Ocean Dynamics (2009) 59:83–95
DOI 10.1007/s10236-008-0167-0

High-resolution ensemble forecasting for the Gulf of Mexico eddies and fronts

François Counillon · Laurent Bertino

Received: 10 March 2008 / Accepted: 7 November 2008 / Published online: 6 December 2008

Abstract High-resolution models and realistic boundary conditions are necessary to reproduce the mesoscale dynamics of the Gulf of Mexico (GOM). In order to achieve this, we use a nested configuration of the Hybrid Coordinate Ocean Model (HYCOM), where the Atlantic TOPAZ system provides lateral boundary conditions to a high-resolution (5 km) model of the GOM. However, such models cannot provide accurate forecasts of mesoscale variability, such as eddy shedding event, without data assimilation. Eddy shedding events involve the rapid growth of nonlinear instabilities that are difficult to forecast. The known sources of error are the initial state, the atmospheric condition, and the lateral boundary condition. We present here the benefit of using a small ensemble forecast (10 members) for providing confidence indices for the prediction, while using a data assimilation scheme based on optimal interpolation. Our set of initial states is provided by using different values of a data assimilation parameter, while the atmospheric and lateral boundary conditions are perturbed randomly. Changes in the data assimilation parameter appear to control the main position of the large features of the GOM in the initial state, whereas changes in the boundary conditions (lateral and atmospheric) appears to control the propagation of cyclonic eddies at their boundary. The ensemble forecast is tested for the shedding of Eddy Yankee (2006). The Loop Current and eddy fronts observed from ocean color and altimetry are almost always within the estimated positions from the ensemble forecast. The ensemble spread is correlated both in space and time to the forecast error, which implies that confidence indices can be provided in addition to the forecast. Finally, the ensemble forecast permits the optimization of a data assimilation parameter for best performance at a given forecast horizon.

Keywords Ensemble forecasting · Ensemble prediction system (EPS) · Lateral boundary perturbation · Gulf of Mexico · Eddy resolving · EnOI

Responsible Editor: George Robert Halliwell

F. Counillon (✉) · L. Bertino
Mohn Sverdrup Center/Nansen Environmental and Remote Sensing Center Thormøhlensgate 47, Bergen, Norway
e-mail: francois.counillon@nersc.no

1 Introduction

The dynamics in the Gulf of Mexico (GOM) are dominated by the powerful northward Yucatan Current flowing into a semienclosed basin. This current forms a loop, called the Loop Current (LC) that exits through the Florida Straits, and in turn becomes the Gulf Stream. At irregular intervals (Vukovich 1988; Sturges and Leben 2000), the LC sheds large eddies that propagate westward across the GOM. As oil production moves further into deeper waters, the costs related to strong current hazards are increasing accordingly, and accurate 3-dimensional forecasts of currents are needed. High-resolution models are necessary to reproduce the dynamics of the area and their variability (Chassignet et al. 2005), and a horizontal resolution of 5 km appears to be sufficiently high to resolve the mesoscale features, such as the eddy shedding, considering the first mode (baroclinic) of the Rossby radius ($R_o \simeq 30$ km in the area; Oey et al. 2005a).

However, such models cannot provide accurate forecasts of mesoscale current features because the initial state is unknown, the model is imperfect, and the input data is in error. Data assimilation methods address this issue and estimate the optimal initial state considering the model, the available observations, and their respective uncertainty. The widely furnished altimetry data set represents the eddy front and is used here. The computational cost of high-resolution models has so far constrained the choice of data assimilation methods to simple optimal interpolation based schemes (Oey et al. 2005b; Chassignet et al. 2005). At the time of writing the article, we could afford the model integration of 10 members in real-time at such resolution. A 10-member ensemble is still too small for using the ensemble Kalman filter (EnKF, Evensen 2006), which usually requires $\mathcal{O}(100)$ dynamical members for oceanic application (Natvik and Evensen 2003). Therefore, we perform data assimilation with the computationally cheap ensemble optimal interpolation (EnOI, see, e.g., Oke et al. 2002; Evensen 2003), which allows for 3-dimensional multivariate update and appears as relatively suited for assimilation of altimetry data in the GOM (Counillon and Bertino 2008). We then analyze the capacity of a 10-dynamical-members ensemble for the purpose of ensemble prediction system (EPS) as often done in atmospheric applications (Molteni et al. 1996).

The benefit of EPS is twofold: the ensemble mean provides a more accurate prediction than each ensemble member separately; the ensemble spread can be used for providing confidence indices for the prediction. High-resolution ocean EPS is very recent. Yin and Oey (2007) use a high-resolution bred-ensemble forecast with perturbations of the initial state in the GOM. They found that the ensemble mean provides a closer agreement to the observations than a conventional single forecast. We analyzed here the capacity of the ensemble spread for providing confidence indices for the prediction. In order to achieve this, the correct sources of error should be perturbed and the ensemble should be sufficiently large. In Counillon and Bertino (2008), it was shown that the efficiency of the EnOI is sensitive to a data assimilation parameter, called α. We therefore use different values of α for generating a set of initial states. In Oey et al. (2003), the boundary conditions (lateral and atmospheric) appear to influence the LC stability, and recently, Lugo-Fernández (2007) showed that the eddy shedding period is nonlinearly dependent on the initial state, the atmospheric perturbation, and the lateral boundary conditions. Our EPS thus uses different values of α as a proxy for the perturbation of the initial state and random perturbations of the latter two sources of error.

The outline of this paper is as follows. Section 2 presents the data assimilative system. Section 3 analyzes the model sensitivity to each of the error sources considered, with focus on the spatial scale of the anomalies generated, their amplitudes, and growth rate. Section 4.1 presents a sensitivity study for a data assimilation parameter and examines the Gaussian properties of a 10-member perturbed ensemble. Section 4.2 compares the 10-member ensemble front position to ocean color data and to the front calculated from altimetry data for the Eddy Yankee shedding event (2006). Section 4.3 investigates the benefits of the 10-member ensemble by evaluating the spread–skill correlation both in space and time, for sea surface height (SSH). Conclusions are given in Section 5.

2 The data assimilative system

A DA system provides an optimal model state, given a dynamic model and a set of measurements, and their respective error statistics. The circulation in the GOM is mainly quasigeostrophic, and the dynamics provide clear SSH signals. Furthermore, near real-time altimetry observations of the ocean are achieved by combining data from different satellites.

The sea-level anomaly (SLA) data used for assimilation are the maps provided in near real-time by SSALTO/DUACS on a 1/3° Mercator grid (Le Traon et al. 2003). The standard deviation of the measurements is assumed to be constant, and it is using the average value specified by the provider in the GOM area (3 cm). The measurements are less accurate in the coastal area; therefore, measurements are selected only in regions deeper than 300 m, which correspond in the GOM to distances of at least 50 km from the coast. Accordingly, a Gaussian covariance with a decorrelation radius of 50 km is used for the observation error. Observations near the model boundary are not assimilated. Observations are assimilated weekly.

