A STUDY OF THE IMPACT OF ALTIMETRY DATA ASSIMILATION ON SHORT-TERM PREDICTABILITY OF THE HYCOM OCEAN MODEL IN REGIONS OF THE TROPICAL AND SOUTH ATLANTIC OCEAN

Leonardo Nascimento Lima1 and Clemente Augusto Souza Tanajura1,2

ABSTRACT. In this study, assimilation of Jason-1 and Jason-2 along-track sea level anomaly (SLA) data was conducted in a region of the tropical and South Atlantic ($7^\circ$N-$36^\circ$S, 20°W up to the Brazilian coast) using an optimal interpolation method and the HYCOM (Hybrid Coordinate Ocean Model). Four 24 h-forecast experiments were performed daily from January 1 until March 31, 2011 considering different SLA assimilation data windows (1 day and 2 days) and different coefficients in the parameterization of the SLA covariance matrix model. The model horizontal resolution was $1/12^\circ$ and the number of vertical layers was 21. The SLA analyses added to the mean sea surface height were projected to the subsurface with the Cooper & Haines (1996) scheme. The results showed that the experiment with 2-day window of along-track data and with specific parameterizations of the model SLA covariance error for sub-regions of the METAREA V was the most accurate. It completely reconstructed the model sea surface height and important improvements in the circulation were produced. For instance, there was a substantial improvement in the representation of the Brazil Current and North Brazil Undercurrent. However, since no assimilation of vertical profiles of temperature and salinity and of sea surface temperature was performed, the methodology employed here should be considered only as a step towards a high quality analysis for operational forecasting systems.

Keywords: data assimilation, optimal interpolation, Cooper & Haines scheme, altimetry data.

RESUMO. Neste estudo, a assimilação de dados de anomalia da altura da superfície do mar (AASM) ao longo da trilha dos satélites Jason-1 e Jason-2 foi conduzida em uma região do Atlântico tropical e Sul ($7^\circ$N-$36^\circ$S, 20°W até a costa do Brasil) com o método de interpolação ótima e o modelo oceanico HYCOM (Hybrid Coordinate Ocean Model). Foram realizados quatro experimentos de previsão de 24 h entre 1 de janeiro e 31 de março de 2011, considerando diferentes janelas de assimilação de AASM (1 dia e 2 dias) e diferentes coeficientes na parametrização da matriz de covariância dos erros de AASM do modelo. A resolução horizontal empregada no HYCOM foi $1/12^\circ$ para 21 camadas verticais. As correções de altura da superfície do mar devido à assimilação de AASM foram projetadas abaixo da camada de mistura através da técnica de Cooper & Haines (1996). Os resultados mostram que o experimento com assimilação de dados ao longo da trilha dos satélites com a janela de 2 dias e com parametrizações da matriz de covariância específicas para sub-regiões da METAREA V foi o mais acurado. Ele reconstruiu completamente a altura da superfície do mar e também proporcionou melhorias na circulação oceânica reproduzida pelo modelo. Por exemplo, houve substancial melhoria da representação nos campos de correntes e subcorrentes do Brasil. Entretanto, tendo em vista que não foi realizada a assimilação de perfis verticais de temperatura e de salinidade e da temperatura da superfície do mar, a metodologia apresentada deve ser considerada apenas como um passo na conquista de uma análise oceânica e de um sistema previsor de qualidade para fins operacionais.

Palavras-chave: assimilação de dados, interpolação ótima, técnica de Cooper & Haines, dados de altimetria.

1Rede de Modelagem e Observação Oceanográfica (REMO), Centro de Pesquisa em Geofísica e Geologia, Universidade Federal da Bahia (CPGG/UFBA), Campus de Ondina, Instituto de Geociências, Sala C-312, Travessa Barão de Jeremoabo, s/n, 40170-115 Salvador, BA, Brasil. Phone: +55(71)3283-6625; Fax: +55(71)3283-6501 – E-mail: leonasclima@gmail.com

2Departamento de Física da Terra e do Meio Ambiente, Instituto de Física, Universidade Federal da Bahia (UFBA), Campus de Ondina, Travessa Barão de Jeremoabo, s/n, 40170-280 Salvador, BA, Brasil. Phone: +55(71)3283-6685; Fax: +55(71)3283-6606 – E-mail: cast@ufba.br
INTRODUCTION

During the past decade, operational ocean forecasting systems have been installed in various agencies and institutions (Dombrowsky et al., 2009). Developed by different countries, these systems provide real-time forecasts and thirteen of them operate with recommendations and standards suggested by the GODAE OceanView project, successor to the successful Global Ocean Data Assimilation Experiment (GODAE) project that ended in 2008 (Le Traon et al., 2010). Systems associated with the GODAE OceanView are established in Australia, Canada, China, United States, France, England, Italy, Japan, Norway, and Europe as a whole. Brazil, through the Modeling and Oceanographic Observation Network (Rede de Modelagem e Observação Oceanográfica, REMO) (www.rederemo.org), and India joined this international effort in December 2010.

REMO, in Brazil, aims at technological innovation in operational and physical oceanography while acting intensely on ocean modeling for short-term simulation and forecasting. The numerical ocean models have contributed to a better understanding of ocean circulation and its variability. They have been continuously improved and increasingly allow the exploitation studies of ocean physical processes on synoptic scales, a task not easily accomplished due to relatively sparse in situ measurements. Despite the fact that the quite sophisticated numerical models include a wide range of physical processes and accurate numerical techniques, data assimilation methods are crucial to ocean forecasting (Chassignet et al., 2009). They are responsible for generating the initial condition of the forecast models and for much of their predictability, especially in the mesoscale processes and circulation.

