Correspondence to:
L. N. Lima,
leonasclima@gmail.com

An Investigation of Ocean Model Uncertainties Through Ensemble Forecast Experiments in the Southwest Atlantic Ocean

L. N. Lima[1] , L. P. Pezzi[1,2] , S. G. Penny[3,4] , and C. A. S. Tanajura[5]

[1]Center for Weather Forecast and Climate Studies (CPTEC), National Institute for Space Research (INPE), Sao Jose dos Campos, Brazil, [2]General Earth Observation Coordination (OBT), National Institute for Space Research (INPE), Sao Jose dos Campos, Brazil, [3]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA, [4]National Oceanic and Atmospheric Administration, National Centers for Environmental Prediction, College Park, MD, USA, [5]Department of Earth and Environmental Physics, Physics Institute, Federal University of Bahia (UFBA), Salvador, Brazil

Abstract Ocean general circulation models even with realistic behavior still incorporate large uncertainties from external forcing. This study involves the realization of ensemble experiments using a regional model configured for the Southwest Atlantic Ocean to investigate uncertainties derived from the external forcing such as the atmosphere and bathymetry. The investigation is based on perturbing atmospheric surface fluxes and bathymetry through a series of ensemble experiments. The results showed a strong influence of the South Atlantic Convergence Zone on the underlying ocean, 7 days after initialization. In this ocean region, precipitation and radiation flux perturbations notably impacted the sea surface salinity and sea surface temperature, by producing values of ensemble spread that exceeded 0.08 and 0.2 °C, respectively. Wind perturbations extended the impact on currents at surface, with the spread exceeding 0.1 m/s. The ocean responded faster to the bathymetric perturbations especially in shallow waters, where the dynamics are largely dominated by barotropic processes. Ensemble spread was the largest within the thermocline layer and in ocean frontal regions after a few months, but by this time, the impact on the modeled ocean obtained from either atmospheric or bathymetric perturbations was quite similar, with the internal dynamics dominating over time. In the vertical, the sea surface temperature exhibited high correlation with the subsurface temperature of the shallowest model levels within the mixed layer. Horizontal error correlations exhibited strong flow dependence at specific points on the Brazil and Malvinas Currents. This analysis will be the basis for future experiments using ensemble-based data assimilation in the Southwest Atlantic Ocean.

Plain Language Summary The numerical models are powerful tools to provide knowledge about the ocean state concerning currents eddies, meanders, and other ocean dynamic and thermodynamic processes on a range of temporal and spatial scales. An accurate numerical model makes possible to get a tridimensional ocean representation with some confidence during time. Even though the ocean numerical models have been incorporating improvements, mainly due to a growing evolution of the computational resources, they are still somewhat limited and bring uncertainties on their simulations due many reasons that are related to the applied physical parameterization, atmospheric forcing, bathymetry, and some other issues. It is crucial to investigate and to know these uncertainties. This study goes further on the uncertainty investigations in order to create the basis (prior step) for an ensemble-based data assimilation system for the Southwest Atlantic Ocean. Our results indicated that uncertainty in wind forcing plays a major role in the determination of uncertainty in the ocean state. Compared to atmospheric forcing, the uncertainty in bathymetry produced a larger impact on the ocean representation, especially in shallow waters, though this may be in part due to excited waves at the initial time.

1. Introduction

Ocean models used for operational forecasts have evolved to the point that their higher horizontal resolutions now resolve processes such as ocean eddies, jets, and Rossby waves, processes that are only partially described by a purely observation-based assessment, for example, of satellite and in situ data. Despite the advances in horizontal resolution, these models may still introduce large uncertainties from different sources: physical parameterizations, atmospheric forcing, lateral boundary and initial conditions, bathymetry, river

runoff, and so forth. Due to a trend toward coupled modeling at operational centers (e.g., European Centre for Medium-Range Weather Forecasts Roadmap to 2025), understanding the uncertainties present in the ocean model component is essential for diagnosing errors in the full coupled forecast system.

In ensemble-based ocean data assimilation, a set of ocean models will quickly converge if forced with identical forcing, causing the filter to diverge from the observed state. In order to counter this effect, Hoffman et al. (2012) and Penny et al. (2013) applied random perturbations in the wind fields in order to account for the uncertainties derived from the atmospheric forcing into the forecast error covariance matrix. Penny et al. (2015) applied a more dynamic estimate of uncertainty by using perturbations derived from the ensemble of surface conditions made available from the 20th Century Reanalysis (Compo et al., 2011). Momentum and heat fluxes play an important role on the ocean surface, and ocean models exhibit sensitivity to these fluxes. For example, Jung et al. (2014) showed that the spatial resolution of the atmospheric fields has a significant impact on the results of the modeled ocean and found that the small-scale atmospheric phenomena (e.g., fronts, mesoscale cyclones, and topographic jets) are important contributors in driving the mean horizontal wind-driven ocean circulation. Further, the bulk formulations that translate these atmospheric surface conditions to fluxes provide additional uncertainties to the ocean model (Brunke et al., 2011, 2002).

Besides the surface fluxes, uncertainties exist in the model dynamics and physics parameterizations as well. In recent years, increasing computing capabilities have allowed ocean models to be run at higher-resolution configurations (Chassignet & Xu, 2017), by tending to a continuous improvement in the representation of the ocean physical processes. However, there are still open questions regarding the application of physical parameterizations. While parameterizations have evolved, they have also mainly been based on deterministic formulations derived from measurements that were conducted in specific regions such that they may not be optimal for the global ocean (Miller et al., 2017). In order to avoid uncertainties from employing deterministic parameter estimations, Brankart (2013) showed that the application of stochastic parameterization in the seawater equation of state can produce a notable effect on the average large-scale circulation of the ocean, with a more prominent signal in regions of high mesoscale activity. Next, Brankart et al. (2015) investigated unsolved processes and scales of the ocean model through stochastic temperature (T) and salinity (S) fluctuations and noted that the small scales constantly modify the structure of the large-scale density, and thus the pathway of the large-scale circulation, as a result of the nonlinearity of the equation of state. Brankart et al. (2015) also mentioned but did not investigate the uncertainties derived from external forcing such as the atmospheric forcing, river runoff, or lateral boundary conditions.

While few studies have been conducted to investigate the uncertainties from the lateral boundaries in regional ocean modeling (e.g., Counillon & Bertino, 2009; Sandery et al., 2014), this impact has more frequently been addressed in regional atmospheric modeling (e.g., Davies, 2014; Gustafsson et al., 1998; Nuissier et al., 2012; Wu et al., 2005). Regional models require lateral boundary conditions from a global model, which are usually of coarser resolution. As a result of the change in resolution, downscaling errors can propagate through the inner domain from the boundaries. Sandery et al. (2014) showed the sensitivity to different lateral boundaries in a nested configuration by employing different configurations for a regional ocean model. Their results using more accurate fields as lateral boundary reduced the bias of ocean areas into the inner domain.

There are other sources of ocean model uncertainties that have been minimally explored in the literature. For instance, an appropriate river discharge is crucial for physical and biogeochemical representations in the model. This has immediate impacts in the coastal regions but for larger sources such as the Amazon outflow may have significant impacts on general circulation (Huang & Mehta, 2010; Jahfer et al., 2017). Unfortunately, most of the rivers (e.g., Amazon and La Plata) have not been well monitored, with time series containing gaps due to the difficulty of maintaining the measuring platforms. This deficiency has increased the need for techniques that supplement the measurements by estimating missing segments in the river time series (Elshorbagy et al., 2000). However, even after applying such correction techniques, uncertainties remain that must be adequately represented in data assimilation.

Another important source of uncertainty is the ocean model bathymetry. High-resolution estimates reconstructed from satellite altimeters and gravity measurements (e.g., Sandwell et al., 2006; W. Smith & Sandwell, 2004) must be smoothed and interpolated in order to obtain adequate physical and numerical stability for a given model resolution and physics requirements. However, this smoothing and subsequent interpolation are often based on generic tools as described in Penven et al. (2008). Nevertheless, the adequate
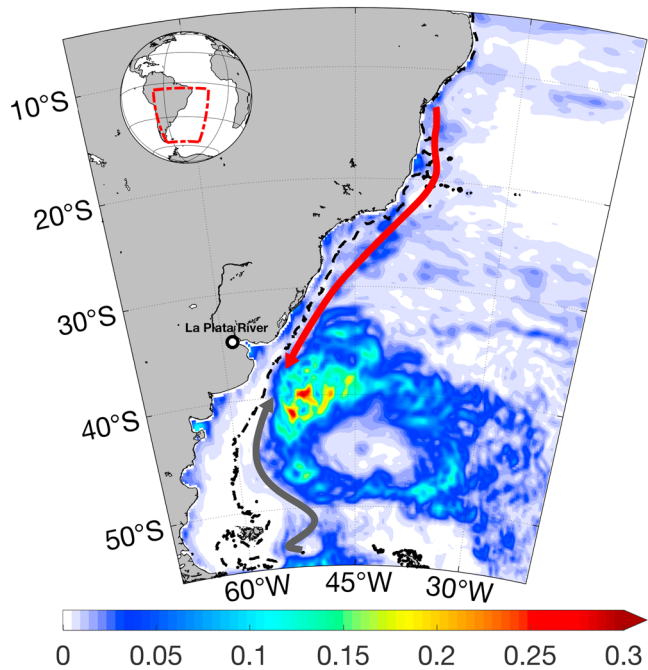
Figure 1. Time-averaged eddy kinetic energy in square meters per square seconds (1 January to 1 July 2009) in the Southwest Atlantic Ocean, using the fields of geostrophic velocity anomalies that were produced and distributed by the Copernicus Marine and Environment Monitoring Service (http://marine.copernicus.eu/). The black-dashed contour represents the isobath of 200 m. The red (gray) arrow schematically indicates the Brazil (Malvinas) current at surface.

creation of the model bathymetry is a turning point for an appropriate ocean representation and requires the use of improved tools like proposed by Sikirić et al. (2009) in order to prevent the terrain-following coordinate models from horizontal gradient-pressure errors.

In general, these many sources of uncertainties are difficult to assess. However, for applications such as data assimilation and ensemble forecasting, it is important that these uncertainties are properly represented. Ensemble forecasts produce simultaneous model integrations, ideally taking into account model uncertainties in order to indicate how both the initial condition and systematic model errors project onto the forecast uncertainty. For atmospheric models, perturbations are often derived also from methods like breeding that allows studying the growth rates and saturation of the initial condition errors in time (Toth & Kalnay, 1993, 1997). The result of the breeding method is a set of normalized fields that serve as "dynamically conditioned" initial perturbations for the ensemble. In the Tasman Sea, O'Kane et al. (2011) applied the breeding method on ensemble experiments using a regional model to investigate the spatial pattern of the fastest errors according to a given initial state. Their results showed that over 7 days forecast errors arise due to dynamic instabilities and that these forecast errors can be expected to dominate analysis errors. Counillon and Bertino (2009) assessed the skill of a regional ensemble prediction system in the Gulf of Mexico by perturbing the known sources of error in a model and found the displacement of large ocean features due to the perturbations of the initial state, whereas perturbations of both lateral boundary conditions and atmospheric forcing stimulated the propagation of smaller-scale instabilities.

To our knowledge, there have not been studies that have investigated the uncertainties of ocean models in the Southwest Atlantic Ocean (SWA). There are important oceanic phenomena in the SWA that need improved comprehension (Figure 1). The Brazil Current (BC) is a western boundary current that flows southward transporting warm waters adjacent to the Brazilian coast. The vertical structure of the water masses around the BC is complex. For example, it is possible to find waters flowing southward or northward depending on the latitude and depth of the South Equatorial Current bifurcation (Stramma & England, 1999). Also in the SWA is the Malvinas Current (MC), which flows northward and carries cold waters from the Southern Ocean. The vertical structure of the MC is barotropic, generating a northward flux that is practically continuous through all depths. At approximately 35–40°S, the meeting of the BC and MC generates the so-called Brazil-Malvinas Confluence (BMC; Combes & Matano, 2014), a frontal region where a highly complex ocean circulation develops with the generation of meanders and eddies resulting in a region of high mesoscale variability (Mason et al., 2017). BMC plays also a key role on modulate the marine atmospheric boundary layer (Pezzi et al., 2009, 2005).