2.1 Data assimilation

The data assimilation problem consists of accommodating a dynamical model with measurements considering their respective error statistics. We use the EnOI, a data assimilation method based on the EnKF, which uses a stationary ensemble composed of model states as a square root representation of the covariance matrix. As a consequence, this method is computationally

cheap, but is still 3-dimensional and multivariate and conserves the linear properties such as the geostrophic balance (Oke et al. 2002; Evensen 2006). The EnOI analysis is computed in Eq. 1.

$$\boldsymbol{\psi}^{\mathrm{a}} = \boldsymbol{\psi}^{\mathrm{f}} + \alpha \mathbf{A}'\mathbf{A}'^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\left(\alpha \mathbf{H}\mathbf{A}'\mathbf{A}'^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathrm{N}_S\mathbf{R}\right)^{-1} \times (\mathbf{d} - \mathbf{H}\boldsymbol{\psi}^{\mathrm{f}}). \quad (1)$$

The matrix **A** represents a large historical ensemble composed of model states sampled over a long time integration. Here, we have used 2.5 years of weekly model output (122 members in total). This ensemble is kept unchanged through the assimilation cycle and is referred as "static ensemble." $\mathbf{A}'$ is the centered historical ensemble (i.e., $\mathbf{A}' = \mathbf{A} - \overline{\mathbf{A}}$), where the overbar denotes ensemble averaging. $\boldsymbol{\psi}^{\mathrm{a}}$ and $\boldsymbol{\psi}^{\mathrm{f}}$ are the model analysis and forecast states, **d** is a vector of measurements, **R** is the measurement error covariance matrix, N_S is the size of the static ensemble, and **H** is the measurement operator relating the prognostic model state to the measurements. An ensemble of model states sampled over a long time period may have a variance that is inadequate to represent the instantaneous forecast error variance, so that a factor α is introduced to rescale it.

The EnOI assumes that the temporal variability is representative of the instantaneous forecast error. This assumption is analyzed in detail in Counillon and Bertino (2008). The EnOI appears to be relatively suitable for assimilation of altimetry in the GOM but shows some limitations that are circumvented using localization.

2.2 The nested model system

The forward model skill is very important in a DA system because it contributes to the error growth from the initial state, and even more so in the EnOI since the error statistics depend on its ability to reproduce the dynamics. Chassignet et al. (2005) demonstrate the skill of HYCOM for the GOM and highlight the importance of horizontal resolution for the representation of mesoscale dynamics. The inflow through the Yucatan Straits also has a strong influence on the northward penetration of the LC; on its stability after separation from the Campeche Bank; and, therefore, on the timing of the shedding event (Oey et al. 2003; Abascal et al. 2003). A nested configuration can satisfy these two requirements using reasonable computing facilities. A coarser Atlantic system (TOPAZ) provides lateral boundary conditions to a high-resolution model of the GOM (Fig. 1) using lateral boundary techniques described in Browning and Kreiss (1982). For the barotropic components (velocities and pressure), the boundary conditions are computed exactly, while taking into consideration both the waves propagating into the regional model from the external solution and the waves propagating out through the boundary from the regional model. For the slow varying variables, i.e., baroclinic velocities, temperature, salinity, and layer interface, a simple relaxation technique is used. This constitutes the standard nesting procedure with HYCOM, with an additional horizontal interpolation to the nested model grid.

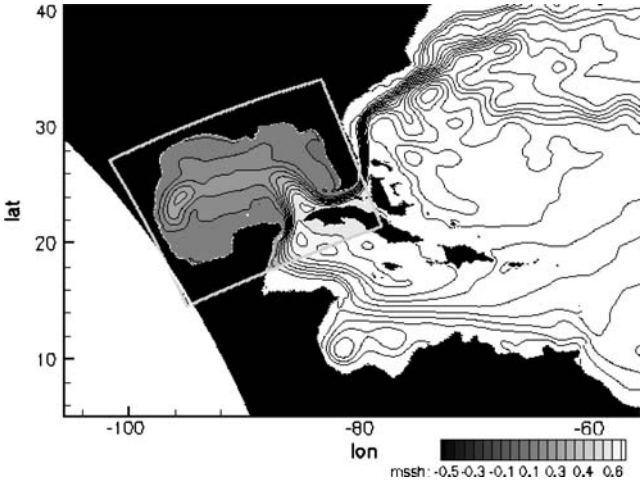

Fig. 1 Mean SSH of TOPAZ3 interpolated into the high-resolution local model grid delimited by the *gray box*

TOPAZ is a real-time forecasting system for the Atlantic and Arctic basins using HYCOM (Evensen 2006, Chapter 15; see also http://topaz.nersc.no/), which uses advanced data assimilation techniques (i.e., EnKF). Due to a reduced inflow in the GOM in the current operational forecasting system TOPAZ2, we use here TOPAZ3 prototype. This prototype greatly improves the accuracy of the boundary condition but does not include data assimilation. The TOPAZ3 model grid has a horizontal resolution between 11 and 16 km (approximately 1/8°) created using a conformal mapping of the poles to two new locations by the algorithm outlined in Bentsen et al. (1999). TOPAZ3 is initialized from the GDEM3 climatology (Teague et al. 1990) and spun up for 16 years. The TOPAZ3 system transports 19.5 Sv into the GOM, instead of the 23.8 Sv measured in CANEK program during the same 10-month period (Sheinbaum et al. 2002). Note that TOPAZ3 9-year average net transport is 22 Sv.

Our high-resolution model is set-up with a 5-km horizontal resolution, which is sufficient to resolve the

features such as mesoscale eddy considering the first-mode (baroclinic) Rossby radius (30–40 km). It uses a fourth-order numerical scheme for treating the advection of momentum in the primitive equations (Winther et al. 2007). To minimize the necessary spin-up time, the initial state is interpolated from an equilibrated state of TOPAZ3 and spun up for 3 years.

In HYCOM, the vertical coordinates are isopycnal in the open, stratified ocean, but smoothly revert to z-level coordinates in the mixed layer and/or unstratified seas (Bleck 2002). Both models use 22 hybrid layers with the minimum thickness of the top layer of 3 m. The bathymetry is specified using the General Bathymetric Chart of the Oceans database with 1' resolution, interpolated to the model grid. The models are forced by the 6-hourly and 0.5° analyzed fields from the European Center for Medium-range Weather Forecasting (ECMWF). The models use monthly average river discharge value taken from Dai and Trenberth (2002), Dümenil et al. (1993).

The diagnosed model SSH is the steric height anomaly that varies due to the barotropic pressure mode, the deviations in temperature, and salinity and does not include the inverse barometer effect (atmospheric pressure) for consistency with the SLA measurements. As the SLA needs to be referred to a mean SSH, a 2-year average of TOPAZ3 SSH is interpolated to the high-resolution grid (Fig. 1). It shows a maximum value induced by the resident LC base and a positive track induced by the passage of eddies that drift westward. Qualitatively, it compares well with the mean dynamic topography based on satellite and in-situ measurements (Rio and Hernandez 2004).

3 Model sensitivity to different sources of error

In ensemble forecasting, the ensemble spread can be representative of the forecast error if the correct sources of error are perturbed and the ensemble is sufficiently large. Although the model is able to reproduce the dynamics of the region, it is in error because of an inaccurate initial state and inaccurate lateral/atmospheric boundary conditions. This error grows with the unstable mode of the flow. Lugo-Fernández (2007) shows that, in the GOM, the eddy shedding period is nonlinearly dependent on the initial state, the lateral boundary conditions, and the atmospheric forcing. Therefore, these three sources of error are considered in the following. We outline the method and the assumptions made for simulating each source of error, analyze their response in the model using twin experiment, and quantify their contribution in the ensemble spread by integrating a small ensemble of 10 members.