The assimilation methods mathematically combine in an optimal or sub-optimal way physical fields simulated or predicted by models with observed data, and produce new fields, the so-called analyses, with errors smaller than those of the models (Daley, 1991; Kalnay, 2003). Once the analyses are more accurate than the fields purely produced by the models, they can be used to provide initial condition for ocean forecasting models, to contribute to a better understanding of the physical mechanisms responsible for the variability of the oceans, and to complement information from monitoring systems. Another use of assimilation is on the design and optimization of monitoring systems, by conducting studies of the analysis sensitivity to the observations.

A major obstacle in ocean data assimilation is the relatively small number of observed ocean data available for assimilation and assessment. For example, oceanic fields that are measured globally, such as sea surface height (SSH) or sea surface temperature (SST) derived from satellites, provide information only on the surface and are incomplete. For example, the altimeter measurements are performed only along satellite tracks and the data are not accurate in coastal regions (Chelton et al., 2001). As for in situ data, since 2003 the Argo monitoring system has offered a huge contribution to operational oceanography (Gould et al., 2004). Currently, about 3600 floats collect vertical profiles of temperature and salinity up to 2000 m depth every 10 days throughout the global ocean. However, some regions such as the South Atlantic are still poorly sampled. The observed data have, therefore, great heterogeneity in time and space, and an operational forecasting system requires a data assimilation scheme that takes into account these aspects (Ezer & Mellor, 1994; Fukumori, 2001).

In recent decades, many satellites with altimeters were launched to monitor SSH changes. This observation system can be considered the most important for operational oceanography (Le Traon, 2011), because the sea level is directly related to ocean circulation through the geostrophic approximation. According Hurlburt et al. (2008), satellite derived SSH data is key to map the state of ocean waters. Thus, the sea surface height anomaly (SSHA) or the sea level anomaly (SLA), which depicts the spatial and temporal variability of the sea level, has been successfully used in data assimilation systems, because the anomaly imposes dynamic restrictions throughout the modeled water column. The assimilation of this variable in ocean models adjusts the thermohaline fields in surface and subsurface (Fox et al., 2000; Zheng et al., 2007). Also, synthetic temperature and salinity profiles may be derived statistically from the variations of SSHA, climatologies, in situ observations and outputs of ocean models (Ezer & Mellor, 1994; Fox et al., 2002). In another approach, the technique developed by Cooper & Haines (1996) (CH96), starting from a given SSH field, imposes corrections of the thickness of the isopycnal layers and adjusts model circulation while maintaining potential vorticity.

Thus, the assimilation of the SLA into numerical ocean models has become an important tool to improve predictability and representation of ocean physical state. Recent studies have investigated the importance of spatial and temporal coverage of altimetry data used in data assimilation into ocean models. Benkiran (2007) investigated the assimilation impact of SLA data derived from four satellites in the North Atlantic. Pujol et al. (2010) estimated the assimilation impact of altimetry data from different satellites in the Mediterranean. However, few studies have been conducted in the tropical western portion of the South Atlantic. As part of REMO activities, this work aims at performing short-term ocean forecasting experiments using the Hybrid Coordinate...
Ocean Model (HYCOM) over a region of the tropical and South Atlantic Ocean, and investigating model predictability from different initial conditions. The HYCOM is the model used operationally in REMO since 2010 to produce large-scale prediction in the Atlantic Ocean at the Navy Hydrographic Center (Centro de Hidrografia da Marinha, CHM).

The present work aims at contributing to the construction of an operational system for altimetry data assimilation into HYCOM. More specifically, it aims to demonstrate the correct implementation of the optimal interpolation method (OI) (Ezer & Mellor, 1994) to assimilate SLA data derived from Jason-1 and Jason-2 satellites in conjunction with the CH96 method, in order to improve the model SSH estimation over a region in the equatorial and southwest Atlantic Ocean, specifically the so-called METAREA V (7°N-35.5°S, 20°W up to the coast of Brazil). This work also investigates different approaches to assimilate SLA considering spatial and temporal decorrelation scales in the region of interest and assess whether the analysis and the 24 h model forecasts are sensitive to these different approaches. This paper focuses only on the months of January, February and March 2011, given that it is considered a preliminary work to test the sensitivity of the model to the assimilation of SLA and CH96 method. Longer periods may be further investigated in pre-operational mode. After the presentation of the study area, the paper describes briefly the dynamic model and its configuration, the processing of altimetry data for assimilation, assimilation methodology and experiments. Then, the results and discussion are presented, followed by the conclusions.

Study area
The region of interest of this study is the METAREA V, between 35.5°S and 7°N and from 20°W up to the coast of Brazil, excluding sovereignty areas from neighboring countries. This area is under the responsibility of the Brazilian Navy according to the country’s commitments as a member of the International Convention for the Safety of Life at Sea (SOLAS). Among these commitments are the development of procedures related to maritime meteorology activities that contribute to the safety of navigation, safety of life and prevention of pollution in Brazilian Jurisdictional Waters (NORMAM 19/04H, 2004).

The main water masses in the region were discussed in several papers (e.g. Stramma & England, 1999; Stramma & Schott, 1999; Silveira et al., 2000; Cirano et al., 2006). In most of the South Atlantic and along the Brazilian coast, they are distributed as follows: the surface layer is occupied by Tropical Water (TW). Just below the TW, according to Stramma & Schott (1999), two types of the South Atlantic Central Water (SACW) are found: a lighter water generated in the southwestern region of the tropical gyre and another generated in the southern region of the South Atlantic. Silveira et al. (2000) considered the SACW as part of the subtropical gyre which splits when it meets the coast of Brazil. The Antarctic Intermediate Water (AIW), present in METAREA V, originates in the Antarctic Convergence Zone with the sinking of Antarctic Surface Water and Subantarctic Water in the Circumpolar Zone between 50 and 60°S, especially in the northern Drake Passage and Malvinas Current (Stramma & England, 1999). The AIW, located below the SACW, is identified by a salinity minimum and oxygen maximum, whereas the North Atlantic Deep Water (NADW), just below the AIW, crosses the equator and flows southward along the western boundary of the basin.