In this study, we investigate uncertainties in the ocean state using a pure ensemble method, performed using the Regional Ocean Modeling System (ROMS) in the SWA. The main motivation is to investigate uncertainties in the representation of the ocean state due to uncertainties in the atmospheric forcing and bathymetric fields. This investigation is based on perturbing atmospheric surface fluxes and bathymetry through a series of ensemble experiments. Ensemble-based error correlations are also estimated to compare the ocean flow-dependent patterns that are generated due to these two sources of perturbations. These patterns are spatial structures that evolve over time according to the ocean dynamics and partly reveal the "errors of the day" (Kalnay, 2003). A proper study of the model errors and their spatial evolution in time is important for the estimation of the forecast error covariance matrix in ensemble data assimilation. This paper is a starting point for creating a data assimilation system based on the local ensemble transform Kalman filter (Hunt et al., 2007) in the SWA that is going to be shown in a follow study. By investigating error correlations between the ensembles of surface atmosphere and ocean states, we also provide a complement to new studies in coupled data assimilation (e.g., Sluka et al., 2016).

The manuscript is organized as follows. Section 2 provides information about the regional ocean model used in this study and the configuration used in the ensemble experiments as well as the statistical metrics used to assess the results. Section 3 provides a description of the results in three subsections discussing (1) the influence of atmospheric perturbations on the ocean simulations, (2) the comparisons between the impact of bathymetric and atmospheric perturbations, and (3) the estimation of error correlations in the ocean ensemble to investigate flow-dependence. Section 4 presents the conclusions and future work.

2. Methodology

2.1. The Ocean Model and Configuration

ROMS is a three-dimensional, free-surface, terrain-following vertical coordinate ocean model. It solves the Navier-Stokes equation under simplifying assumptions using Reynolds averaging, the Boussinesq approximation, and the hydrostatic approximation. The model solves its equations through a split-explicit model in which depth-integrated (barotropic) equations are resolved with a shorter time step than the full 3-D (baroclinic) equations. The depth-integrated momentum and continuity equations are below:

$$\frac{\partial(D\overline{u})}{\partial t} + \frac{\partial(\overline{u}D\overline{u})}{\partial x} + \frac{\partial(\overline{v}D\overline{u})}{\partial y} - fD\overline{v} = -D\frac{\partial P}{\partial x} + \tau_{sx} - \tau_{bx}, \tag{1}$$

$$\frac{\partial(D\overline{v})}{\partial t} + \frac{\partial(\overline{u}D\overline{v})}{\partial x} + \frac{\partial(\overline{v}D\overline{v})}{\partial y} + fD\overline{u} = -D\frac{\partial P}{\partial y} + \tau_{sy} - \tau_{by}, \tag{2}$$

$$\frac{\partial\zeta}{\partial t} + \frac{\partial(D\overline{u})}{\partial x} + \frac{\partial(D\overline{v})}{\partial y} = 0, \tag{3}$$

where $\overline{u}$ is the depth-integrated velocity in the x direction, $\overline{v}$ is the depth-integrated velocity in the y direction, f is the Coriolis parameter, P is the pressure, τ_{sx} is the surface stress and τ_{bx} is the bottom stress in the x direction, τ_{sy} is the surface stress and τ_{by} is the bottom stress in the y direction, $D = h + \zeta$ is the total depth, h is the depth of the sea floor below mean sea level, and zeta (ζ) is the free-surface elevation. All equations and a more detailed description of the coupled barotropic-baroclinic system are shown in Shchepetkin and McWilliams (2005), Haidvogel et al. (2008), and Shchepetkin and McWilliams (2009).

The model is configured here for the domain 55°S to 5°S and 70°W to 20°W encompassing the SWA ocean with horizontal resolution of 1/12°. The number of vertical sigma levels is 30. The experiments are forced by 6-hourly atmospheric fields of net shortwave radiation, downward longwave radiation with the upward longwave radiation being a function of the model sea surface temperature (SST), precipitation, atmospheric surface pressure, 2-m specific humidity, 2-m air temperature, and 10-m wind speed at ~0.31° spatial resolution from the Climate Forecast System Reanalysis (CFSR) of the National Centers for Environmental Prediction (Saha et al., 2010). The next section describes how these CFSR fields are perturbed to generate ensemble forcing. The atmospheric fluxes of sensible heat (H_s), latent heat (H_l), and momentum (τ) at the surface are calculated using the equations (4)–(6) according to a bulk formulation as described in Fairall et al. (1996) and Fairall et al. (2003):

$$H_s = \rho c_p C_h U(\theta_{\text{air}} - \text{SST}), \tag{4}$$

$$H_l = \rho L_e C_e U(q_s - q), \tag{5}$$

$$\tau = \rho C_d U^2, \tag{6}$$

where C_h, C_e, and C_d are, respectively, the heat transfer, humidity transfer, and drag coefficients. The variable θ_{air} is the potential temperature, SST is the sea surface temperature, q_s is the specific humidity at sea level, q is the 2-m specific humidity, and U is the 10-m wind speed relative to the sea surface.

The ROMS initial and lateral open ocean boundary conditions are derived from the Simple Ocean Data Assimilation (SODA) reanalysis version 2.2.4 (Carton & Giese, 2008). SODA assimilates T and S profiles from the Word Ocean Database (Boyer et al., 2013) into the Parallel Ocean Program model version 2.1 (R. D. Smith et al., 1992). This version of SODA uses the Parallel Ocean Program ocean model with an average 0.25° × 0.4° horizontal resolution and 40 vertical levels. As surface forcing, it employs wind stress from the

mean state of the 20th Century Reanalysis version 2 (20CRv2; Compo et al., 2011) and heat fluxes derived from bulk formulae. The ROMS initial and boundary input files are prepared from the T, S, u, v, and sea surface height fields of the SODA outputs in a monthly averaged form, mapped onto a uniform $0.5° \times 0.5° \times 40$ level grid (http://sodaserver.tamu.edu/assim/SODA_2.2.4/). Our ROMS spin-up run commences on 1 January 1980, so the model is freely run from a monthly SODA field of January 1980 as initial state and forced by the CFSR 6-hr fields at the surface and the SODA monthly fields at the boundaries. After a spin-up of 29 years, the ensemble experiments are initialized.

2.2. The Ensemble Experiments

Experiments are conducted using an ensemble of ocean integrations forced by perturbed boundary conditions. We investigate the impacts of perturbations applied to multiple atmospheric forcing fields and also model bathymetry. Experiments commence on 1 January 2009, and all 28 ensemble members start from identical initial conditions so that the initial ensemble spread is zero.

2.2.1. Atmospheric Perturbations

Time-dependent perturbations are added to the atmospheric forcing fields, and each member is integrated independently so that the perturbations are the primary driver for increasing the ensemble spread over the time.

The atmospheric perturbations are applied to the CFSR fields that are inputs to the surface flux calculations of the ROMS bulk formulation. Perturbations are sampled from the 56-member ensemble of the 20th Century Reanalysis version 2c (20CRv2c; Compo et al., 2011). The 20th Century Reanalysis is an international effort with the aim to provide a comprehensive global atmospheric circulation data set for the 20th century, and version 2c spans 1850–2014 with a spatial resolution of $2° \times 2°$. The 20CRv2c uses the Global Forecast System from National Centers for Environmental Prediction for the atmosphere with prescribed sea ice boundary conditions from the COBE-SST2 (Hirahara et al., 2014), new pentad Simple Ocean Data Assimilation with sparse input (SODAsi.2) SST fields (Giese et al., 2016), and assimilates additional pressure observations from the International Surface Pressure Databank version 3.2.9.

In order to compute the atmospheric perturbations, we first randomly choose 28 members of the 20CRv2c. The mean of the 28-member sample is removed to generate perturbations $\delta X_k^R = X_k^R - \overline{X}^R$, where k is the ensemble member and the superscript R indicates that they are derived from the 20CRv2c. The members are interpolated to the higher CFSR resolution, and the perturbations are recentered at the CFSR values using the patch interpolation method from a software that is bundled with the Earth System Modeling Framework (www.earthsystemcog.org/projects/regridweightgen/). The perturbed CFSR fields are $\widetilde{X}_k^C = \overline{X}^C + \delta X_k^R$, where k is the ensemble member and the superscripts C indicates the CFSR. The anomaly fields are prepared at every 6 hr, so that the atmospheric perturbations vary with time. Six ensemble experiments are performed applying perturbations separately to the wind, radiation fluxes, specific humidity and precipitation with a duration of 12 months, and surface pressure and surface air temperature with a duration of 3 months. A final 6-month experiment is conducted applying perturbations to all atmospheric forcing fields simultaneously.

2.2.2. Bathymetric Perturbations

The bathymetry specified for in ROMS is typically generated using an integrated toolbox as described by Penven et al. (2008). The regional bathymetry field for the SWA regional domain is extracted from the ETOPO high-resolution global topography (Amante & Eakins, 2009). Two smoothing filters are then applied to this regional field (h). The first filter (F_1) is applied through an averaging procedure to avoid isolated seamounts. The justification for this smoothing is that geological features sometimes have a smaller spatial dimension than the horizontal resolution of the model, and their presence may cause undesired instabilities in the model. Next, the regional field h is reduced to a "slope parameter" (r), where $r = |h_{+1/2} - h_{-1/2}|/|h_{+1/2} + h_{-1/2}|$. ROMS employs a staggered Arakawa C-grid with the bathymetric values at the central position of each grid cell, such that the indices "1/2" indicate the value of bathymetry according to the half distance of each grid cell and its adjacent one. An iterative procedure is applied with the application of a modified Shapiro smoother on $\log(h)$ whenever $r > r_{max}$. This iterative step is required to prevent pressure gradient errors for the model integrations. After the iterations conclude, a second filter ($F2$) is applied to remove very steep gradients that could cause model integrations to either produce unrealistic fluxes or fail.
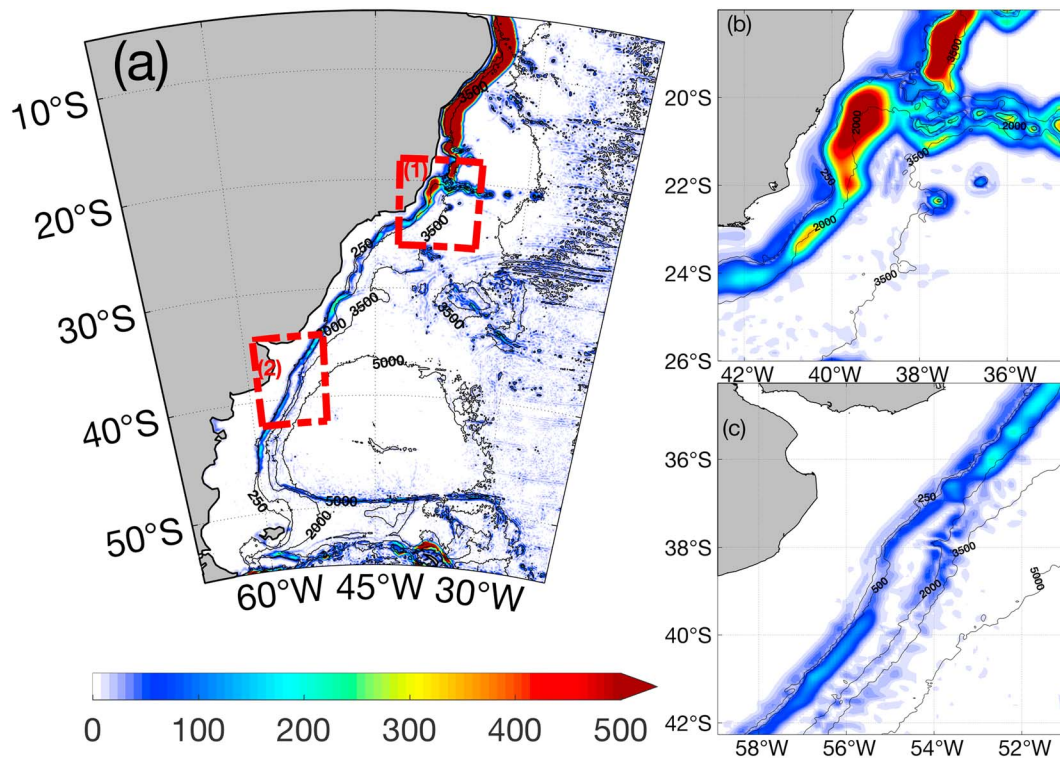
Figure 2. Standard deviation in meters (shade) estimated from the 28 members of perturbed bathymetries for the (a) model numerical domain and both (b) region 1 and (c) region 2 that are demarcated by red rectangles in (a). The black contours represent the bathymetric values in meters from the ETOPO field.

We generate ensemble perturbations of the bathymetry by randomly indicating a different number of passages to apply F_1 on each of the 28 different bathymetric fields. The parameter r_{max} is also randomly defined for each ensemble member such that the Shapiro smoother ends at different iterations for each ensemble member. A different number of passages is also randomly determined to apply F_2 on each ensemble member to finally create an ensemble of perturbed bathymetries.

