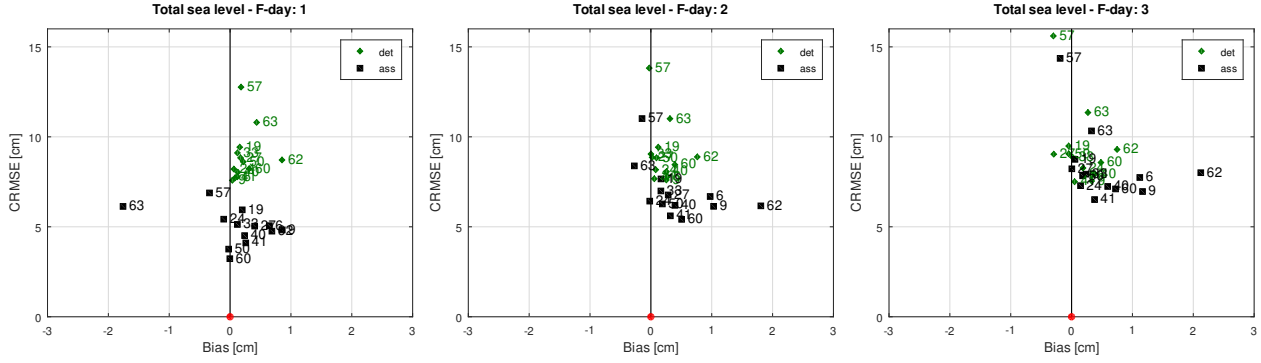


**Figure 7.** Bias diagrams for the first (left), the second (centre) and the third (right) -day forecast of the total sea level simulations. The deterministic results (green diamonds) are plotted with the assimilation ones (black squares). The red dot is the perfect agreement, while positive bias means an overestimation of the model.

contrary, in the Adriatic Sea they were deeply studied; being much more energetic, their reproduction is essential for the correct reproduction of the sea level and of flooding events.

In December 2019 (period included in our simulations), significant seiche events, among the most energetic ever recorded in the Adriatic, took place (Fig. 8). Despite their intensity, they were not preceded by any strong storm surge, which could have triggered them. A possible explanation could be that these oscillations were triggered by an oscillation of the meteorological forcing at a frequency similar to that of the normal modes of the basin (which are also the resonant frequencies).

These events were generally poorly predicted by storm surge models operating in Venice (none with DA), the city most affected by flooding in the northern Adriatic. Fig. 9 shows the total sea level recorded in station 56 (Venice) and the first three days of forecast from the surge simulations (with the addition of the astronomical tide). The oscillations observed in the figure are therefore a superposition of the astronomical tide on the surge signal, which is dominated by the seiche oscillation. At the beginning of the forecast, the DA corrects an error of about 30 cm and maintains a continuous improvement over time, which can also be appreciated after three days of forecast. Although in the previous section we have seen how the statistical improvement at three days is not very appreciable, in the most extreme seiche cases and therefore of greater importance, the error of the initial state of the model tends to be larger and the DA provides a greater contribution.

This event demonstrates the particular effectiveness of DA in correcting the sea level in the case of seiche fluctuations in the Adriatic Sea. To better highlight this feature and see if it is also present in the rest of the Mediterranean Sea, we carried out spectral analyses of the NTR from the observations and of the modelled surge, in all the stations for December 2019. Before examining the performances of the model with and without DA in the reproduction of spectra, we report below the results obtained for the periods in the Adriatic and Mediterranean basins. Although the periods of the main modes are quite well known in the Adriatic Sea (Cerovecki et al., 1997; Vilibić et al., 2005; Vilibić, 2006; Bajo et al., 2019), there are, to our knowledge, no works that report the main periods of the Mediterranean Sea, based on observations. However, Schwab and Rao