In the first layers of the western region of METAREA V, the circulation presents the Brazil Current (BC) and the North Brazil Current (NBC), a result of the bifurcation of the South Equatorial Current (SEC) in its westward flow along the South Atlantic. According to Silveira et al. (2000), the bifurcation position varies seasonally at the TW level, but the average position is around 10°S. There are controversies about the average location of the bifurcation, which was the research target in the study by Rodrigues et al. (2007). At the level of SACW, the bifurcation of SEC occurs at about 20°S latitude. Thus, one organized flow is generated when the SACW and TW meet and flow toward the equator, resulting in the North Brazil Under-Current (NBUC). South of Vitória-Trindade Ridge, the BC has a fairly energetic pattern, which often results in the formation of strong cyclonic and anti-cyclonic meanders. One of these features is the cyclonic vortex of Victoria, first described by Schmid et al. (1995). At the AIW level, the bifurcation reaches latitudes further south and generates the Intermediate Western Boundary Current (IWBC) that flows north. Stramma & England (1999) locate the AIW bifurcation axis near 25°S. Next to the 2000 meter depth, the Deep Western Boundary Current of the South Atlantic (DWBC) transports NADW to the Southern Hemisphere (Silveira et al., 2000). South of 8°S, the DWBC disintegrates into a sequence of migratory vortices, adjusting the flow toward the south of the NADW (Dengler et al., 2004; Schott et al., 2005).

METHODOLOGY
Ocean Model
To investigate altimetry data assimilation, experiments were performed using HYCOM (Bleck, 2002) configured for the domain 45°S-10°N, 18°W-67°W with horizontal resolution of 1/12°,
corresponding to 601 and 733 points in the zonal and meridional directions, respectively. The number of vertical layers was 21. This domain encompasses METAREA V. The HYCOM in the 1/12° grid was nested in HYCOM configured for the entire Atlantic Ocean (78°S-50°N, 100°W-20°E) with horizontal resolution of 1/4° and 21 layers (Fig. 1). In this domain, a simplified ice model already installed in HYCOM was enabled and lateral boundary conditions for temperature and salinity were climatological. At 50°N, the mass flow was zero, while constant barotropic flow was imposed on the eastern and western side walls. An inflow of 110 Sv on the west and 10 Sv on the east, just south of South Africa, were imposed. An outflow of 120 Sv was prescribed in west to the south of the Agulhas Current region until Antarctica.

Unlike altimeter derived SSH, HYCOM SSH is obtained from the Montgomery potential at the surface. The hydrostatic balance allows expressing the state variables in terms of the barotropic component of pressure and the layer thicknesses. Chin et al. (2002) described the SSH calculation scheme in the Miami Isopycnal Coordinate Ocean Model (MICOM), the isopycinal layered model that served as the basis for the HYCOM.

The bathymetry used in the integrations originated from the TerrainBase (TERRAINBASE, 2007) of the National Geophysical Data Center and has a spatial resolution of 5 minutes (~8.5 km at 25°S). The surface atmospheric forcings employed here were 24 h forecasts produced by the Global Forecast System (GFS) of the National Centers for Environmental Prediction/National Oceanic and Atmospheric Administration (NCEP/NOAA) with 0.5° × 0.5° resolution. The predicted atmospheric fields were precipitation, 10 m wind, long-wave and short-wave surface radiation fluxes, air temperature and mixing ratio at 2 m and they were given at every 3 hours to produce a sequence of 24 h ocean forecasts. The integrations included the discharge of the main rivers present on the coast of Brazil.

An important aspect of SLA data assimilation is that the mean SSH over the altimeter observation period must be provided. Unfortunately, the Earth’s geoid is not currently known with sufficient accuracy to provide a mean SSH in mesoscale. Therefore, it is of utmost importance to have a mean of the model that is reasonably accurate, since many ocean fronts and current mean paths cannot be well defined only from hydrographic climatologies. In this study, the field of mean SSH was calculated by averaging daily SSH produced from the integration between 2003 and 2009.

**Altimetry data and the adjustment for assimilation**

This study uses the Jason-1 and Jason-2 along-track (SLA) data obtained from the French Archivage system, Validation Interprétation des données et des Satellites Océanographiques (AVISO). Data were processed using a convolution filter or a moving average of seven points, distant approximately 6 km from one another. This filter eliminates wavelengths up to 40 km and prevents the insertion of small scale ocean features that cannot be captured by the model. The Jason-1 and Jason-2 SLA data assimilated into HYCOM were calculated by the AVISO system on relation the mean period from 1993 to 1999. In this period, the height corresponding to the average ocean circulation in relation to geoid variations, called mean dynamic topography (MDT), was also calculated by AVISO. The MDT was estimated through objective analysis with the inclusion of altimetry, hydrology, drifters and a geoid model (Rio & Hernandez, 2004). The addition of the
MDT and the SLA gives the sea level height which restricts ocean circulation, that is, the instantaneous dynamic topography.

In general, a great difficulty regarding SLA data assimilation is to account for the bias that exists between the observed data and the result produced purely by the ocean model. In this work, the model SLA was calculated by subtracting from the instantaneous model SSH a mean SSH produced by a free model run with atmospheric forcing and NCEP reanalysis from 2003 to 2009.

The differences observed between the model SLA and observation are probably associated with the different mean SSHs used to calculate them. Thus, when the bias is high, as in the present study, an alternative commonly used is the SLA correction. Several schemes have been proposed for ocean prediction when the bias is present (Chepurin et al., 2004; Balmaseda et al., 2007). In this study, a procedure proposed by Tanajura et al. (2013) for adjusting the SLA data is employed. The details are described below.