Once generated, one 12-month ensemble experiment includes exclusively perturbations in bathymetry so that each perturbed field is held constant for its respective ensemble member over the entire experiment. The standard deviation field estimated from the 28-member ensemble of bathymetries exhibits the highest values on the regions with the highest vertical gradient such as over the continental slope and sparse seamounts on the deep ocean (Figure 2).

2.3. Statistical Metrics

Ensemble data assimilation methods typically use the ensemble mean as the best estimation of the ocean state. The ensemble spread is the standard deviation of the ensemble members from the ensemble mean, so it provides a quantitative estimate of the flow-dependent uncertainty. The ensemble mean and spread of a given variable x calculated from k members is

$$\overline{x} = \frac{\sum_{i=1}^{k} x_i}{k}, \tag{7}$$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{k} (x_i - \overline{x})^2}{k-1}}. \tag{8}$$

The error covariances are estimated from the ensemble of ocean variables and evolve in time incorporating flow-dependent patterns. If the ensemble mean is a reasonable approximation of the true state, then the following equation gives the error covariance for two model variables x and y:

$$C_{x,y} = \frac{1}{k}\Sigma_{i=1}^{k}(x_i - \overline{x})(y_i - \overline{y}). \tag{9}$$

The normalized version of the error covariance, the error correlation, is the ratio of covariance by the standard deviation of x and y. Let X be a vector containing all model variables such as the oceanic T and S, zonal (U), and meridional (V) current components, and free-surface (zeta), zonal (U_{bar}), and meridional (V_{bar}) barotropic current components. In data assimilation methods, the forecast error covariance matrix incorporates the estimates of the model error covariances that are calculated from the model variables as in equation (10). The forecast error covariance matrix is a square, symmetric matrix in which the diagonal terms are variances of the model variables and the off-diagonal terms are covariances between the model variables (Kalnay, 2003).

$$P^f = \frac{1}{k-1}\Sigma_{i=1}^{k}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)^T. \tag{10}$$

3. Results and Discussion

The following sections discuss the influence of atmospheric perturbations and bathymetric perturbations.

3.1. The Influence of Atmospheric Perturbations on the Ocean Simulations

We commence with an investigation of the ocean response to perturbations applied to different atmospheric variables. We examine the impacts 7 days after the initialization of the ensemble experiments, with all members starting from identical initial conditions. We first focus on the qualitative features of the resulting ensemble spread of each oceanic variable that can be attributable to each perturbed atmospheric variable. The regions where the perturbed atmospheric ensemble forcing obtains high spread (Figure 3) generally result in high values of spread for the ocean (Figure 4). The characteristics of the ensemble spread around the SWA region is dependent on how the ROMS bulk formulation uses each atmospheric variable to calculate the fluxes (equations (4)–(6)).

The largest values of ensemble spread in the ocean are generated when all atmospheric variables are perturbed (Figure 4a). The experiment that included perturbations on wind only produced similar results (Figure 4b). We thus infer that most of the impact produced in the ocean surface derives from uncertainty in the wind field. The experiments applying perturbations to either radiation fluxes or specific humidity generate relatively high values of ensemble spread for SST (Figures 4c and 4d), and we suspect the impact on SST may be larger in clear conditions. Unlike the other atmospheric forcing perturbations, precipitation perturbations produce a stronger impact on sea surface salinity (SSS) (Figure 4f); however, this is also affected by the presence of a large convective system at this time. The experiment with perturbed 2-m air temperature shows a significant impact along coast (Figure 4e) but not overcoming the impact of other perturbations as in the wind. This suggests that a longer integration time is needed for 2-m air temperature to impact offshore regions.

In January 2009, the climate monitoring bulletin CLIMANALISE (Climanálise, 2009) from the Center for Weather Forecast and Climate Studies (CPTEC) documented one South Atlantic Convergence Zone (SACZ) occurred from 4 to 8 January 2009 that can be characterized as an intense and well-configured oceanic SACZ case (Pezzi et al., 2016). Our results show a clear impact of this SACZ case on the oceanic ensemble simulations as revealed by the large SSS spread in Figure 4f. The date of 7 January was chosen to be analyzed because at this day, the SACZ was on its mature phase and well developed as an oceanic case. This large ensemble spread shows a clear impact on the ocean surface due to high spread of the precipitation along the SACZ region (Figure 3f). Those atmospheric perturbations in the SACZ region produce in the underlying ocean a NW-SE band of relatively high spread that reaches the South American continent almost perpendicularly, between 20°S and 30°S (Figure 4). The atmospheric perturbations produce a SACZ signal in the underlying ocean depending on the perturbed atmospheric variable. The experiments that include perturbations in either radiation fluxes or precipitation produce higher spread, respectively, for T (Figure 4c) and S (Figure 4f), exceeding 0.2 °C and 0.08. Because of the momentum transfer at surface, wind perturbations have the largest impact on ocean surface currents (Figure 4b), with spread greater than 0.1 m/s.

Next, we concentrate our attention on the experiment with perturbations applied to all atmospheric fields. Figure 5 shows the ensemble average and spread calculated from the main flux components of the heat
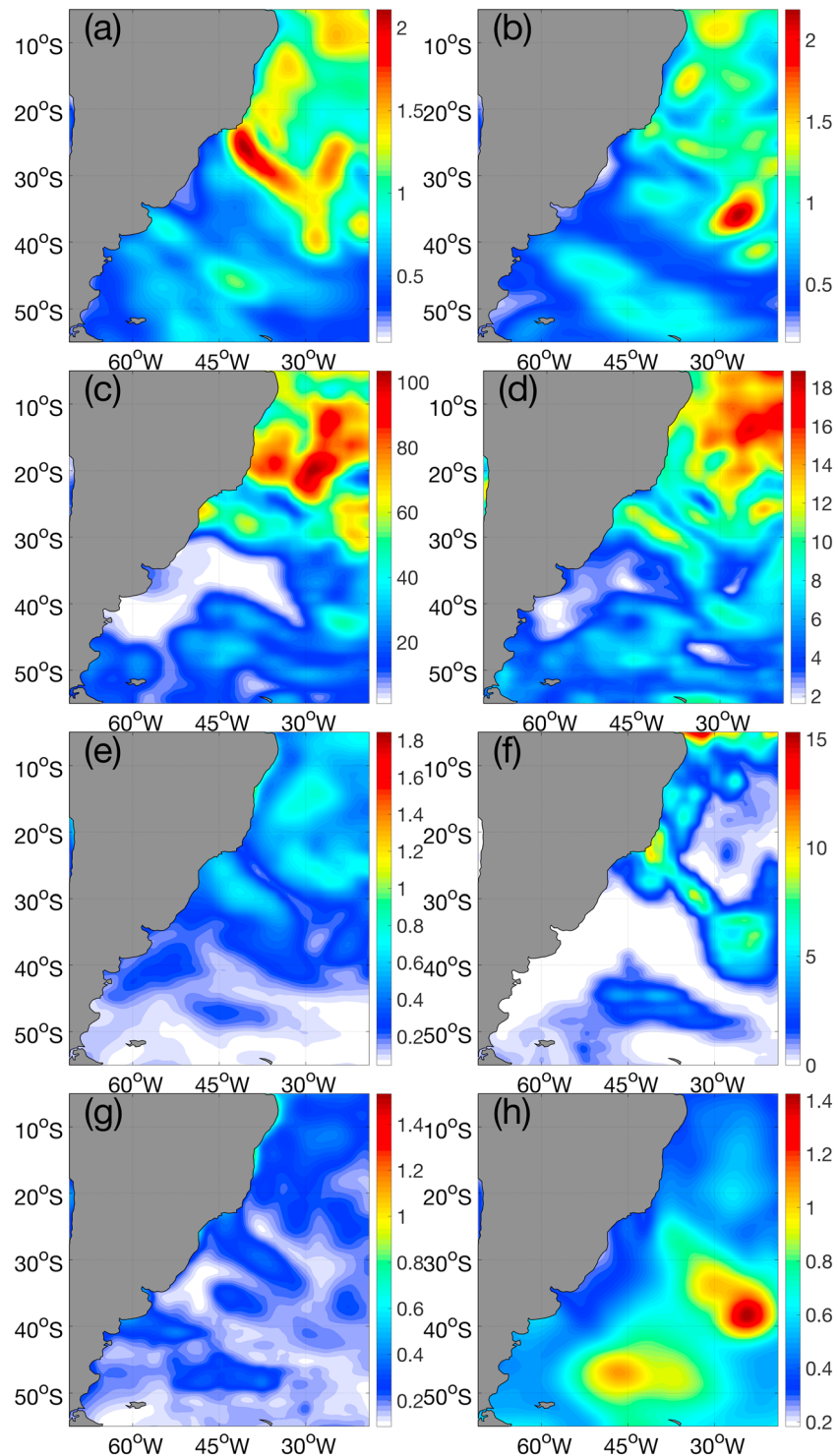
Figure 3. Ensemble spread fields of the (a) wind v-component (m/s), (b) wind u-component (m/s), (c) net shortwave radiation (W/m^2), (d) downward longwave radiation (W/m^2), (e) 2-m specific humidity (g/kg), (f) precipitation (mm/day), (g) 2-m air temperature (°C), and (h) pressure (mbar) on 7 January 2009.

