Journal of Operational Oceanography, 2015
http://dx.doi.org/10.1080/1755876X.2015.1014646

Assimilation of sea-level anomalies and Argo data into HYCOM and its impact on the 24 hour forecasts in the western tropical and South Atlantic

F.B. Costa[a,b*] and C.A.S. Tanajura[b,c,d]

[a]*Graduate Program in Geophysics, Physics Institute and Geophysics Institute, Federal University of Bahia (UFBA), Salvador, Brazil;* [b]*Oceanographic Modeling and Observation Network (REMO), Center for Research in Geophysics and Geology/UFBA, Salvador, Brazil;* [c]*Department of Earth and Environmental Physics, Physics Institute/UFBA, Salvador, Brazil;* [d]*Ocean Sciences Department, University of California, Santa Cruz (UCSC), USA*

Argo and along-track sea-level anomaly (SLA) data from satellites were assimilated into the Hybrid Coordinate Ocean Model in the western tropical and South Atlantic Ocean. An optimal interpolation method was employed in the assimilation, and the Cooper and Haines scheme projected the altimetry information into the subsurface. The run with assimilation of SLA and Argo data reduced the root mean square deviation of the 24 h forecasts of SLA, sea surface temperature, and subsurface temperature and salinity by 21.4%, 11.5%, 28.1%, and 15.8%, respectively, with respect to the control run without assimilation. Important improvements were also observed in the circulation.

Introduction

Ocean models in simulation and prediction mode generate forecasts with errors owing to limitations in physical parameterizations, numerical algorithms, resolution, and uncertainties in the atmospheric forcing and boundary conditions (Chassignet et al. 2009). An approach to minimize these errors involves the use of data assimilation methods. They combine observational data with model results to produce the so-called objective analysis with smaller errors than the model alone (Daley 1991). The analysis can be employed as model initial condition and lead to a gain in short- and medium-range predictability, since the latter is quite dependent on the quality of the initial condition (Chassignet et al. 2009; Kalnay et al. 1996).

The analysis depends on the data-assimilation method, on the model errors, and on the quality, availability, and quantity of oceanic observations. Among the assimilation methods, there are the Kalman filters, optimal interpolation (OI), three-dimensional variational (3D-Var), and the four-dimensional variational (4D-Var) schemes. The first three are sequential, and the best estimate is produced at a single time, but an observational time window can be used by considering different errors of observation. They have a relatively low computational cost and are commonly used in operational systems. The 4D-Var produces the best estimate over a finite time interval by minimizing a cost function considering the trajectory of the non-linear model and its linearized adjoint, as well as all observations available in the period. They have a relatively high computational cost (Evensen 2003; Moore et al. 2011b; Xie et al. 2011).

Ocean observations had significant advances in the past two decades. For the sea surface, satellites collect remotely sensed data of sea surface temperature (SST), sea surface height (SSH), and, recently, sea surface salinity (SSS), with high temporal and spatial resolution. Despite the high coverage of the remotely sensed data, it is restricted to the surface and is insufficient to characterize the subsurface thermodynamic variability (Chassignet et al. 2006). In the subsurface, the Argo system plays an important role with over 3500 free-drifting profiling floats measuring temperature and salinity in the upper 2000 m providing many *in situ* data for assimilation systems and other purposes (Oke et al. 2008; Xie and Zhu 2010).

The large quantity of observational data available in near-real time allowed several groups and institutions to implement operational ocean forecasting systems (OOFSs) with data-assimilation techniques that efficiently constrain the surface and vertical structure of their ocean models. The Global Ocean Data Assimilation Experiment (GODAE) and its continuation, the GODAE OceanView, have helped organizing collaboration among the international community dedicated to developing operational oceanography in both global and regional scales. Among the operational systems in GODAE OceanView are the

*Corresponding author. Email: fbcosta.oceano@gmail.com

American Hybrid Coordinate Ocean Model (HYCOM) +Navy Coupled Ocean Data Assimilation (NCODA), the French Mercator Océan (Dombrowsky et al. 2009), and the Brazilian Navy Hydrographic Center (in portuguese CHM) in partnership with the Oceanographic Modeling and Observstion Network (in portuguese REMO). The HYCOM+NCODA employs the HYCOM based on a 1/12° global domain and 32 vertical hybrid layers initialized by the US NCODA system. Altimetry data, SST, and vertical profiles of temperature and salinity from moored buoys and from the Argo floaters are assimilated with a multivariate OI scheme. The French system Mercator Océan employs the Nucleus for European Modeling of the Ocean (NEMO) with 1/4° and 1/12° resolution in global and regional domains, respectively, with 50 levels in z-coordinate. Its assimilation system is based on the singular evolutive extended Kalman (SEEK) filter that assimilates all data available (Cummings et al. 2009; Dombrowsky et al. 2009). The CHM/REMO system is based on HYCOM with 1/4° and 1/12° of horizontal resolution over the whole Atlantic and the South Atlantic Ocean, respectively, and 21 vertical hybrid layers. The CHM/REMO system assimilates SST and sea-level anomaly (SLA) from HYCOM+NCODA analyses with an OI scheme that rely on fixed co-variances between the surface and subsurface according to Ezer and Mellor (1997). These systems are in permanent development. For instance, the HYCOM+NCODA is being upgraded to a 3D-Var scheme into a 1/25° resolution model (Pat Hogan, personal communication).