The observed and the model SLA were different along the satellite tracks even when the mean model SSH was carefully chosen. We soon realized that if the observed SLA was assimilated directly, a large ridge or trough would be produced along the track rendering a totally unrealistic analysis. In order to capture the location and magnitude of the mesoscale ridges and troughs and gently pass this information to the model, the following data adjustment was imposed: (1) First, it was necessary to calculate the HYCOM SLA, subtracting from the instantaneous SSH model the mean SSH; (2) with the HYCOM SLA interpolated to the position of the observed SLA track, the difference HYCOM SLA minus observed SLA was calculated; (3) then, the average of this difference was calculated; (4) lastly, the adjusted SLA was obtained by adding the observed SLA to the average calculated in step 3 (Fig. 2). This adjustment technique improved SLA fields in previous experiments with HYCOM under REMO (Tanajura et al., 2013). Thus, the adjusted observed SLA was assimilated according to the data assimilation scheme discussed in the section below.

**Optimal Interpolation (OI)**

At time $t$, SLA data from Jason-1 and Jason-2 satellite tracks that passed the numerical domain over a period $\Delta t$ must be assimilated into the model. Thus, the analysis ($x^a$) of SLA on a model grid point is given by

$$x^a = x^f + K[y - H(x^f)],$$  \hspace{1cm} (1)

where $y$ represents the $N$ satellite data that were assimilated; $x^f$ the model SLA or background; $H$, the observation operator that interpolates the model space into the observations; and $K$, the optimal weight matrix also called the gain matrix (Kalnay, 2003). To calculate $K$, fixed and pre-determined error covariance matrices of the model and the observations were used. Unlike more recent methods that require bigger computational infrastructure, the suggested OI does not excessively increase computational cost compared to the model run without assimilation (Ezer & Mellor, 1994). This method is based on minimizing the analysis quadratic error. The minimization of this error from
Figure 3 – SLA (10^-2 m) standard deviation calculated with data from AVISO, showing the seven sub-regions of the numeric domain defined for applying the decorrelation parameters. The standard deviation was calculated for the months of January, February and March 2007 to 2010.

Eq. (1) yields the following linear system for the optimal weight (Eq. 2).

$$K = BH^T (HBH^T + R)^{-1}.$$  (2)

$R$ is the covariance matrix of the errors of observations and $B$ the covariance matrix of the errors of the model. $R$ was assumed to be the diagonal with values of (0.02 m)^2. The proposed OI parameterizes the covariance of the model errors as follows: $B = \sigma^2 G$. The variance ($\sigma^2$) was calculated with the daily outputs of the free integration from 2003 to 2009, time interval in which the mean SSH was extracted. To impose localization to the model error covariance matrix, the parameter $G$ (Eq. 3) selects the satellite points that are used in the analysis of a given model grid point, taking into account zonal, meridional and temporal decorrelation parameters, $k_x, k_y$ and $k_t$, respectively, the zonal distance $\Delta x$, the meridional distance $\Delta y$ between the observation points and the analysis point, and the time period $\Delta t$ between the observation time and the analysis time according to:

$$G = \exp \left[ -(k_x \Delta x)^2 - (k_y \Delta y)^2 - (k_t \Delta t)^2 \right].$$  (3)

In studies of altimetry assimilation with the Princeton Ocean Model (POM) in the Gulf Stream, Ezer & Mellor (1994) determined the parameters of spatial decorrelation, $k_{x}^{-1}$ and $k_{y}^{-1}$, equal to 170 km and temporal decorrelation, $k_{t}^{-1}$, equal to 22 days. These values were comparable to those found by Auer (1987).

In this study, however, the time and length scales were estimated from SLA data from AVISO, in 1/3° spatial resolution for the months of January, February and March from 2007 to 2010. Thus, from these data and based on the similarity of the standard deviation (Fig. 3), the model domain was subdivided into seven regions. For each sub-region, ten grid points were selected at random and, for each of the selected points the SLA autocorrelation was calculated in space and time (Figs. 4 and 5). Thus, decorrelation parameters $(k_x, k_y, k_t)$ were determined as a function of spatial and temporal e-folding decay of the autocorrelation (Table 1).

Table 1 – Spatial and temporal decorrelation parameters estimated with the SLA data from AVISO for January, February and March from 2007 to 2010.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>$k_x^{-1}$ (km)</th>
<th>$k_y^{-1}$ (km)</th>
<th>$k_t^{-1}$ (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>310</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>440</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>370</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>450</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>140</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>120</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Following the estimation of the gain matrix according to the OI method, the SLA analysis was added to the mean model SSH. The CH96 method was used to compare this newly created SSH instantaneous field with the background field model. The result obtained after correcting the model with CH96 was the initial condition of the 24 h forecasts in the assimilation run.

Revista Brasileira de Geofísica, Vol. 31(2), 2013
The Cooper and Haines technique (1996)

The CH96 technique is very important and widely used operationally in ocean forecasting centers (Ferry et al., 2007). However, it cannot be considered data assimilation. As mentioned above, two variables are important for calculating the HYCOM SSH: barotropic pressure and layer thicknesses. The CH96 technique requires a correction in the thickness of the model layers using a given SSH field. Fox et al. (2000) list several advantages of this technique, including the preservation of the T, S relation of the water column. The paper (Chassignet et al., 2007) talks the basic principle of CH96 technique, which is conservation of the subsurface potential vorticity during the moment of assimilation.

The technique is based on a water column lifting-lowering procedure provided by the projection of the SSH increments, which is converted in increments of temperature, salinity and velocity. From the SLA analysis produced according to Eq. (1), for each model grid point, if the reconstituted SSH analysis is less than the background SSH, the isopycnical layers closest to the surface are displaced upward. Thus, certain mass of surface light water is removed and the equivalent mass of denser water is inserted in the bottom, so that the SSH decreases. In contrast, if the SSH derived from the SLA analysis is greater than the background SSH, the thermocline is lowered by removing water from the dense bottom and inserting lighter and hot water in the surface layers. The vertical displacement happens, however, without any modification of the barotropic pressure.