balance at the ocean surface like the surface net heat flux (Figures 5a and 5b), perturbed downward shortwave radiation used to force ROMS (Figures 5c and 5d), and the derived fluxes of the model bulk formulation such as the net longwave radiation (Figures 5e and 5f) and the net sensible heat (Figures 5g and 5h) as well as the net latent heat (Figures 5i and 5j). We emphasize the SACZ region (black trapezium in
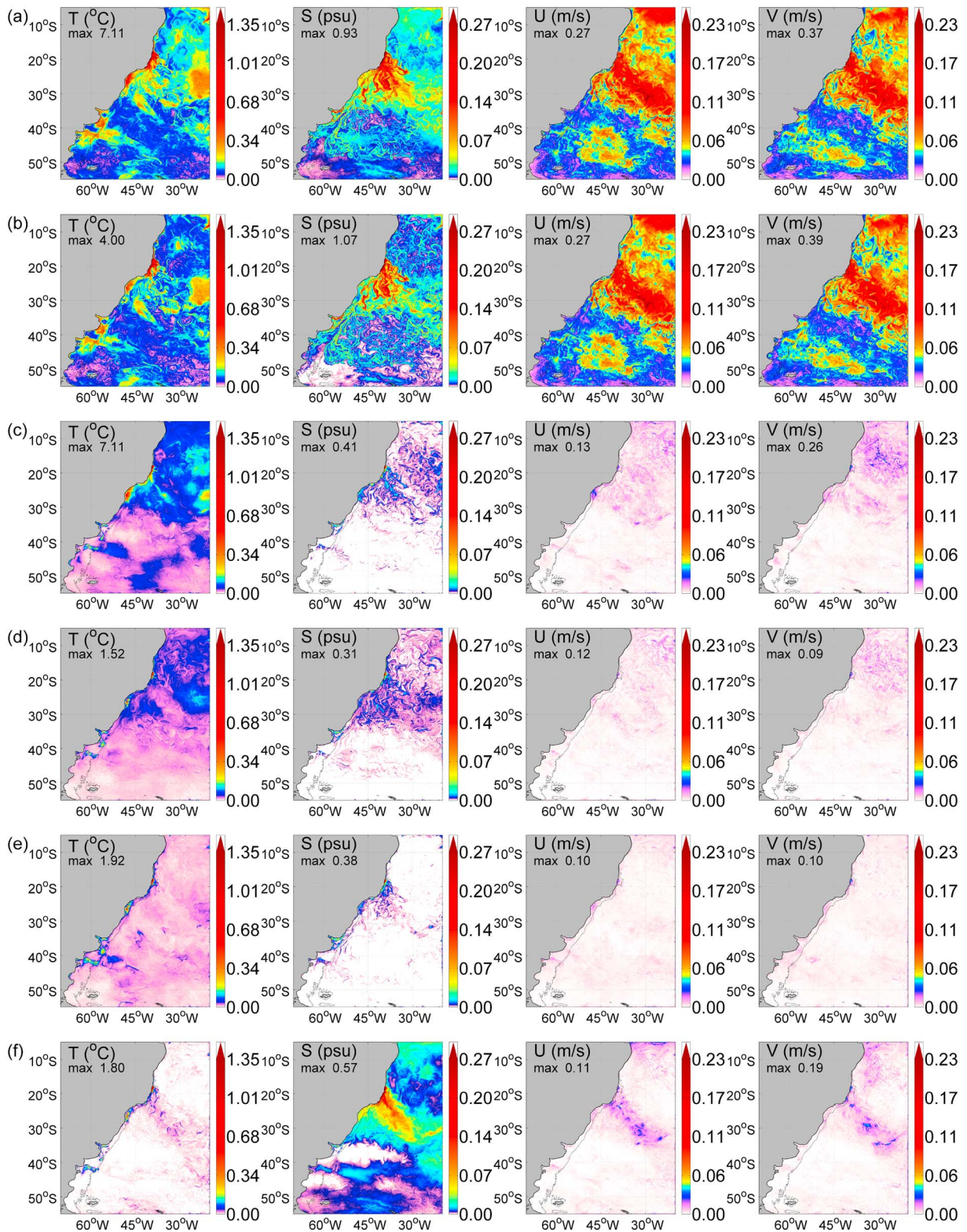
Figure 4. Ensemble spread at the surface of prognostic ocean model variables, sorted by impact on sea surface temperature. These results were computed on 7 January 2009 for the ensemble experiments with perturbations applied to (a) all forcing, (b) wind components, (c) longwave and shortwave radiation fluxes, (d) 2-m specific humidity, (e) 2-m air temperature, and (f) precipitation.
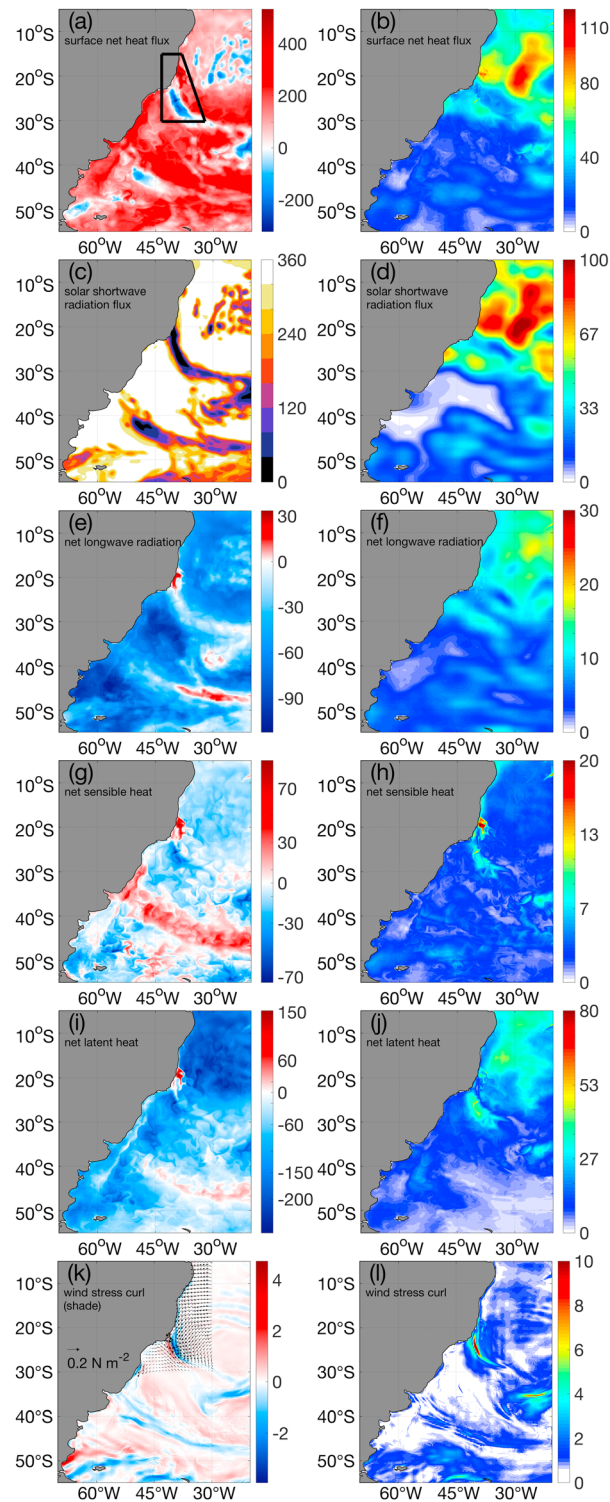
Figure 5. The left panels are the ensemble averages and the right panels are the ensemble spreads of surface net heat flux in watts per square meters (a and b), solar shortwave radiation flux in watts per square meters (c and d), net longwave radiation flux in watts per square meters (e and f), net sensible radiation flux in watts per square meters (g and h), and net latent radiation flux in watts per square meters (i and j). Next, (k) shows the ensemble average of wind stress curl flux ($\times 10^{-6}$ N/m^3) in which the blue (red) color indicates upwelling (downwelling), and the overlapping vectors are the ensemble mean of wind stress (N/m^2). (l) shows the ensemble spread of wind stress curl ($\times 10^{-7}$ N/m^3). The black trapezium in (a) indicates the South Atlantic Convergence Zone region. All results were computed on 7 January 2009 from the experiment with perturbations applied to all atmospheric variables.

Figure 5a) because the areas of high spread are coincident for both atmospheric and oceanic variables, as shown in previous figures, and due to its importance for producing precipitation in some areas of South America during the austral summer. Upward fluxes (negative values) of sensible heat, latent heat, and net longwave radiation are present in the CFSR according to the model bulk formulae. The solar shortwave radiation, on the contrary, has only positive values, indicating ocean heating; though a band of reduced downward shortwave radiation associated with the SACZ is apparent. Cloud cover present during SACZ episodes has been recognized as an important mechanism that reduces incoming solar radiation in the ocean (Chaves & Nobre, 2004; De Almeida et al., 2007). Because of the heat balance at surface, there is an upward net heat flux (i.e., an ocean cooling; Figure 5a). Considering the entire SWA domain, the ensemble spread of the heat flux seems to follow the pattern of the perturbed solar shortwave radiation spread in the SWA (Figure 5d), like in the north region, with the exception of the one computed for the sensible heat (Figure 5h). However, if we look only at the SACZ region, the ensemble spread patterns of the sensible and latent heat (Figures 5h and 5j) do not follow the pattern of the solar shortwave radiation.

Figure 5k shows the ensemble average of the wind stress curl (shade) and wind stress (vector), and Figure 5l exhibits the spread of wind stress curl, all calculated from the perturbed wind components. The ensemble average suggests that the presence of SACZ alters the wind pattern commonly seen in the summer by altering the direction of the northeasterly winds flowing parallel to the Brazilian coast. Castelao and Barth (2006) reported a common coastal upwelling due to the northeasterly winds, but our results suggest that the alteration of the wind direction may have reduced the intensity of this phenomenon. Our results show two regions where the ensemble average of the wind stress curl exhibits different patterns: immediately north and underlying the SACZ (Figure 5k). North near the coast, negative values of wind stress curl indicate the occurrence of upwelling. Instead, the wind stress curl is positive underlying the SACZ, indicating that the upwelling is ceased. Chaves and Nobre (2004) also mentioned alterations of Ekman pumping in the Brazilian coast under an occurrence of the SACZ.

We compute error correlations (equation (9)) between SST and atmospheric variables that participate in the heat and momentum flux calculation at the ocean-atmosphere interface (Figure 6). Due to an upward heat flux, positive correlations larger than 0.5 are found between the resultant heat balance and SST in the SACZ area (Figure 6a). This indicates a tendency of ocean cooling, with decreasing SST. Examining the same region, the error correlations computed between SST and solar shortwave radiation flux (Figure 6b), meridional wind stress (Figure 6c), and wind stress curl (Figure 6d) suggest that the perturbed wind fields have a more significant influence on SST than the perturbed solar shortwave radiation flux. The ensemble spread of wind stress curl exhibits higher values in the SACZ area (Figure 5l), which may determine the peculiar spatial shape of the sensible and latent flux spreads (Figures 5h and 5j) along the SACZ rather than the perturbed solar shortwave radiation flux. From experiments using the breeding method, Hoffman et al. (2009) established a conversion of kinetic to potential ocean eddy energy and suggested that the growing instabilities found in this area are primarily forced by winds.

### 3.2. The Bathymetric and Atmospheric Influences on Ocean

We next examine uncertainty in the ocean state due to uncertainty in the representation of bathymetry in the ocean model. Figure 7 presents the spatial average over time of the ensemble spread of T, S, ζ, and SST at the observation locations in order to assess the impact in the ocean model from perturbations applied at surface and bottom interfaces via atmospheric forcing and bathymetry, respectively. We use T and S profiles from the Pilot Research Moored Array in the tropical Atlantic (Bourlès et al., 2008), expendable bathythermographs (XBTs) (Cheng et al., 2016), and Argo (Riser et al., 2016). Unlike moorings and XBTs that make measurements at fixed points, Argo floats flow along with the currents (typically at 1,000 m), so their positions vary over time and are fairly distributed around the SWA. In order to assess the ζ and SST ensemble spread, we choose, respectively, the locations of the absolute dynamic topography (ADT; ftp://ftp.aviso.oceanobs.com/global/delayedtime/along-track) and SST (https://data.nodc.noaa.gov/pathfinder/Version5.2/; Casey et al., 2010) observations of the ocean surface derived from remote sensing, which have relatively good coverage since at 1/12-degree model resolution, the satellite tracks cover most of the SWA in an interval of a few days. We note that ROMS uses the Boussinesq approximation, in which the model surface elevation, represented by the variable ζ, does not consider the steric effect. However, this real effect is measured by the altimetry
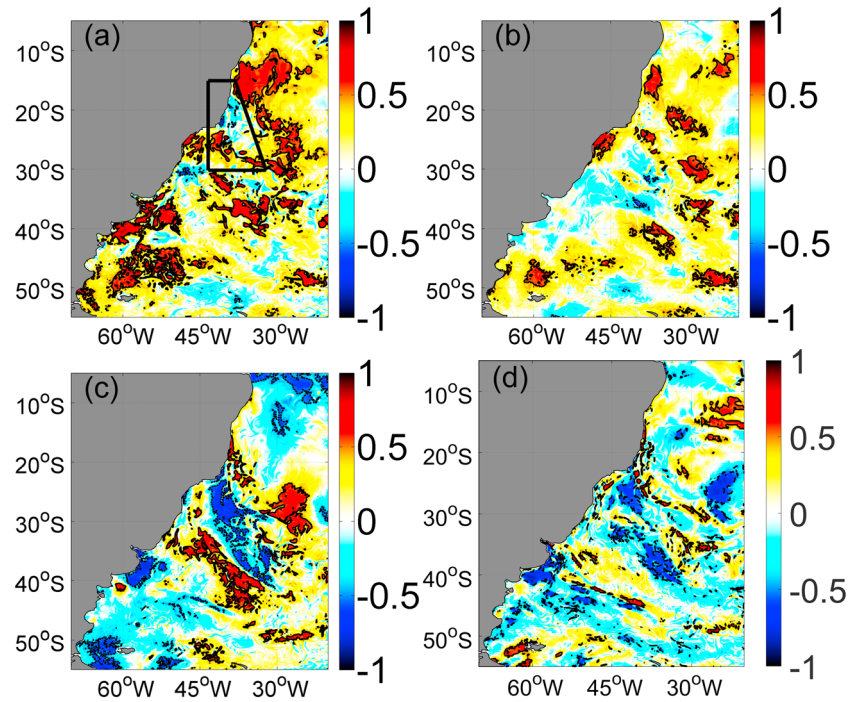
Figure 6. Ensemble error correlation computed between sea surface temperature and (a) surface net heat flux, (b) perturbed solar shortwave radiation flux, (c) meridional wind stress component, and (d) wind stress curl on 7 January 2009.

such that it is considered for the ADT measurements. We do not perform any direct comparisons of the model ζ and observed ADT.