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


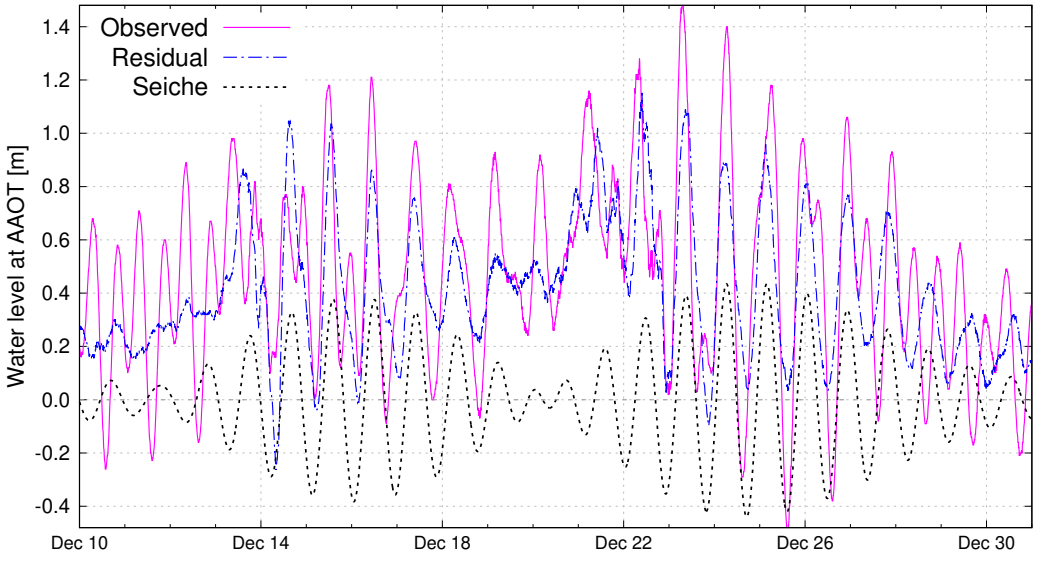


**Figure 8.** Seiche event happened on December 2019, recorded at the AAOT station (n. 57). The measured total water level is plotted with a pink solid line, the NTR with a blue dashed-dotted line and the seiche contribution, extracted from the NTR with a bandpass filter, with a green dotted line.

(1983), using a simple barotropic model, foresees some of them and describes the shape of the main barotropic modes. Below is a brief description of the form and the expected period, comparing it with what we found from the observations.

Schwab and Rao (1983), by calculating the eigenvalues of a simplified numerical model of the Mediterranean Sea, found four modes of oscillation. The first mode (M1) relates to an oscillation with a single positive amphidromic node in the Gulf of Sicily, with an expected period of 38.5 hours. This mode, which should have maximum amplitude at both the western and eastern ends of the Mediterranean, has not been identified by our observations, probably because it has not been solicited by any forcing in the period considered by us.

The second mode (M2) has a more complex shape with a negative amphidromic node in the western basin, a positive one in the eastern basin and a third in the Adriatic. This barotropic oscillation mode has an expected period of 11.4 hours. A similar peak, with a period of 12.8 hours, is present in several stations of the western basin analysed by us (Fig. 11). The difference from the expected peak can be explained by considering the various simplifications and the low resolution of the model in the paper, which dates back many years ago.

The third mode (M3) has three positive amphidromic points in the Mediterranean and two, one positive and one negative, in the Adriatic. This mode has a period of 8.4 hours and maximum amplitude near Gibraltar and along the west coast of the Adriatic Sea. Indeed, from our measurements, a peak at 8-8.3 hours is quite evident in some stations in the western Mediterranean basin (Fig. 11) and a hinted peak is also present in Trieste (Fig. 10) and in other stations on the western coast of the Adriatic Sea.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


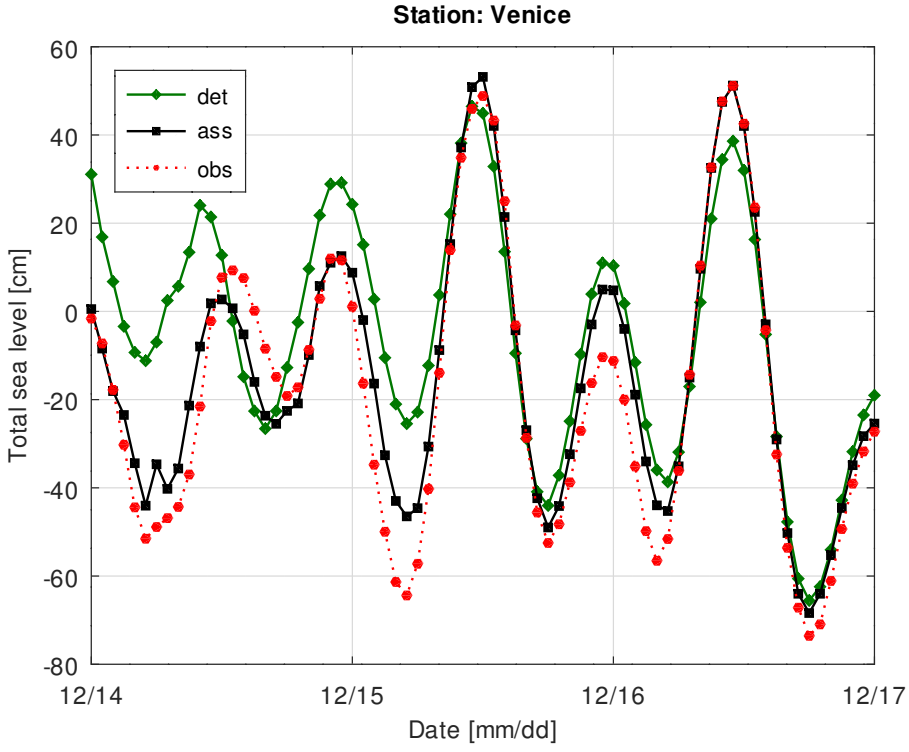


**Figure 9.** Forecast of 14 December 2019 extracted from surge simulations (adding the tide). The observations (obs) are compared with the normal forecast (det) and with the DA forecast (ass).