In Brazil, the first effort to implement an operational ocean forecasting system was initiated with REMO (Lima and Tanajura 2013). The present work describes a pre-operational forecasting system that helped REMO to join the GODAE OceanView. The system focuses on the assimilation of along-track SLA data and Argo temperature and salinity vertical profiles with a simplified OI and the Cooper and Haines scheme (C&H) (Cooper and Haines 1996) into HYCOM. The area of interest is the Atlantic METAREA V, between 36°S and 7°N from 20°W up to the Brazilian coast. This area is the responsibility of the Brazilian Navy for search and rescue missions as an integrant of the International Convention for the Safety of Life as Sea (SOLAS). Also, this area has high economic and environmental relevance owing to petroleum extraction. The general goals of the present work are to contribute to a better understanding of the impacts of data assimilation into HYCOM and to build an intermediate step towards the construction of a more sophisticated data-assimilation system for the CHM/REMO operational ocean forecasting system.

The next sections describe the model configuration, the data-assimilation schemes employed for SLA and Argo data, and the numerical experiments. They are followed by the Results and discussion, and the Conclusions.

Methodology

Ocean model

HYCOM uses isopycnal coordinates for the open stratified ocean, terrain-following coordinates in shallow coastal regions, and fixed pressure-level coordinates in the mixed layer and over unstratified ocean regions (Chassignet et al. 2007). The adjustment of the vertical interfaces occurs after solving the five prognostic equations: two for the horizontal motion, the continuity equation, and two for the thermodynamic conservation that can be salinity and potential temperature (Bleck 2002; Chassignet et al. 2009).

In the present work, HYCOM was configured with horizontal resolution of 1/12° with 601 by 733 grid points in the zonal and meridional directions, respectively, for a subregion of the South and tropical Atlantic, namely from 46°S to 10°N, and from 60°W to 20°W, which contains the METAREA V. This domain was nested in another HYCOM configuration with horizontal resolution of 1/4° over the region 78°S–55°N, 100°W–20°E, excluding the Pacific Ocean and the Mediterranean. Both domains were configured with 21 vertical layers, from which 18 were hybrid, and the top 3 were fixed as z-level coordinate layers with a minimum thickness of 3 m. The full hybrid coordinate system was used, so that 18 isopycnic layers could be transformed into sigma- or z-coordinate layers.

The 1/4° grid was initialized with the World Ocean Atlas (WOA) (WOA 2001) temperature and salinity fields at rest. For spin-up, it was first integrated for 40 years with climatological monthly mean fields of atmospheric forcing from the Comprehensive Ocean-Atmosphere Data Set (COADS) (COADs 2007) (available at http://icoads.noaa.gov). The 1/12° grid was initialized with the interpolated output from the year 31 of the 1/4° HYCOM spin-up run and then integrated for 10 years with COADS forcing. Afterwards, both domains were integrated for eight years from 1 January 2003 to 31 December 2010 employing the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al. 1996). The last output of this run was the initial condition for this work.

For all runs, HYCOM was forced with wind stress, precipitation, surface air temperature, surface specific humidity – transformed into surface mixing ratio – and net longwave and shortwave fluxes. Freshwater fluxes from the main rivers were imposed by a parameterization based on precipitation minus evaporation from climatological monthly mean data.

Altimetry data

SLA data from the Jason-1 and Jason-2 satellites stored by the Geophysical Data Record (GDR) and distributed by the

French *Archivage, Validation et Interprétation des données des Satellites Océanographiques* (AVISO) were used in this work. These data were treated following Tanajura et al. (2013), and regions shallower than 1000 m were excluded owing to uncertainties in tidal corrections. Since the SLA data employed in this work are available for the same location approximately every 10 days, a 1-day observation window may be insufficient to produce significant model constraint. Therefore, a 3-day observation window in the past with respect to the analysis time was used in each assimilation cycle. The latter was realized every 3 days. This strategy reduces the computational cost and allows the model to better diffuse the analysis increments along the integration.

The model SLA interpolated to the points of observation presented discrepancies with respect to the AVISO observations, since the reference surfaces employed in the model and in the observations to calculate the anomalies differ. Therefore, an adjustment was imposed on the observed SLA following Tanajura et al. (2013) and according to the following steps: (1) first, the model SLA was extracted by subtracting the mean SSH from the instantaneous SSH (the mean SSH was calculated from the model free run from January 2004 to December 2010); (2) after interpolating the model SLA to the observation space, the mean along the track was calculated for the model, and the observations and the offset of these means subtracted from the observed SLA; and (3) the adjusted observed SLA was finally used to calculate the innovation in the assimilation algorithm (Figure 1). In previous experiments (Lima and Tanajura 2013; Tanajura et al. 2013) this adjustment has proved to produce good results in correcting the mesoscale structure of the HYCOM SSH field towards observations and having a positive impact in the ocean surface circulation. The along-track adjustment procedure respects the model SSH mean and focuses on the correction of the model mesoscale structures.

SLA assimilation scheme

For the assimilation of SLA, the following OI scheme was employed:

$$X_a = X + K\,(Y - \mathrm{HX}) \qquad (1)$$

where X_a represents the SLA analysis, X is the model background SLA or first guess or *prior* state, Y is the observational vector, H is the observation operator, and K is the gain matrix. The gain matrix depends on model error covariance matrix, B, and the observation error covariance matrix, which was assumed as a diagonal matrix with constant squared errors equal to 0.02 m (Daley 1991). B was parameterized as $B = \sigma^2 G$, with the variance (σ^2) extracted from the 2004–2010 free run, and G is given by the formula

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

For each model grid point, G selects which observations will be used in the analysis considering the zonal distance Δx and the meridional distance Δy between the observation points and the analysis point, and the time period Δt between the observation time and the analysis time. k_x and k_y represent the spatial decorrelation parameters, while k_t represents the temporal decorrelation parameter. In previous SLA assimilation studies, fixed parameters were used by Tanajura et al. (2013) in the same domain of this work. Their results showed substantial improvements in the SLA and the Brazil Current system when compared with the free model run.