3.1 The initial state: data assimilation parameter

Oey et al. (2005b) observe that the major source of SSH errors over a 4-week forecast horizon is due to error in the initial state. A common approach to perturb the initial state is to perturb the assimilated measurements. In the EnOI, a static ensemble represents the forecast error and a parameter α is introduced to rescale their variances, see Section 2.1. As the unknown instantaneous forecast error evolves with time, it is unclear which value of α is optimal. Therefore, we select different values of α, reported in Table 1, where lower values result in weaker assimilation and reversely. Too large values of α can initiate noise and perturb the balance of the system, whereas too low values can lead to a loss of accuracy. The sample encompasses the previously estimated optimal value (i.e., ≈ 0.09, see Counillon and Bertino (2008)), and the extreme values of the sample are chosen such that the assimilation produces efficient updates with limited assimilation noise. This approach provides a set of different initial states, and additionally allows for a sensitivity study to this parameter, although the associated changes of the initial state are not strictly random.

To estimate the model sensitivity to the parameter α in the initial perturbation, we compared a free model run to the run with the strongest assimilation ($\alpha = 0.36$) on the 6th of June. Figure 2a is the SSH of the control run on the 4th of July, and Fig. 2b is the SSH anomalies (i.e., perturbed run minus control run) developed after 4 weeks. The anomalies are large in scale and amplitude in the vicinity of the LC and become smaller and weaker in the western GOM. The anomalies obtained here are similar in shape and in size to those obtained in Yin and Oey (2007) with perturbation of the assimilated measurements. The changes in the parameter α initiate changes in the position and orientation of the mesoscale features such as the LC and associated eddies.

3.2 The lateral boundary conditions

Many studies show the importance of the inflow through the Yucatan Straits for the timing of the eddy

Table 1 Value of α used for each member

	Member									
	1	2	3	4	5	6	7	8	9	10
α	0.004	0.014	0.032	0.058	0.09	0.13	0.17	0.23	0.29	0.36

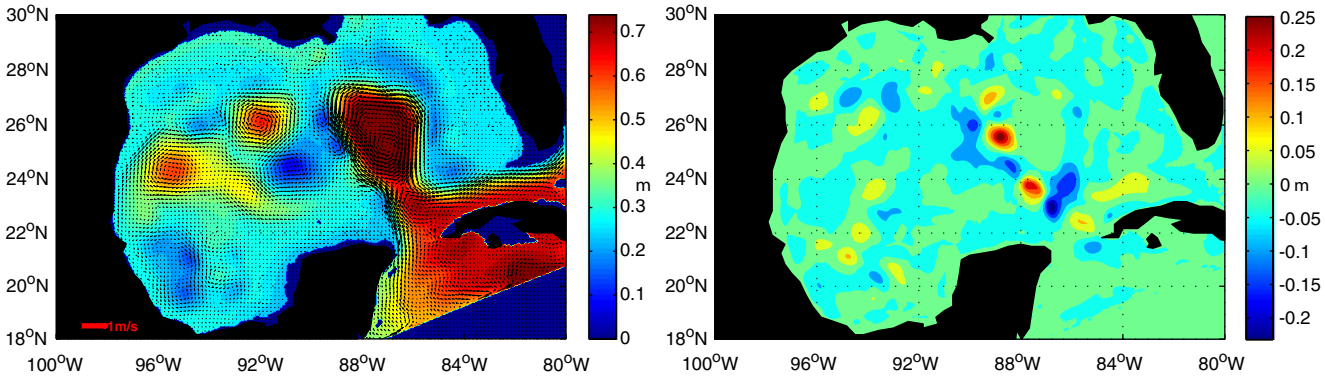

Fig. 2 **a** Control run on the 4th of July 2006; **b** anomalies 4 weeks after assimilation with $\alpha = 0.36$

shedding. Maul (1977) and, more recently, Bunge et al. (2002) showed that the vertical structure of the Yucatan straits inflow influences the life cycle of LC. In Murphy et al. (1999), Candela et al. (2002), the flux of potential vorticity is characterized by the passage of eddies that influence the timing of the shedding. Following this, Cherubin et al. (2006) show that the vorticity gradients (vertical and horizontal) enhance fast-growing instabilities at the boundary of the LC. Those instabilities evolve as cyclonic eddies that chop anticyclonic eddies off from the LC.

As mentioned in Section 2.2, the TOPAZ3 system provides lateral boundary conditions in dynamical balance and relatively consistent with observations of the area. As the current TOPAZ3 system does not use data assimilation, the inflow cannot be expected to be in phase with the reality. Perturbation of the inflow timing is used here to simulate this source of error, with a time lag to the boundary conditions. Alternatively, one could apply 3-dimensional random perturbations to the lateral boundary conditions, but preserving their physical consistency would be technically challenging.

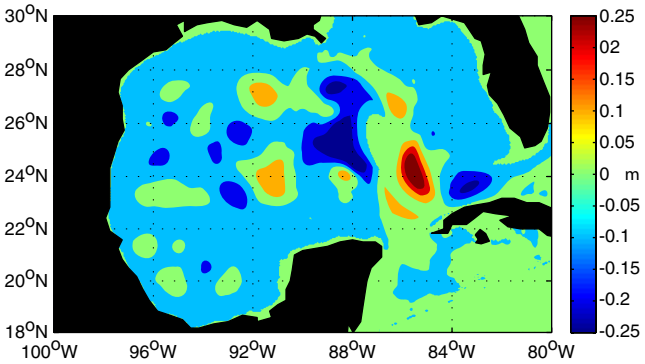

Fig. 3 Model SSH anomalies after 4 weeks of perturbations of lateral boundary perturbations

Abascal et al. (2003) analyzed the variability of the energy and the transport through the Yucatan Straits from measurements and observed peaks in the 5–10-, 20–40-, and 50–100-day bands. However, the time lag applied for the boundary condition should be kept in a small range to avoid spurious effects from seasonal variations. The lag is set randomly at the beginning of each model run between −37 and +37 days from the actual date.

The sensitivity of the model to the proposed perturbation is analyzed in a twin experiment, where one run is forced by lateral boundary conditions that are 37 days older than the other one (the latter being the control run, Fig. 2a). Figure 3 shows the anomalies from the control run on the 4th of July developed after running both models for 4 weeks. We observe dipoles of negative and positive anomalies at the boundary of the LC. They have radii varying from 35 to 60 km, which agree relatively well with the first baroclinic mode (30–40 km). These anomalies appear first at the southern boundary of the model. Deviations from the main position of the Yucatan Current result in bands of positive/negative anomalies that propagate through the Cayman Sea. The narrowing of the Yucatan Straits and then the interaction with the Campeche Bank induce small anomalies that propagate clockwise at the boundary of the LC, approximately at 30 km/day. They fade out when passing to the eastern side of the LC, probably due to interaction with the Florida Shelf. Those anomalies correspond to small cyclonic and/or LC meanders (Schmitz 2005).