In the integrations of this study, the CH96 technique was configured to operate below the mixed layer, therefore, in the isopycnal layers of HYCOM.
A STUDY OF THE IMPACT OF ALTIMETRY DATA ASSIMILATION ON SHORT-TERM PREDICTABILITY OF THE HYCOM OCEAN MODEL

Assimilation Experiments

Four experiments were conducted, three with and one without assimilation, also called control. For all the experiments, 24-hour predictions were made during the period from January 1 to March 31, 2011 using different initial conditions on each day. The model altimetry was corrected daily for each forecast initial condition with data assimilation on the sea surface of SLA from the Jason-1 and Jason-2 along-tracks. The assimilation method used was the OI proposed by Mellor & Ezer (1991) and Ezer & Mellor (1994). The SLA was assimilated, then, the analysis of the newly created SLA was added to mean model SSH, reconstructing the instant model SSH. From this field, the surface information was passed to the subsurface using the CH96 technique. The algorithm used is briefly shown in Figure 6.

The prediction experiments were performed based on choosing different decorrelation parameters. In experiment 1, the spatial parameters \((k_{x}^{-1}, k_{y}^{-1})\) were similar to those proposed by Ezer & Mellor (1994) for the Gulf Stream (170 km). In experiment 2, the spatial parameters were calculated from the AVISO SLA gridded data with 1/3° resolution, as shown in Table 1. In these first two experiments, each initial condition from 01/01/2011 to 03/31/2011 was obtained with the assimilation of all SLA data of Jason-1 and Jason-2 satellites available in the day of the analysis (Fig. 7).

The third experiment included a larger SLA database. Each initial condition from 01/01/2011 to 03/31/2011 was obtained with assimilation of SLA data on the current and on the previous day to the analysis (Fig. 7). In this case, temporal decorrelation was included via the parameter \(k_{t}\) (Table 1).

Table 2 shows an example of computational processing time necessary to integrate the model and to perform the OI and the CH96 technique. The processing time varies depending on the performance of the machine.

In all experiments of SLA data assimilation, surface information was passed to the subsurface using the CH96 technique. This step differed from the creation and assimilation of synthetic profiles of temperature and salinity used in Ezer & Mellor (1994).

Metrics

The quality of the experiments was objectively investigated by comparing the model results with the observed SLA data from AVISO and with the results derived from HYCOM-Navy Coupled Ocean Data Assimilation System (HYCOM+NCODA) (Chassignet et al., 2009; Dombrowsky et al., 2009). The centered root-mean-square-error (CRMSE) and correlations (R) were calculated following the metrics suggested by GODAE OceanView (Chassignet et al., 2006). These assessments were made using the following formulas:
CRMSE = \left\{ \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y} - (X_i - \bar{X}))^2 \right\}^{1/2} \quad (4)

R = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})/\left(\sigma_x \sigma_y\right) \quad (5)

where \( X_i (i = 1, 2, 3, \ldots, n) \) is the set of model results linearly interpolated to the \( n \) observation point and \( Y_i (i = 1, 2, 3, \ldots, n) \) is the set of available observations, \( \bar{X} \) and \( \bar{Y} \) are the mean and standard deviation of the model result interpolated to the observation points (available observations), respectively.

The CRMSE is an important statistical metric widely used to compare results obtained from different sources, as the case of SLA produced by AVISO and HYCOM. The bias between the observed data and the model output is eliminated in this metric.

**RESULTS AND DISCUSSION**

Figure 8 shows that the assimilation of the SLA data derived from Jason-1 and Jason-2 satellites into HYCOM was properly formulated and implemented. The analysis increment, the difference between the analysis \( (x^a) \) minus the background \( (x^b) \), in each experiment reflects the characteristics of the decorrelation parameters set forth in Table 1. For experiment 1 (Fig. 8a), using the fixed spatial decorrelation \( k_{x}^{-1} = k_{y}^{-1} = 170 \) km, the analysis increment around the satellite track shows a constant width region throughout the numeric domain. In experiment 2 (Fig. 8b), the spatial decorrelation parameters vary and provide differentiated increment with different widths support for the various chosen sub-regions. For example, for sub-region 3, where the decorrelation radius is \( k_{x}^{-1} = k_{y}^{-1} = 370 \) km, the increment has a greater diameter than sub-region 7, where

\[ k_{x}^{-1} = k_{y}^{-1} = 120 \text{ km}. \]

Experiment 3 had SLA data assimilation of the current and previous day to the analysis. Thus, Figure 8c also shows that the SLA data assimilation of two days and two satellites imposes corrections in much greater number of domain points.

**Evaluation of the SLA forecasts**

The SLA means derived from HYCOM forecasts are different from the means derived from AVISO and from the HYCOM+NCODA analysis (Fig. 9) due to the difference between the means of SSH used to calculate the AVISO SLA and the experiments performed in this work. In much of the numerical domain, the mean SLA derived from AVISO reaches values between –0.10 and 0.15 m, the mean SLA from the HYCOM+NCODA analysis reaches values from –0.25 to 0.15 m, while the mean SLA derived from the forecast experiments achieve relatively lower values, around –0.25 to 0 m.

The average calculated for the HYCOM+NCODA (Fig. 9e) partially reproduces the mean SLA from AVISO (Fig. 9f). South of 22.5°S, the mean from HYCOM+NCODA displays regions with more negative SLA values, following the trend shown in the means from AVISO. To the north of this latitude, while the AVISO data has a positive mean for almost the entire domain, the HYCOM+NCODA still shows negative values for most of the region. The means of the forecasts for each 24-h experiment (Figs. 9a, 9b and 9c) do not exhibit this behavior. However, corrections are verified in the assimilation experiments compared to control. For example, the 24-h forecast average of the control experiment shows a large region with negative anomalies less than –0.2 m between about 16 and 27°S. In this region, the assimilation experiments produced forecasts with higher mean
values, a behavior closer to the AVISO field. Another correction provided by the assimilation happens in the region west of 30°W and between 15°-22°S.