However, Lima and Tanajura (2013) delimited seven sub-regions in the HYCOM 1/12° domain based on the standard deviation of the 1/3° resolution AVISO gridded data, and they calculated one spatial decorrelation parameter for each subregion. Their results showed an improvement with respect to the assimilation in which the same decorrelation scale was employed in the whole domain. They employed 3 months (January, February, and March) in their experiments, and the decorrelation parameters were based on these specific months. For the present work, a similar strategy was used, but in addition to the use of different space decorrelation scales for each subregion, the decorrelation parameters also considered temporal variability. Here, they were calculated for each month centred in a 3-month window. The data used in this calculation were the AVISO gridded data from 2007 to 2010.

The constructed SLA analysis is added to the mean model SSH, and the C&H is applied. The analysis increment is projected into the model subsurface by the adjustment of the model layer thicknesses. In order to do this, the C&H scheme takes low- (high)-density water parcel from the upper (bottom) isopynic layers and inserts it as high (low) density water in the bottom (upper) layers to decrease (increase) SSH. This procedure acts only in the isopycnic layers below the mixed layer and has an impact on the temperature vertical profile because of the new layer depths after the adjustment.

Argo assimilation scheme

The Argo data-assimilation scheme was based on Bergthorsson and Doos's (BD) method where the gain matrix, K, is prescribed (Daley 1991). Therefore, K, will depend only on the number of observations inside the radius of influence, and it is parameterized as a function of the distance of each observation to the analysis grid point. Considering i as a model grid point, the method can be described

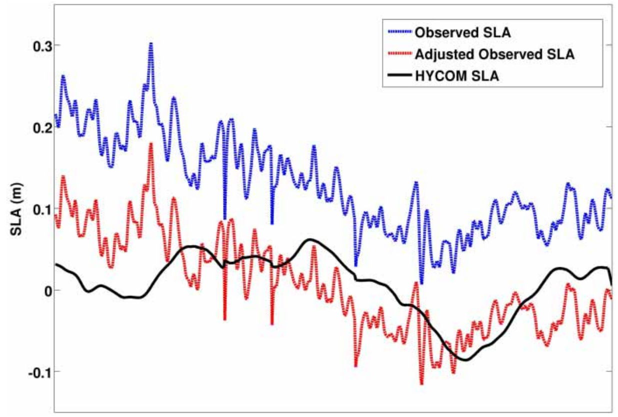

Figure 1. Along-track SLA (m) observed data (blue), interpolated model SLA to the track (black) and adjusted SLA data (red).

by the equation:

$$X_{ai} = X_{bi} + \sum_{j=1}^{n} K_{ij}(Y_j - H_j X_b) \tag{3}$$

where X_{ai} is temperature or salinity analysis at i, X_{bi} is the background at i, n is the number of observations inside the influence radius around point i, Y_j is the observation value at point j, and H_j is the observation operator that projects the model background into the observation point j. The parameterization of K_{ij} follows the formula:

$$K_{ij} = \frac{\exp^{-r_{ij}^2/\alpha R^2}}{\sum_{k=1}^{n} (\exp^{-r_{ij}^2/\alpha R^2} + \varepsilon_{jik}^2)} \tag{4}$$

where, for every K_{ij}, R is the radius of influence, α is a parameter controlling the relevance and spreading of the observation, and ε_{ij}^2 is an estimate of the ratio between the observation and the model square error. The term r_{ij}^2 denotes the distance between each observation and the analysis point, and is given by:

$$r_{ij}^2 = (y_i - y_j)^2 + (x_i - x_j)^2 + (t \times \mathrm{d}t_j)^2 \tag{5}$$

The first and second terms represent the squared distance in each direction between the model grid point i and observation point j. In the third term, t is the time window in days, which includes the analysis day and 2 days in the past, and $\mathrm{d}t_j$ corresponds to the temporal difference between the observation and analysis day.

The usage of the time window increases the spatial coverage of observations. This is especially important regarding the Argo profilers, since each one sends data every 10 days, and the mean global spatial coverage is about 300 km (Sakov et al. 2010). Previous experiments conducted in REMO by Carneiro and Tanajura (personal communication) led to the choice of the following parameters for the present work: constant radius of influence equal to 10 grid points; α equal to 0.2; salinity observational error equal to 0.02 psu; and temperature observation error equal to 0.05°C, for the first 3 months of integration, and then reduced to 0.025°C for the rest of integration period in order to increase the impact of the temperature observations.

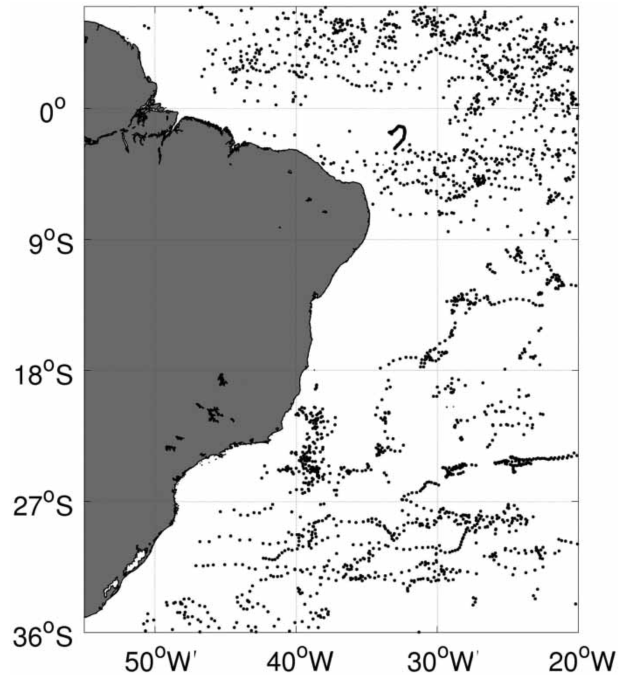

Figure 2. Position of the 2427 Argo profilers assimilated.