We observed additional barotropic waves on the first day following the model restart, which are caused by the sudden change in the boundary conditions. However, these waves do not seem to disturb the model as they propagate out of the domain or are damp out within a day.

3.3 The atmospheric forcing

Oey et al. (2003) show that the atmospheric forcing also has an influence on the eddy shedding frequency. They use a larger model and attribute the change of frequency to the remote influence of the perturbations on the water transported through the Caribbean. In our system, such information is contained in the lateral boundary condition and we analyze here the impact of perturbing the atmospheric forcing fields on our high-resolution model domain only (Cayman Sea and GOM, see Fig. 4).

The perturbations of the atmospheric fields are simulated with a spectral method (Evensen 2003). For simulating the residual error, we set the spatial decorrelation radius to 50 km, which is the resolution of the atmospheric forcing fields used, and corresponds to perturbations that stimulate our high-resolution ocean model. The decorrelation time-scale is of 3 days. The standard deviations of the fields perturbed are: 0.095 N m^{-2} for the eastward and northward drag coefficient; 1.6 m/s for the wind speed; 22% for the cloud cover, and 3°C for the air temperature.

The sensitivity of the model to the proposed perturbation system is analyzed in a twin experiment, where one run is the control run and the other is forced by the perturbed atmospheric forcing fields. Figure 4 shows the anomalies from the control run after 4 weeks. The dominant signal is similar to the one observed with perturbations of the lateral boundary condition, characterized by dipoles of positive and negative anomalies (approximately 40 km), which propagate clockwise around the LC. Contrary to perturbations of the lateral boundary, other anomalies of smaller amplitude are observed throughout the whole domain.

The anomalies develop first uniformly throughout the basin and then intensify in specific areas such as in shelf areas (northern shelf and Campeche Bank) and at the boundary of the large-scale features (LC, anticyclonic eddies). After 2 weeks, the perturbations induce a small displacement of the Yucatan Current that produces similar anomalies as with perturbations of the lateral boundary condition. They dominate in amplitude compared to the rest of the anomalies after approximately 20 days.

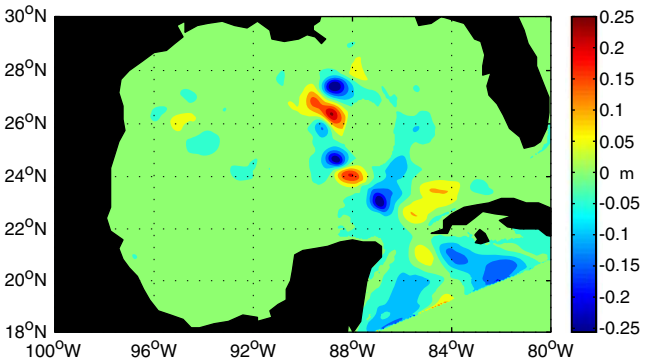

Fig. 4 Model SSH anomaly after 4 weeks of perturbations of the forcing fields

3.4 Contribution from each source of error

The above results indicate that each source of error stimulates a different model response in terms of anomaly growth and pattern. In order to quantify their respective contributions to the ensemble spread, we have processed three separate ensemble runs of 10 dynamical members each, where every ensemble run uses one source of error, as described above.

Figure 5 shows with thick lines the daily anomaly in surface elevation between the 10-member ensemble mean $\overline{\eta_p}$ and the control run η_c, calculated as follows:

$$\delta\eta = \sqrt{\int_\Omega (\overline{\eta_p} - \eta_c)^2}, \tag{2}$$

where Ω is the model domain. The assimilation update produces an initial deviation of 6.5 cm from the control run. With all perturbation systems, the perturbed ensemble mean deviates from the control run.

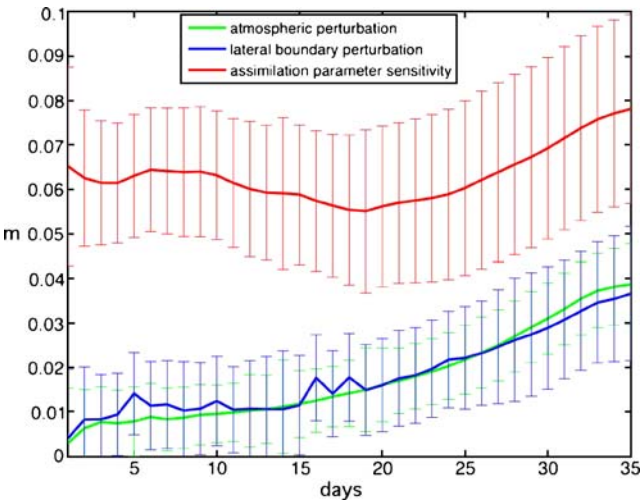

Fig. 5 Ensemble deviation from the control run for each source of error. The *thick lines* represent the deviation from 10-member ensemble mean, and the *error bars* represent the ensemble spread

The ensemble spread $\delta\eta_p$ is represented with error bars in Fig. 5, and is calculated as follows:

$$\delta\eta_p = \sqrt{\int_\Omega \frac{1}{N_d}\sum_{i=1}^{N_d}(\overline{\boldsymbol{\eta}_p} - \boldsymbol{\eta}_p(i))^2}, \quad (3)$$

where N_d is the dynamic ensemble size (i.e., $N_d = 10$).

All sources of error induce a growing $\delta\eta$ and $\delta\eta_p$ with time, which indicates that the model is sensitive to all of them. After 35 days, the ensemble spread $\delta\eta_p$ is: 4.3 cm with varying α, 2.8 cm with a perturbation of the lateral boundary, and 1.8 cm with a perturbation of the atmospheric forcing fields. Varying the parameter α produces the largest ensemble spread. This is expected because α controls the position and the size of the LC and associated eddies, whereas the two other types of perturbations control the propagation of smaller-scale cyclonic eddies.

The perturbations of the lateral boundary produce a large spread initially, which is caused by the barotropic adjustment consequently to the sudden change in the barotropic term of the boundary condition (respectively, assimilation noise). The same occurs with large values of α that produce data assimilation noise. In both situations, the spread is rapidly damped (within a day) and does not seem to create major disturbances in the model.

4 Ensemble forecast results

Our ensemble forecast is tested for the shedding of Eddy Yankee (2006). It is relatively clear from MODIS Ocean Color data that the shedding occurs around the 19th of July, reattaches to the LC from the east about a week later, and then remains attached for 2 months.

The ensemble runs are started 7 weeks prior to the shedding of Eddy Yankee in order to spin-up the perturbation system. In the first ensemble run (referred to as run 0 in Fig. 6), only the lateral and atmospheric boundary conditions are perturbed. The six following ensemble runs include weekly assimilation of SLA and are hereafter referred to as run 1 to run 6. The first assimilation is applied on the 7th of June 2006 and the last one on the 12th of July. After each assimilation, the ensemble is run forward 14 days, which corresponds to a 7-day forecast horizon with respect to the availability of the near real-time SLA altimeter data (SALTO/DUACS maps are available with 1 week of delay).