There are other punctual corrections, however, the comparison of the mean SLA forecasts alone does not allow identifying them easily. Therefore, the next figures show in more detail the impact of the SLA assimilation on the HYCOM predictability.

As an example, SLAs are taken for two points in regions with different variability. At 29°S, 38°W (Fig. 10a), it is clearly seen a bias of approximately 0.20 m between the HYCOM predictions and the AVISO observations. Due to assimilation, the difference between the SLA of experiment 3 and AVISO decreases from approximately 0.20 m in early January to values around 0.10 m in late March. Experiments 1 and 2 also converge towards AVISO data at that point, but not as effectively as experiment 3. The predictions derived from the control remain very different from the AVISO SLA. At 39°S-43°W (Fig. 10b), the AVISO variability is well captured by experiments 2 and 3. This point is located in a high variability region in the mid-latitudes. Experiments 2 and 3 present minimum SLA value around the 16/feb and a maximum around 08/mar, close in magnitude and phase to the AVISO values. Experiment 1 is very close to experiments 2 and 3, but on days 06/feb and 08/mar its values departed considerably from the data and the other assimilation experiments. The control does not have the same quality of the assimilation experiments. The SLA of the control displays intraseasonal variability and amplitude over time much smaller than AVISO and the experiments at the analyzed point.

Figure 11 shows the CRMSE calculated in relation to the observed AVISO SLA for the entire METAREA V contained in the model domain. This metric is centered. It removes the mean of SLA from each data and produces values much smaller than both the uncentered root-mean-square-error and the simple difference between model and observation. The errors of 24-h forecast of assimilation experiments are lower than the control and over much of the period are lower than the HYCOM+NCODA. The errors of assimilation forecasts are about 0.050 m, so that this metric does not show the forecast sensitivity in relation to different decorrelation parameters. The CRMSE mean of the experiment 3 over the studied period is equal to that found in experiment 2, 0.052 m, slightly better than the experiment 1, 0.054 m, and much better than the control, 0.075 m.

The horizontal CRMSE maps (Fig. 12) for METAREA V show that the highest values are located in the mid-latitudes and in the equatorial regions with intense dynamics and intraseasonal variability. Forecasts with assimilation reduce the errors in relation

![Figure 9](image-url)
Figure 10 – Comparison at (a) 29°S-38°W and (b) 39°S-43°W from 01/01/2011 to 03/31/2011 of the AVISO SLA data (m) and the SLA results predicted from the control experiments, 1 (e1), 2 (e2) and 3 (e3).

Figure 11 – CRMSE curves (m) of the SLA forecasts from 01/01/2011 to 03/31/2011 for the experiment control (red), 1 (black), 2 (brown), 3 (blue dashed) and the HYCOM+NCODA analyses (green) in relation to the AVISO SLA data for the METAREA V domain.

to the control run (Fig. 12d) and produce values similar to those of the HYCOM+NCODA (Fig. 12e). The control has CRMSE close to 0.15 m in the region 50°W-40°W north of 6°N, and experiments with assimilation reduce the error to less than 0.01 m. In the latitudes between 15 and 29°S, in the region adjacent to the Brazilian continental shelf, the experiments with assimilation also reduce the errors of the control run between 0.10 and 0.15 m to less than 0.10 m. In the region along the continental shelf, including the mouth of the Amazon River, the best results are given by HYCOM+NCODA and few differences are observed between the assimilation experiments. There is no assimilation of altimetry data on the continental shelf since the available data are for regions with water deeper than 1000 m. It would depend on the model dynamics to extrapolate information from the regions with observation to regions without observation. Perhaps a longer experiment would be able to produce corrections on the continental shelf.

It is possible to verify that the spatial correlation of the SLA forecasts with the AVISO data displays similar or even better quality than the HYCOM+NCODA and far superior than the control run (Fig. 13). The correlations remain most of the time between 50% and 70%, while the control reaches a minimum of

Brazilian Journal of Geophysics, Vol. 31(2), 2013
A STUDY OF THE IMPACT OF ALTIMETRY DATA ASSIMILATION ON SHORT-TERM PREDICTABILITY OF THE HYCOM OCEAN MODEL

Figure 12 – Horizontal maps of average CRMSE of the 24 h forecasts of SLA (m) from the predictions of the experiments (a) 1, (b) 2, (c) 3, (d) control and (e) the analyses of the HYCOM+NCODA in relation to AVISO from 01/01/2011 to 03/31/2011.

10%, showing no SLA predictability. This clearly indicates that the model without assimilation is unable to simulate and predict the ocean weather, that is, the features and vortices of mesoscale and high variability so important for predicting surface currents and the thickness of the mixed layer, among others.

Figure 13 – Time series of SLA forecasts correlation from 01/01/2011 to 03/31/2011 for the experiment control (red), 1 (black), 2 (brown), 3 (blue dashed) and the HYCOM+NCODA analyses (green) in relation to the AVISO SLA data for the METAREA V domain.

In general, the results presented up to here show great contribution of the altimetry data assimilation to improve the 24-h SLA forecast of HYCOM. In comparison with the AVISO SLA, forecast errors of experiments 1, 2 and 3 decreased and the correlations increased substantially compared to the control. The results, however, did not clearly indicate which method of parameterization of the error covariance matrix of the model produces better SLA forecasts, that is, whether the dependence of the decorrelation parameters by sub-region brings benefits.

Seeking to investigate which experiment yields the best 24-h SLA forecasts, Taylor diagrams were produced (Taylor, 2001) displaying the CRMSE, correlation and standard deviation of the SLA experiments with respect to AVISO gridded data (Fig. 14). The assimilation experiments produce average correlation of approximately 60%, error of 0.050 m and standard deviation of approximately 0.070 m, higher than the approximately 0.050 m standard deviation of the observations. The forecast errors of experiments 2 and 3 are similar, 0.054 m, and slightly lower than experiment 1, 0.055 m. In general, the results of SLA experiments of this work are slightly better than the results of HYCOM+NCODA.