The assimilation was made independently for each variable. Also, the assimilation of temperature and salinity was performed in the standard levels employed by the Levitus climatology. Therefore, observations and background were first vertically interpolated to the Levitus levels. After the assimilation was realized, the temperature and salinity analysis were interpolated back to the centre of isopycnic layers, and the increments were added to the background state. HYCOM's hybrid coordinate generator recalculates new layer thicknesses considering these increments. All 2427 Argo profilers assimilated in this work are shown in Figure 2.

Numerical experiments and evaluation metrics

From 1 January 2011 to 31 March 2012, three integrations were performed, employing predicted atmospheric forcing from National Oceanic and Atmospheric Administration (NOAA)/NCEP Global Forecast System (GFS) every 3 h with resolution of 0.5°. The first run had no assimilation, and this is called the control run. In the second run, only the SLA was assimilated, and this is called as the A_SLA run. In the third run, first temperature and salinity than SLA were assimilated; this is called as the A_TS_SLA run. As mentioned above, the observational time window was 3 days – the present day and two in the past – for all the assimilation experiments, which were conducted only every 3 days, starting on 1 January 2011.

The experiments were evaluated objectively by the root mean square deviation (RMSD) defined by:

$$\mathrm{RMSD} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n} (Y_i - X_i)^2} \tag{6}$$

where n is the number of observations, Y_i is the observation value at point i, and X_i is the model value interpolated to same point i. Also, for SLA, temperature, and salinity, Taylor diagrams were constructed. In these diagrams, the analysis from the GODAE OceanView systems HYCOM +NCODA, Mercator-Océan, and CHM/REMO was also included to help comparing the skills of the 24 h forecasts produced in the present work. These analysis systems are referred to hereafter as NCODA, Mercator, and CHM/REMO, respectively. Integrated volume transport of the Brazil Current was also evaluated. This is a key feature for REMO and the Brazilian oceanographic community in general. All the evaluations of the experiments focused on the 24 h forecasts. Therefore, independent observations are used in the evaluation.

Results and discussion

In order to evaluate the assimilation impact objectively, the RMSD of the 24 h SLA forecasts with respect to AVISO were calculated over the METAREA V from 1 April 2011 to 31 March 2012 as shown in Figure 3. Comparing the control run with the A_SLA run, the reduction in RMSD in the latter is substantial in practically all the METAREA V. For instance, in some regions to the north of 0°N, the RMSD is reduced 60%, from 0.10 m to 0.04 m, and around 27°S it is reduced almost 85%, from 0.14 m to 0.02 m. In regions in which the control run was closer to the observations, such as that between 9°S and 18°S east of 30°W, the reduction in the error was smaller, 40%, from 0.05 m in the control run to 0.03 m in the A_SLA. In coastal regions, the model was able to efficiently extrapolate assimilation information and produce RMSD reductions of more than 0.2 m in regions in which the control run had the largest errors, as in the Amazonas river mouth.

For the A_TS_SLA [Figure 3(e)], the RMSD reduction is also substantial with respect to the control run. However, the RMSD is slightly larger than the A_SLA run by about 0.003 m. Even in regions where there were no Argo profilers, particularly from 9°S to 18°S west of 30°W, RMSD was modified. Therefore, HYCOM propagated Argo information to other variables, such as SLA. Despite the fact that the assimilation of Argo increased the 24 h forecast error for SLA, it did not compromise the quality of the forecasts. The overall RMSD mean was 0.070 m, 0.051 m, and 0.055 m for the control run, A_SLA run, and A_TS_SLA run, respectively.

For SST, the RMSD with respect to the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) of the A_SLA [Figure 3(d)] shows that in some regions the error increased, while in others it decreased when compared with the control run [Figure 3(b)]. For instance, along 27°S between 40°W and 25°W, the control run presents an error larger than 2.5°C, while in the A_SLA run the RMSD in the region is reduced to 1.5°C. On the other hand, south of 18°S, excluding the region at 27°S mentioned above, the RMSD was 1.0°C greater than the control run. When the Argo data were included in the assimilation [Figure 3(f)], a large reduction in the RMSD was produced in practically all the METAREA V when compared with the control run and A_SLA run results. The greatest corrections of 1.0°C and 1.45°C were observed north of 0°N and south of 18°S, respectively. Also, with respect to the A_SLA, there were positive impacts produced by the A_TS_SLA between 9°S and 18°S west of 30°W of about 0.4°C, despite the absence of Argo data in this region. These results indicate that the Argo information propagated to other locations and produced a change in the whole SST field along time. The overall RMSD mean was 1.22°C, 1.27°C, and 1.08°C for the control run, A_SLA run, and A_TS_SLA run, respectively.

The Taylor diagram for SLA (Figure 4) corroborates what was shown previously. Compared with the control run, the A_SLA run shows a significant impact in the SLA 24 h forecast towards AVISO gridded data by reducing the centred root mean square deviation (CRMSD) from 0.070 m to 0.052 m and increasing the correlation from 0.44 to 0.71. The A_TS_SLA run presented similar results, reducing the CRMSD from 0.070 m to 0.055 m and increasing the correlation from 0.44 to 0.68. These results indicate that the A_TS_SLA forecast was not as efficient as the A_SLA forecast when SLA is assessed. However, as mentioned above, this lower skill did not compromise the substantial positive impact in SLA when compared with the control run. However, the two assimilation runs have the same standard deviation of SLA and show a slight increase with respect to the control run, from 0.070 m to about 0.073 m, while the observation shows 0.062 m.