Considering the depth-averaged ensemble spread computed at the T (Figure 7a) and S (Figure 7b) vertical profile locations, the experiments with perturbations applied on all atmospheric forcing or the bathymetry
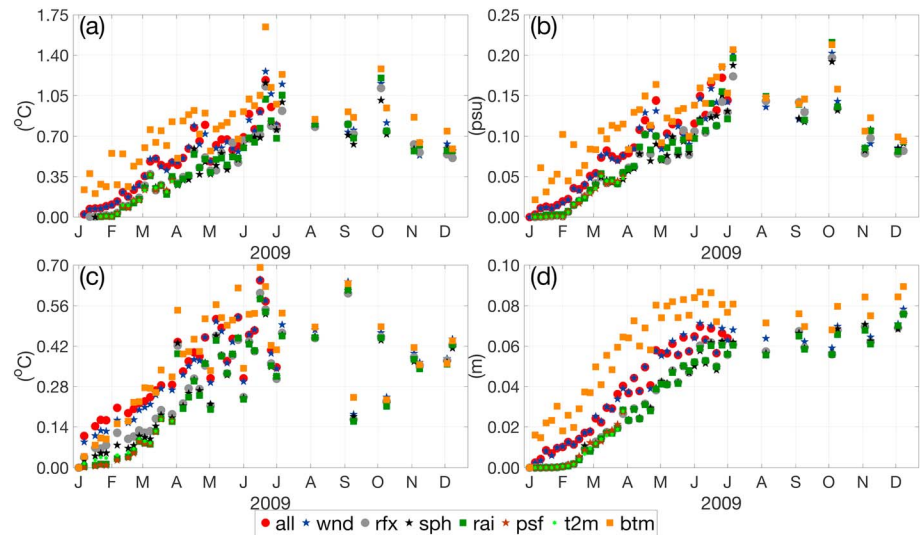
Figure 7. The spatial average of the ensemble spread at the observations locations of (a) T vertical profiles, (b) S vertical profiles, (c) SST, and (d) ADT. For the vertical profiles, (a) and (b) also consider the depth-averaged values of ensemble spread. Each symbol represents a different ensemble experiment in which perturbations were applied in all forcing fields (all), wind (wnd), radiation fluxes (rdx), specific humidity (sph), rain (rain), pressure field (psf), air temperature (t2m), and bathymetry (btm). The latter half of the results are sparse due to reduced storage of model output. S = salinity; T = temperature; SST = sea surface temperature; ADT = absolute dynamic topography.
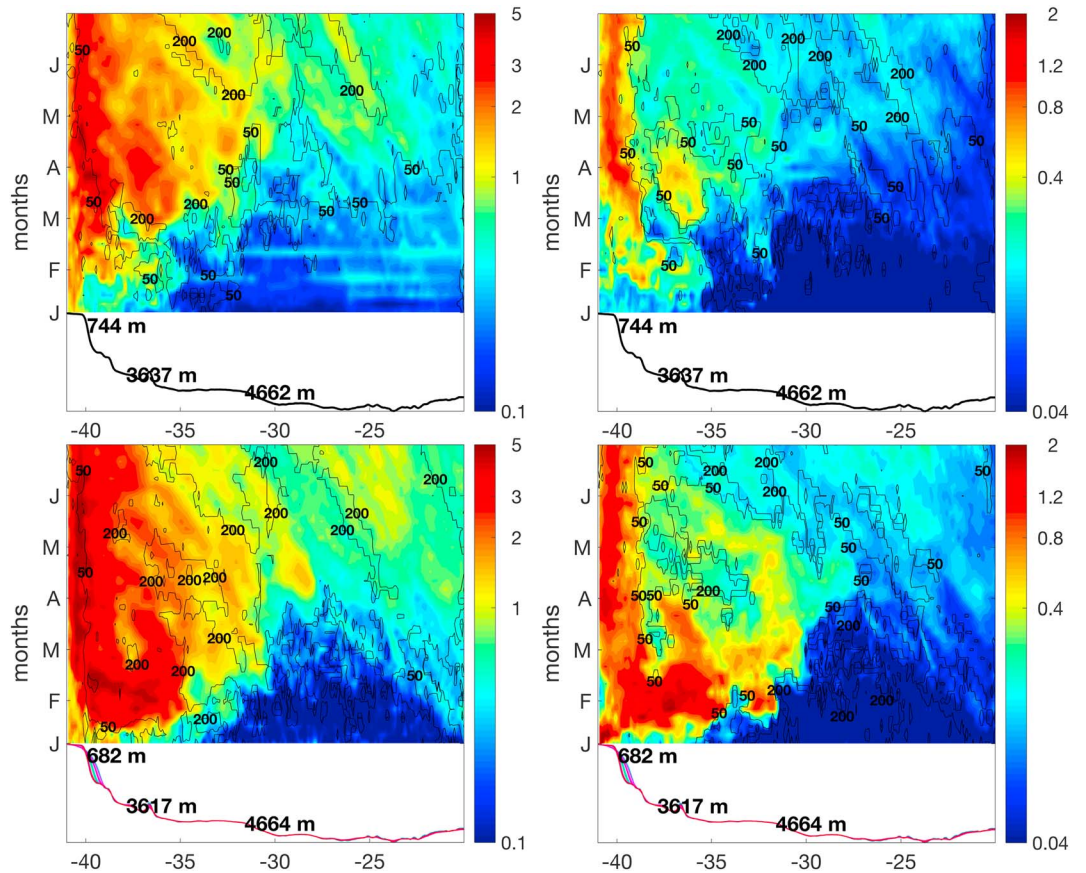
Figure 8. Hovmöller (longitude versus time) diagrams at 22°S showing the maximum spread value in the vertical water column for the temperature in degrees Celsius (left panels) and salinity (right panels) derived from the first 6 months of the experiments including perturbations in all atmospheric forcing (top panels) and bathymetric (bottom panels) fields. The black contours indicate the isolines of the depths in which those maximum spread values were found, and the color curves on the bottom exhibit the bathymetry at the chosen latitude.

present relatively close values over time. The T and S seem to be slightly more sensitive to the bathymetric perturbations. However, the perturbations of the bathymetry are unrelated to the perturbations from the 20CR and cannot be meaningfully calibrated to make a direct comparison. The spread increases continuously until June and July, reaching values of approximately 1.5 °C (0.2) for T (S). In some cases, S is more sensitive to perturbations in precipitation such as at the beginning of August and October. This illustrates the anticipated importance of the precipitation to determine the ocean S variability. In the first 2 months, the spatial average of the SST spread exhibits higher values when perturbations are applied to all atmospheric forcing fields or only on the wind fields (Figure 7c). In contrast, the SST spread has a lesser response to the perturbations in bathymetry. Bathymetric, wind, and all-atmospheric perturbations produced the highest SST spread from March to June, exceeding 0.6 °C in June. During this time, specific humidity perturbations and precipitation perturbations produce the lowest spread. Perturbed radiation fluxes generated intermediate and sometimes high values of SST spread. The highest values of ζ spread occurred due to bathymetric perturbations, exceeding 0.08 m near May (Figure 7d). Considering only atmospheric forcing, the largest spread occurred due to perturbations applied either to all atmospheric fields or only to wind fields, reaching almost 0.08 m late in December. The other atmospheric perturbations have a lower impact on ζ. The ζ spread reaches values comparable to those found for surface elevation by Counillon and Bertino (2009), who used a 10-member ensemble experiment with the HYbrid Coordinate Ocean Model in the Gulf of Mexico.

Hovmöller diagrams (longitude versus time) show the maximum spread in the vertical water column (Figure 8). For the experiment of perturbed bathymetry (bottom panels in Figure 8), both T and S diagrams identify a faster increase of spread in shallow waters on the continental shelf and slope. As each member
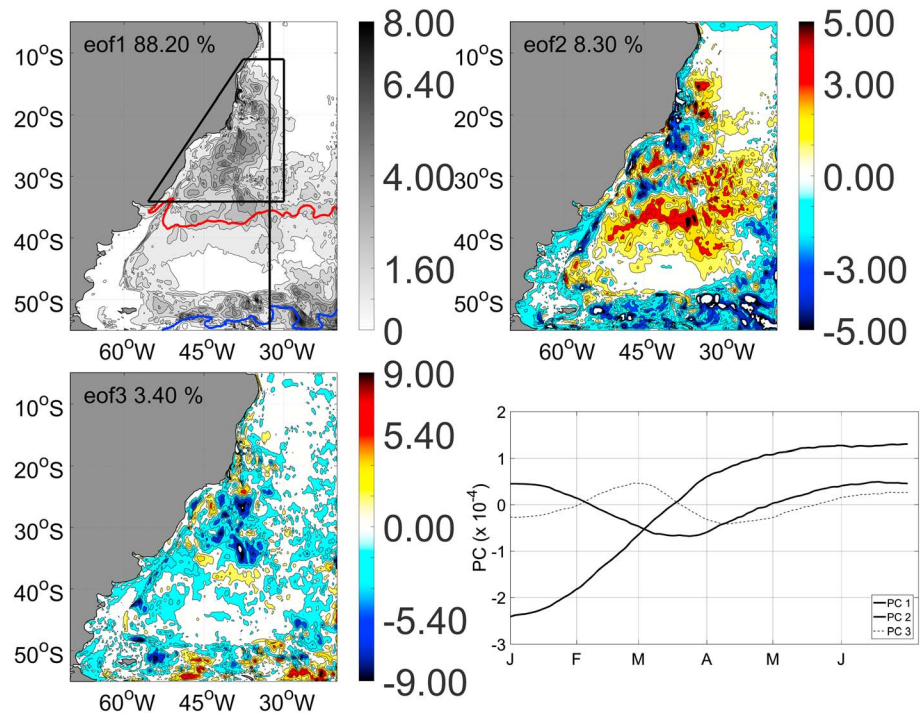
Figure 9. EOF1–EOF3 of the daily spread calculated for the depth-integrated acceleration term obtained from the experiment with perturbations in all forcing. The curves are the time series of the first three principal components (PC) that are related to each EOF. The black trapezium encompasses the region for which the spread of the depth-integrated momentum equation terms is exhibited in Figure 10. The 20 (red line) and 5 °C (blue line) isotherms represent the warm and cold oceanic fronts on 1 July 2009. The meridional line (black line) is the section in 33°W that is showed in Figure 11. EOF = empirical orthogonal function.

starts to be integrated with a perturbed bathymetry, there is an initial adjustment (via rapid waves) to the change in bathymetry. During the ensemble experiment, the ocean response to such adjustment can be viewed by looking at the highest values of maximum spread for S between 37.5°W and 35°W in February 2009. Next, the impact from perturbed bathymetry is quickly dispersed during March 2009, whereas the maximum spread values of S decrease. After April 2009, the remaining spread comes from changes in the circulation near the coast and is also linked to the model internal dynamics. A slower increase is shown on deeper waters. Perturbing all atmospheric forcing also produces a significant impact over shallow regions (top panels in Figure 8) but to a lower extent and with lower values for both T and S. Examining the isolines of depth, the maximum spread reaches larger depths more rapidly under bathymetric perturbations. The occurrence of structured patterns in maximum spread that extend in the NW-SE direction is an indication of flow-dependent uncertainty (i.e., the maximum values propagate from east to west with the ocean circulation).

Next, we investigate the impact on the ensemble of the depth-integrated momentum equation terms (equations (1) and (2)) caused by the atmospheric and bathymetric perturbations. The meridional and zonal components of this equation behave similarly. We present the empirical orthogonal function (EOF) for the daily 2-D spread of the meridional acceleration term (first term of the left-hand side of the equation (2)) considering the period from 1 January to 30 June 2009 (Figure 9). The first EOF (EOF1) captures by far the largest part of the total variance, about 88% (Table 1). The two types of perturbations exhibit a similar EOF1 pattern, so we only show the structure of EOF1 from the experiment with the perturbed atmosphere. The EOF1 shows a stronger signal adjacent to the continental shelf break along the east coast of Brazil (see black trapezium in Figure 9), suggesting an association with the BC pathway. There is also a large signal related to the ocean frontal and BMC regions where high mesoscale eddy activity has been reported (Mason et al., 2017). The EOF1 exhibits no signal around the northeast of
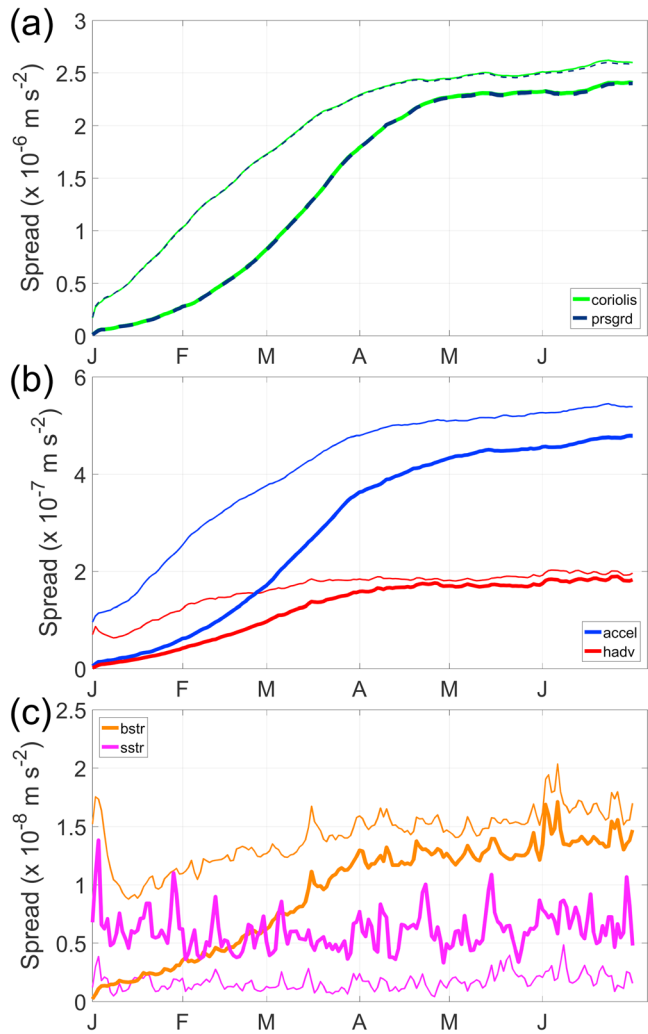
Table 1
Percentage That Each EOF Explains the Total Variance of the Meridional Acceleration Term Spread Over Time From the Ensemble Experiments With Perturbations in All Forcing and Bathymetry

EOF	All forcing	Bathymetry
1 (%)	88.20	80.20
2 (%)	8.30	13.40
3 (%)	3.40	6.20

Note. EOF = empirical orthogonal function.