Finally, the fourth Mediterranean mode (M4) of 7.4 hours should be related to the main oscillation of the Tunisian bight, where we have no observations and therefore we cannot check its presence. From the spectra that we have analysed, there seems to exist a fifth mode, here called M5, which is very evident in the stations of the western Mediterranean basin and with a period of 6.2 hours (Fig. 11). However, we have no information of this oscillation from the scientific literature of our knowledge.

Regarding the Adriatic Sea, the fundamental mode, here referred to as A1, is an oscillation extended to the entire basin, with a nodal line south of the Strait of Otranto, near the bathymetric line of 1000 m, and has a period of about 21.2 hours. This oscillation is the most energetic among those analysed and is clearly visible from the observations in our possession, from which a period of 21.3 hours is calculated (Fig. 10).

The second mode (A2) has a nodal line that cuts the Adriatic basin north of Ancona and a second line south of the nodal line of the fundamental mode, near the 2000 m bathymetric line. This oscillation is quite energetic, albeit less than the main one, and has a period of about 10.7 hours, which is perfectly confirmed by our observations (Fig. 10). Finally, the third Adriatic mode (A3) has a nodal line under the Po delta, one just above the Gargano peninsula and a third line coinciding with that of the main mode. This oscillation has a period of about 6.7 hours, but is not present in our observations.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


**Table 1.** Periods of the barotropic modes in the Adriatic and Mediterranean Seas. The mode identification is written in the first column, while the second column shows the estimation from literature, based on observations, the third column shows the model estimation from Schwab and Rao (1983) and the last column shows our estimation from observations.

| Mode ID | $T_{ol}[h]$ | $T_s[h]$ | $T_{op}[h]$ |
|---|---|---|---|
| A1 | 21.2 | 20.1 | **21.3** |
| A2 | 10.7 | 9.3 | **10.7** |
| A3 | 6.7 | 6.8 | **-** |
| M1 | - | 38.5 | **-** |
| M2 | - | 11.4 | **12.8** |
| M3 | - | 8.4 | **8.3** |
| M4 | - | 7.4 | **-** |
| *M5* | - | - | **6.2** |

In summary, Tab. 1 reports all the periods here described. For a more accurate description of their form and propagation, see Schwab and Rao (1983).

As already discussed, the seiche propagation should depend on the initial state of the dynamical system and not on its boundary conditions. For this reason, their correct reproduction is not affected by the error of the forcing/boundary conditions, but only by the error in the initial state. The errors, that in DA are called model and representative errors, are considered small. Consequently, DA should be very effective in reducing the error in the reproduction of seiche oscillations.

Fig. 10 and Fig. 11 show the power spectral density from the NTR and the first-day forecast of the surge simulations, with and without DA. Fig. 10 shows the power spectra for two stations in the Adriatic Sea, Trieste, in the north, and Bari near the end of the basin in the south. Both the peaks of the main mode, A1 and that of the second mode, A2, are clearly visible. Note that both peaks are much more energetic in the first station than in the second, which is located near the nodal lines of the two modes. The two peaks are both underestimated by the model without DA, while with the DA the main peak is reproduced very well, especially in the north. The A2 peak remains slightly underestimated at both stations but improves significantly. Finally, in the Trieste station, a peak corresponding to the period of the third mode of the Mediterranean Sea (M3) is slightly visible in the observations (as predicted by Schwab and Rao, 1983). However, the spectra of the model, both with and without DA, are noisy in this part of frequencies and do not reproduce it. Also in Trieste, and in other Adriatic stations (not shown), there is a peak at 5.2 hours of which we do not know the origin and which is generally not reproduced with or without DA.