Also, the assimilation runs predicted SLA with a quality similar to the other three aforementioned GODAE OceanView operational systems. As expected, the control run had the worst skills, showing that atmospheric forcing is not able to constrain the SLA field. Nevertheless, it is important to highlight that there are differences regarding the assimilation methods and remotely sensed data employed, so there are plenty of

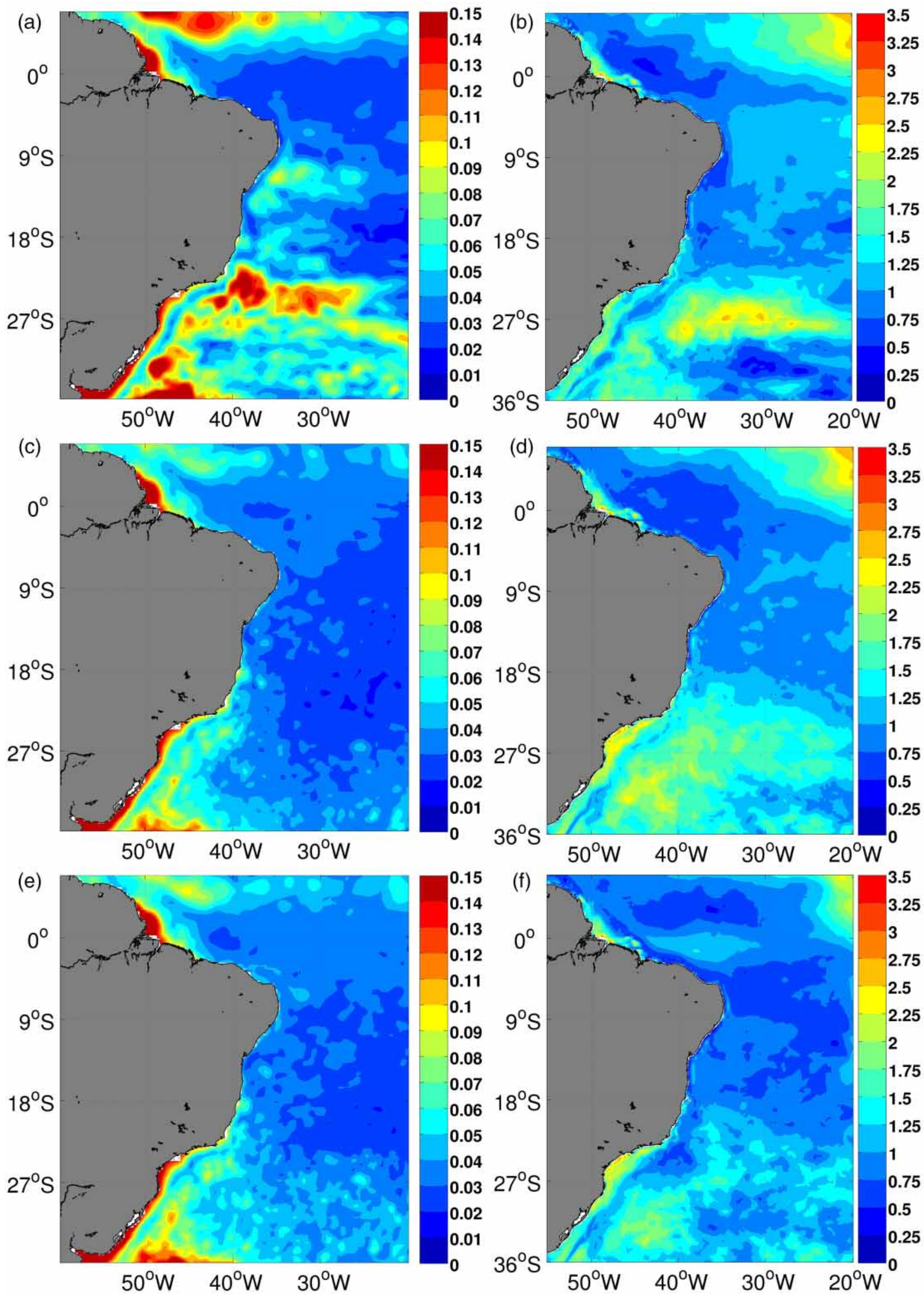

Figure 3. SLA RMSD (m) with respect to AVISO for the (a) control run, (c) A_SLA run, (e) A_TS_SLA run, and SST RMSD (°C) with respect to OSTIA for the (b) control run, (d) A_SLA run, and (f) A_TS_SLA from 1 April 2011 to 31 March 2012.

improvements to be implemented in the system constructed in this work.

To evaluate the impacts of the assimilation in the forecast of the subsurface temperature and salinity, the vertical mean profile of RMSD with respect to Argo data was calculated from 1 April 2011 to 31 March 2012 (Figure 5). The assimilation of only SLA presents a small impact in the temperature forecast, by reducing the RMSD profile between 75 and 1000 m from 1.5°C to 1.4°C. This reduction is more pronounced in the top 500 m. On the other hand, when the Argo data were assimilated, a much more significant impact was achieved, especially in the first 750 m. The RMSD of the 24 h temperature forecast was reduced from 1.6°C to 1.1°C. Around 100 m, all three runs attained the largest error. This is associated with the thermocline region, which presents a sharp