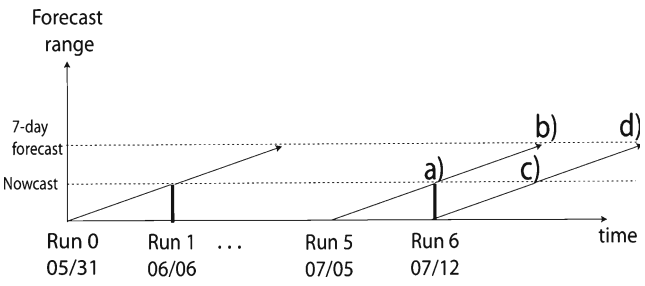

Fig. 6 Schematic of the ensemble forecast experiment. The *arrows* represent the model integration of the 10 members. The *solid lines* represent data assimilation of each these members with the EnOI. The letters *a–d* correspond to the panels in Fig. 8

4.1 Influence of the data assimilation parameter α

This study has allowed for a sensitivity study of the parameter α. This parameter, often referred as "strength" parameter (see Eq. 1), is a critical parameter controlling the performance of the EnOI. A large value will increase the forecast error relative to observation error and favor the initial agreement of the model forecast with the available observations, but it may cause side effects such as artificial gravity waves (Counillon and Bertino 2008).

In our experiment, each dynamic member keeps the same value of α over the successive ensemble runs (see Table 1). Their respective daily RMS error is computed against altimeter SLA track data, as follows:

$$\varepsilon(\boldsymbol{\psi}_i) = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\mathbf{H}\boldsymbol{\psi}_i(j) - \mathbf{d}(j))^2}. \quad (4)$$

The daily RMS error of each member $\varepsilon(\boldsymbol{\psi}_i)$ is averaged over the six successive ensemble runs to reduce the variability caused by the random boundary perturbations. The errors in Fig. 7a are split between the analysis period (from day 0 to day 7; blue bars) and the forecast period (from day 7 to day 14; red bars).

Members 1–3 have significantly higher errors than the other members. This indicates that the value of α for these three members is too low to maintain as high an accuracy as the other members, over successive assimilation cycles. Looking at members 4 to 10, the benefits of increasing α are no longer obvious for the forecast period (blue bars). Additional runs would have reduced the variability caused by random boundary conditions.

Oke et al. (2006)[1] indicate that a high value of α might produce a better estimate initially but might deteriorate the forecast accuracy with time. In Fig. 7b,

[1] Our notation of α follows that of Evensen (2003) but corresponds to α^2 in Oke et al. (2006)

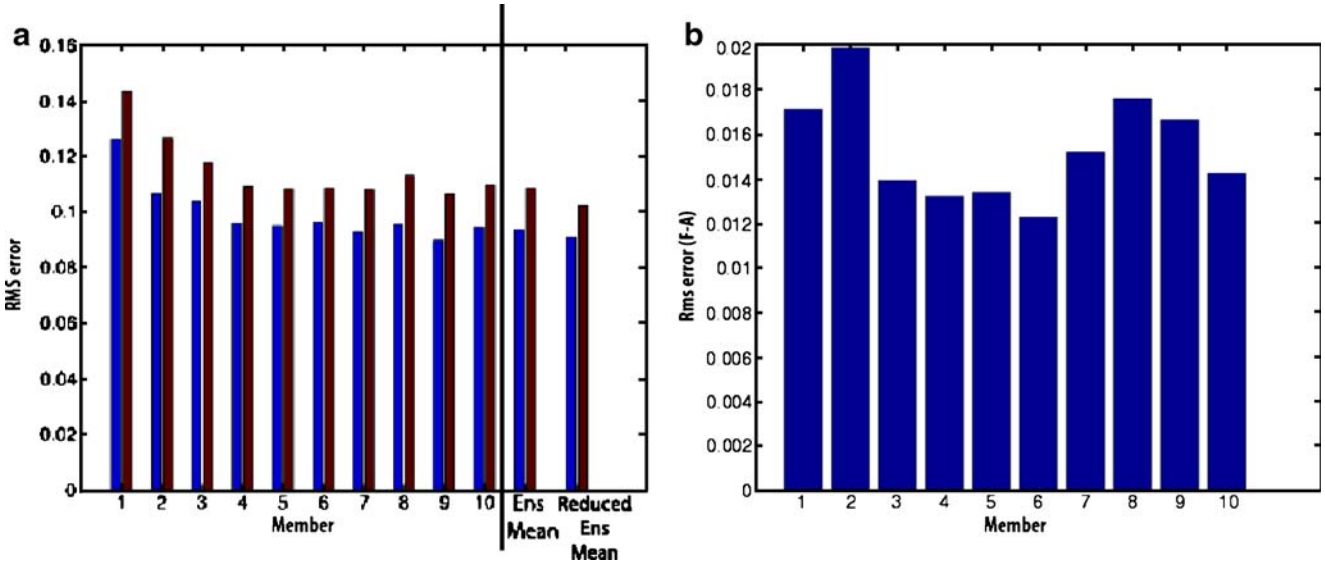

Fig. 7 **a** Mean RMS error of each ensemble member compared to SLA track data for the analysis period (0 to 7 days) in *blue* and forecast period (7 to 14 days) in *red*; **b** difference between the two periods

we have calculated the difference of SLA RMS error between the forecast period and the analysis period. From this plot, we can see that the central members (4–6) have smaller RMS error growth. This confirmed that increasing the value of α is beneficial up to a value of 0.09, but any further increase (in particular values higher than 0.17) produces a faster increase of the error.

We now look for the best estimator based on the ensemble forecast. In the Gaussian case, the ensemble mean is the best estimator, as in Yin and Oey (2007). Here, the three first members are outliers of the ensemble, and make the ensemble non-Gaussian. As a consequence, the ensemble mean is not the best guess, although it is fairly close, see Fig. 7a. As the other members appear as equally likely to give the best solution, we have computed the reduced ensemble mean (computed without the three outliers), which presents a lower RMS error than any other member. Although the runs that use low value of α over several assimilation cycles are, on average, less accurate than others, they may perform well on a single assimilation cycle. In order to maintain the Gaussian properties of the ensemble, one should rather reset the value of α randomly for each dynamical ensemble member at the beginning of each assimilation cycle.

4.2 Assessment of ensemble frontal positions

Chassignet et al. (2005) show that ocean color (OC) data are useful for identifying the position of the fronts of the LC and the eddies. Furthermore, OC data provide an independent source of validation and have higher resolution than the altimetry data. In Fig. 8, the deep blue contour area represents the low chlorophyll water (< 0.3 mg/m^3) that originates from the Caribbean Sea. The light green areas characterize the water with higher chlorophyll concentration (> 0.5 mg/m^3), which is usually found in areas of high biological production along the coast or within cyclonic eddies. The high chlorophyll water often propagates along the outer edge of the LC eddies and clearly defines the front.

To evaluate the accuracy of the ensemble forecast, we compare the ensemble front (10 cm SSH isoline) spaghetti plot with the OC data. We also add the front characterized by the 10 cm isoline in the altimeter SSH maps (hereafter referred to as SSH data). Although OC data has a higher resolution, the SSH data may be a useful indicator of the error in the assimilated SSH data or be used to locate the front when clouds mask the OC data. The model and the SSH data fronts are tuned to fit best with the OC data.[2] Only run 5 and run 6 are presented here because they cover the shedding and prereattachment of eddy Yankee and are thus the most interesting (and complex).

