



Assimilation of HF radar surface currents to optimize forcing in the northwestern Mediterranean Sea

J. Marmain^{1,2}, A. Molcard^{1,2}, P. Forget^{1,2}, A. Barth³, and Y. Ourmières^{1,2}

¹Université de Toulon, UMR7294, CNRS/INSU, IRD – Mediterranean Institute of Oceanography (MIO), UM 110, 83957 La Garde, France

²Aix Marseille Université, CNRS/INSU, IRD, Mediterranean Institute of Oceanography (MIO), UM 110, 13288 Marseille, France

³GeoHydrodynamics and Environment Research (GHER), AGO/MARE, University of Liège, Liège, Belgium

Correspondence to: J. Marmain (julien.marmain@univ-tln.fr)

Received: 25 July 2013 – Revised: 21 March 2014 – Accepted: 28 April 2014 – Published: 11 June 2014

Abstract. HF radar measurements are used to optimize surface wind forcing and baroclinic open boundary condition forcing in order to constrain model coastal surface currents. This method is applied to a northwestern Mediterranean (NWM) regional primitive equation model configuration. A new radar data set, provided by two radars deployed in the Toulon area (France), is used. To our knowledge, this is the first time that radar measurements of the NWM Sea are assimilated into a circulation model. Special attention has been paid to the improvement of the model coastal current in terms of speed and position. The data assimilation method uses an ensemble Kalman smoother to optimize forcing in order to improve the model trajectory. Twin experiments are initially performed to evaluate the method skills. Real measurements are then fed into the circulation model and significant improvements to the modeled surface currents, when compared to observations, are obtained.

1 Introduction

The ocean coastal dynamics of the northwestern Mediterranean (NWM) Sea has been studied over the last decades using observations (e.g., Millot, 1999; Petrenko, 2003; Sammarti et al., 1995; Allou et al., 2010), modeling (Echevin et al., 2003; Casella et al., 2011; Ourmières et al., 2011; Pairaud et al., 2011), or a combination of both (Schaeffer et al., 2011b; Guihou et al., 2013). The NWM is a micro-tidal sea (e.g., Alberola et al., 1995) and its dynamic is mainly governed by wind forcing and by the Northern Current (NC),

which is the northern branch of the general circulation of the NWM Sea (Millot, 1999). The NC is a quasi-permanent density current that originates in the eastern part of the NWM and forms a large-scale loop. Bathymetric constraints operate on the northern, western and southwestern sides of its path. This geostrophic current can experience a strong variability in this particular region, due to the steep topography and the irregular coastline, as well as the atmospheric forcing. The NC varies in strength and width, its meandering can generate eddies, and its separation in the vicinity of Toulon (Fig. 1) controls water fluxes at the entrance of the Gulf of Lion (GoL). Several of the above studies were carried out in the GoL or off Nice. However, the dynamics of the NC in the intermediate area, typically the region off Toulon, is poorly documented. Since May 2010, an HF radar system has been deployed in this region in an attempt to fill this void, allowing the gathering of continuous and synoptic observations of surface currents up to 100 km off the coastline with typical temporal and spatial resolutions of 1 h and 3 km, respectively.

Scientific interest in the assimilation of HF radar measurements in hydrodynamical models has been growing over the last few years. Various assimilation methods have been developed based on nudging (Lewis et al., 1998), Kalman filtering using parametrized (Paduan and Shulman, 2004; Shulman and Paduan, 2009) or ensemble (Breivik and Sætra, 2001; Oke et al., 2002b; Barth et al., 2008a) error covariances, and the application of adjoint-based methods (Kurapov et al., 2003; Hoteit et al., 2009).

The method applied herein consists in optimizing the wind-forcing and lateral boundary-forcing components of a

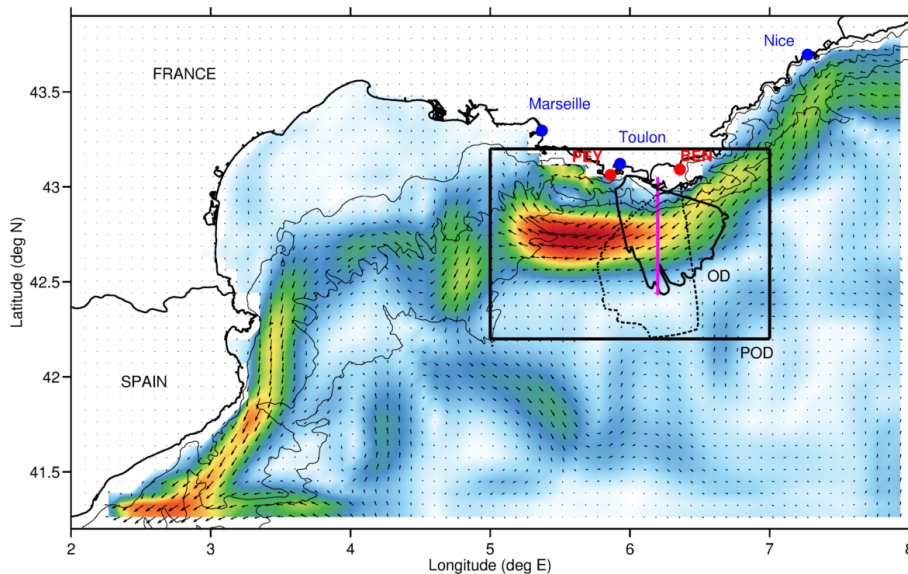


Figure 1. Depiction of the studied area. The larger area is the model domain (MD). Red points mark radar stations PEY and BEN. The black contour lines show areas where HF radar data are available more than 20 % over the assimilation period for PEY (full) and BEN (dashed) radars and is called the observation domain (OD). The black box sketches the pseudo-observation domain (POD). The meridional magenta line crossing the coastal current will be used for Hovmöller diagrams. Surface currents of free simulation averaged over the assimilation period are shown. Color bars range from 0 (white) to 0.5 (red) m s^{-1} . The 100, 1000 and 2000 m isobaths are drawn (black thin lines).

numerical regional primitive equation model in order to minimize differences between the model and the radar-derived surface currents.

Wind optimization is an important issue as, in coastal zones, surface wind errors can explain a large part of the model error regarding surface currents (He et al., 2004; Barth et al., 2008b). Such optimization can be achieved by correcting the surface wind field (Barth et al., 2011) or the parameters governing the wind–sea momentum transfer, e.g., the drag coefficient (Ten Brummelhuis et al., 1993; Peng et al., 2012).

Similarly, open boundary condition (OBC) optimization can be achieved by correcting the OBC values themselves or by correcting the OBC’s formulation and discretization. This last approach has been addressed several times (e.g., Orlandi, 1976; Flather, 1976; Marchesiello et al., 2001; Blayo and Debreu, 2005). It consists in a numerical treatment of the oceanic boundaries allowing the removal of the perturbations generated inside the model domain without deteriorating the inner model solution (Røed and Cooper, 1986). Research in this field is still ongoing. Instead of pursuing this approach further, we shall address the first approach, i.e., the optimization of time-dependent OBC values. This has been achieved previously in barotropic models (e.g., Seiler, 1993; Ten Brummelhuis et al., 1993; Gunson and Malanotte-Rizzoli, 1996; Bogden et al., 1996; Taillandier et al., 2004) and more rarely in baroclinic models (e.g., Shulman et al., 1999; Taillandier et al., 2009).

Variational methods (e.g., Le Dimet and Talagrand, 1986) are often preferred to tackle forcing optimization problems. The adjoint model technique is the one most often used, but depending on the model it can sometimes be difficult to formulate (Shulman et al., 1999). Another way of solving optimization problems is by the application of a Kalman filter (e.g., Gelb, 1974). However, in its original form, this technique has a prohibitively high computational burden to be used in geophysical applications. The development of reduced rank methods such as ensemble methods (Evensen, 1994) allows for the reduction of this computational cost because only a limited number of model states ($\mathcal{O}(100)$) needs to be stored in order to obtain an accurate estimate of the error covariance matrices. Moreover, ensemble methods are efficient for dealing with error propagation in non-linear models.

Recently, the Ensemble Perturbation Smoother (EnPS) has been developed to optimize stationary tidal boundary conditions of a hydrodynamical model of the German Bight (Barth et al., 2010), a very shallow area (maximum depth of about 50 m) where winds and tides are the main sources of variability. The EnPS has also been used successfully to optimize surface wind estimates (Barth et al., 2011) in the same area, but the resulting model surface current was not directly evaluated. This beckons the following questions: firstly, whether optimizing wind can improve model surface currents, and secondly, whether it is possible to optimize non-stationary baroclinic OBC by applying the EnPS with the aim of improving model surface currents in a highly dynamic area.

These issues are addressed by using the EnPS in a regional baroclinic model of the NWM Sea and by using the HF radar-derived surface currents measured off Toulon.

Section 2 describes the hydrodynamical model configuration and the experimental data set. The data assimilation method is explained in Sect. 3. Results of wind optimization are shown in Sects. 4 and 5, showing the results of OBC optimization. Finally, results are discussed in the conclusion.

2 Model and data

2.1 Model

The simulations described below were obtained using the Regional Ocean Modeling System (ROMS, Shchepetkin and McWilliams, 2003, 2005), which solves the ocean primitive equations, featuring a free surface formulation and hydrostatic approximation. It uses terrain-following coordinates in the vertical dimension and curvilinear coordinates in the horizontal dimension. ROMS incorporates advanced features and high-order numerics, allowing efficient and robust resolution of mesoscale dynamics in the oceanic and coastal domains. ROMS has been widely applied to different oceanic regions for a variety of ocean dynamics studies (Penven et al., 2001; Capet et al., 2008; Casella et al., 2011). The model domain (MD) as depicted in Fig. 1 extends from 2.10 to 7.93° E and from 41.26 to 43.96° N. Horizontal resolution is 1/12°, and 32 vertical σ levels are employed with vertical refinement near the surface to obtain a satisfactory representation of the surface layer.

Two active open boundaries connecting the MD with the NWM basin are incorporated. The OBC treatment is a well-known ill-posed problem (e.g., Chapman, 1985; Marchesiello et al., 2001; Blayo and Debreu, 2005), and the goal here is to implement the optimal conditions in order to allow inflow from the open sea, to conserve mass and energy and to reduce any wave reflection phenomena. A radiation–relaxation condition for 3-D momentum and tracers is combined with a Flather condition for the 2-D momentum (Marchesiello et al., 2001), and a sponge layer is added to avoid spurious reflections. The typical relaxation coefficients have been adjusted to 1 and 10 days (for tracers and baroclinic velocities, respectively) for inflow and to 360 days for outflow.