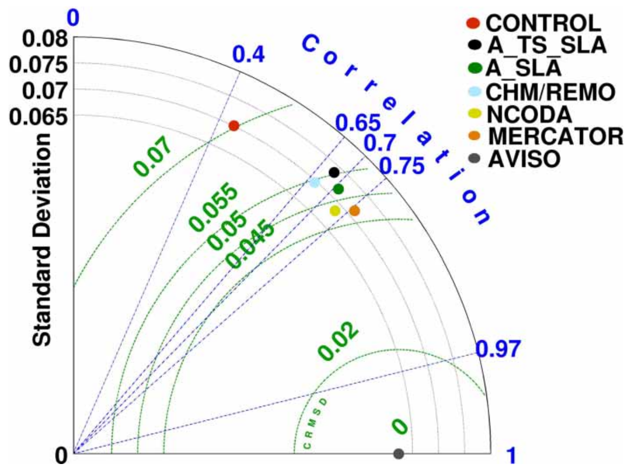

Figure 4. Taylor diagram of SLA with respect to AVISO for 1 April 2011 to 31 March 2012.

temperature gradient, and many models have difficulties to reproduce and forecast (Oke and Schiller 2007; Xie and Zhu 2010). However, it is in this region around 100 m that the corrections by assimilation are the largest along the vertical profile, reducing RMSD from 2.74°C for the control run to 2.59°C and to 1.99°C for the A_SLA run and A_TS_SLA run, respectively.

For the subsurface salinity [Figure 5(b)], both assimilation runs also presented positive impacts, but the assimilation of Argo imposed a much stronger correction. The most significant improvements were from 0.290 psu to 0.276 psu (4.8%) between 50 and 250 m and from 0.267 psu to 0.209 psu (21.7%) between 50 and 500 m for the A_SLA and the A_TS_SLA runs, respectively, with respect to the control run. Similarly to the temperature, the greatest errors and corrections of salinity occurred in the region with intense gradient associated with the halocline.

The Taylor diagram for the 24 h forecast of temperature with respect to Argo data [Figure 6(a)] shows that assimilation of SLA decreased the CRMSD from 1.59°C to 1.50°C and increased the correlation from 0.986 to 0.987 with respect to the control run. On the other hand, the assimilation of Argo reduced the CRMSD to 1.15°C and increased the correlation to 0.991 with respect to the control run. Positive changes were also verified in the standard deviation, which decreased from 9.05°C (control run) to 9.01°C and to 8.61°C for A_SLA and A_TS_SLA respectively, towards observations (8.49°C). For salinity [Figure 6(b)], SLA assimilation reduced the error from 0.232 psu to 0.223 psu and increased the correlation by 0.005, while the Argo assimilation reduced the error to 0.204 psu and increased the correlation by 0.010, with respect to the control run. For the standard deviation, the A_SLA run had a negative impact of 0.012 psu, while A_TS_SLA had a positive impact of 0.011 psu.

Compared with the three aforementioned GODAE OceanView operational systems, the accuracy of the A_TS_SLA run prediction was much closer than the A_SLA prediction and for the salinity was better than CHM/REMO. It was also verified above that errors of temperature and salinity are greater in the thermocline and halocline regions. Consequently, an increase in the model vertical resolution may be required to improve the accuracy of the background, the analysis, and the forecasts in future versions of the forecasting system.

Impacts of the assimilation of altimetry and Argo data in the zonal and meridional currents were also observed. These variables were not assimilated and were adjusted according to the model equations. For instance, the position of the Bifurcation of the South Equatorial Current (BiSEC) was greatly altered by assimilation. Considering the surface mean velocity from 1 April 2011 to 31 March 2012 (Figure 7), the BiSEC in the control run is positioned around 16°S while in the two other runs it is around 11°S. This impact was exclusively due to the SLA assimilation, since there was no difference between A_SLA and A_TS_SLA results at this specific location. The change in the position is towards the observation around 10°S from Stramma et al. (1990) and model results from Silva et al. (2009) and Rodrigues et al. (2007), who indicated 8° S–13°S and 10°S–14°S, respectively, considering the interannual variability. For the 100 m depth (not shown), the

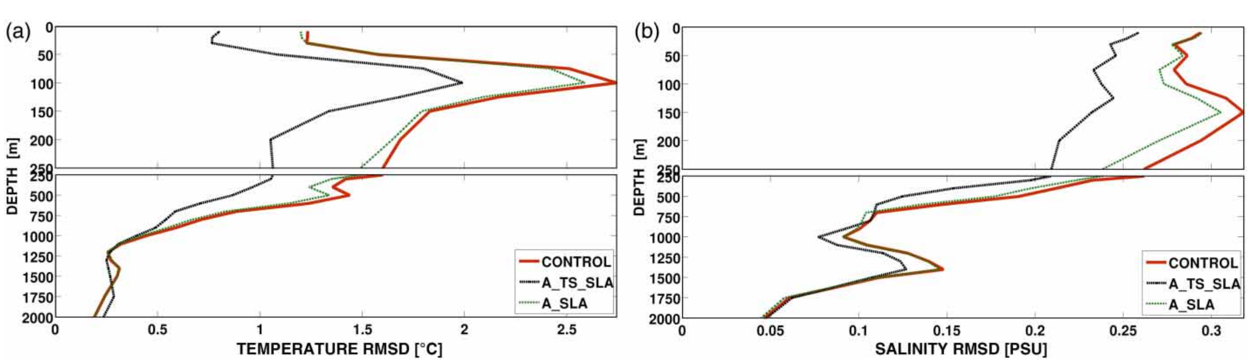

Figure 5. Vertical profile of RMSD with respect to Argo from 1 April 2011 to 31 March 2012 for (a) temperature (°C) and (b) salinity (psu).