On the 12th of July (Fig. 8a), the LC (dark blue in the OC) has cyclonic eddies on either side of its neck (lighter blue in the OC map) with an eastern cyclonic eddy penetrating deeply, which indicates a near shedding scenario. The front calculated from SSH data lies within the ensemble front envelope except in restricted

[2]The front delimitation is calculated in a subjective manner from SSH and can produce slightly different results depending on the threshold used.

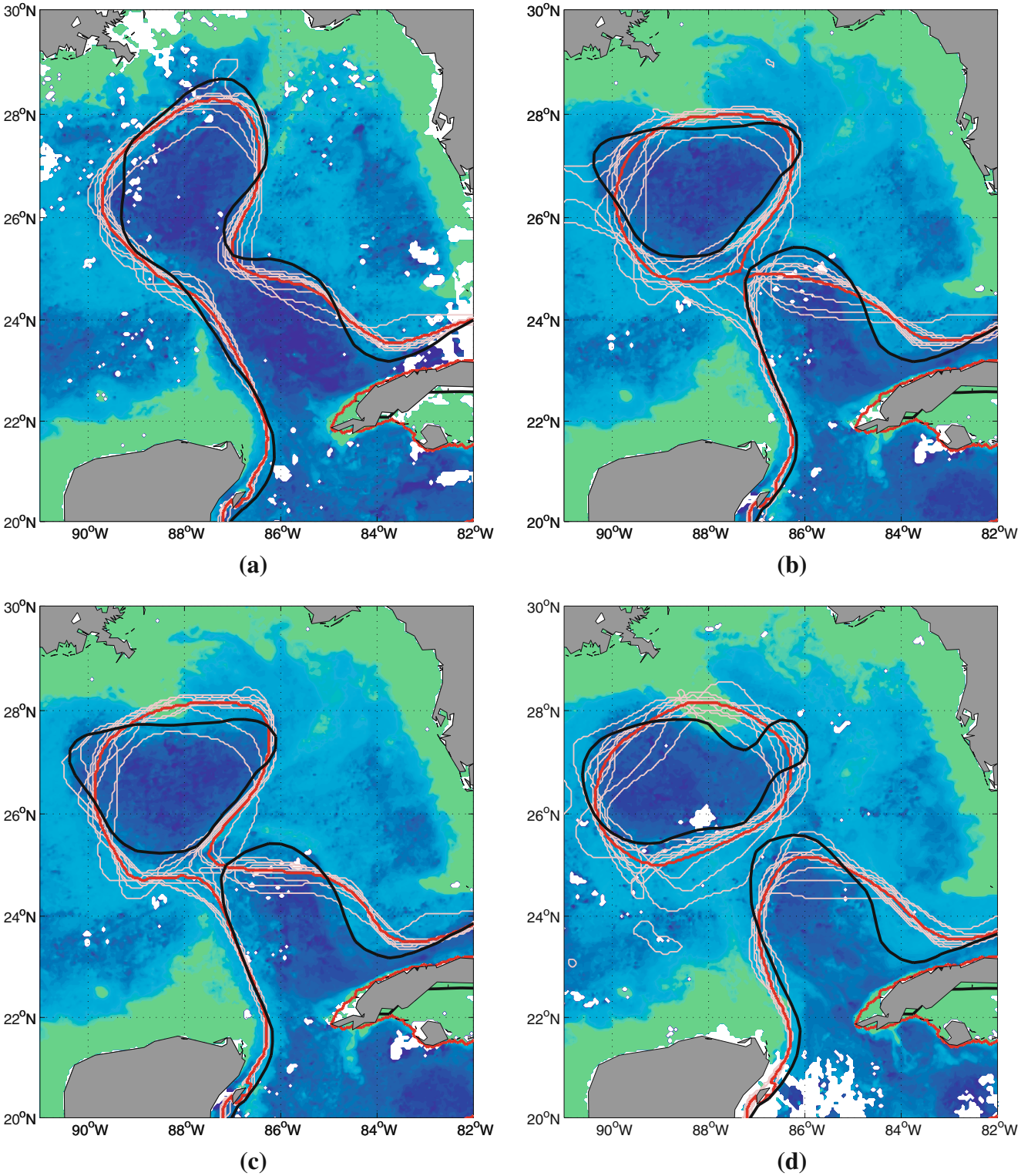

Fig. 8 Overlay of model ensemble fronts (*pink line*) with the nonassimilated OC map (*contour*); for the nowcast on the 12th of July (**a**); for the 7-day forecast on the 19th of July (**b**); for the nowcast on the 19th of July (**c**), and for the 7-day forecast on the 26th of July (**d**). *Blue color* (respectively, *green*) indicates low (respectively, *high*) concentration of chlorophyll, and cloud-covered areas are in *white*. The *thick black line* represents the front derived from SSH altimeter maps, and the *thick red line* is the ensemble mean

areas: the northeastern and the northwestern tips of the LC, and south of the eastern cyclonic intrusion (at about 25° N). At the western tip, the ensemble spread is narrow and agrees relatively well with the OC front, whereas the SSH data front does not extend far enough to the west. On the contrary, in the northeastern tip of the LC the ensemble spread is large, and the SSH data front provides a better agreement with the OC front. The model also has a tendency to locate the eastern cyclonic intrusion too far to the south and has a large spread there.

On the 19th of July (where Fig. 8b is the 7-day forecast and Fig. 8c the nowcast), one can observe a high concentration of chlorophyll (green in the OC) advected around the remaining LC. The eddy seems on the verge of separation, since only a small filament connects the eddy to the LC from the west. The ensemble forecast shows how ambiguous the situation is, as some members have already shed the eddy, and others are still connected. Furthermore, the ensemble front envelope is suddenly much wider than on the previous run pointing out the complexity of the situation. The triangular shape of the eddy is observed in most of the members, but not as pronounced as in the measurements. The SSH data front lies within the ensemble front envelope (both in the 7-day forecast and in the nowcast), except in the northeastern tip of the remaining LC (at about 25° N), the southeastern and northwestern tip of the eddy, and its northern front. On the northeastern tip of the remaining LC, the ensemble fronts are shifted to the south but the ensemble spread is large. Compared to the 7-day forecast, the nowcast has slightly corrected the front and reduced the ensemble spread, but this area is problematic in all panels of Fig. 8, which might indicate a model bias. Once again, the southeastern front of the eddy seems better located in the model than in SSH data, especially in the nowcast, where the spread is narrow and the ensemble isolines follow the OC well. On the northwestern tip of the eddy, the spread is large and some members capture the complex shape of the front, especially in the nowcast. On the northern front, the model does reproduce the complex shape of the eddy, induced by a strong cyclonic eddy that interacts with Eddy Yankee. This cyclonic eddy was previously too close to the coast, and was therefore misrepresented in the assimilated SSH maps. This has consequences in the nowcast (started the 12th of July), and to a lesser extent in the 7-day forecast (started on the 5th of July). This might explain why the northern front is better located in the 7-day forecast than in the nowcast.