The data for the relaxation boundary conditions were derived from the PSY2V4R2 operational configuration developed by MERCATOR-OCEAN (<http://www.mercator-ocean.fr>), based on the NEMO (Nucleus for European Modelling of the Ocean) model. The PSY2V4R2 configuration features a horizontal resolution of 1/12° with 50 vertical z levels, and includes a sequential data assimilation system (SAM2V1) based on the Kalman filter. Satellite sea surface temperature (SST) and sea level anomaly (SLA) are assimilated, in addition to available in

situ temperature and salinity profiles. Daily averages of the velocity, mass fields and surface height at the MD open boundaries are used. The forcing is made off-line.

At the air–sea interface, the oceanic model is forced by the ALADIN operational regional model from Météo-France. ALADIN runs at a resolution of 1/10° and features data assimilation and state-of-the-art atmospheric physics (Fischer et al., 2005). The importance of such high spatial and temporal resolution in terms of oceanic response has been widely studied and demonstrated over the last few years, particularly to achieve a more realistic representation of the oceanic mesoscale circulation (Langlais et al., 2009; Schaeffer et al., 2011a).

The initial condition of the simulation is based on the daily averaged baroclinic and barotropic velocities, temperature, salinity and surface elevation derived from PSY2V4R2 for 1 July 2011. A 4 month spin-up period is considered to be adequate for reaching a state of equilibrium. The model output corresponds to 3-hour averaged fields.

The average surface current for the period 15 November–15 December 2012 is shown in Fig. 1. The model simulations give a maximum mean current of 0.5 m s⁻¹, in accordance with the observations.

2.2 Observations

Since May 2010, HF radars have been monitoring surface currents off Toulon (Fig. 1). The radars employed are WERAs (Wellen Radar, Gurgel et al., 1999) and operate at 16.15 MHz with 50 kHz bandwidth. A radar system is installed at Fort Peyras in Cape Sicié (hereafter PEY) and operates in a monostatic configuration, i.e., transmitter and receiver are very close together (~ 100 m). The second system operates in a bistatic configuration. The transmitting antennas are those of the PEY radar, whereas the receiving antennas consist of a linear array of eight antennas installed near Cape Bénat (hereafter BEN), 40 km to the east of PEY. Integration time, i.e., the time over which the temporal signals are processed, is one hour. Velocity accuracy is 4 cm s⁻¹, and the spatial sampling is 3 km in range. An azimuthal discretization of 2° was achieved through direction finding using the MUSIC processing technique (Schmidt, 1986). This technique is now in common use in the HF radar community (Lipa et al., 2006; Sentchev et al., 2013). The radial component of the current is measured along the look direction of the radar. With the given radar frequency, the radial velocity is theoretically measured at a depth of 74 cm (Stewart and Joy, 1974).

The MUSIC technique involves spatial holes, temporal gaps and false alarms (outliers) on radial velocity maps. For the present study, a series over a 30 day time period was used at every radar cell in order to remove outliers. Outlier removal was performed using the histogram of the temporal gradients of the current. Spatial holes and temporal gaps were not filled, in order to avoid the addition of spu-

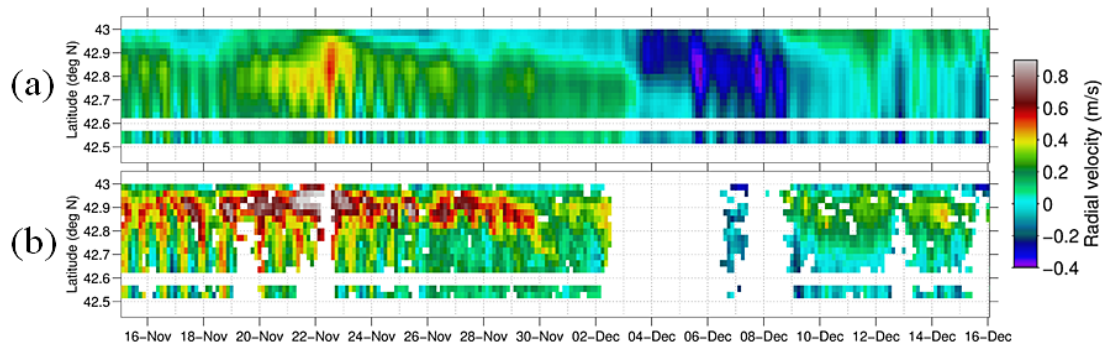


Figure 2. Hovmöller diagram of radial velocities from the PEY station along a south–north transect at 6.2° N (see Fig. 1), (a) from free simulation and (b) from observation.

rious and correlated information in radial velocities, which can result from interpolations. Only the radial component maps were taken into consideration, and not the current vectors that can be deduced by the geometric combination of the two sets of radar measurements. The first reason for this choice is to prevent the introduction of the geometric dilution of precision, which is inherent in the vector mapping process (Chapman et al., 1997). The second reason is to dispose of the most extended observation area and to avoid wasting valid observations.

HF radar measurements from PEY and BEN (y_{pey} , y_{ben}) gathered during the period 15 November 2011 00:00 UTC to 15 December 2011 23:00 UTC were taken into consideration for this study. A spatial mask rejecting the input from radar cells having a temporal coverage below 20 % over the period was applied to the data. The 20 % temporal coverage level is depicted in Fig. 1 and is hereafter referred to as the observed domain (OD). When measurements from both stations are available, variability in the total coverage is mainly due to the degradation of PEY performance owing to high wind speed.

Twenty-eight temperature and salinity profiles from conductivity–temperature–depth (CTD) measurements were also taken into consideration for the analysis of assimilation results. CTD measurements were carried out in the area within the geographical grid defined by coordinates 5.8 – 7.35° E and 42.6 – 43.3° N, over the period 11–14 December 2011, during a cruise undertaken in the framework of the European TOSCA (Tracking Oil Spill and Coastal Awareness network, <http://www.tosca.med.eu>) program. Note that this data set is independent data that can be used for validation of results.

2.3 Model validation

The Hovmöller diagram (time versus latitude) of both model-derived and observed radial velocities is depicted for the PEY station (Fig. 2). The radial velocity is the component of the total surface velocity of the current \mathbf{U} in the radial direction. By convention its sign is positive (negative) for velocity going toward (away from) the station. High posi-

tive values in the PEY radial component are a distinct signature of the coastal current (Marmain et al., 2011). The model clearly underestimates the intensity of the current radial components, and the position of the maximum seems slightly shifted off-shore. However, the temporal variability is well simulated: the change in sign at the beginning of December, as well as the high-frequency modulations (near-inertial oscillations), and the maximum intensity observed from PEY around 23 November are reproduced well by the model.

The validity of the vertical structure of the model was evaluated using CTD measurements. Model temperature and salinity exhibit low departures from observations. According to the definition of the root mean square error (RMSE), E_r and the bias B given in Sect. 3, $E_r = 0.780^{\circ}\text{C}$ and $B = 0.002^{\circ}\text{C}$ for temperature and $E_r = 0.084$ and $B = -0.054$ for salinity. Significant correlations appear (0.79 and 0.86 respectively). The main errors come from the inaccurate representation of the thermocline by the model. However, densities are in good agreement.

3 Data assimilation methodology

HF radar-derived velocities have already been used to correct M2 tidal boundary conditions (Barth et al., 2010) and surface wind fields (Barth et al., 2011) using the so-called EnPS, a fixed-interval sequential smoother (Cosme et al., 2012). In this study, we investigate, firstly, whether optimizing wind can improve model surface currents, and secondly, whether it is possible to optimize non-stationary baroclinic OBC by applying the EnPS with the aim of improving model surface currents in a highly dynamic area. Uncertainties in these atmospheric and oceanic fields are represented by an ensemble of wind forcing or an ensemble of OBC forcing based on perturbations of the original (or background) fields (\mathbf{x}^b).

3.1 Ensemble setup

The perturbation method is detailed in Barth et al. (2011). Briefly, the spatial structure of the forcing perturbations is

obtained by using a Fourier decomposition of the forcing field. The Fourier coefficients correspond to the spatial field. For each temporal frequency subjected to perturbation, a complex random time series is constructed with a temporal correlation scale according to the perturbed frequency, with nil mean and unit variance (Evensen, 1994). Such a series is then multiplied by the corresponding Fourier coefficient and the resulting products are summed. The real part of this quantity corresponds to the forcing perturbation \mathbf{x}_p^l , with l as the member index. \mathbf{x}_p^l is added to \mathbf{x}^b , resulting in

$$\mathbf{x}^l = \mathbf{x}^b + \alpha \mathbf{x}_p^l. \quad (1)$$

With this method, forcing errors are assumed to have similar spatial and temporal scales to the original, non-perturbed forcing. By using this approach, it is assumed that the uncertainty in the forcing field is mainly caused by phase and amplitude errors of the underlying processes rather than errors in the variability. Perturbations are estimated using the time variability of the forcing field. \mathbf{x}_p^l is weighted by the α coefficient in order to achieve realistic perturbations. This approach ensures that a spatial structure is multiplied by a compatible temporal scale. The forcing ensemble must verify $\mathbf{x}^b = \langle \mathbf{x}^l \rangle$, with $\langle \cdot \rangle$ as the ensemble average operator.

An ensemble of 100 perturbations for zonal and meridional components of the wind vector \mathbf{W} was generated over the MD and over a period extending from 1 November to 31 December 2011, taking into consideration temporal scales ranging from 3 h to 30 days. The minimum period is the temporal resolution of the original forcing and the maximum is set to 30 days, assuming that higher periods do not contain uncertainties.

An ensemble of 100 OBC perturbations was generated, taking into consideration temporal scales ranging from 1 to 30 days, over the same time period, for temperature, salinity, surface elevation, zonal velocities and meridional velocities (respectively, T^{obc} , S^{obc} , η^{obc} , U^{obc} , V^{obc}) at the eastern boundary of the MD. Only this boundary was taken into consideration because it is located upstream of the observation area along the NC flow. Perturbations were introduced in the northern part of the OBC (42.5–44° N) and over the sub-surface ocean layer (–200–0 m) as the prognostic variables at higher depths do not exhibit any significant correlation with the surface current in the OD. In fact, such a restriction on the perturbation domain can be seen as a spatial localization (Hamill et al., 2001; Houtekamer and Mitchell, 2001). Therefore, we consider that errors in the model solution originate only from the perturbed area.