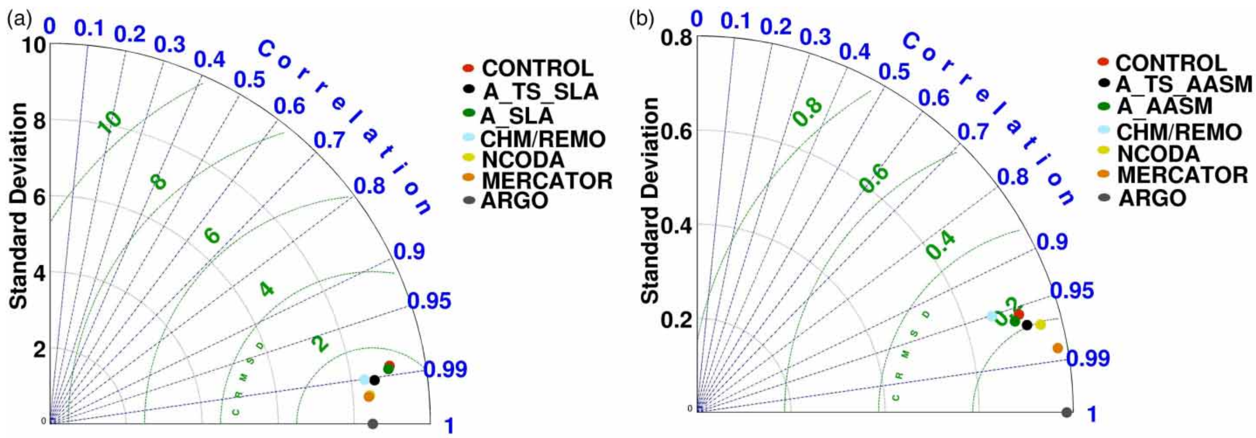

Figure 6. Taylor diagram with respect to Argo from 1 April 2011 to 31 March 2012 for (a) temperature (°C) and (b) salinity (psu).

BiSEC is at 18°S according to the control run and at 16°S according to the two assimilations runs. The change in BiSEC position towards the south with depth has been well documented, and the models were able to capture this pattern. However, the assimilation runs are able to better represent this change with depth, when compared with the observations made by Stramma and England (1999). They observed the BiSEC at 15°S for the 0–100 m layer. The assimilation of Argo data had no significant impact in the currents in the BiSEC region probably owing to the complete absence of Argo data in this region, as shown in Figure 2.

Crucial modifications were also observed in the subsurface currents in other regions, for instance along 24°S (Figure 8). A clear signature of the southward Brazil Current (BC) can be seen in all runs. It is represented by two branches, one on the continental shelf and the other on the continental slope. However, the control run predicted a BC relatively weak and shallow, with a maximum speed of about 0.2 m/s and limited to the upper 250 m on the continental slope. According to Silveira et al. (2004), the BC has a maximum velocity equal to 0.5 m/s and reaches a depth of 400 m. When SLA data were assimilated into HYCOM, the BC became weaker, 0.1 m/s, and shallower, 100 m, showing that at this specific latitude, the SLA alone was not able to improve the representation of the BC. When Argo data were assimilated together with SLA in the A_TS_SLA run, a substantial improvement was verified in the prediction of the BC extension and intensity. The BC in the A_TS_SLA run presented a maximum mean velocity of 0.35 m/s and reached a vertical extension of 400 m. Therefore, this run showed that the assimilation of Argo data had a crucial role in predicting the BC at 24°S.

The mean southward BC transport predicted by all three runs from 1 April 2011 to 31 March 2012 was evaluated with respect to literature along multiple latitudes as shown in Figure 9. In general, the predicted BC transport was less than 10 Sv except at 31°S. The increase in transport with higher latitudes was well captured by the model runs, and this is in agreement with observations. From 13°S to 17°S, the assimilation runs produced a larger and more realistic transport than the control run. In this latitude range, there is no relevant difference between the two assimilation runs, probably owing to the complete absence of Argo profilers in adjacent regions. Therefore, the transport increase can be attributed to the SLA data assimilation. Between 21°S and 24°S, there is a large increase in the BC transport. At 24°S, the largest difference in the mean BC transport is observed between A_TS_SLA, with 10 Sv, and A_SLA, with less than 5 Sv. The latter underestimated the BC transport when compared with most observations, which indicate values around 10 Sv. On the other hand, the A_TS_SLA run achieved very similar results to those found by Stramma (1989) (9.4 Sv and 10.1 Sv), and its standard deviation includes all estimated values in the literature, except for the lowest value according to Evans et al. (1983) (4.1 Sv). These results corroborate that the Argo system assimilation has a crucial role in BC prediction with the present system. Its impact might be stronger with an increase in the amount of data available for assimilation. This would justify the realization of new assimilation experiments to support the development of observing systems in the region.

Conclusions

Vertical profiles of temperature and salinity from Argo and along-track SLA satellite data from Jason-1 and Jason-2 were assimilated for 1 year and 3 months into HYCOM in the Atlantic Ocean METAREA V in two different runs. In the first (A_SLA), only SLA data were assimilated, and in the second (A_TS_SLA), Argo and SLA data were assimilated. In both runs, assimilation was realized every 3 days using a 3-day observational window. For the assimilation of SLA, an OI scheme