In Fig. 11 two stations in the Mediterranean Sea near Gibraltar are considered, one on the European coast and one on the African coast. In both stations the second and third barotropic modes of the Mediterranean basin are visible (M2, M3). Their energy is much lower than that of the Adriatic modes (about 1,000 times) and, probably for this reason, they are corrected less by the DA; although some improvement is visible, especially for M2. Both stations and many others in the western Mediterranean basin show a third, more energetic peak, which could be a fifth barotropic mode (M5). We can exclude that this

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


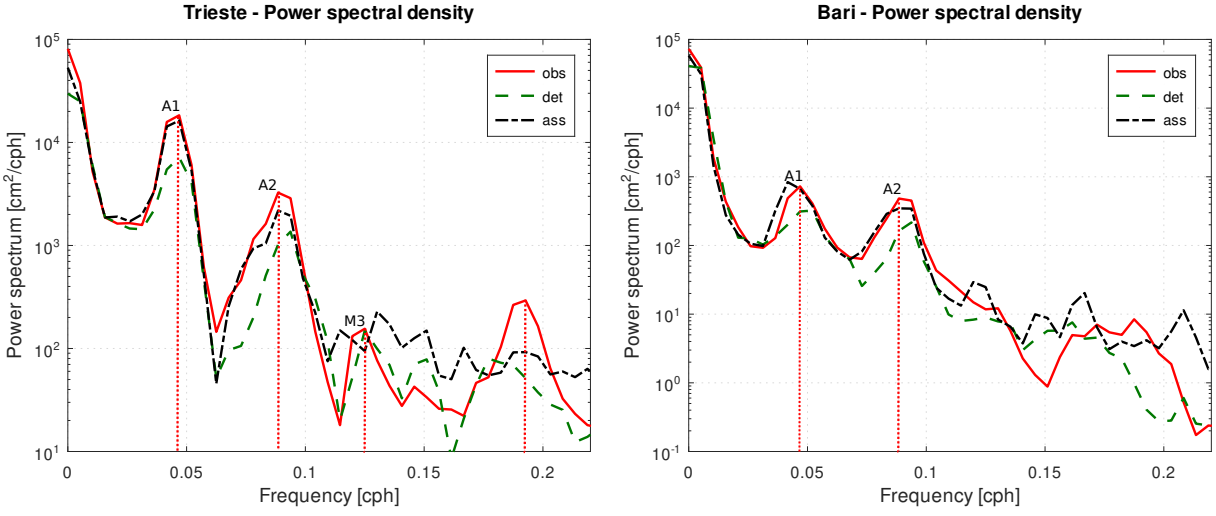


**Figure 10.** Power spectral density of the sea levels in Trieste and Bari, Adriatic Sea.

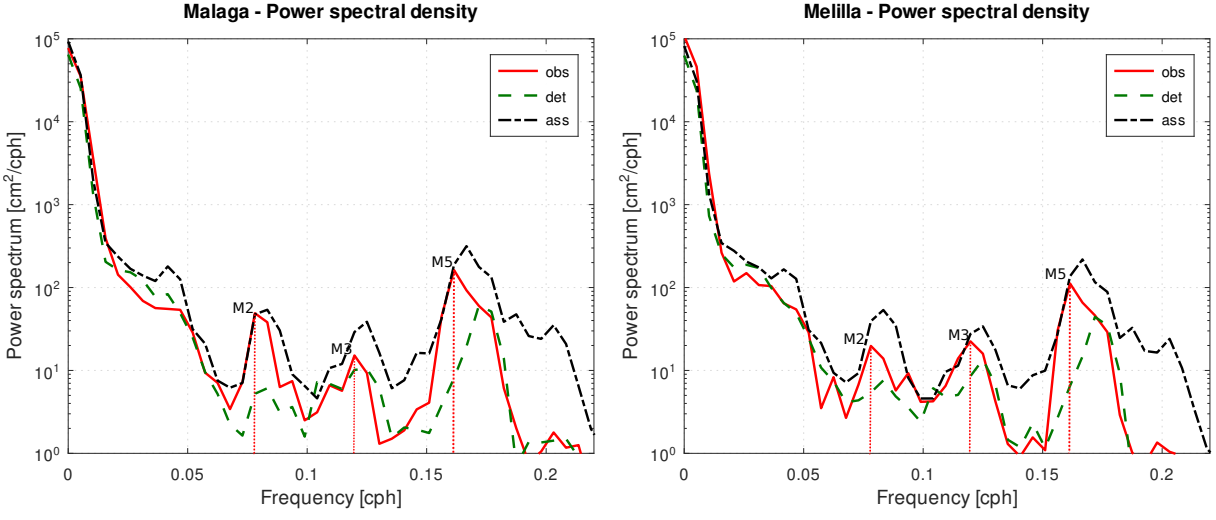


**Figure 11.** Power spectral density of the sea level in Malaga and Melilla, western Mediterranean Sea.

peak is a spurious signal from a partial subtraction of the astronomical tide from the NTR, as it is also present in the surge signal of the model results without DA (det). This peak is corrected by the DA even though it is broadened in frequency.