On the 26th of July (Fig. 8d), the ensemble front gives a relatively good location of the eddy except that members 1 and 2 place the northern front too far south. However, the shape of the eddy in the model is too circular and does not reproduce the complex northeastern front seen in the OC data. This is once again due to the misrepresentation of the northern cyclonic eddy. None of the members have reattached to the eddy yet, as in the OC, but the rotation of the eddy indicates a reattachment in the following days.

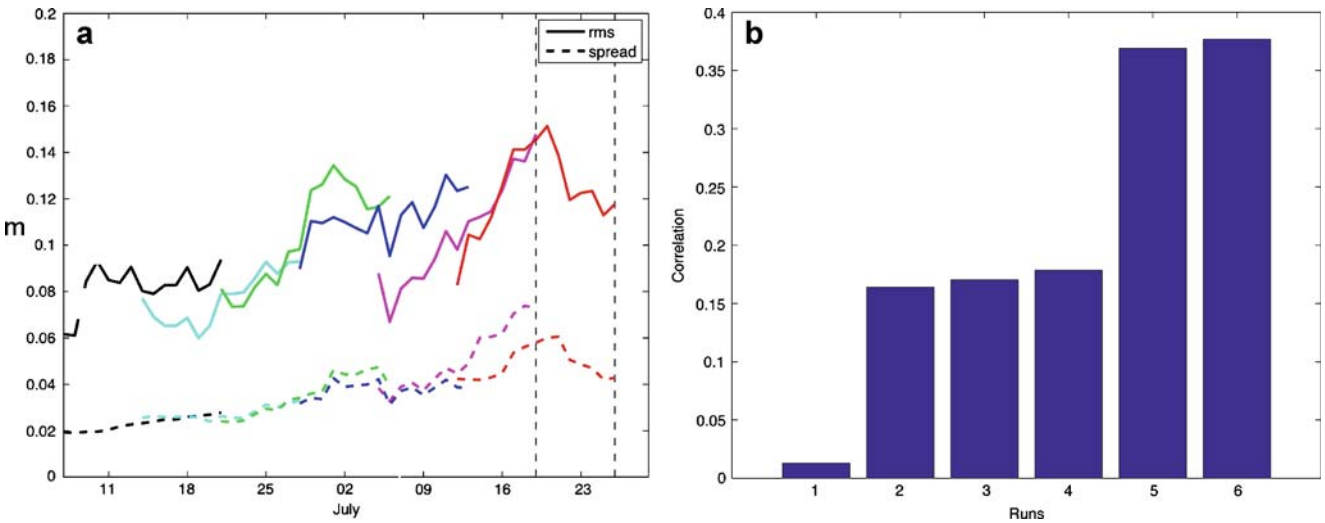

Fig. 9 **a** Evolution of the ensemble spread (*dashed line*) and RMS errors (*solid line*). Every ensemble run is plotted with a different color. The *vertical dashed lines* correspond to the shedding date of eddy Yankee on the 19th of July and to the reattachment of Eddy Yankee on the 26th of July. **b** Spatial correlation for each run

4.3 Comparison of ensemble vs actual errors

In Section 4.2, the ensemble fronts show relatively good agreement both with the OC and SSH data. We now quantify the accuracy of the ensemble mean forecast and compare it with the ensemble spread. This diagnostic, often referred to as spread–skill correlation, is commonly used in meteorology in order to provide confidence in the prediction (Molteni et al. 1996). We analyze the correlation between the two quantities evaluating whether the ensemble spread can predict episodes of higher/lower accuracy and areas of larger/smaller error.

For this purpose, we compute the daily values of the ensemble mean RMS error (i.e., $\varepsilon(\overline{\boldsymbol{\psi}})$ with respect to Eq. 4) and the ensemble spread $\delta\eta$ (see Eq. 5) at the location of the SLA track data. During a shedding event, the fast dynamics of the eastern GOM contrast with the slower activity of the remaining domain, and the RMS errors vary from day to day, due to irregular sampling of satellite tracks. Over 5 days, the SSH tracks cover the domain more uniformly and provide a more stable estimate of the RMS error. In Fig. 9, the ensemble mean RMS error (solid line) and the ensemble spread (dashed line) are averaged spatially over 5-day windows, except for the first and last days. The nesting area is removed from the calculations because the perturbation of the boundary condition implies an artificial correlation there.

$$\delta\eta = \sqrt{\frac{1}{mN_d}\sum_{j=1}^{m}\sum_{i=1}^{N_d}\left(\mathbf{H}\overline{\boldsymbol{\psi}(j)} - \mathbf{H}\boldsymbol{\psi}_i(j)\right)^2} \tag{5}$$

The dynamical ensemble spread is two to three times smaller than the forecast error. This underestimate is a common weakness of a small-sized EPS (Buizza et al. 2005). It indicates that our dynamical ensemble is suboptimal, either because of the small ensemble size or because the perturbation system still does not fully represent all of the model errors. A pragmatic way to palliate for this is the use of ensemble inflation. However, we aim at providing confidence index with a small ensemble size and focus on the correlation.

In the first two runs, the variability of the ensemble spread is almost zero, as a spin-up time of the ensemble is necessary (see Section 3.4). For the last three runs, the model produces a larger spread, and the two curves vary in good agreement. In particular, both the spread and the ensemble mean error present a maximum for the shedding of the Eddy Yankee on the 19th of July (dashed line), and then reduce when the eddy gets close to reattached on the 26th. Over the six ensemble runs, the two curves present a correlation of 0.83 at a 99% confidence level. This indicates that the ensemble spread can be useful for providing time-confidence indices of our forecast.

The ensemble mean RMS error is usually larger near the SSH front and reaches a maximum where cyclonic and anticyclonic eddies interact with each other. This is because, in these regions, the SSH gradients are larger and the dynamics more chaotic. Here, we analyzed whether the ensemble spread can predict the area where the error will grow faster. For this purpose, Fig. 9b shows the spatial correlation between the run average of the ensemble mean error and ensemble spread. There is almost no correlation initially, but the correlation increases with successive runs and has a maximum during the last two runs ($R = 0.37$).[3] The correlation is not large, but it is probably impaired by the error in the ensemble mean (when the front is misplaced). Still, this result indicates that the ensemble can provide additional spatial information regarding the forecast accuracy, in particular, at the time during which the model is the most inaccurate.

5 Conclusion

This work evaluates the skill of an EPS with a high-resolution HYCOM model for providing confidence indices for the prediction. For this purpose, we aim at perturbing the known sources of error in our model, which are the initial state, the atmospheric conditions, and the lateral boundary conditions. Our EPS uses different values of α as a proxy for perturbing the initial state and random perturbations of the latter two sources of error. Changes of α control the displacement of large features in the GOM (e.g., LC, warm core eddies), whereas perturbations of both lateral and atmospheric boundary conditions stimulate the propagation of smaller-scale instabilities, such as cyclonic eddies that circulate around the LC. It takes about 3 weeks for these cyclonic eddies to develop and propagate around the LC. The variation of the parameter α makes the largest contribution in the ensemble spread, but the other two kinds of perturbations are also important because the growth of cyclonic eddies plays a key role in the shedding process (Schmitz 2005).