Figure 10. Spatial average of the spread computed for the region demarcated by the black trapezium in Figure 9. The thin (thick) curves correspond to the spread obtained from bathymetric (atmospheric) perturbations for each term of the meridional component of depth-integrated momentum equation: (a) $fD\overline{u}$ (Coriolis term—dashed green) and $D\frac{\partial P}{\partial y}$ (pressure gradient—dark blue), (b) $\frac{\partial(D\overline{v})}{\partial t}$ (acceleration—clear blue) and $\frac{\partial(\overline{uD\overline{v}})}{\partial x} + \frac{\partial(\overline{vD\overline{v}})}{\partial y}$ (horizontal advection—red), and (c) τ_{by} (bottom stress—orange) and τ_{sy} (surface stress—pink).

the SWA, which reflects the small ensemble spread of the meridional acceleration term (not shown) and can also indicate small spread values for the other model state variables since the region is dominated by the relatively low variability circulation of the eastern branch of the South Atlantic gyre not very much affected by the BC system. Due to our regional model configuration, this area may require perturbations applied to the lateral boundary conditions in order to increase the spread to adequately model uncertainty in the ocean state. The EOF1 pattern has one sign in the whole domain, showing that the spread of meridional acceleration evolves as a monopole in the region. Its principal component increases until it saturates in the end of April (Figure 9). Both EOF2 and EOF3 have more complex structures so that their combination with EOF1 and with their respective principal components practically cancels the spread of meridional acceleration in the beginning of the integration. After March, EOF1 completely dominates the other modes.

We now extend the evaluation to other terms of the meridional depth-integrated momentum equation: Coriolis force, pressure gradient, horizontal advection, surface stress, and bottom stress (Figure 10). The spatial mean of the estimated spread for the region demarcated by the black trapezium in Figure 9 shows that most of the terms exhibit values with a similar growth pattern as that presented by the principal component 1 of the meridional acceleration spread over time (Figure 9). Because of the geostrophic approximation, the ensemble spread of the Coriolis force and pressure gradient present very close averaged values, so that their curves overlap over time (Figure 10a). The estimated spread values from the experiment of perturbed bathymetries increase faster and present higher values, except for the surface stress term (Figure 10).

Figure 11 evaluates the ensemble spread of T on different dates along 33°W. It presents a rapid increase under bathymetric perturbations, like was reported for S and T at 22°S (Figure 8) and the depth-integrated momentum equation terms (Figure 10). The vertical water column over some seamounts, for example, around 10°S, exhibits values greater than 0.9 °C, which is not obtained with the experiment of perturbed atmosphere on 7 January 2009. Atmospheric perturbations concentrate their impact on the ocean surface and only generate comparable values of spread to that obtained from bathymetric perturbations on 6 February 2009 in some regions above 500 m. Even so, the atmospheric perturbations do not generate ensemble spread exceeding 0.9 °C. Perturbing the bathymetry generates larger ensemble spread within the thermocline layer. However, both atmospheric and bathymetric perturbations result in similar characteristics in the ensemble spread after 6 months of integration (e.g., on 1 July 2009), when the highest values of the ensemble spread concentrate in the thermocline and the ocean zone around the warmer ocean front, which is located in the range from 35°S to 30°S. However, even at 6 months, the bathymetric perturbations produce a slightly larger impact. The ensemble spread derived from other experiments that apply perturbations to a single atmospheric forcing field exhibits an analogous pattern (not shown) on 1 July 2009. This common pattern suggests a dominance of the internal model dynamics on the ocean state uncertainty after a few months of model integration, regardless of the source of perturbation.

### 3.3. The Investigation of Error Covariances in the Ocean Ensemble

To investigate error covariances in the ocean ensemble, vertical and horizontal correlations are computed between surface and subsurface model fields derived from the 28-member ensemble experiments. This investigation is in support of future data assimilation experiments using ensemble-based methods.
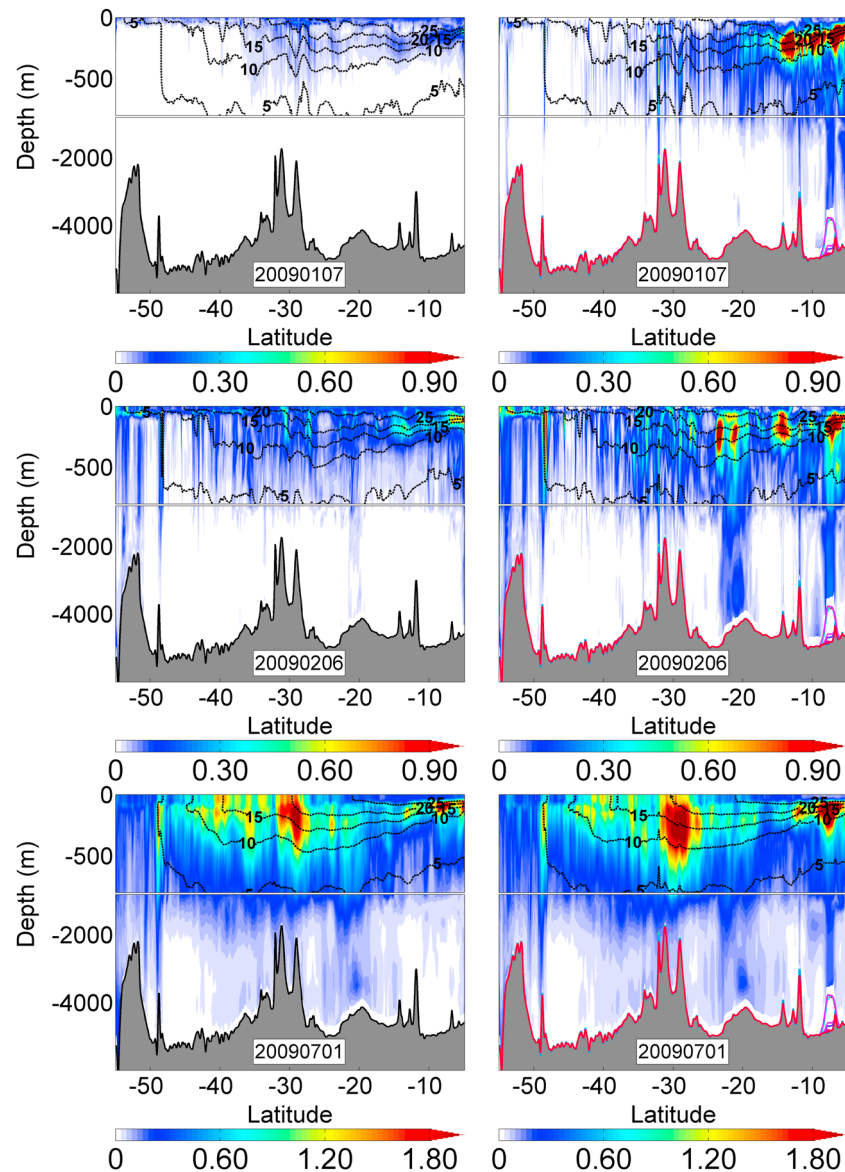
Figure 11. Temperature spread (in °C) along 33°W for the experiment with perturbations in all atmospheric forcing (left) and bathymetry (right) on 7 January 2009 (top), 6 February 2009 (middle), and 1 July 2009 (bottom). The ensemble distribution of bathymetry is shown by different colors in the right column.

In the vertical, there are minor differences between the error correlations from the ensemble of bathymetric or atmospheric perturbations. For both cases, SST error correlations with respect to the subsurface T are highest above the mixed layer depth (MLD; Figure 12). At 50-m depth, error correlations reach values close to 1 over almost the entire numeric domain (left panels in Figure 13). Error correlations decrease at 100 m; the main reduction appears approximately northward of 25°S, where the MLD values are shallower than 100 m. There are not large differences in correlations due to the two types of perturbation, but bathymetric perturbations appear to produce slightly larger correlations at deeper levels (e.g., northward of 25°S at 100-m depth). SST error correlations with the subsurface S (not shown) exhibit similar patterns, in which the error correlations are largest toward the surface. However, regions of higher values are more limited, and there are also large areas of poor error correlations between SST and S.

Error correlations between ζ and T at depth show a different behavior (right panels in Figure 13). In this case, it appears that there is not a dependence on MLD as was mentioned above in the SST and T error correlations.
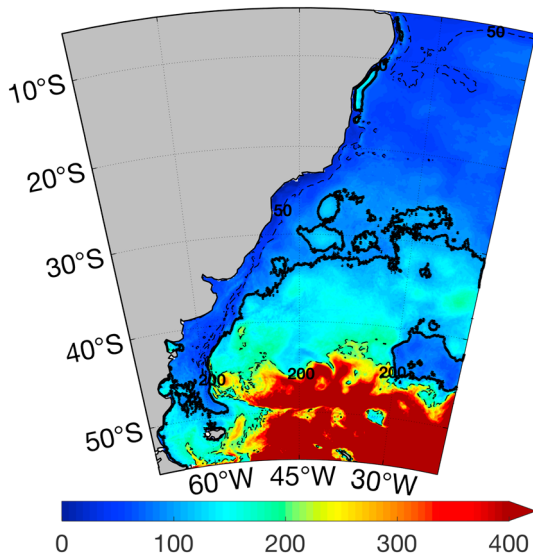
Figure 12. Ensemble mean of the MLD (in meters) on 1 July 2009 computed from the temperature results of the experiment that applied perturbations in all atmospheric forcing and based on the criterion in which the temperature changes by 0.5 °C relative to the temperature of the shallowest model level. The black contours are the MLD isolines of 50, 100 (thick line), and 200 m. The MLD derived from the experiment with bathymetric perturbations presented a similar pattern on this date. MLD = mixed layer depth.

Rather, the error correlations increased toward the thermocline depths. The ζ error correlations with S (not shown) also reached higher values at 100 and 200 m than that found at 50 m, but they were lower as compared with that obtained between ζ and T. Small changes in the surface often correspond to larger changes in the thermocline depth. We believe the error correlations between ζ and T and S at depth can be attributed to the placement of the thermocline and halocline, respectively. Large uncertainty in the positioning of these sharp vertical gradients will lead to corresponding uncertainty in the sea surface height. Assessing the impact of different perturbations, the atmospheric forcing fields generated slightly larger values of error correlations between ζ and subsurface fields.

To illustrate horizontal error correlations of T, we examine two specific points located in the BC and MC regions, respectively, at 38.6°W to 22.0°S and 54.9°W to 38.3°S (Figure 14). The error correlations are computed between the SST of each point with respect to the SST of surrounding points and the model T in the subsurface levels.