For each forcing perturbation, the model was run using the same initial condition as the free simulation. A 15 day ensemble spin-up was selected to attain a satisfactory variance between surface current members. The ensemble simulations for the time period 15 November–15 December 2011 were taken into consideration.

3.2 Assimilation scheme

A general description of the EnPS (Ensemble Perturbation Smoother) is given in Appendix A. For more information, the reader is referred to Barth et al. (2010, 2011).

Since the EnPS is a non-sequential method, all available observations over the period of analysis are used to compute the optimal state. The time dimension is consequently embedded in the observation and optimization vectors. The model trajectory is not estimated directly. Rather, the optimal forcing trajectory is derived and then run in the model. Furthermore, the advantage of estimating the forcing first is that the final model trajectory satisfies (per construction) the model equations (and thus all non-linear balances between model variables), which is not the case if the model trajectory is analyzed directly, and correcting the forcing instead of the model state also avoids unrealistic transients (Malanotte-Rizzoli et al., 1989).

The optimized forcing \mathbf{x}^a is computed using the Kalman filter analysis step with a non-linear observation operator:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y}^o - h(\mathbf{x}^b)), \quad (2)$$

where \mathbf{K} is the Kalman gain that varies according to ensemble and observation error covariance matrices (Appendix A). The non-linear observation operator $h(\cdot)$ is applied to the forcing \mathbf{x} and $h(\mathbf{x})$ represents the surface currents within the integration period.

The analyzed model trajectory is obtained in a final step, by performing a new run of the model using the optimized forcing field.

The analysis \mathbf{x}^a is unbiased if both background estimate \mathbf{x}^b and observations \mathbf{y}^o are unbiased. If the observation operator is non-linear, it can be verified that the analysis is unbiased using Eq. (3) (Barth et al., 2011):

$$\langle h(\mathbf{x}^l) \rangle = h(\langle \mathbf{x}^l \rangle). \quad (3)$$

The degree of linearity of the observation operator was verified using Eq. (3). The overall root mean square difference between both sides of Eq. (3) was computed for wind and OBC ensembles. A value of 0.03 m s^{-1} was found, which is small in comparison with other errors and bias. However, the standard analysis methods provide often useful results even if the background is biased.

The observation vector \mathbf{y}^o is associated with an observation error covariance matrix \mathbf{R} . For the sake of simplicity, \mathbf{R} is assumed diagonal and it is also assumed that \mathbf{R} is the sum of the measurement error covariance $R_{obs}\mathbf{I}$ and the representativity error covariance $S_{obs}^2\mathbf{I}$ where \mathbf{I} is the identity matrix. R_{obs} represents the measurement errors that arise from noise effects imparted by data processing as well as by the radar itself. S_{obs} is introduced to characterize processes and scales present in the observation but not resolved by the model. The representativity error could be extended to include processes that cannot be corrected using these forcings only (Barth et al., 2011).

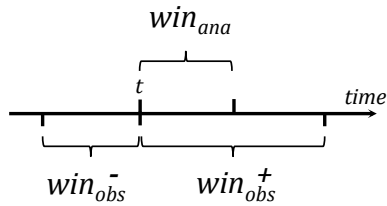


Figure 3. Schematic representation of the optimization window (win_{ana}), the backward (win_{obs}^-) observation window and the forward (win_{obs}^+) observation window. t is the beginning of win_{ana} .

To increase the effective dimension of the error subspace, to discard distant observations involving unrealistic correlations and to improve computational capabilities, the Kalman filter analysis step can be performed in reduced spatial domains, zone by zone, which is known as spatial localization (Hamill et al., 2001; Houtekamer and Mitchell, 2001). A similar approach can be taken in the temporal dimension.

Typically, the model integration period is divided into a series of shorter, non-overlapped, intervals called optimization windows with a duration of win_{ana} . Considering an observation window of duration win_{obs} , such as $win_{ana} = win_{obs}$, forcing fields within the time range $[t; t + win_{ana}]$ are updated using all available observations in such observation windows, with t as the beginning of the optimization window (Barth et al., 2010, 2011). Here, the definition of the observation window was modified to take into account backward (win_{obs}^-) or forward (win_{obs}^+) observations within the time range $[t - win_{obs}^-; t + win_{obs}^+]$ (Fig. 3), considering the delay of the oceanic response to the lateral and surface forcing that may result in non-instantaneous correlations between forcing and surface currents (e.g., Seiler, 1993). This way, the standard EnPS is retrieved with $win_{obs}^+ = win_{ana}$ and $win_{obs}^- = 0$.

3.3 Experimental setup

The assimilation methodology was first applied to a set of radar pseudo-observations in identical twin experiments (TE) to assess the method, and then applied to experimental radar observations.

In TE, forcing and model control states (\mathbf{x}^{true} , $h(\mathbf{x}^{true})$) are extracted from the ensemble members. These control states are used to generate a set of pseudo-radial velocity fields by adding a Gaussian noise. The free run corresponds to running the model using the original forcing. The assimilation of these pseudo-observations results in an analyzed forcing that can be used in the model to simulate the resulting analyzed model surface current. The true state is known and the skill of the assimilation scheme can be quantified.

Pseudo-observation values are given for a geographical grid defined by coordinates 5–7° E and 42.2–43.2° N, the pseudo observation domain (POD), and are defined as the projection of model surface velocities in the radial directions

at each model node. The measurement noise is assumed to be nil.

The experimental observations consist in the radial velocity measurements derived from the two PEY and BEN radars (Sect. 2.2) with R_{obs} is 0.04 m s⁻¹, corresponding to the velocity accuracy.

3.4 Diagnostics

The skill of the assimilation method is evaluated by comparing the control state (control) to the free (b) or analyzed (a) states in terms of RMSE (e.g., Oke et al., 2002a). In the case of the twin experiments, the control state is the true one, true. In the real case experiments, a cross-validation-like technique is used (e.g., Brankart and Brasseur, 1996): the control state is a random set of observations, o, discarded at each time step from the initial set of observations.

Let n be the spatial index of dimension N , and k the temporal index of dimension K . The RMSE of a variable Φ is defined by

$$Er^{a,b} = \sqrt{\frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (\Phi_{n,k}^{a,b} - \Phi_{n,k}^{control})^2} \quad (4)$$

When the summation is applied to both spatial and temporal dimensions of Φ , Er corresponds to an overall scalar measurement of the assimilation performance. If the sum is computed in the spatial dimension only, the error is evaluated during the experiment time evolution ($Er_N(k)$), and when the average is calculated in the temporal dimension, the performance is evaluated in terms of spatial structure ($Er_K(n)$).

The bias B can also be computed to determine the mean difference between model solution and control:

$$B^{a,b} = \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (\Phi_{n,k}^{a,b} - \Phi_{n,k}^{control}) \quad (5)$$

The skill score “Sc” is an additional metric that can measure the performance of the data assimilation (e.g., Oke et al., 2002a), and is defined by

$$Sc = 1 - \left(\frac{Er^a}{Er^b} \right)^2. \quad (6)$$

A positive score “Sc” indicates that the assimilation is efficient, while a negative score indicates a degradation. In particular, $Sc = 1$ is obtained for a perfect analyzed field, and $Sc = 0$ reflects a nil impact on average.

Eventually, $\Delta = Er_K^b(n) - Er_K^a(n)$ will also be computed. $\Delta > 0$ ($\Delta < 0$) indicates an area of improvement (degradation).

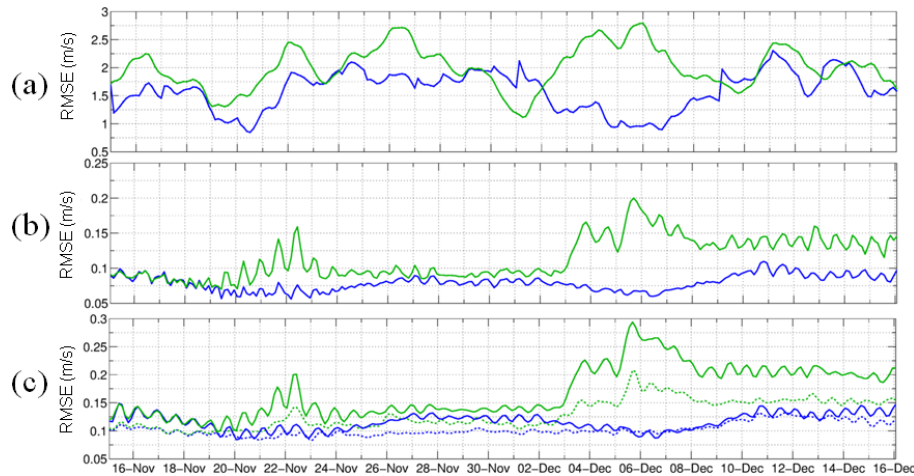


Figure 4. Root mean square error Er time evolution of (a) wind (W), (b) radial velocity (y) and (c) surface current (U) for *WindTE*, computed over the POD (full lines) or the MD (dashed lines). Er^b (Er^a) is in green (blue).

4 Impact of the wind optimization

4.1 Results of a twin experiment

The first sensitivity study concerns the wind forcing. Wind perturbations are introduced in such a way that the resulting wind ensemble represents the wind uncertainties. Insufficient wind perturbations may induce only a small variance of the model surface current ensemble and the model control space will be misrepresented by this ensemble of model states. On the contrary, excessive wind perturbations can generate unrealistic patterns in the wind ensemble or in the corresponding model ensemble. We found that, for the present model setting, an error in wind forcing corresponding to 50 % of its temporal variability ($\alpha = 0.5$) is the most efficient in terms of the overall RMSE between control and analyzed states for wind and model solutions during the assimilation period.

The representativity error, S_{obs} , is also investigated. We look for the value of S_{obs} that provides the most accurate optimized wind and model solutions for a given optimization and observation window. A high value of S_{obs} can give insufficient weighting to the observations relative to the background state, while a low value of S_{obs} can produce unrealistic optimized forcings. Therefore, the analysis of the forcing and the model solution must be investigated concurrently to find the optimal compromise. For the present experiment, S_{obs} was set at 1 m s^{-1} .