## 4 Discussion

Overall, the DA has shown that it can effectively correct barotropic sea level signals in the Mediterranean Sea. In the case of the astronomical tide, the correction is very effective and can be used to determine the tidal signal spatially and in coastal areas

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


without tide gauge stations. In these areas, it is possible to locally extract the results of the DA ensemble average for a year or more and perform a harmonic analysis to determine, with a small error, the amplitudes and phases of the tidal components. From these quantities, it is then possible to calculate the astronomical tide also in the forecast and use it to predict the total sea level, by adding it to the results of surge simulations. In the case of the astronomical tide, more than for the other components, the DA has shown that it can provide an excellent correction of the simulated level even in areas very far from the assimilated stations. This benefit is due to the good number of ensemble members that we used and the fact that the perturbations used to create it were created correctly. Furthermore, using localisation techniques, this fact would not have been possible, since these techniques limit the correction to areas much closer to the assimilated stations. Finally, this is also due to the fact that the barotropic sea levels, especially the astronomical tide, have very large spatial correlations.

Reanalysis simulations (hindcast with DA) with the total sea level and the surge showed great accuracy. A further improvement would perhaps be possible using an ensemble Kalman smoother, which is easily applicable to a simulation with the EnKF, such as those performed in this paper. Such simulations can be performed to create sea level reanalyses for very long periods, to be used in climatological studies. In these cases, especially in basins surrounded by a complex orography such as the Mediterranean, the meteorological forcings are often underestimated. The error caused by this underestimation is difficult to correct by the DA in the forecast, however, in hindcast, the DA with hourly assimilation manages to remedy these underestimates very well. Furthermore, since components that are not strictly barotropic are also present in the assimilated level (coming from the NTR measures), the reanalysis of the level will also contain this part, providing a good reproduction of the sea level in all its components, as demonstrated by the good results in the reanalysis we have obtained. From the point of view of the computational load, although there is a need to use a significant number of ensemble members, the simulations are perfectly parallelisable and are not a big issue on a modern server. Each simulation is two-dimensional and barotropic so it is performed very quickly, being able to use high time-steps.

The impact of the DA in the forecast is linked to the reduction of the error of the initial state, although it would be possible to investigate the possibility of using the DA as a parameter estimator. In the forecast, the error, which initially is due only to the initial state, amplifies over time, mainly due to the error of atmospheric forcings. Analysing the results statistically, the simulations without DA do not show much deterioration from the first to the third day of the forecast. However, this is not true in the case of extreme events, where meteorological forcings generally have a greater error. In the case of the forecast, the improvement given by the DA decreases fairly quickly moving away from the initial state. On the first day of the forecast, the improvement is still very strong, but it decreases in the following days. However, persistence is greater if seiche oscillations are triggered, which are reproduced correctly only if the initial state has a small error. In any case, by decreasing the error of the initial state, the DA also improves the reproduction of all the phenomena taking place at the time of the analysis, and therefore both the error linked to the astronomical component (in the case of the forecast of the total sea level), both the error linked to surge phenomena in formation. Finally, as in the case of reanalysis, the observed assimilated level may contain non-barotropic components, which are included in the initial state of analysis. However, these components cannot be correctly evolved in the forecast with a barotropic model.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


Finally, the study of seiche oscillations is a field in which DA is certainly not very used. In the forecast, we have seen that the DA can lead to a significant improvement, especially where these fluctuations are very energetic, as in the Adriatic Sea. The reanalysis of the surge can also be used for an in-depth study of their spatial propagation. As previously mentioned, while the modes of the Adriatic Sea are known and have been studied in numerous papers, the modes of the Mediterranean Sea have not been analysed much. The observations in our possession confirm and partially correct the periods found by the Schwab's simulations, as far as the M2 and M3 modes are concerned. However, the main mode of the Mediterranean Sea was not observed, probably due to the short length of the period analysed by us (one month). It can be assumed that during this period the M1 mode was not excited, but more studies are needed. There is also an oscillation at 6.2 hours, which we called M5, that is not present in the literature but is evident in many validation and calibration stations (not shown), along the coasts of the western Mediterranean basin. This oscillation, which is more energetic than the M2 and M3, is underestimated by the model without DA, but even with the use of the DA, it is not reproduced correctly. Considering that oscillations with a longer period are reproduced better even if less energetic, it is possible that the DA has more difficulty in correcting the higher frequency oscillations. This may be due to the frequency of assimilated data, every hour, which may be too low to define these modes.

## 5 Conclusions

In this paper we investigated the impact of DA in reproducing the barotropic sea level in the Mediterranean Sea. We analysed the performances both in hindcast and forecast, simulating only the astronomical component, the surge component and the total sea level. The results are good in all cases, particularly for the reanalysis of the astronomical tide, but also for the other components. The error in the forecast simulations is still very low, especially on the first day, gradually increasing the following days until it reaches the error of the simulations without DA. However, sometimes the improvement can last for the following days, up to five days in some cases, especially when seiche oscillations are triggered. Statistically, the total sea level simulations are more accurate overall than those of the surge, even without DA, as they have a lower bias error.