The skill of our EPS is tested on Eddy Yankee and shows good capabilities. The perturbation system used here is able to produce a significant spread of the front

[3] All the correlations are significant to 99% confidence level.

position and mimic the dynamics of the eddy shedding. Qualitatively, the front observed from ocean color is almost always found within the ensemble spread. Quantitatively, the ensemble spread still underestimates the RMS error but is correlated with the RMS error both spatially and temporally, after a spin-up time of 2–3 weeks. It implies that additional uncertainty indices can be provided for the prediction.

The Gaussian properties of the ensemble are impaired by three outliers that have undergone data assimilation too weakly. In order to obtain all the members equally likely to give the best solution and to maintain the Gaussian properties of the ensemble, the parameter α should be randomized.

A study of the influence of the magnitude of the parameter α on the forecast horizon skill has been carried out. It appears that an optimal average value of this parameter can be estimated for a given forecast horizon. Considering a 7-day forecast horizon, we found that increasing the value of α over 0.17 induces a faster error growth and was not improving the result posterior to the nowcast stage, although the initial state is closer to the observations.

This study could greatly benefit from increasing the number of members, and extending the period of study over other shedding event, in order to gain more stability and, thus, more confidence in our results. This will be possible with increasing computing power. If the spread–skill correlation is confirmed, it should be possible to use such information for data assimilation purposes.

Finally, there is a promising perspective with the use of the outer TOPAZ3 model with the EnKF and the ECMWF EPS, which can provide ensemble boundary conditions to the high-resolution model. This should avoid the initial barotropic instability and there should be a useful synergy in using an ensemble both in the inner and in the outer model.

Acknowledgements The authors acknowledge funding from the Research Council of Norway, Shell E&P, partial funding from the EU MERSEA project (SIP3-CT-2003-502885), and a grant of CPU time from the Norwegian Supercomputing Project (NOTUR). Useful comments were received from two anonymous reviewers.

References

Abascal AJ, Sheinbaum J, Candela J, Ochoa J, Badan A (2003) Analysis of flow variability in the Yucatan Channel. J Geophys Res 108:3381–3399

Bentsen M, Evensen G, Drange H, Jenkins AD (1999) Coordinate transformation on a sphere using conformal mapping. Mon Weather Rev 127:2733–2740

Bleck R (2002) An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. Ocean Model 4:55–88

Browning GL, Kreiss HO (1982) Initialization of the shallow water equations with open boundaries by the bounded derivative method. Tellus 34:334–351

Buizza R, Houtekamer P, Toth Z, Pellerin G, Wei M, Zhu Y (2005) A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. Mon Weather Rev 133:1076–1097

Bunge L, Ochoa J, Badan A, Sheinbaum J (2002) Deep flows in the Yucatan Channel and their relation to changes in the Loop Current extension. J Geophys Res 107(C12):256. doi:10.1029/2001JC001

Candela J, Sheinbaum J, Ochoa J, Badan A, Leben R (2002) The potential vorticity flux through the Yucatan Channel and the Loop Current in the Gulf of Mexico. Geophys Res Lett 29:2059. doi:10.1029/203GL015587

Chassignet EP, Hulburt HE, Smedstad OM, Barron CN, Ko DS, Rhodes RC, Shriver JF, Wallcraft AJ, Arnone AR (2005) Assessment of data assimilative ocean models in the Gulf of Mexico using Ocean Color. Circ Gulf Mex Obs Models 161:87–100

Cherubin LM, Morel Y, Chassignet EP (2006) The role of cyclones and topography in the Loop Current ring shedding. J Phys Oceanogr 31:569–591

Counillon F, Bertino L (2008) Ensemble optimal interpolation: multivariate properties in the Gulf of Mexico. Tellus (in press)

Dai A, Trenberth KE (2002) Estimates of freshwater discharge from continents: latitudinal and seasonal variations. J Hydrometeorol 3:660–687

Dümenil L, Isele K, Liebscher HJ, Schröder U, Schumacher M, Wilke K (1993) Discharge data from 50 selected rivers for GCM validation. Max-Planck-Institut für Meteorologie, Hamburg

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 53: 343–367

Evensen G (2006) Data assimilation: the ensemble Kalman filter. Springer, New York

Le Traon PY, Dibarboure G, Dorandeu J (2003) SSALTO/DUACS and operational altimetry. Geoscience and remote sensing symposium, 2003. IGARSS'03. In: Proceedings. 2003 IEEE International, vol. 2. IEEE, Piscataway

Lugo-Fernández A (2007) Is the Loop Current a chaotic oscillator? J Phys Oceanogr 37:1455–1469

Maul GA (1977) The annual cycle of the Gulf Loop Current, 1: observations during a one-year time series. J Mar Res 35: 29–47

Molteni F, Buizza R, Palmer TN, Petroliagis T (1996) The ECMWF ensemble prediction system: methodology and validation. Q J R Meteorol Soc 122:73–119

Murphy SJ, Hurlburt HE, O'Brien JJ (1999) The connectivity of eddy variability in the Caribbean Sea, the Gulf of Mexico, and the Atlantic Ocean. J Geophys Res 104:1431–1453

Natvik LJ, Evensen G (2003) Assimilation of ocean colour data into a biochemical model of the North Atlantic Part 2. Statistical analysis. J Mar Syst 40:155–169

Oey LY, Lee HC, Schmitz Jr WJ (2003) Effects of winds and Caribbean eddies on the frequency of Loop Current eddy shedding: a numerical model study. J Geophys Res 108: 3324

Oey LT, Ezer T, Lee HC (2005a) Loop Current, rings and related circulation in the Gulf of Mexico: a review of numerical models. Circ Gulf Mex Obs Models 161:31–56

Oey LY, Ezer T, Forristall G, Cooper C, DiMarco S, Fan S (2005b) An exercise in forecasting Loop Current and eddy frontal positions in the Gulf of Mexico. Geophys Res Lett 32:L12,611

Oke PR, Allen JS, Miller RN, Egbert GD, Kosro PM (2002) Assimilation of surface velocity data into a primitive equation coastal ocean model. J Geophys Res 107:3122–3147

Oke PR, Sakov P, Corney SP (2006) Impacts of localisation in the EnKF and EnOI: experiments with small model. Ocean Dyn 57:32–45

Rio MH, Hernandez F (2004) A mean dynamic topography computed over the world ocean from altimetry, in situ measurements, and a geoid model. J Geophys Res 109:1–19

Schmitz WJ (2005) Cyclones and westward propagation in the shedding of anticyclonic rings from the Loop Current. Geophys Monogr 161:241–261

Sheinbaum J, Candela J, Badan A, Ochoa J (2002) Flow structure and transport in the Yucatan Channel. Geophys Res Lett 29:10.1–10.4

Sturges W, Leben R (2000) Frequency of ring separations from the Loop Current in the Gulf of Mexico: a revisited estimate. J Phys Oceanogr 30:1814–1819

Teague W, Carron MJ, Hogan PJ (1990) A comparision between the generalized digital environmental model and levitus climatologies. J Geophys Res 95:7167–7183

Vukovich F (1988) Loop Current boundary variation. J Geophys Res 93:585–615

Winther N, Morel Y, Evensen G (2007) Efficiency of high order numerical schemes for momentum advection. J Mar Syst 67:31–46

Yin XQ, Oey LY (2007) Bred-ensemble ocean forecast of Loop Current and rings. Ocean Model 17:300–326