The assimilation process also requires the selection of the duration of the temporal optimization window, win_{ana} , used to localize the analysis. This duration should minimize the discontinuities between two consecutive windows. Typically, the amplitude of these discontinuities must be on the same order as the temporal variability of the model, and again a compromise must be found to reduce the error of optimized

forcing and the resulting model solution. 1 day was found to be adequate for win_{ana} .

Traditionally, observations are considered within the optimization window. However, depending on the application, the covariances between observations and ensemble simulations at a given time may be low due to a weak causality between forcing and model solution at the observation points. At such a stage we suggest taking into consideration a temporal shift between the observation and optimization windows. Given a time reference that corresponds to the beginning of win_{ana} , the observations before (after) this time are included in win_{obs}^- (win_{obs}^+). The values $win_{obs}^- = 0$ day and $win_{obs}^+ = 3$ days provide the best results in terms of the overall Er for wind and surface currents. The experiment performed with this data set is hereafter referred to as *WindTE*.

Figure 4a shows the improvement in wind field estimation using idealized radar measurements in terms of the time evolution of $Er_N(k)$. The assimilation is effective over the total period except at some dates (e.g., 30 November–2 December 2011), representing only 8.6 % of the overall duration. These dates mainly correspond to weak covariances between surface currents on the POD and wind over the MD, corresponding to the presence of mesoscale circulation patterns that are not wind induced.

The resulting velocity field in the radial component shows an improvement during the assimilation period after a 3-day adjustment period (Fig. 4b). The overall Sc of radial velocities is 0.54. Oscillations can be observed to have a period of 17 h, which is close to the inertial period at the considered latitudes, but cannot be directly correlated with the wind solutions (Fig. 4a). The presence of such oscillations indicates that free and control surface current near-inertial oscillations are not in phase and that the wind optimization has only a limited impact on the near-inertial oscillations.

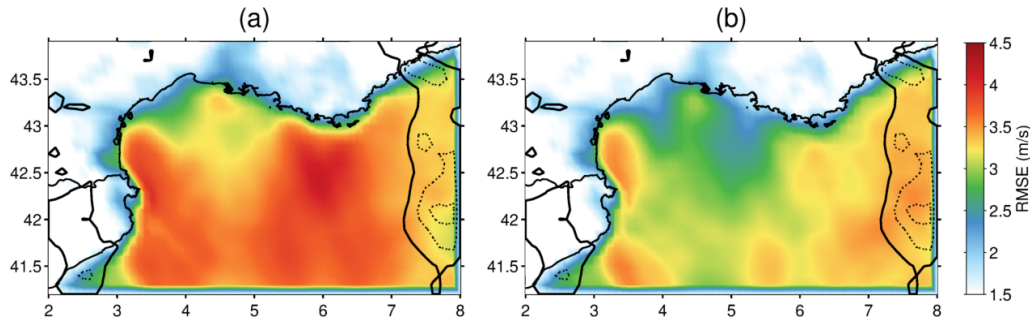


Figure 5. Spatial pattern of (a) Er^b and (b) Er^a of the wind (W) computed over the assimilation period for *WindTE*. Δ is indicated by contour lines and only 0 m s^{-1} (bold black lines) and negative (-0.2 and -0.1 m s^{-1} , dashed black lines) values are shown.

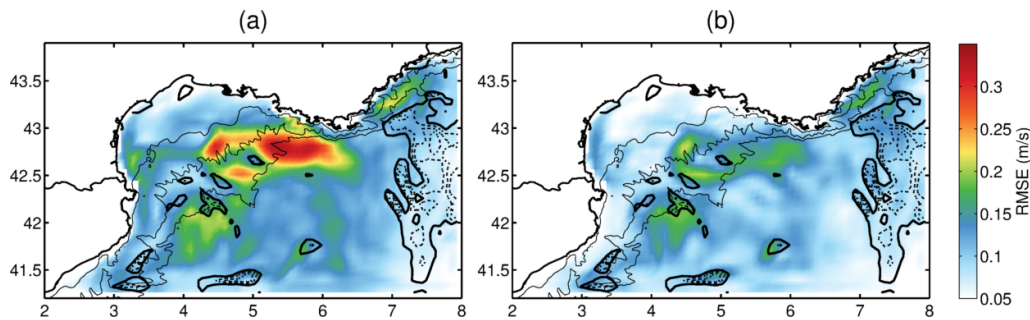


Figure 6. Spatial pattern of (a) Er^b and (b) Er^a of the surface current U for *WindTE*. Δ is indicated by contour lines and only 0 m s^{-1} (bold black lines) and negative (-0.03 , -0.02 and -0.01 m s^{-1} , dashed black lines) values are shown. The 100, 1000 and 2000 isobaths are drawn (black thin lines).

The spatial patterns of $Er_K(n)$ for the forcing wind are depicted in Fig. 5, superimposed on the negative values of Δ to highlight the regions where the optimized wind is degraded compared to the original wind. Over most of the MD, assimilation of surface currents improves the wind estimation. The area in the eastern part of the MD, where wind estimation is degraded, corresponds to a region where original and control winds are in good agreement before assimilation.

The model response to the analyzed wind forcing in terms of Er of surface velocity field is shown in Fig. 6. The surface current vector field resulting from analysis is generally more accurate than the free run. Maximum improvement is observed at the entrance of the GoL and along the shelf break. Along the eastern open boundary the assimilated solution is less accurate, corresponding to the degradation of the analyzed wind forcing (Fig. 5).

Hovmöller diagrams of radial velocities along a south–north transect at 6.2° N across the POD are shown in Fig. 7 for the PEY radar station. The efficiency of the optimized wind forcing on the radial velocity model response is visible in the high intensity patterns. For the positive radial velocities (red), the assimilation reduces the overestimation of the free run in the period 20–24 November. Later in December, the free run clearly underestimates the radial velocities, and again the optimized wind intensifies these features, for outward components (4–10 December) and inward components

(10–16 December). Moreover, the use of the optimized wind preserves the near-inertial oscillations. Results for the BEN radar station (not shown) are similar but weaker.

4.2 Results of a real case experiment

The *WindReal* experiment is performed to correct the same free model run with real radar observations, from the two PEY and BEN radar stations (Fig. 1). Observations are averaged over consecutive 3 h windows to improve computational capabilities.

The best compromise in terms of lowest overall Er and Sc is obtained for $\alpha = 0.3$, $S_{\text{obs}} = 1 \text{ m s}^{-1}$, $\text{win}_{\text{ana}} = 1 \text{ day}$, $\text{win}_{\text{obs}}^- = 0 \text{ day}$ and $\text{win}_{\text{obs}}^+ = 3 \text{ days}$, according to the findings from sensitivity tests performed on these parameters.

For this experiment, contrary to the TE where the true state is known in terms of wind and model solution, no ground truth data are available to validate the optimized wind forcing. Meteorological stations in the region did supply wind observations effectively, but all were assimilated into the ALADIN meteorological forcing, and no additional independent wind observations were available. To ensure realistically optimized winds, even if this study does not aim at improving wind fields, the original and optimized winds have been compared to in situ winds from four land stations located near Toulon and from two moorings located in the GoL and off

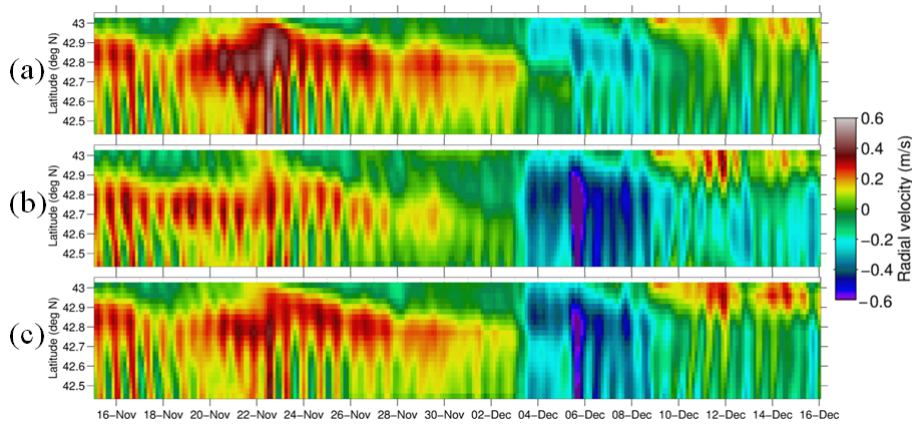


Figure 7. Hovmöller diagram of radial velocities from PEY along the south–north transect at 6.2° N (see Fig. 1), simulated by (a) the free simulation, (b) the control and (c) the analysis for *WindTE* simulation.

Table 1. Radial velocity statistics for the PEY, BEN and both (PEY + BEN) radar stations computed for the free, *WindReal* and *ObsReal* runs in relation to the observation over the entire assimilation period. *B* and *Er* are in m s^{-1} . *Sc* is without unit.

		Free	<i>WindReal</i>	<i>ObsReal</i>
PEY + BEN	<i>B</i>	−0.101	−0.051	−0.078
	<i>Er</i>	0.191	0.158	0.169
	<i>Sc</i>	–	0.31	0.23
PEY	<i>B</i>	−0.133	−0.074	−0.099
	<i>Er</i>	0.225	0.179	0.192
	<i>Sc</i>	–	0.36	0.27
BEN	<i>B</i>	−0.056	−0.022	−0.053
	<i>Er</i>	0.138	0.127	0.134
	<i>Sc</i>	–	0.13	0.06

Nice. No significant changes appear in correlations between observed and original or *WindReal* analyzed winds. The optimized meridional wind component approximates the original one. The optimized zonal wind component *Er* increases by 0.6 m s^{-1} compared to the original wind. The original and optimized wind maps show some differences, but the optimization does not introduce any unrealistic features.

The time evolution of $Er_N(k)$ computed from the radial surface current components measured by the PEY and BEN radars is plotted in Fig. 8.

Current estimation by using wind optimization is better for the PEY station than for the BEN station, with an *Sc* of 0.36 and 0.13, respectively (Fig. 8b, c, and Table 1). The two radar stations have different viewing geometries, in particular regarding the NC vein, which is the most dynamically active region where errors are the highest (Fig. 9a, d). While the PEY radar station is well placed to identify the westward coastal current velocity corresponding to the largest component of the NC, the BEN station mainly measures its mean-

dering and oscillations, as the NC flows perpendicularly to the look direction of the BEN radar.