In the BC (Figures 14a–14h), correlations are larger than 0.5 surrounding a small region at different depths of the evaluated point. Both atmospheric and bathymetric perturbations produced similar correlation patterns in the shallowest levels. However, atmospheric perturbations generated larger correlations in more distant regions from the evaluated point at deeper levels such as 100 and 200 m. At these depths, atmospheric perturbations also have high correlation extending along the ζ isolines, which indicates a dependence on the BC flow. Calado et al. (2010) noted

a generation of meanders and eddies in the BC at 22°S. These features can generate complexity in the error correlation within the area around the point located at 38.6°W to 22.0°S and may require a larger ensemble to be adequately resolved.
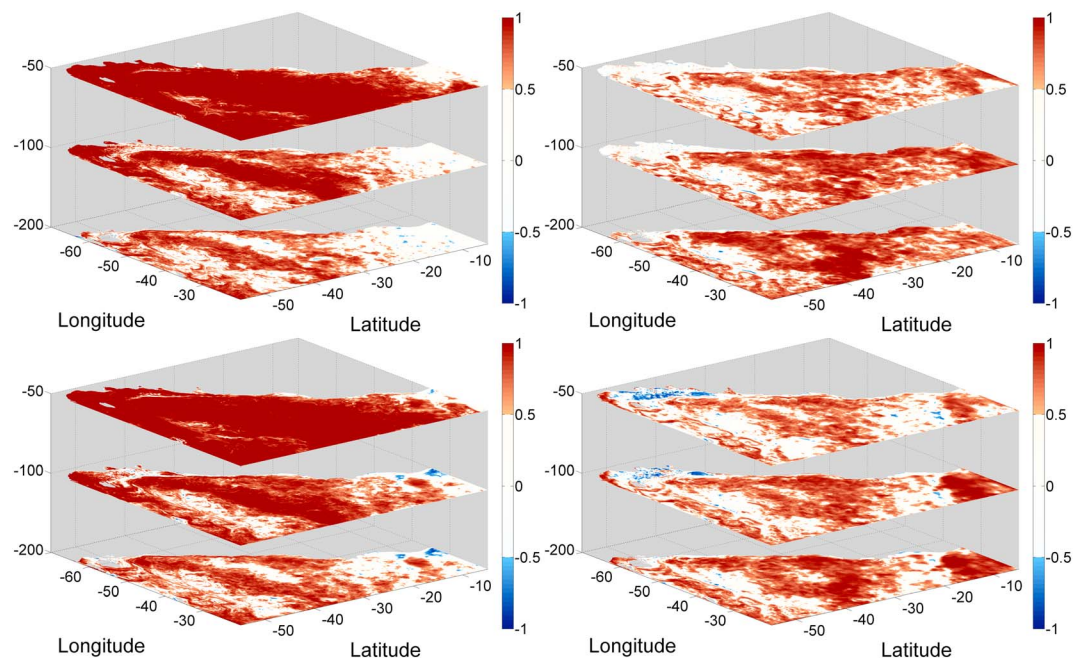
Figure 13. Correlations derived from an ensemble of atmospheric (top panels) and bathymetric (bottom panels) perturbations. Statistics are computed for the sea surface temperature (left panels) and ζ (right panels) with respect to the temperature in 50, 100, and 200 m after 6 months of model integration (on 1 July 2009).
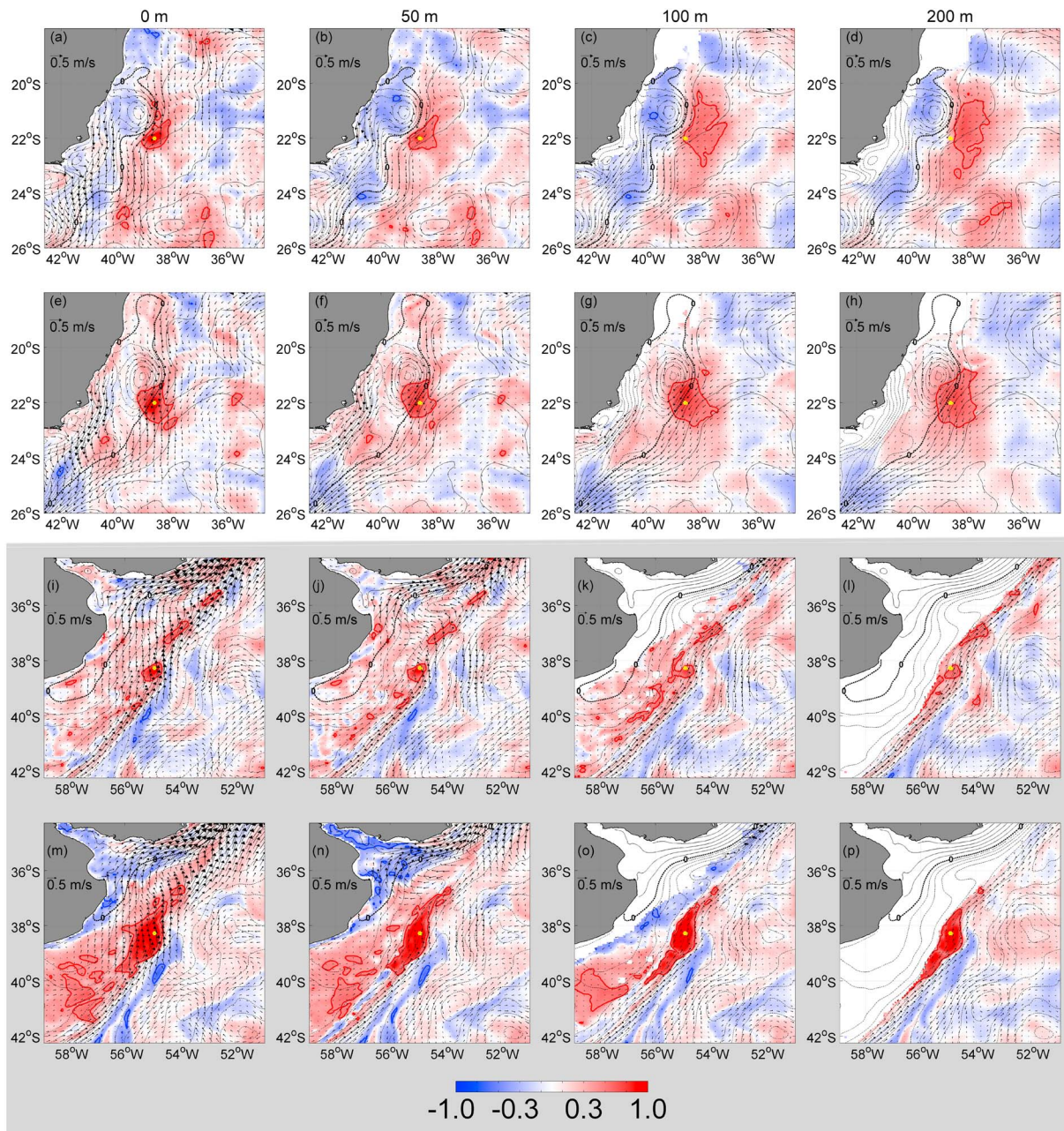
Figure 14. Correlations after 6 months of model integration (on 1 July 2009) between sea surface temperature and the surrounding points of the temperature at different depths for the ocean ensemble derived from atmospheric (a to d; i to l) and bathymetric perturbations (e to h; m to p). Analyses were done in two points located at 54.9°W to 38.3°S and 38.6°W to 22.0°S (yellow dots) over the Brazil Current (white background) and Malvinas Current (gray background) regions, respectively. See the correlations greater than 0.5 (red lines), lower than −0.5 (blue lines), and the ζ contours (black lines, with the isoline of 0 m thicker). The vectors represent the flow of the ensemble mean ocean currents on that corresponding date.

For the MC (Figures 14i–14p), the SST error correlations exhibit flow dependency. The regions of higher error correlation follow the current flow toward the north, extending latitudinally along its pathway. Comparing the source of the perturbation, the bathymetric disturbances generate high error correlation in a large region surrounding the study point, while atmospheric perturbations generate high error correlations in a relatively small region, though significant values are also found along the current pathway in northern regions. The error correlations between SST and T at depth of surrounding grid points are relatively large and follow

the direction of the subsurface current. This behavior is related to the barotropic flow that is inherent to the MC (Piola et al., 2013) and is characterized by a flux following the same pathway in different vertical levels.

The flow-dependent patterns of error correlations mentioned above are important for DA methods that utilize ensemble-derived error covariance estimates. Advanced methods like the ensemble Kalman filter evolve these patterns during the assimilation process. For each cycle, the ensemble of forecasts provides a time-dependent estimate of the forecast error covariance matrix. State-of-the-art methods currently merge this dynamically varying forecast error covariance estimate with a climatological average in order to account for model bias and mitigate the risks of ensemble degeneracy (e.g., Penny et al., 2015).

4. Summary and Conclusions

Ensemble experiments were performed to illustrate the influences of various sources of perturbations on the ocean model simulation. This was made perturbing the surface fluxes and bathymetry in order to investigate uncertainties in the modeled ocean, which are relevant to ensemble data assimilation. The results indicated that uncertainty in wind forcing plays a major role in the determination of uncertainty in the ocean state. Perturbations applied only to atmospheric wind fields produced an ocean ensemble spread quite similar to that acquired when all atmospheric variables were perturbed. Most of this effect is dictated by the importance of winds as used in the atmospheric surface model formulation (bulk model) for the computation of heat and wind stress fluxes. The results indicated a strong influence of the SACZ on the underlying ocean. In this oceanic region, the analysis of heat fluxes indicated that wind perturbations might have a greater congruence with the SST than other atmospheric variables like the radiation fluxes. The results suggest that it is preferable to apply the perturbations to all atmospheric fields used in the bulk flux calculations in order to achieve an appropriate estimate of uncertainty due to inaccuracies in the atmospheric fields. Alternatively, in the absence of a balanced set of atmospheric perturbations from a source such as the 20CRv2c used here, the application of only wind perturbations might be enough.

Other individual perturbations in different forcing fields provoked a minor impact, by significantly affecting only one ocean variable. The air temperature, radiation fluxes, and specific humidity perturbations generated a larger impact on the SST than on other ocean variables at surface. Perturbations to the precipitation highly impacted the SSS, as was shown in the evaluation of S spread at the S profile locations.

Compared to atmospheric perturbations from reanalysis, bathymetric perturbations via slight adjustments to the model bathymetry generation algorithm produced a larger impact on the ensemble of model simulations, especially on shallow waters, where the dynamics is largely dominated by barotropic processes. The ocean above the continental shelf and continental slope experienced the highest impacts, though this may also be in part due to excited waves at the initial time. The impact was minor in the deep ocean, by a slower evolving spread during the integration. Patterns were identified showing an evolution of the spread following specific directions along the same depth, indicating that internal model dynamics influenced the evolution of the spread for the ocean variables.

The impact on the ensemble spread of the modeled ocean obtained from either perturbing the atmosphere or bathymetry was quite similar after a few months. For instance, these experiments produced comparable ensemble spread as verified on 1 July 2009 (Figure 11), after 6 months of integration. This suggests that the model internal dynamical instabilities may dominate over these longer time scales. When applying data assimilation, all members are drawn closer to the mean at every cycle (e.g., every 1 day). As a result, the time scales of the growing ensemble spread will behave more like the early part of our experiments, where uncertainty in external forcing dominates. Nevertheless, we recommend further examination perturbing other model imperfections. We planned to include perturbations for the La Plata river discharge, but the difference between the integrations with and without river showed that its plume was limited to the continental shelf. It is possible that perturbations to the river discharge in other parts of the world may have varying impacts on the uncertainty of the general ocean circulation. We further suggest applying perturbations to the lateral open ocean boundary conditions in regional model configurations in order to increase the spread in the inner domain, mainly in the northeast region of the SWA.

A study of the ocean ensemble allowed the extraction of useful information from error correlations between the ocean variables. In the vertical, SST and subsurface T showed good agreement within the MLD, increasing

in correlation toward the surface. In contrast, ζ and T showed higher correlation values going down to deeper levels toward the thermocline, at least until 200 m. Furthermore, horizontal error correlations identified patterns that extended to deeper levels following the pathway of the BC and MC. These flow-dependent patterns are important characteristics for DA experiments based on ensemble-derived error covariance matrices since these methods aim to represent the real uncertainties derived from the model state, estimating the errors of the day.