Although the standard deviation of the control run is similar to the experiments with assimilation, the correlation with AVISO data is about 30% and the error is 0.070 m. This demonstrates that the model has the variability needed to yield a good forecast of the dynamic topography, but without SLA assimilation, the model is unable to yield good predictions. Certainly, the quality of the control run and assimilation experiments, as well as the free run used to generate the initial condition on January 1, 2011, depends on the atmospheric forcing, boundary con-
ditions, model resolution and all the physical parameterizations employed. However, the above result indicates that even if the model had an excellent representation of the mean SSH, it would not be able to capture the features of the observed daily dynamic topography. Only the SLA constrains imposed by the assimilation can determine the correct location of the features of the observed SLA.

Figure 14 – Taylor diagram summarizing the statistics of 24-h forecasts of the control (red), 1 (black), 2 (brown), 3 (blue) experiments and HYCOM+NCODA analyses (green) in relation to AVISO SLA (m) from 01/01/2011 to 03/31/2011 for METAREA V. The gray dot corresponds to the statistics of the AVISO SLA with respect to itself. The blue dashed half-circles indicate CRMSE isolines.

Impact on the Circulation off the Southwest Atlantic

The results below compare the velocity and circulation produced by the forecasts from the control run, the experiment 3 and the HYCOM+NCODA analysis. As discussed in the introduction, the sea surface elevation is a very important field for dynamic oceanography, because the sea level is directly related to ocean circulation and thermohaline state of the entire water column. Among the forecast experiments, experiment 3 was chosen because the statistical results showed that the SLA forecast was slightly better than the other experiments with assimilation. Thus, experiment 3 will be mentioned from here on simply as the SLA experiment with assimilation (SLAA).

The depths of 100 m, 300 m and 1000 m were chosen to investigate the impact of SLA data assimilation and the CH96 technique in the 24 h forecast of the velocity. These depths were chosen following the depth ranges defined by Cirano et al. (2006) for the main water masses found off the western boundary of the South Atlantic, using $z$-level outputs of the Ocean Circulation and Climate Advanced Modeling Project (OCCAM). The analyses are limited to 1000 m because below this depth there were no major changes between the forecast of the control and of SLAA experiment.

Figure 15 – Average of 24-h velocity (m/s) forecasts produced from 01/01/2011 to 03/31/2011 at 100 m for the experiment (a) control, (b) SLAA, and (c) average of the HYCOM+NCODA analyses. The red solid line denotes the position of SEC bifurcation at the TW, according to Stramma & England (1999).

TW is found at a depth of 100 m. At this depth, north of 10°S, all experiments represent a northward average flow with velocities above 1 m/s associated with NBUC along the coast of Brazil (Figs. 15a, b, c). This flow begins further south and is associated with the northern branch of the bifurcation of SEC. The average of the SLAA forecasts (Fig. 15a) establishes a more intense branch to west of SEC about 10°S, but it is possible to identify the SEC branching north and south only at approx-
imately 15°S. The HYCOM+NCODA also display the average position of the SEC bifurcation at 15°S (Fig. 15c). These results agree with the available literature. Using data from the World Ocean Circulation Experiment (WOCE), Wienders et al. (2000) located the SEC bifurcation at 14°S on several surface isopycnals. Stramma & England (1999) established the SEC bifurcation on the surface at 15°S.

Unlike the results found in the literature, the control run mean (Fig. 15a) has a northward flow, south of 15°S. At 100 m, south of 15°S, the BC flows south following the morphology of the continental shelf. So, except for the results of the control, the BC flows south reaching the approximate speed of 0.4-0.5 m/s in some parts. Probably due to the strong baroclinicity and effects of topography in the region from the Royal Charlotte Bank until the São Tomé Cape, between 15°S and 22°S, the BC energy pattern displays frequent formation of cyclonic and anticyclonic

---

**Figure 16** – Average of velocity (m/s) forecasts produced from 01/01/2011 to 03/31/2011 at 300 m for the experiment (a) control, (b) SLAA and (c) the average of the analyses HYCOM+NCODA. The red solid line denotes the position of SEC bifurcation at the SACW according to Stramma & England (1999).

**Figure 17** – Average of velocity (m/s) forecasts produced from 01/01/2011 to 03/31/2011 at 1000 m for the experiment (a) control, (b) SLAA, and (c) the average of the analyses of the HYCOM+NCODA.
meanders (Schmid et al., 1995). Some of these structures are represented in the results of the experiments.

Below the TW, the SACW is represented at the depth of 300 m. At the SACW level, several authors have placed the SEC bifurcation at about 20-24°S (Stramma & England, 1999; Wienders et al., 2000; Rodrigues et al., 2007). North of 22°S, the mean velocity of the control run, of the SLAA and of HYCOM+NCODA show a continuous northward flow along the western boundary. However, while the SLAA (Fig. 16b) and the control run (Fig. 16a) show a flow with velocities above 0.6 m/s associated with the NBUC, mainly north of 10°S, the HYCOM+NCODA (Fig. 16c) shows much lower velocities. South of 22°S, the HYCOM+NCODA and SLAA show a discontinuous flow, indicating that the bifurcation is located around this latitude. However, even south of 22°S, the average control shows a northward flow; therefore, in this experiment the bifurcation takes place south of 22°S. So even establishing a higher northward flow associated with NBUC, the average SLAA shows that the corrections imposed on the initial condition probably produced more accurate forecasts of currents at 300 m, mainly due to the repositioning of the SEC bifurcation at the SACW level.