Table 1 shows radial velocities *Er*, *B* and *Sc* for PEY, BEN and both radar stations. When both stations are considered together, *Er* reaches values of up to 0.15 m s^{-1} (Table 1), and the overall bias is halved.

The spatial patterns of $Er_K(n)$ are limited to the radar regions (Fig. 9) to be compared with the observations. Free radial velocities from the PEY station show higher discrepancies with the observations (Fig. 9a) than free radial velocities from the BEN station (Fig. 9d). The assimilation performs well in most of the domain, especially north of 42.7° N . The best improvement (red areas of the figure) corresponds to the NC vein, the most energetic feature of the region where errors are the highest.

Hovmöller diagrams from free (Fig. 2a) and analyzed (Fig. 11a) radial velocities from the PEY station show that optimized wind forcing acts on the NC vein intensity. During the period 15–26 November 2011, the intensity of the NC vein is strengthened, better approximating the observations (Fig. 2b), and is weakened during the period 4–9 December 2011. A southward shift of the NC vein was also found during the period 10–16 November 2011.

Data assimilation performance reduces when moving southward and degradation regions may appear. This may be due to the scarcity of observations when moving away from the radar stations.

The norm of vector speed difference $U^b - U^a$ between free currents U^b and analyzed currents U^a averaged over the assimilation period and at the surface (Fig. 10a) and integrated over the first 200 m below the surface (Fig. 10c) shows that the corrections are primarily located on the surface in the NC vein all along the French coast and over the GoL shelf break and can occur over the whole MD. However, they vanish with depth.

Once we have verified that the wind optimization obtained through the assimilation of surface velocities does

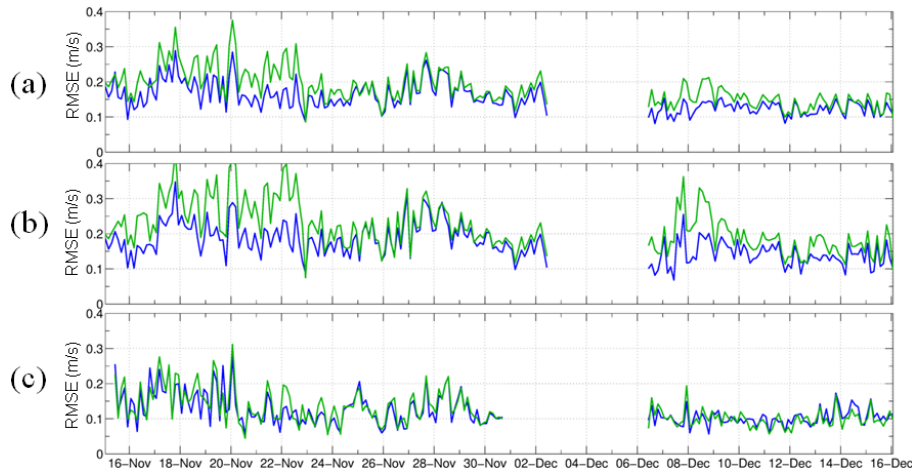


Figure 8. (a) Er time evolution of radial velocity from (a) both radar sites (y), (b) PEY (y_{peg}) and (c) BEN (y_{ben}) for *WindReal*. Er^b (Er^a) is in green (blue).

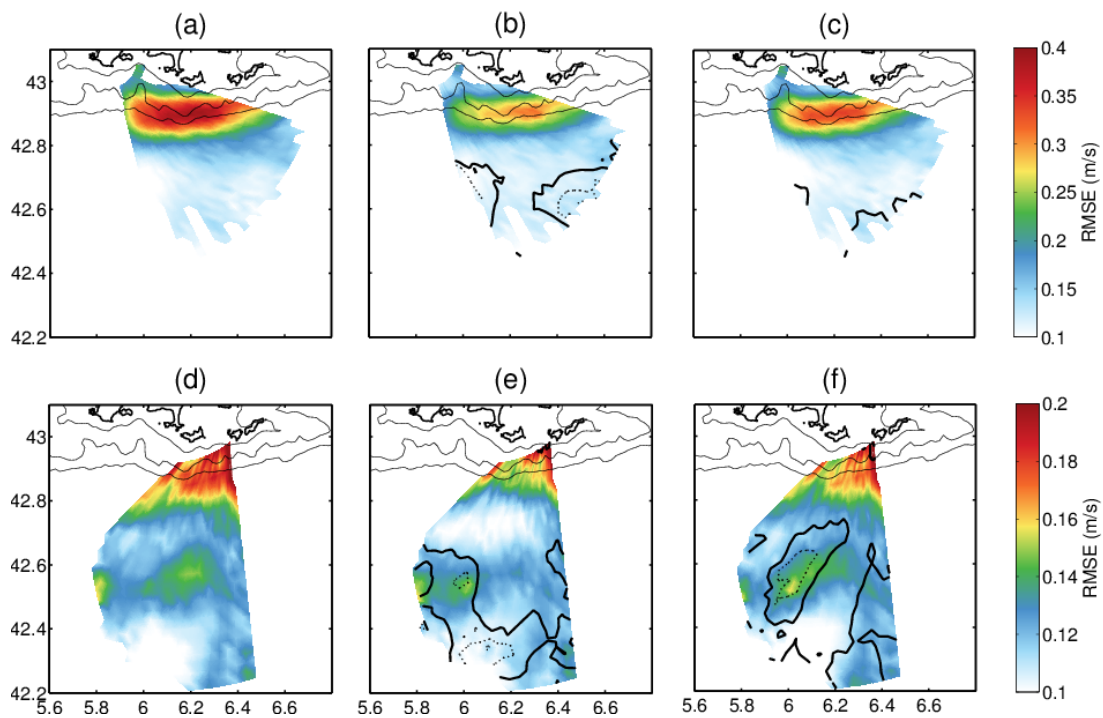


Figure 9. Spatial distributions of (a) Er^b (y_{peg}), (b) Er^a (y_{peg}) for *WindReal* and (c) Er^a (y_{peg}) for *ObcReal* (respectively d, e, f for y_{ben}). Δ is indicated by contour lines and only 0 m s^{-1} (bold black lines) and negatives (-0.02 and -0.01 m s^{-1} , dashed black lines) values are shown. Colorbars differ for PEY and BEN. The 100, 1000 and 2000 m isobaths are drawn (black thin lines).

not degrade the model solution in the observation area, the question of its impact off the OD, and in the whole water column, in terms of hydrological variables, arises. Temperature and salinity from CTD measurements and free/analyzed simulations are compared. The wind optimization method employed only has a weak effect on the temperature ($B = 0.017^\circ$, $Er = 0.818^\circ$) and salinity ($B = -0.049$, $Er = 0.081$) profiles. No significant changes in correlation be-

tween observed and free or *WindReal* analyzed temperature and salinity are noted. This result is not surprising, as the primary wind effect on the ocean dynamics is surface current response, and only for longer timescales should the turbulent processes transfer the surface instant impact to deeper layers. However, assimilation of radial velocities may not enable the corresponding wind forcing to impact the whole water column or any hydrological variables.

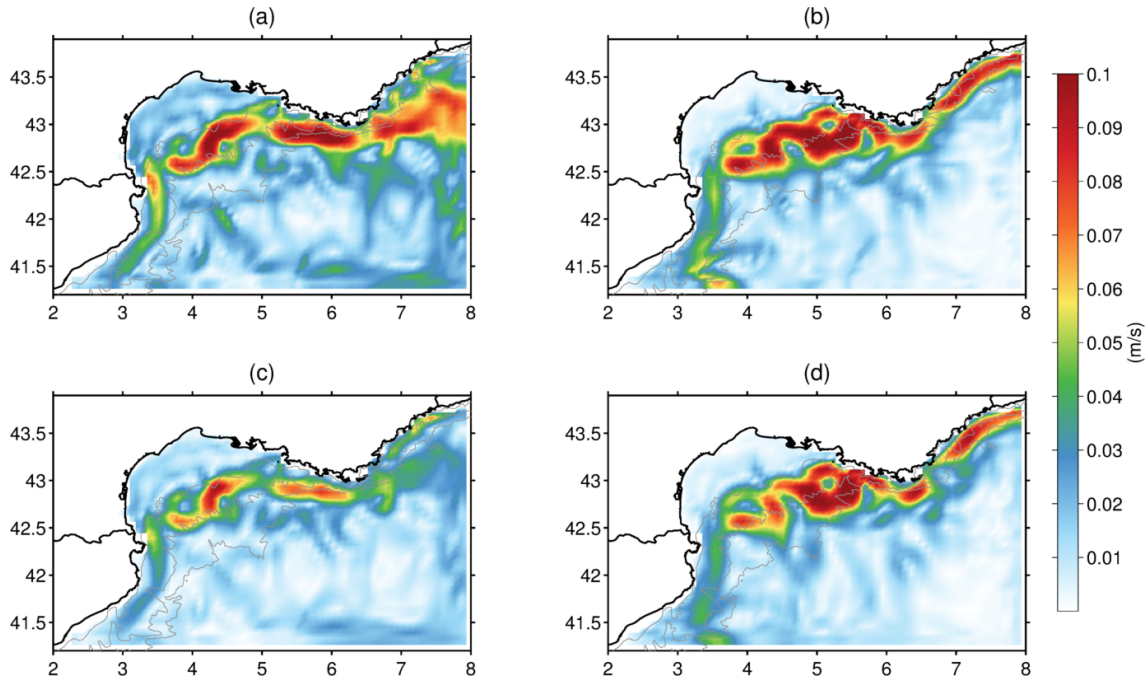


Figure 10. Vector speed difference computed over the assimilation period between free surface current and analyzed surface current from (a) *WindReal* or (b) *ObcReal*, and free current and analyzed current integrated over the first 200 m below the surface from (c) *WindReal* or (d) *ObcReal*. The 100, 1000 and 2000 m isobaths are drawn (gray lines).