References

Amante, C., & Eakins, B. W. (2009). ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24, (March), 19. https://doi.org/10.1594/PANGAEA.769615

Bourlès, B., Lumpkin, R., McPhaden, M. J., Hernandez, F., Nobre, P., Campos, E., et al. (2008). The Pirata Program: History, accomplishments, and future directions. *Bulletin of the American Meteorological Society*, *89*(8), 1111–1126. https://doi.org/10.1175/2008BAMS2462.1

Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., et al. (2013). World Ocean Database 2013. Sydney Levitus, Ed.; Alexey Mishonoc, Technical Ed., NOAA Atlas (209 pp.). https://doi.org/10.7289/V5NZ85MT

Brankart, J. M. (2013). Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. *Ocean Modelling*, *66*, 64–76. https://doi.org/10.1016/j.ocemod.2013.02.004

Brankart, J. M., Candille, G., Garnier, F., Calone, C., Melet, A., Bouttier, P. A., et al. (2015). A generic approach to explicit simulation of uncertainty in the NEMO ocean model. *Geoscientific Model Development*, *8*(5), 1285–1297. https://doi.org/10.5194/gmd-8-1285-2015

Brunke, M. A., Wang, Z., Zeng, X., Bosilovich, M., & Shie, C. L. (2011). An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalyses, satellite-derived, and combined global datasets. *Journal of Climate*, *24*(21), 5469–5493. https://doi.org/10.1175/2011JCLI4223.1

Brunke, M. A., Zeng, X., & Anderson, S. (2002). Uncertainties in sea surface turbulent flux algorithms and data sets. *Journal of Geophysical Research*, *107*(C10), 3141. https://doi.org/10.1029/2001JC000992

Calado, L., da Silveira, I. C. A., Gangopadhyay, A., & de Castro, B. M. (2010). Eddy-induced upwelling off Cape São Tomé (22°S, Brazil). *Continental Shelf Research*, *30*(10–11), 1181–1188. https://doi.org/10.1016/j.csr.2010.03.007

Carton, J. A., & Giese, B. S. (2008). A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Monthly Weather Review*, *136*(8), 2999–3017. https://doi.org/10.1175/2007MWR1978.1

Casey, K. S., Brandon, T. B., Cornillon, P., & Evans, R. (2010). The past, present, and future of the AVHRR pathfinder SST program. In *Oceanography from Space: Revisited*. https://doi.org/10.1007/978-90-481-8681-5_16

Castelao, R. M., & Barth, J. A. (2006). Upwelling around Cabo Frio, Brazil: The importance of wind stress curl. *Geophysical Research Letters*, *33*, L03602. https://doi.org/10.1029/2005GL025182

Chassignet, E. P., & Xu, X. (2017). Impact of horizontal resolution (1/12° to 1/50°) on Gulf Stream separation, penetration, and variability. *Journal of Physical Oceanography*, *47*(8), 1999–2021. https://doi.org/10.1175/JPO-D-17-0031.1

Chaves, R. R., & Nobre, P. (2004). Interactions between sea surface temperature over the South Atlantic Ocean and the South Atlantic Convergence Zone. *Geophysical Research Letters*, *31*, L03204. https://doi.org/10.1029/2003GL018647

Cheng, L., Abraham, J., Goni, G., Boyer, T., Wijffels, S., Cowley, R., et al. (2016). XBT science: Assessment of instrumental biases and errors. *Bulletin of the American Meteorological Society*, *97*(6), 924–933. https://doi.org/10.1175/BAMS-D-15-00031.1

Climanálise (2009). Boletim de Monitoramento e Análise Climática (CPTEC/INPE). Climanálise ISSN: 0103–0019 (Vol. 18). Retrieved from http://climanalise.cptec.inpe.br/~rclimanl/boletim/pdf/pdf03/jan03.pdf

Combes, V., & Matano, R. P. (2014). Trends in the Brazil/Malvinas confluence region. *Geophysical Research Letters*, *41*, 8971–8977. https://doi.org/10.1002/2014GL062523

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Counillon, F., & Bertino, L. (2009). High-resolution ensemble forecasting for the Gulf of Mexico eddies and fronts. *Ocean Dynamics*, *59*(1), 83–95. https://doi.org/10.1007/s10236-008-0167-0

Davies, T. (2014). Lateral boundary conditions for limited area models. *Quarterly Journal of the Royal Meteorological Society*, *140*(678), 185–196. https://doi.org/10.1002/qj.2127

De Almeida, R. A. F., Nobre, P., Haarsma, R. J., & Campos, E. J. D. (2007). Negative ocean-atmosphere feedback in the South Atlantic Convergence Zone. *Geophysical Research Letters*, *34*, L18809. https://doi.org/10.1029/2007GL030401

Elshorbagy, A. A., Panu, U. S., & Simonovic, S. P. (2000). Group-based estimation of missing hydrological data: I. Approach and general methodology. *Hydrological Sciences Journal*, *45*(6), 849–866. https://doi.org/10.1080/02626660009492388

Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., & Edson, J. B. (2003). Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. *Journal of Climate*, *16*(4), 571–591. https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Giese, B. S., Seidel, H. F., Compo, G. P., & Sardeshmukh, P. D. (2016). An ensemble of ocean reanalyses for 1815–2013 with sparse observational input. *Journal of Geophysical Research: Oceans*, *121*, 6891–6910. https://doi.org/10.1002/2016JC012079

Gustafsson, N., Källen, E., & Thorsteinsson, S. (1998). Sensitivity of forecast errors to initial and lateral boundary conditions. *Tellus Series A: Dynamic Meteorology and Oceanography*, *50*(2), 167–185. https://doi.org/10.3402/tellusa.v50i2.14519

Haidvogel, D. B., Arango, H., Budgell, W. P., Cornuelle, B. D., Curchitser, E., di Lorenzo, E., et al. (2008). Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System. *Journal of Computational Physics*, *227*(7), 3595–3624. https://doi.org/10.1016/j.jcp.2007.06.016

Hirahara, S., Ishii, M., & Fukuda, Y. (2014). Centennial-scale sea surface temperature analysis and its uncertainty. *Journal of Climate*, *27*(1), 57–75. https://doi.org/10.1175/JCLI-D-12-00837.1

Hoffman, M. J., Kalnay, E., Carton, J. A., & Yang, S. C. (2009). Use of breeding to detect and explain instabilities in the global ocean. *Geophysical Research Letters*, *36*, L12608. https://doi.org/10.1029/2009GL037729

Hoffman, M. J., Miyoshi, T., Haine, T. W. N., Ide, K., Brown, C. W., & Murtugudde, R. (2012). An advanced data assimilation system for the Chesapeake Bay: Performance evaluation. *Journal of Atmospheric and Oceanic Technology*, *29*(10), 1542–1557. https://doi.org/10.1175/JTECH-D-11-00126.1

Huang, B., & Mehta, V. M. (2010). Influences of freshwater from major rivers on global ocean circulation and temperatures. *Advances in Atmospheric Sciences*, *27*(3), 455–468. https://doi.org/10.1007/s00376-009-9022-6.1

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008

Jahfer, S., Vinayachandran, P. N., & Nanjundiah, R. S. (2017). Long-term impact of Amazon river runoff on northern hemispheric climate. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-10750-y

Jung, T., Serrar, S., & Wang, Q. (2014). The oceanic response to mesoscale atmospheric forcing. *Geophysical Research Letters*, *41*, 1255–1260. https://doi.org/10.1002/2013GL059040

Kalnay, E. (2003). *Atmospheric modeling, data assimilation, and predictability*. Cambridge, UK: Cambridge University Press.

Mason, E., Pascual, A., Gaube, P., Ruiz, S., Pelegrí, J. L., & Delepoulle, A. (2017). Subregional characterization of mesoscale eddies across the Brazil-Malvinas confluence. *Journal of Geophysical Research: Oceans*, *122*, 3329–3357. https://doi.org/10.1002/2016JC012611

Miller, A. J., Collins, M., Gualdi, S., Jensen, T. G., Misra, V., Pezzi, L. P., et al. (2017). Coupled ocean–atmosphere modeling and predictions. *Journal of Marine Research*, *75*(3), 361–402. https://doi.org/10.1357/002224017821836770

Nuissier, O., Joly, B., Vié, B., & Ducrocq, V. (2012). Uncertainty of lateral boundary conditions in a convection-permitting ensemble: A strategy of selection for Mediterranean heavy precipitation events. *Natural Hazards and Earth System Sciences*, *12*(10), 2993–3011. https://doi.org/10.5194/nhess-12-2993-2012

O'Kane, T. J., Oke, P. R., & Sandery, P. A. (2011). Predicting the East Australian Current. *Ocean Modelling*, *38*(3–4), 251–266. https://doi.org/10.1016/j.ocemod.2011.04.003

Penny, S. G., Behringer, D. W., Carton, J. A., & Kalnay, E. (2015). A Hybrid Global Ocean Data Assimilation System at NCEP. *Monthly Weather Review*, *143*(11), 4660–4677. https://doi.org/10.1175/MWR-D-14-00376.1

Penny, S. G., Kalnay, E., Carton, J. A., Hunt, B. R., Ide, K., Miyoshi, T., & Chepurin, G. A. (2013). The local ensemble transform Kalman filter and the running-in-place algorithm applied to a global ocean general circulation model. *Nonlinear Processes in Geophysics*, *20*(6), 1031–1046. https://doi.org/10.5194/npg-20-1031-2013

Penven, P., Marchesiello, P., Debreu, L., & Lefèvre, J. (2008). Software tools for pre- and post-processing of oceanic regional simulations. *Environmental Modelling and Software*, *23*(5), 660–662. https://doi.org/10.1016/j.envsoft.2007.07.004

Pezzi, L. P., de Souza, R. B., Acevedo, O., Wainer, I., Mata, M. M., Garcia, C. A. E., & de Camargo, R. (2009). Multiyear measurements of the oceanic and atmospheric boundary layers at the Brazil-Malvinas confluence region. *Journal of Geophysical Research*, *114*, D19103. https://doi.org/10.1029/2008JD011379

Pezzi, L. P., de Souza, R. B., & Quadro, M. F. L. (2016). Uma revisão dos processos de interação oceano-atmosfera em regiões de intenso gradiente termal do oceano atlântico sul baseada em dados observacionais. *Revista Brasileira de Meteorologia*, *31*(4), 428–453. https://doi.org/10.1590/0102-778631231420150032

Pezzi, L. P., Souza, R. B., Dourado, M. S., Garcia, C. A. E., Mata, M. M., & Silva-Dias, M. A. F. (2005). Ocean-atmosphere in situ observations at the Brazil-Malvinas confluence region. *Geophysical Research Letters*, *32*, L22603. https://doi.org/10.1029/2005GL023866

Piola, A. R., Franco, B. C., Palma, E. D., & Saraceno, M. (2013). Multiple jets in the Malvinas current. *Journal of Geophysical Research: Oceans*, *118*, 2107–2117. https://doi.org/10.1002/jgrc.20170

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., et al. (2016). Fifteen years of ocean observations with the global Argo array. *Nature Climate Change*, *6*(2), 145–153. https://doi.org/10.1038/nclimate2872

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058. https://doi.org/10.1175/2010BAMS3001.1

Sandery, P. A., Sakov, P., & Majewski, L. (2014). The impact of open boundary forcing on forecasting the East Australian Current using ensemble data assimilation. *Ocean Modelling*, *84*, 1–11. https://doi.org/10.1016/j.ocemod.2014.09.005

Sandwell, D. T., Smith, W. H. F., Gille, S., Kappel, E., Jayne, S., Soofi, K., et al. (2006). Bathymetry from space: Rationale and requirements for a new, high-resolution altimetric mission. *Comptes Rendus Geoscience*, *338*(14–15), 1049–1062. https://doi.org/10.1016/j.crte.2006.05.014

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, *9*(4), 347–404. https://doi.org/10.1016/j.ocemod.2004.08.002

Shchepetkin, A. F., & McWilliams, J. C. (2009). Correction and commentary for "Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the regional ocean modeling system" by Haidvogel et al., J. Comp. Phys. 227, pp. 3595-3624. *Journal of Computational Physics*, *228*(24), 8985–9000. https://doi.org/10.1016/j.jcp.2009.09.002

Sikirić, M. D., Janeković, I., & Kuzmić, M. (2009). A new approach to bathymetry smoothing in sigma-coordinate ocean models. *Ocean Modelling*, *29*(2), 128–136. https://doi.org/10.1016/j.ocemod.2009.03.009

Sluka, T. C., Penny, S. G., Kalnay, E., & Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophysical Research Letters*, *43*, 752–759. https://doi.org/10.1002/2015GL067238

Smith, R. D., Dukowicz, J. K., & Malone, R. C. (1992). Parallel ocean general circulation modeling. *Physica D: Nonlinear Phenomena*, *60*(1-4), 38–61. https://doi.org/10.1016/0167-2789(92)90225-C

Smith, W., & Sandwell, D. (2004). Conventional bathymetry, bathymetry from space, and geodetic altimetry. *Oceanography*, *17*(1), 8–23. https://doi.org/10.5670/oceanog.2004.63

Stramma, L., & England, M. (1999). On the water masses and mean circulation of the South Atlantic Ocean. *Journal of Geophysical Research*, *104*(C9), 20,863–20,883. https://doi.org/10.1029/1999JC900139

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, *74*(12), 2317–2330. https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2

Toth, Z., & Kalnay, E. (1997). Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, *125*(12), 3297–3319. https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2

Wu, W., Lynch, A. H., & Rivers, A. (2005). Estimating the uncertainty in a regional climate model related to initial and lateral boundary conditions. *Journal of Climate*, *18*(7), 917–933. https://doi.org/10.1175/JCLI-3293.1