We analysed also the periods of the seiches (barotropic modes) in the Mediterranean Sea. Since these oscillations are mostly studied in the Adriatic, where they are more energetic, we have reported the periods of the modes in both the Adriatic and the Mediterranean basins, calculated from the observations in our possession. In Adriatic, we detected the two main modes (A1, A2) and the third Mediterranean mode (M3). In the Mediterranean basin, outside the Adriatic sub-basin, we detected the second and third modes (M2, M3) and a mode that we called M5, of 6.2 hours. The effect of the DA in the reproduction of the seiches in forecast mode was analysed as well, considering the first-day forecast. The analysis of the power spectra shows that the DA improves the reproduction of the peaks, especially for the low-frequency modes, while some difficulties are found in the reproduction of those at higher frequencies.

Further developments foresee the use of the modelling configuration tested here to develop an operational system for forecasting the sea level on the Mediterranean coasts, with a focus on the Italian coasts. This system will be installed at the ISPRA Centre and will use the assimilation of the stations along the Italian coast managed by the Centre, providing a five-day forecast

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


of the total sea level. Regarding the hindcast simulations, these will be extended over time to obtain reanalyses of past periods, trying to retrieve data from other stations in the Mediterranean Sea and exploiting the altimetric data.

The study of the seiche oscillations needs further investigations, especially in the Mediterranean Sea, where some modes are not well understood, to better determine the forms, periods and decay times. In such works, the DA can provide a more accurate spatial description of the surge, from which the seiche components can be extracted.

*Code availability.* The hydrodynamic model can be downloaded at: https://github.com/SHYFEM-model/shyfem. The modified version of the model, with the data assimilation code at: https://github.com/marcobj/shyfem

## Appendix A: In-situ coastal stations

In this appendix we report the table with the in-situ stations where we retrieved the data used in this paper.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


**Table A1.** List of stations with sea-level measurements. The stations with an asterisk are those used in the validation, while the others have been assimilated. The numbering is the one used in the paper and the geographical coordinates of their position are reported.

| ID | Lon | Lat | Station | ID | Lon | Lat | Station |
|---|---|---|---|---|---|---|---|
| 1 | -2.930 | 35.290 | Melilla | 35 | 14.750 | 40.676 | Salerno |
| 2 | -4.417 | 36.711 | Malaga | 36 | 15.275 | 40.029 | Palinuro |
| 3 | -3.520 | 36.720 | Motril | 37 | 15.190 | 38.785 | Ginostra |
| 4 | -2.478 | 36.830 | Almeria | 38 | 8.403 | 40.842 | Porto-Torres |
| 5 | -1.899 | 36.974 | Carboneras | 39 | 9.114 | 39.210 | Cagliari |
| 6* | -0.973 | 37.596 | Murcia | 40* | 8.309 | 39.147 | Carloforte |
| 7 | -0.481 | 38.338 | Alicante | 41* | 13.371 | 38.121 | Palermo |
| 8 | -0.310 | 39.440 | Valencia | 42 | 13.076 | 37.504 | Sciacca |
| 9* | 1.419 | 38.734 | Formentera | 43 | 13.526 | 37.285 | Porto-Empedocle |
| 10 | 1.450 | 38.917 | Ibiza | 44 | 15.093 | 37.498 | Catania |
| 11 | 3.117 | 39.867 | Alcudia | 45 | 12.604 | 35.499 | Lampedusa |
| 12 | 1.213 | 41.078 | Tarragona | 46 | 17.137 | 39.083 | Crotone |
| 13 | 2.160 | 41.340 | Barcelona | 47 | 17.223 | 40.475 | Taranto |
| 14 | 3.107 | 42.520 | Port-Vendres | 48 | 18.497 | 40.147 | Otranto |
| 15 | 3.699 | 43.397 | Sete | 49 | 16.866 | 41.140 | Bari |
| 16 | 4.893 | 43.405 | Fos-sur-Mer | 50* | 16.177 | 41.888 | Vieste |
| 17 | 5.914 | 43.122 | Toulon | 51 | 15.501 | 42.119 | Tremiti |
| 18 | 6.717 | 43.359 | Port-Ferreol | 52 | 14.414 | 42.355 | Ortona |
| 19* | 6.933 | 43.483 | La-Figueirette | 53 | 13.890 | 42.960 | San-Benedetto-del-Tronto |
| 20 | 7.421 | 43.728 | Monaco | 54 | 13.506 | 43.624 | Ancona |
| 21 | 9.350 | 42.967 | Centuri | 55 | 12.282 | 44.492 | Ravenna |
| 22 | 8.938 | 42.635 | Ile-Rousse | 56 | 12.426 | 45.418 | Venezia |
| 23 | 8.760 | 41.920 | Ajaccio | 57* | 12.511 | 45.313 | AAOT |
| 24* | 9.374 | 41.836 | Solenzara | 58 | 13.757 | 45.649 | Trieste |
| 25 | 8.018 | 43.878 | Imperia | 59 | 21.319 | 37.640 | Katakolo |
| 26 | 8.870 | 44.380 | Genova | 60* | 23.621 | 37.935 | Peiraias |
| 27* | 9.857 | 44.096 | La-Spezia | 61 | 24.941 | 37.438 | Syros |
| 28 | 10.299 | 43.546 | Livorno | 62* | 35.653 | 34.242 | Batroun |
| 29 | 10.238 | 42.742 | Marina-di-Campo | 63* | 29.879 | 31.209 | Alexandria |
| 30 | 11.789 | 42.093 | Civitavecchia | | | | |
| 31 | 12.634 | 41.446 | Anzio | | | | |
| 32 | 12.965 | 40.895 | Ponza | | | | |
| 33* | 13.589 | 41.209 | Gaeta | | | | |
| 34 | 14.269 | 40.841 | Napoli | | | | |