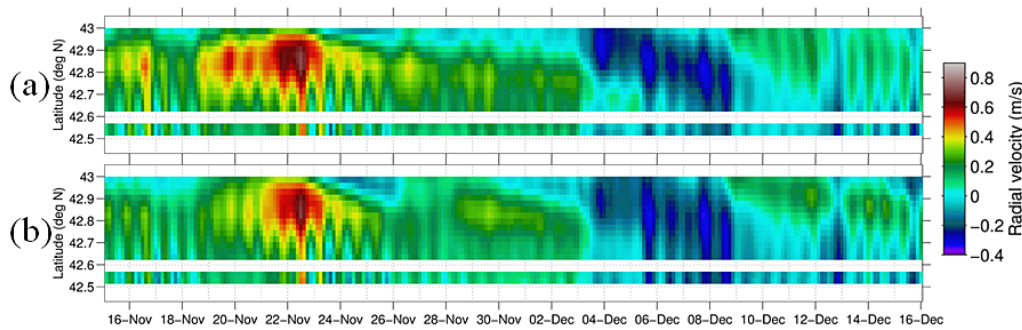


Figure 11. Hovmöller diagram of radial velocities from the PEY station along a south–north transect at 6.2° N (see Fig. 1) (a) from *WindReal* and (b) *ObcReal* analysis.

5 Impact of the OBC optimization

5.1 Results of a twin experiment

An attempt to correct the OBC values using the same method as for the wind correction is evaluated first on the TE. These experiments are performed with 1 day-smoothed atmospheric forcing to weaken the influence of the high-frequency wind variability on surface currents and focus on the OBC influence only.

The method requires the tuning of the same parameters as for wind optimization: α , S_{obs} , win_{ana} , win_{obs}^+ , win_{obs}^- . The reference control state is a member of the ensemble with $\alpha = 1$. After several numerical tests, the best compromise

in terms of the overall Er and Sc was found with $\alpha = 1.5$, $S_{obs} = 2 \text{ m s}^{-1}$, $win_{ana} = 1 \text{ day}$, $win_{obs}^- = 0 \text{ day}$ and $win_{obs}^+ = 15 \text{ days}$. The experiment performed with this data set is hereafter referred to as *ObcTE*.

It should be noted that the choice $win_{obs}^{+,-}$ is crucial. As the coastal circulation flows westward, win_{obs}^+ allows the temporal localization to take into account the delay necessary for propagating OBC information up to the observation area. This delay is about 15 days for an advection propagation of 25 cm s^{-1} (e.g., Fig. 1). The nil value obtained for win_{obs}^- denotes that no valuable information propagates from the observation area to the eastern OBC.

The assimilation method leads to a significant improvement in optimized OBC variables U^{obc} and η^{obc} . The method

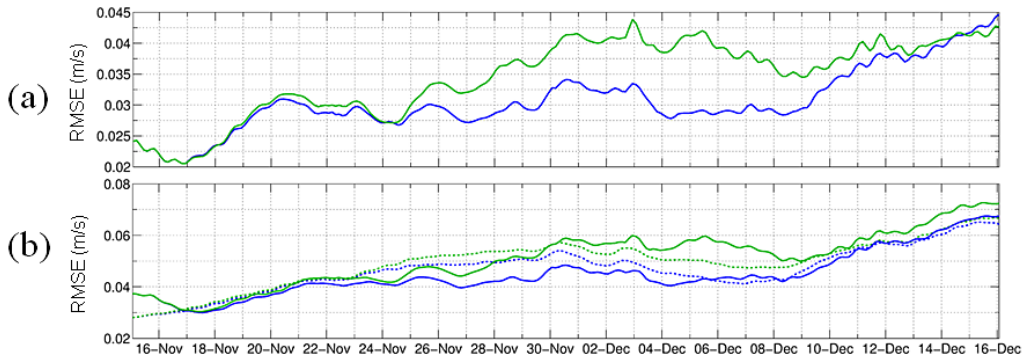


Figure 12. Time evolution of (a) $Er^{a,b}(y)$ and (b) $Er^{a,b}(U)$ computed over the POD (full lines) or the MD (dashed lines) for *ObcTE*. Er^b (Er^a) is in green (blue).



Figure 13. Temporal evolution of (a) $Er^{a,b}(y)$, (b) $Er^{a,b}(y_{pey})$, and (c) $Er^{a,b}(y_{ben})$ for *ObcReal*. Er^b (Er^a) is in green (blue).

Table 2. T^{obc} , S^{obc} , U^{obc} , V^{obc} and η^{obc} statistics computed for *ObcTE* between control and free or analyzed OBC values over the assimilation period. Er unit depends on physical parameters. Sc is without unit.

	Er^b	Er^a	Sc
T^{obc} (°C)	0.147	0.142	0.06
S^{obc}	0.014	0.014	0.06
U^{obc} (ms ⁻¹)	0.022	0.019	0.27
V^{obc} (ms ⁻¹)	0.017	0.019	-0.24
η^{obc} (m)	0.009	0.008	0.20

has only a limited effect on T^{obc} and S^{obc} , while V^{obc} is degraded (Table 2). This suggests that zonal velocity and surface elevation at the eastern boundary are the key parameters controlling model surface currents in the POD. The first parameter imposes the fluxes at the entrance of the MD, whereas the second parameter governs the geostrophic component of the NC, which is the main large-scale feature of the region (e.g., Birol et al., 2010). The degradation of the

meridional velocity reflects the weak correlation that exists between surface currents in the POD and V^{obc} . The spatial pattern of $Er_K^{a,b}(n)$ at the eastern boundary of the model shows that the maximum improvement for U^{obc} and η^{obc} occurs north of 43.5° N (not shown).

In the POD, the optimized OBC values improve the model radial velocity (Fig. 12a). From 15 November to 18 November 2011, no effect on the surface currents is apparent. This corresponds to the time necessary for the OBC information to reach the POD. From 25 November 2011, $Er_N^a(k)$ is smaller than $Er_N^b(k)$, with differences of up to 1 cm s⁻¹, except at the end of the time period. The lower efficiency of assimilation on surface currents from 14 December 2011 is a direct consequence of the degradation of U^{obc} . The overall Sc of radial velocities is 0.22. Similar comments can be made for the vectorial surface currents (Fig. 12b), except that no degradation occurs at the end of the time period.

Despite several regions of degradation, model radial surface currents are globally improved, as well as surface temperature and salinity (not shown).

5.2 Results of a real case experiment

The same free run and the same real radar observations as in *WindReal* are used to correct OBC values. In order to set the optimal value for α , the nearest observations were used to get an estimate of the forcing error relative to its temporal variability at the eastern boundary. α ranged from 0.5 to 1. In this study, the lowest overall Er were obtained for $\alpha = 1$. The value $S_{\text{obs}} = 2 \text{ m s}^{-1}$ was the best compromise for the representativity error. The other parameters are identical to those of the *ObcTE* experiment. The experiment performed with this data set is hereafter referred to as *ObcReal*.

Figure 13 shows the temporal evolution of $\text{Er}_N(k)$ for radial surface currents, for each individual radar site, and for both. When both stations are considered, the analyzed currents better approximate observations, with differences between $\text{Er}_N^a(k)$ and $\text{Er}_N^b(k)$ of up to 7 cm s^{-1} . The adjustment delay at the beginning of the period is needed for the optimized information to be propagated from the eastern boundary to the OD. Statistical analysis reveals that the absolute value of the overall B is reduced by 2.3 cm s^{-1} for the analyzed surface currents compared to the free surface currents (Table 1). As in *WindReal*, greater improvements are obtained for the PEY station (Fig. 13b) and the impact of data assimilation is small on BEN radial velocities (Fig. 13c).

$\text{Er}_K(n)$ maps for PEY and BEN are shown in Fig. 9c, f. Modeled radar currents are mainly improved north of 42.8° N in the NC vein. For PEY, no significant degradation is observed. On the contrary, for BEN, areas of degradation up to 1.6 cm s^{-1} are observed around 42.6° N , which corresponds to the southern edge of the NC. The analysis of vectorial surface currents (not shown here) shows that, using the optimized OBC, the NC vein is more marked due to higher surface velocities along the whole NC pathway.

Hovmöller diagrams of free (Fig. 2a) and analyzed (Fig. 11b) radial velocities from the PEY station show that the use of optimized OBC values has a similar effect on radial velocities to the use of optimized wind fields but to a lesser extent.

The structure of corrections is investigated by computing the norm of $U^b - U^a$ averaged over the assimilation period, at the surface (Fig. 10b) and integrated over the first 200 m below the surface (Fig. 10d). It appears that corrections are primarily located in the NC vein all along the French coast and over the GoL shelf break and affect a large part of the water column. This result proves clearly that the NC strongly depends on its upstream dynamic.

The effect of OBC optimization on the water column is also studied using CTD measurements. Correlations between observed and free or *OBCreal* analyzed temperature and salinity remain unchanged. A bias difference between *ObcReal* analyzed and free runs of 0.12° C is observed in temperature values, which indicates that the effect of assimilation on the water column is more important than in *Win-*

dReal. A slight degradation is obtained for salinity values, with a bias difference of -0.006 .

6 Conclusions

For the first time, HF radar measurements in the NWM Sea are assimilated into a circulation model. The region is characterized by a steep topography and complex coastline, and its ocean dynamics includes a geostrophic coastal current associated with strong mesoscale activity. The EnPS is used to correct two different sources of error acting on the model surface current: high-resolution wind forcing and baroclinic open boundary conditions.

The performance of the assimilation method can be assessed by evaluating either the optimized forcing or the resulting model response. Since no independent wind measurements were available, the assessment of the wind optimization could be done only on the TEs, which clearly demonstrate the strong non-local impact of this method. A limited area of HF radar observations can be effective in large-scale wind correction, and it can reasonably be inferred that the resulting model response will be improved over the whole area. This method also improves surface current simulations through OBC optimization. This result is non-trivial because the OBC optimized area is located far from the observation area (150 km). The perturbation of OBC is done in the NC vein area and takes into account uncertainties regarding width, depth and speed of the NC vein since all variables are perturbed.

The results show that the correction of the wind forcing has a stronger impact on the resulting simulated surface current than the OBC forcing correction. *WindReal* Er is 1 cm s^{-1} less than the *ObcReal* one, and in both cases, bias is reduced. Wind optimization effects are felt in surface currents over short time intervals contrarily to OBC optimization, whose improved values are propagated by model equations over time and space. In terms of temperature and salinity along the water column, mixed results are achieved. No significant improvements are obtained and areas of degradations can occur in temperature values for *WindReal* or salinity values for *ObcReal*.

The assimilation of HF radar-derived radial velocities acts on model surface currents over the whole of the MD. The main corrections are depicted by red values in Fig. 10, which represent the norm of $U^b - U^a$ averaged over the assimilation period. These corrections are primarily located in the NC vein all along the French coast and over the GoL shelf break. *WindReal* corrections are less significant off Nice but occur over the whole of the MD. When the same vector speed difference is applied to the 200 m-depth averaged currents, the correction introduced by *ObcReal* can affect model currents of the deeper water masses, whereas the correction introduced by *WindReal* mainly concerns model surface currents.