At 1000 m, the AIW level, both the average 24 h forecast of the control run (Fig. 17a) and SLAA (Fig. 17b) yield speeds of 0.5 m/s associated with northward flow along the IWBC. In contrast, the HYCOM+NCODA mean (Fig. 17c) yields a much weaker flow associated with this current. Moreover, this mean exhibits a discontinuous flow south of 20°S. South of 25°S, it is also possible to identify a southward flow form the HYCOM+NCODA results, an indication that the SEC bifurcation would be at latitudes close to that at intermediate levels. According to Stramma & England (1999), the SEC bifurcation at intermediate levels occurs at 25°S. Thus, only the average of HYCOM+NCODA agrees with the results of these authors. Wienders et al. (2000) place the SEC bifurcation in 26-28°S. This position is consistent with the average velocity field of the 24-h forecast of the control run and SLAA. At 1000 m, the results of the control and SLAA are quite similar showing that the assimilation impact of the SLA and CH96 did not change the thickness of the isopycnal layers at this depth.

At 1000 m, deep ocean circulation occurs on the timescale of 1000 years and the velocity is much smaller than on the surface. Therefore, only three months of integration of HYCOM in the forecast mode and the CH96 method acting basically just below the mixed layer were not enough to impose changes in the circulation. Perhaps a multivariate assimilation method that includes temperature and salinity and the covariance between SLA and layer thicknesses can cause greater impacts at this depth.
and its difference with AVISO reached 0.250 m. The assimilation of along-track Jason-1 and Jason-2 SLA data constrained the model and corrected the two-dimensional field of SLA over almost the entire model domain during the 3-month experiment. Only the regions of the continental shelf did not have the SLA directly corrected due to the lack of observed data. The 24 h forecasts produced daily from initial conditions with the assimilation of SLA showed significant improvements compared to forecasts without assimilation.

The SLA assimilation methodology with the strongest positive impact on the SLA reconstruction was the one that assimilated data in a two-day window using space-dependent decorrelation parameters in the parameterization of the model SLA error covariance matrix, experiment 3. However, these results were only slightly higher than those of experiments 1 and 2 that used data in a 1-day window. The statistics showed that the quality of SLA forecasts also exceeded the HYCOM+NCODA analyses. An explanation can be given considering that the assimilation method employed in the NCODA is multivariate and assimilates nearly all the data of interest available in the global systems of collection and dissemination of observed data. Therefore, it imposes other restrictions on the model from the assimilation of vertical temperature and salinity profiles rendering it unable to obey only the restrictions of SLA data.

While the SLAs derived from AVISO were calculated in relation to the average of SSH of the 1993-1999 period (Rio & Hernandez, 2004), the SLAs of the forecast experiments were calculated using an average produced by a free run of the model with atmospheric forcing of NCEP/NCAR reanalysis from 2003 to 2009. Therefore, different SSH means used for calculating the SLA were responsible for the differences between the mean SLA derived from AVISO and the forecast experiments. Metrics that considered the model SLA bias were employed to compare model results with AVISO data. The CRMSE showed in the Taylor diagram yielded a much smaller error than the differences observed in the mean fields. The CRMSE of experiment 3 was 0.054 m and its difference with AVISO reached 0.250 m. The assimilation experiments were able to reduce the CRMSE of 0.070 m of the control, and produce SLA correlations of 60% on average, much higher than the control run (30%). This result shows that the model without assimilation of SLA is not able to reproduce and predict mesoscale features of SLA in the study area using only the atmospheric forcing, even if the standard deviation of SLA is similar to that observed. The comparison of the assimilation experiments with the HYCOM+NCODA analyses using the CRMSE for SLA showed the present experiments have the same or better skills.

The CH96 technique also favored the quality of the 24 h velocity forecasts. In general, the 24 h forecasts derived with assimilation in the initial condition exhibited significant improvements in the representation of the SEC bifurcation on the surface and pycnocline levels. The best results in the currents were also displayed on the western boundary of the South Atlantic, region where the BC, the NBUC at the TW level and the SACW were more accurately represented. Particularly at 22°S, the control run produced circulation which is not in accordance with the observations and ocean analyses, such as the HYCOM+NCODA.

Thus, since the ocean model is a nonlinear system involving velocity, temperature, salinity and pressure, it is necessary to improve the technique used here with the implementation of a multivariate assimilation system, in which the error covariance matrix of the model contemplates the cross variance between these quantities. Work is currently underway in REMO to implement such a system with an Ensemble Optimal Interpolation scheme. It is expected that the forecasts with this new system will further improve the short-term predictability of the model.

ACKNOWLEDGMENTS

This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior of the Ministério da Educação (CAPES) through a master’s scholarship to the first author and a Senior Research fellowship to the second author (BEX 3957/13-6) in the very end of the revision process of this paper. This work was also supported by PETROBRAS and the Brazilian oil regulatory agency ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis), within the special participation research project Oceanographic Modeling and Observation Network (REMO). The authors are thankful to REMO members Davi Mignac, Giovanni Ruggiero and Luiz Otávio Ramos Gavaza for elaborating the routines for model integration and analysis of the results and to Victor Daher for the support in the processing of the altimetry data used in the assimilation. The authors also thank the
HYCOM Consortium for the HYCOM+NCODA analyses available on www.hycom.org.

REFERENCES


NOTES ABOUT THE AUTHORS

Leonardo Nascimento Lima. Oceanography graduate from the Universidade Federal da Bahia (UFBA). He received a M.Sc. degree in Geophysics with emphasis in Physical Oceanography from the Graduate Program at the UFBA. Currently, he works as a researcher of the Rede de Modelagem e Observação Oceanográfica (REMO) in the field of numerical modeling and ocean data assimilation.

Clemente Augusto Souza Tanajura. Mechanical-nuclear engineer with a Ph.D. in meteorology from the University of Maryland, College Park, USA. He worked as a research associate at the Laboratório Nacional de Computação Científica (LNCC) during 18 years. Currently is a Professor at the Universidade Federal da Bahia (UFBA). He was the technical-scientific coordinator of the Rede de Modelagem e Observação Oceanográfica (REMO) from December, 2006 to March, 2013 and is member of the GODAE OceanView Science Team. He works with data assimilation, oceanic and atmospheric numerical modeling, short-term oceanic and atmospheric predictability, and climate studies.