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


*Author contributions.* Marco Bajo: writing and reviewing the paper, developing the data assimilation code and its binding to SHYFEM, running some simulations, processing the results. Christian Ferrarin: Reviewing the paper, running some simulations, processing the results. Georg Umgiesser: Developing the SHYFEM code. Andrea Bonometto and Elisa Coraci: Providing the atmospheric forcing, financial support.

*Competing interests.* We do not have any competing interests.

*Acknowledgements.* We thank ISPRA for supporting the development of an operational system for forecasting the sea level along the Italian coasts.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


## References

Bajo, M., Biasio, F. D., Umgiesser, G., Vignudelli, S., and Zecchetto, S.: Impact of using scatterometer and altimeter data on storm surge forecasting, Ocean Modelling, 113, 85–94, https://doi.org/10.1016/j.ocemod.2017.03.014, 2017.

Bajo, M., Međugorac, I., Umgiesser, G., and Orlić, M.: Storm surge and seiche modelling in the Adriatic Sea and the impact of data assimilation, Quarterly Journal of the Royal Meteorological Society, 145, 2070–2084, https://doi.org/10.1002/qj.3544, 2019.

Barbariol, F., Pezzutto, P., Davison, S., Bertotti, L., Cavaleri, L., Papa, A., Favaro, M., Sambo, E., and Benetazzo, A.: Wind-wave forecasting in enclosed basins using statistically downscaled global wind forcing, Frontiers in Marine Science, 9, https://doi.org/10.3389/fmars.2022.1002786, 2022.

Byrne, D., Horsburgh, K., and Williams, J.: Variational data assimilation of sea surface height into a regional storm surge model: Benefits and limitations, Journal of Operational Oceanography, 0, 1–14, https://doi.org/10.1080/1755876X.2021.1884405, 2021.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, Wiley Interdisciplinary Reviews: Climate Change, 9, e535, https://doi.org/10.1002/wcc.535, 2018.

Cavaleri, L., Bajo, M., Barbariol, F., Bastianini, M., Benetazzo, A., Bertotti, L., Chiggiato, J., Davolio, S., Ferrarin, C., Magnusson, L., Papa, A., Pezzutto, P., Pomaro, A., and Umgiesser, G.: The October 29, 2018 storm in Northern Italy – An exceptional event and its modeling, Progress in Oceanography, 178, 102 178, https://doi.org/https://doi.org/10.1016/j.pocean.2019.102178, 2019.

Cerovecki, I., Orlić, M., and Hendershott, M. C.: Adriatic seiche decay and energy loss to the Mediterranean, Deep Sea Research Part I: Oceanographic Research Papers, 44, 2007 – 2029, https://doi.org/10.1016/S0967-0637(97)00056-3, 1997.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geographical Research, 99, 10,143 –10,162, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Evensen, G.: Sampling strategies and square root analysis schemes for the EnKF, Ocean Dynamics, 54, 539–560, https://doi.org/10.1007/s10236-004-0099-2, 2004.

Evensen, G.: Spurious correlations, localization, and inflation, pp. 237–253, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-03711-5_15, 2009.

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

Ferrarin, C., Bellafiore, D., Sannino, G., Bajo, M., and Umgiesser, G.: Tidal dynamics in the inter-connected Mediterranean, Marmara, Black and Azov seas, Progress in Oceanography, 161, 102–115, https://doi.org/https://doi.org/10.1016/j.pocean.2018.02.006, 2018.