These results inspire confidence in the implementation of the method on OBC forcing optimization. When dealing with a domain with large open boundaries that strongly influences internal ocean dynamics, a small inconsistency from the outside will certainly trigger unrealistic features in the domain of interest. The availability of local monitoring systems in the area of interest could be used with this method, to correct the OBC input information, in terms of intensity or positioning of incoming/outcoming currents. In addition, the monitoring of the region off Toulon can be of value for the comprehension of the NWM sea circulation, in particular when considering the NC variability, which is still poorly documented. The recent development of HF radars on the Mediterranean coast through national and international projects should be extended to enhance the coastal monitoring and the assimilation of HF radar data.

The methodology was successfully applied and differences between observed and modeled surface currents were significantly reduced. However, further improvements can be envisaged by taking more aspects into consideration. Wind and OBC optimization have been tested independently and the impact on the model by simultaneous forcing optimizations should be considered in future work, which would require the use of a more realistic eddy-resolving model. Only the eastern boundary was addressed here, and the complementary treatment of the southern boundary may also improve the overall result. Furthermore, the regional model used in this study is quite coarse and may contain errors not directly linked to the forcing: for instance, the bathymetry and the coastline contour may not be well resolved, or the parametrization of the OBC treatment or the turbulent schemes may not be accurate enough when considering the assimilation of high-resolution surface currents derived from HF radar. The use of a high-resolution regional model configuration with an advanced physical parametrization (Ourmières et al., 2011; Guihou et al., 2013) could also lead to significant improvements.

Acknowledgements. The experiment at sea and the HF radar installation were supported by ECCOP-LEFE/IDAO (CNRS/INSU) and TOSCA (2G-MED09-425). The TOSCA project is cofinanced by the European Regional Development Fund, in the framework of the MED program. This study is also supported by LEFE-INSU GMMC. The calculations were performed using NEC-SX8 from GENCI-IDRIS resources (grant 2012011707). The authors also wish to thank Mercator Ocean and Météo-France for providing numerical data and observations. The MIO RADAR team (Y. Barbin, C. Quentin, J. Gaggelli, P. Guterman) is also acknowledged for its major contribution to radar field campaigns of this study. The authors thank the two anonymous referees for their fruitful comments.

Edited by: G. Desroziers

Reviewed by: two anonymous referees



The publication of this article is financed by CNRS-INSU.

Appendix A: Ensemble Perturbation Smoother

EnPS was proposed by Barth et al. (2010) to optimize tidal boundary conditions and was successfully applied by Barth et al. (2011) to optimize surface wind forcing. Contrary to the classical sequential method, which intends to derive the best estimation of the model state, this method aims at optimizing forcing values. These forcing fields are then run once again in the model to get the final model solution. Thus, the model trajectory is corrected instead of the model state as in the ensemble smoother (Van Leeuwen, 2001), the 4D-LETKF (Hunt et al., 2004, 2007) or the AEnKF (Sakov et al., 2010). This method also ensures that the model solution satisfies the model equations exactly.

This non-sequential method is derived from the sequential EnKF scheme by including the time dimension in the observation vector (Sakov et al., 2010). Consequently, all observations in the optimization period are gathered in a single observation vector \mathbf{y}^o , with an observation error covariance matrix \mathbf{R} .

The optimization vector \mathbf{x} , called the state vector in the classical method, contains all forcing and model parameters to be improved.

An ensemble of forcing \mathbf{x}^l , where l is the index of a member in the ensemble, is generated by adding a perturbation to the background (or original) forcing \mathbf{x}^b . If perturbed forcings are time dependent, their temporal evolution is gathered in a single vector, in the same way as observations.

For each perturbed forcing, the model is run. Note that at this step ensemble members are not influenced by observations. For each ensemble member, the observed part of the model state $h(\mathbf{x}^l)$ is extracted using a non-linear observation operator $h(\cdot)$. Consequently, model dynamic is included in the observation operator. Each element of the ensemble $h(\mathbf{x}^l)$ can be directly compared to the corresponding one in the observation vector \mathbf{y}^o .

Error covariance matrices can be estimated from the use of an ensemble of forcing state and model state. Using the same notation as in Barth et al. (2010, 2011), we define the deviation of the ensemble members around the ensemble mean of forcing \mathbf{S} and observed part of surface currents \mathbf{E} as

$$(\mathbf{S})_l = (L - 1)^{-\frac{1}{2}} (\mathbf{x}^{(l)} - \langle \mathbf{x} \rangle) \tag{A1}$$

$$(\mathbf{E})_l = (L - 1)^{-\frac{1}{2}} (h(\mathbf{x}^{(l)}) - \langle h(\mathbf{x}) \rangle) \tag{A2}$$

where L is the number of ensemble members, the index l refers to the ensemble member ($l = 1, \dots, N$) and $\langle \cdot \rangle$ is the ensemble average operator. These matrices are scaled in such a way that \mathbf{SE}^T and \mathbf{EE}^T represent, respectively, ensemble forcing covariance and covariance between ensemble forcing and extracted observations:

$$\mathbf{SE}^T = \text{cov}(\mathbf{x}^b, h(\mathbf{x}^b)) \tag{A3}$$

$$\mathbf{EE}^T = \text{cov}(h(\mathbf{x}^b), h(\mathbf{x}^b)) \tag{A4}$$

Note that there is no need to form explicitly and to store the covariance matrices, because analysis is computed in the ensemble space (e.g., Nerger et al., 2005).

The optimal forcing \mathbf{x}^a that minimizes the analysis error is computed using Eq. (2), which is the Kalman filter analysis step with a non-linear observation operator (Chen and Snyder, 2007). It can be developed as

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{SE}^T (\mathbf{EE}^T + \mathbf{R})^{-1} (\mathbf{y}^o - h(\mathbf{x}^b)) \tag{A5}$$

or using the Sherman–Morrison–Woodbury formula:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{S} (\mathbf{E}^T \mathbf{R}^{-1} \mathbf{E} + 1)^{-1} \mathbf{E}^T \mathbf{R}^{-1} (\mathbf{y}^o - h(\mathbf{x}^b)) \tag{A6}$$

The superscripts “a” and “b” refer to the analysis and the background estimates. As in Kalman filtering, background and observation errors are assumed to be unbiased, but in the case of a strongly non-linear observation operator, such assumptions are not sufficient to ensure that the analysis \mathbf{x}^a is unbiased. However, the analysis is unbiased if Eq. (3) is verified (Barth et al., 2011).

The model trajectory is corrected in a final step that consists in running the model with the optimized forcing \mathbf{x}^a . Consequently, observations influence the analyzed model solution using an optimal combination of ensemble forcing.

References

- Alberola, C., Rousseau, S., Millot, C., Astraldi, M., Font, J., Garcia-lafuente, J., Gasparini, G., Send, U., and Vangriesheim, A.: Tidal currents in the western mediterranean-sea, *Oceanol. Acta*, 18, 273–284, 1995.
- Allou, A., Forget, P., and Devenon, J.-L.: Submesoscale vortex structures at the entrance of the Gulf of Lions in the Northwestern Mediterranean Sea, *Cont. Shelf Res.*, 30, 724–732, 2010.
- Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: Assimilation of high-frequency radar currents in a nested model of the West Florida Shelf, *J. Geophys. Res.-Oceans* (1978–2012), 113, C08033, doi:10.1029/2007JC004585, 2008a.
- Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: A nested model study of the Loop Current generated variability and its impact on the West Florida Shelf, *J. Geophys. Res.*, 113, C05009, doi:10.1029/2007JC004492, 2008b.
- Barth, A., Alvera-Azcárate, A., Gurgel, K.-W., Staneva, J., Port, A., Beckers, J.-M., and Stanev, E. V.: Ensemble perturbation smoother for optimizing tidal boundary conditions by assimilation of High-Frequency radar surface currents-application to the German Bight, *Ocean Sci.*, 6, 161–178, doi:10.5194/os-6-161-2010, 2010.
- Barth, A., Alvera-Azcárate, A., Beckers, J.-M., Staneva, J., Stanev, E. V., and Schulz-Stellenfleth, J.: Correcting surface winds by assimilating High-Frequency Radar surface currents in the German Bight, *Ocean Dynam.*, 61, 599–610, 2011.
- Birol, F., Cancet, M., and Estournel, C.: Aspects of the seasonal variability of the Northern Current (NW Mediterranean Sea) observed by altimetry, *J. Mar. Syst.*, 81, 297–311, 2010.
- Blayo, E. and Debreu, L.: Revisiting open boundary conditions from the point of view of characteristic variables, *Ocean Modell.*, 9, 231–252, 2005.
- Bogden, P. S., Malanotte-Rizzoli, P., and Signell, R.: Open-ocean boundary conditions from interior data: Local and remote forcing of Massachusetts Bay, *J. Geophys. Res.-Oceans* (1978–2012), 101, 6487–6500, 1996.
- Brankart, J.-M. and Brasseur, P.: Optimal analysis of in situ data in the Western Mediterranean using statistics and cross-validation, *J. Atmos. Ocean. Technol.*, 13, 477–491, 1996.
- Breivik, Ø. and Sætra, Ø.: Real time assimilation of HF radar currents into a coastal ocean model, *J. Mar. Syst.*, 28, 161–182, 2001.
- Capet, X., McWilliams, J., Molemaker, M., and Shchepetkin, A.: Mesoscale to submesoscale transition in the California Current System. Part I: Flow structure, eddy flux, and observational tests, *J. Phys. Oceanogr.*, 38, 29–43, 2008.
- Casella, E., Molcard, A., and Provenzale, A.: Mesoscale vortices in the Ligurian Sea and their effect on coastal upwelling processes, *J. Mar. Syst.*, 88, 12–19, 2011.
- Chapman, D. C.: Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model, *J. Phys. Oceanogr.*, 15, 1060–1075, 1985.
- Chapman, R., Shay, L., Graber, H., Edson, J., Karachintsev, A., Trump, C., and Ross, D.: On the accuracy of HF radar surface current measurements: Intercomparisons with ship-based sensors, *J. Geophys. Res.-Oceans* (1978–2012), 102, 18737–18748, 1997.
- Chen, Y. and Snyder, C.: Assimilating vortex position with an ensemble Kalman filter, *Mon. Weather Rev.*, 135, 1828–1845, 2007.
- Cosme, E., Verron, J., Brasseur, P., Blum, J., and Auroux, D.: Smoothing Problems in a Bayesian Framework and Their Linear Gaussian Solutions., *Mon. Weather Rev.*, 140, 683–695, doi:10.1175/MWR-D-10-05025.1, 2012.
- Echevin, V., Crepon, M., and Mortier, L.: Interaction of a coastal current with a gulf: application to the shelf circulation of the Gulf of Lions in the Mediterranean Sea, *J. Phys. Oceanogr.*, 33, 188–206, 2003.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.-Oceans* (1978–2012), 99, 10143–10162, 1994.
- Fischer, C., Montmerle, T., Berre, L., Auger, L., and Ștefănescu, S. E.: An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system, *Q. J. Roy. Meteorol. Soc.*, 131, 3477–3492, 2005.
- Flather, R.: A tidal model of the northwest European continental shelf, *Mem. Soc. R. Sci. Liege*, 10, 141–164, 1976.
- Gelb, A.: Applied optimal estimation, The MIT press, 1974.
- Guihou, K., Marmain, J., Ourmières, Y., Molcard, A., Zakardjian, B., and Forget, P.: A case study of the mesoscale dynamics in the North-Western Mediterranean Sea: a combined data–model approach, *Ocean Dynam.*, 63, 1–16, doi:10.1007/s10236-013-0619-z, 2013.
- Gunson, J. R. and Malanotte-Rizzoli, P.: Assimilation studies of open-ocean flows: 1. Estimation of initial and boundary conditions, *J. Geophys. Res.-Oceans* (1978–2012), 101, 28457–28472, 1996.
- Gurgel, K.-W., Antonischki, G., Essen, H.-H., and Schlick, T.: Wellen Radar (WERA): a new ground-wave HF radar for ocean remote sensing, *Coast. Eng.*, 37, 219–234, 1999.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001.
- He, R., Liu, Y., and Weisberg, R. H.: Coastal ocean wind fields gauged against the performance of an ocean circulation model, *Geophys. Res. Lett.*, 31, L14303, doi:10.1029/2003GL019261, 2004.
- Hoteit, I., Cornuelle, B., Kim, S., Forget, G., Köhl, A., and Terrill, E.: Assessing 4D-VAR for dynamical mapping of coastal high-frequency radar in San Diego, *Dynam. Atmos. Ocean.*, 48, 175–197, 2009.
- Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, 129, 123–137, 2001.
- Hunt, B., Kalnay, E., Kostelich, E., Ott, E., Patil, D., Sauer, T., Szunyogh, I., Yorke, J., and Zimin, A.: Four-dimensional ensemble Kalman filtering, *Tellus A*, 56, 273–277, 2004.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D: Nonlinear Phenomena*, 230, 112–126, 2007.
- Kurapov, A. L., Egbert, G. D., Allen, J., Miller, R. N., Erofeeva, S. Y., and Kosro, P.: The M 2 internal tide off Oregon: Inferences from data assimilation, *J. Phys. Oceanogr.*, 33, 1733–1757, 2003.

- Langlais, C., Barnier, B., and Fraunié, P.: Towards high resolution coastal and shelf circulation modeling: impact of the atmospheric forcing resolution in the Gulf of Lions, Mercator. *Ocean Q. Newsl.*, 32, 5–13, 2009.
- Le Dimet, F.-X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus A*, 38, 97–110, 1986.
- Lewis, J. K., Shulman, I., and Blumberg, A. F.: Assimilation of Doppler radar current data into numerical ocean models, *Cont. Shelf Res.*, 18, 541–559, 1998.
- Lipa, B., Nyden, B., Ullman, D. S., and Terrill, E.: SeaSonde radial velocities: derivation and internal consistency, *Ocean. Eng., IEEE J.*, 31, 850–861, 2006.
- Malanotte-Rizzoli, P., Young, R. E., and Haidvogel, D. B.: Initialization and data assimilation experiments with a primitive equation model, *Dynam. Atmos. Ocean.*, 13, 349–378, 1989.
- Marchesiello, P., McWilliams, J. C., and Shchepetkin, A.: Open boundary conditions for long-term integration of regional oceanic models, *Ocean Modell.*, 3, 1–20, 2001.
- Marmain, J., Forget, P., and Molcard, A.: Characterization of ocean surface current properties from single site HF/VHF radar, *Ocean Dynam.*, 61, 1967–1979, 2011.
- Millot, C.: Circulation in the western Mediterranean Sea, *J. Mar. Syst.*, 20, 423–442, 1999.
- Nerger, L., Hiller, W., and Schröter, J.: A comparison of error subspace Kalman filters, *Tellus A*, 57, 715–735, 2005.
- Oke, P. R., Allen, J., Miller, R., Egbert, G., Austin, J., Barth, J., Boyd, T., Kosro, P., and Levine, M.: A modeling study of the three-dimensional continental shelf circulation off Oregon. Part I: Model-data comparisons, *J. Phys. Oceanogr.*, 32, 1360–1382, 2002a.
- Oke, P. R., Allen, J. S., Miller, R. N., Egbert, G. D., and Kosro, P. M.: Assimilation of surface velocity data into a primitive equation coastal ocean model, *J. Geophys. Res.*, 107, 3122, doi:10.1029/2000JC000511, 2002b.
- Orlanski, I.: A simple boundary condition for unbounded hyperbolic flows, *J. Comput. Phys.*, 21, 251–269, 1976.
- Ourmières, Y., Zakardjian, B., Béranger, K., and Langlais, C.: Assessment of a NEMO-based downscaling experiment for the North-Western Mediterranean region: Impacts on the Northern Current and comparison with ADCP data and altimetry products, *Ocean Modell.*, 39, 386–404, 2011.
- Paduan, J. D. and Shulman, I.: HF radar data assimilation in the Monterey Bay area, *J. Geophys. Res.-Oceans (1978–2012)*, 109, C07S09, doi:10.1029/2003JC001949, 2004.
- Pairaud, I., Gatti, J., Bensoussan, N., Verney, R., and Garreau, P.: Hydrology and circulation in a coastal area off Marseille: Validation of a nested 3D model with observations, *J. Mar. Syst.*, 88, 20–33, 2011.
- Peng, S., Li, Y., and Xie, L.: Adjusting the wind stress drag coefficient in storm surge forecasting using an adjoint technique, *J. Atmos. Ocean. Technol.*, 30, 590–608, doi:10.1175/JTECH-D-12-00034.1, 2012.
- Penven, P., Roy, C., Brundrit, G., De Verdière, A. C., Fréon, P., Johnson, A., Lutjeharms, J., and Shillington, F.: A regional hydrodynamic model of upwelling in the Southern Benguela, *South Afr. J. Sci.*, 97, 472–475, 2001.
- Petrenko, A. A.: Variability of circulation features in the Gulf of Lion NW Mediterranean Sea. Importance of inertial currents, *Oceanol. Acta*, 26, 323–338, 2003.
- Røed, L. and Cooper, C.: Open boundary conditions in numerical ocean models, in: *Advanced physical oceanographic numerical modelling*, 411–436, Springer, 1986.
- Sakov, P., Evensen, G., and Bertino, L.: Asynchronous data assimilation with the EnKF, *Tellus A*, 62, 24–29, 2010.
- Sammari, C., Millot, C., and Prieur, L.: Aspects of the seasonal and mesoscale variabilities of the Northern Current in the western Mediterranean Sea inferred from the PROLIG-2 and PROS-6 experiments, *Deep. Sea Res. Part I*, 42, 893–917, 1995.
- Schaeffer, A., Garreau, P., Molcard, A., Fraunié, P., and Seity, Y.: Influence of high-resolution wind forcing on hydrodynamic modeling of the Gulf of Lions, *Ocean Dynam.*, 61, 1823–1844, 2011a.
- Schaeffer, A., Molcard, A., Forget, P., Fraunié, P., and Garreau, P.: Generation mechanisms for mesoscale eddies in the Gulf of Lions: radar observation and modeling, *Ocean Dynam.*, 61, 1587–1609, 2011b.
- Schmidt, R.: Multiple emitter location and signal parameter estimation, *Antennas and Propagation, IEEE Trans.*, 34, 276–280, 1986.
- Seiler, U.: Estimation of open boundary conditions with the adjoint method, *J. Geophys. Res.-Oceans (1978–2012)*, 98, 22855–22870, 1993.
- Sentchev, A., Forget, P., Barbin, Y., and Yaremchuk, M.: Surface circulation in the Iroise Sea (W. Brittany) from high resolution HF radar mapping, *J. Mar. Syst.*, 109–110, S153–S168, doi:10.1016/j.jmarsys.2011.11.024, 2013.
- Shchepetkin, A. F. and McWilliams, J. C.: A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate, *J. Geophys. Res.-Oceans (1978–2012)*, 108, 3090, doi:10.1029/2001JC001047, 2003.
- Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, *Ocean Modell.*, 9, 347–404, 2005.
- Shulman, I. and Paduan, J. D.: Assimilation of HF radar-derived radials and total currents in the Monterey Bay area, *Deep. Sea Res. Part*, 56, 149–160, 2009.
- Shulman, I., Lewis, J. K., and Mayer, J. G.: Local data assimilation in the estimation of barotropic and baroclinic open boundary conditions, *J. Geophys. Res.-Oceans (1978–2012)*, 104, 13667–13680, 1999.
- Stewart, R. H. and Joy, J. W.: HF radio measurements of surface currents, *Deep Sea Res.*, 21, 1039–1049, 1974.
- Taillandier, V., Echevin, V., Mortier, L., and Devenon, J.-L.: Controlling boundary conditions with a four-dimensional variational data-assimilation method in a non-stratified open coastal model, *Ocean Dynam.*, 54, 284–298, 2004.
- Taillandier, V., Echevin, V., Mortier, L., and Devenon, J.-L.: Controlling the open ocean forcing of baroclinic coastal jets with a 4D-var method, *Dynam. Atmos. Ocean.*, 48, 155–174, 2009.
- Ten Brummelhuis, P., Heemink, A., and Van Den Boogaard, H.: Identification of shallow sea models, *Int. J. Num. Method. Fluids*, 17, 637–665, 1993.
- Van Leeuwen, P. J.: An ensemble smoother with error estimates, *Mon. Weather Rev.*, 129, 709–728, 2001.