Ferrarin, C., Bajo, M., Benetazzo, A., Cavaleri, L., Chiggiato, J., Davison, S., Davolio, S., Lionello, P., Orlić, M., and Umgiesser, G.: Local and large-scale controls of the exceptional Venice floods of November 2019, Progress in Oceanography, 197, 102 628, https://doi.org/https://doi.org/10.1016/j.pocean.2021.102628, 2021.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, https://doi.org/https://doi.org/10.1002/qj.49712555417, 1999.

Hersbach, H.: Sea Surface Roughness and Drag Coefficient as Functions of Neutral Wind Speed, Journal of Physical Oceanography, 41, 247 – 251, https://doi.org/10.1175/2010JPO4567.1, 2011.

https://doi.org/10.5194/egusphere-2022-1126
Preprint. Discussion started: 20 October 2022


Järvinen, H. and Undén, P.: Observation screening and background quality control in the ECMWF 3D-Var data assimilation system, p. 33, https://doi.org/10.21957/lyd3q81, 1997.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, https://doi.org/10.1017/CBO9780511802270, 2002.

Kepert, J. D.: On ensemble representation of the observation-error covariance in the Ensemble Kalman Filter, Ocean Dynamics, 54, 561–569, https://doi.org/https://doi.org/10.1007/s10236-004-0104-9, 2004.

Mariani, S., Casaioli, M., Coraci, E., and Malguzzi, P.: A new high-resolution BOLAM-MOLOCH suite for the SIMM forecasting system: assessment over two HyMeX intense observation periods, Natural Hazards and Earth System Sciences, 15, 1–24, https://doi.org/10.5194/nhess-15-1-2015, 2015.

Međugorac, I., Pasarić, M., Pasarić, Z., and Orlić, M.: Two recent storm-surge episodes in the Adriatic, International Journal of Safety and Security Engineering, 6, 589 – 596, https://doi.org/10.2495/SAFE-V6-N3-589-596, 2016.

Pérez, B., Fanjul, E. A., Pérez, S., De Alfonso, M., and Vela, J.: Use of tide gauge data in operational oceanography and sea level hazard warning systems, Journal of Operational Oceanography, 6, 1–18, https://doi.org/10.1080/1755876X.2013.11020147, 2013.

Pugh, D. T.: Tides, surges and mean sea-level (reprinted with corrections), John Wiley & Sons Ltd, 1996.

Schwab, D. and Rao, D.: Barotropic oscillations of the Mediterranean and Adriatic seas, Tellus, 35(A), 417–427, 1983.

Scicchitano, G., Scardino, G., Monaco, C., Piscitelli, A., Milella, M., De Giosa, F., and Mastronuzzi, G.: Comparing impact effects of common storms and Medicanes along the coast of south-eastern Sicily, Marine Geology, 439, https://doi.org/10.1016/j.margeo.2021.106556, 2021.

Smagorinsky, J.: General circulation experiments with the primitive equations: I. the basic experiment, Monthly Weather Review, 91, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2, 1963.

Storto, A.: Variational quality control of hydrographic profile data with non-Gaussian errors for global ocean variational data assimilation systems, Ocean Modelling, 104, 226–241, https://doi.org/https://doi.org/10.1016/j.ocemod.2016.06.011, 2016.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183–7192, https://doi.org/https://doi.org/10.1029/2000JD900719, 2001.

Tsimplis, M. N., Proctor, R., and Flather, R. A.: A two-dimensional tidal model for the Mediterranean Sea, Journal of Geophysical Research: Oceans, 100, 16 223–16 239, https://doi.org/https://doi.org/10.1029/95JC01671, 1995.

Umgiesser, G. and Bergamasco, A.: A staggered grid finite element model of the Venice Lagoon, in: Finite Element in Fluids, edited by Periaux, J., Morgan, K., Ofiate, E., and Zienkiewicz, O., pp. 659–668, Pineridge Press, Swansea, 1993.

Vidard, A., Balmaseda, M., and Anderson, D.: Assimilation of Altimeter Data in the ECMWF Ocean Analysis System 3, Monthly Weather Review, 137, 1393–1408, https://doi.org/10.1175/2008MWR2668.1, 2009.

Vilibić, I.: The role of the fundamental seiche in the Adriatic coastal floods, Continental Shelf Research, 26, 206 – 216, https://doi.org/10.1016/j.csr.2005.11.001, 2006.

Vilibić, I., Domijan, N., and Cupic, S.: Wind versus air pressure seiche triggering in the Middle Adriatic coastal waters, Journal of Marine Systems, 57, 189 – 200, https://doi.org/10.1016/j.jmarsys.2005.04.007, 2005.

Welch, P.: The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Transactions on Audio and Electroacoustics, 15, 70–73, https://doi.org/10.1109/TAU.1967.1161901, 1967.