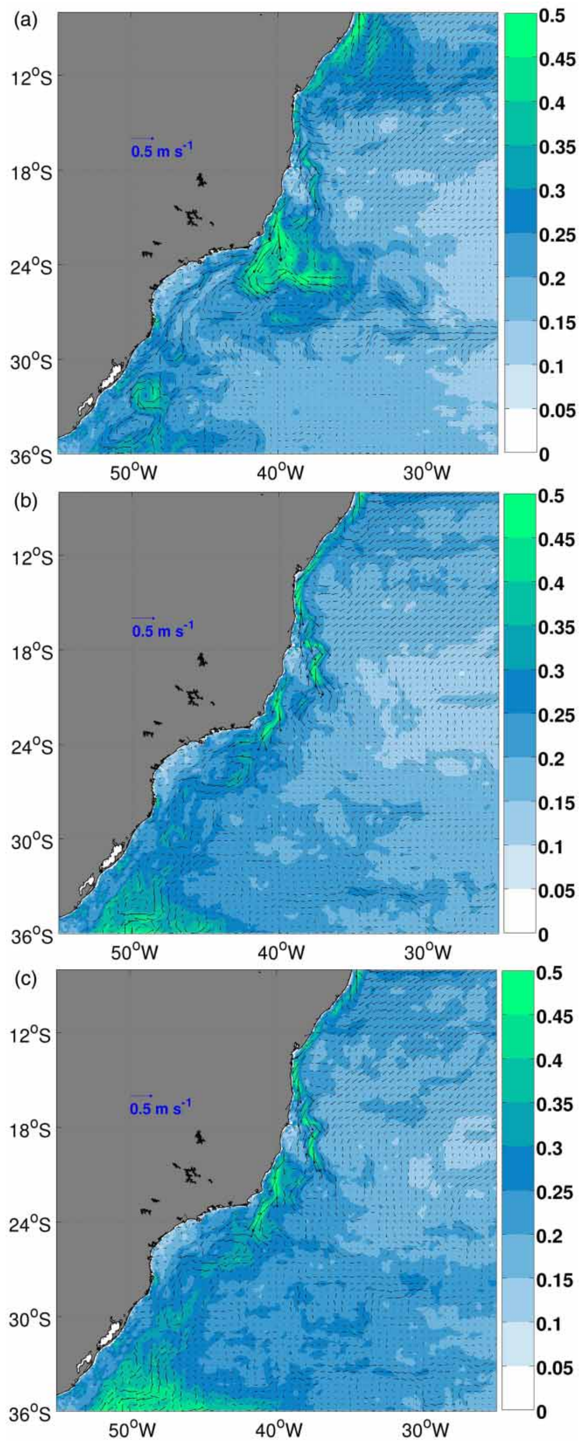

Figure 7. Mean horizontal surface velocity (m/s) from 1 April 2011 to 31 March 2012 for: (a) control run, (b) A_SLA, and (c) A_TS_SLA.

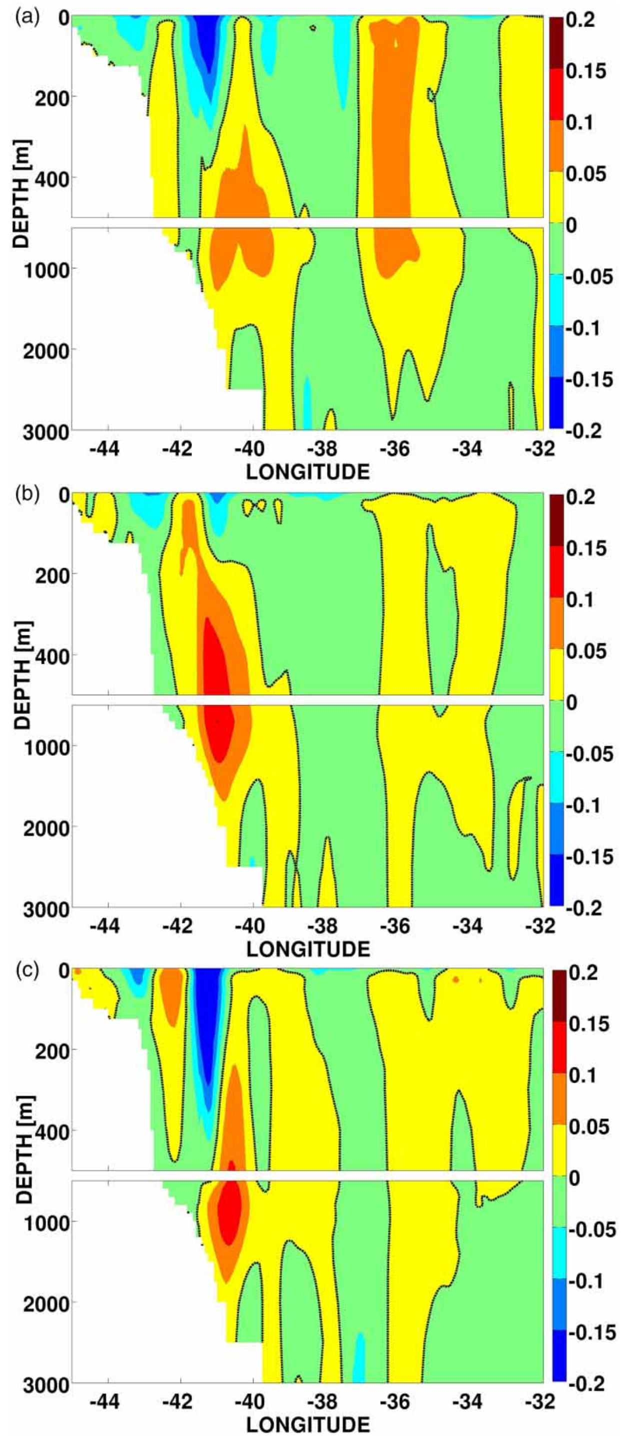

Figure 8. Vertical cross-section of the mean meridional velocity (m/s) from 1 April 2011 to 31 March 2012 at 24°S according to (a) the control run, (b) A_SLA run, and (c) A_TS_SLA run.

that considered seven different localization scales depending on the region and on the month of the year was applied in the surface, and the C&H was employed to project the altimetry information into the subsurface. For the assimilation of Argo data, a projection of the model background state and the Argo data into standard z-levels was conducted first. The assimilation was performed by an OI scheme with prescribed co-variances,

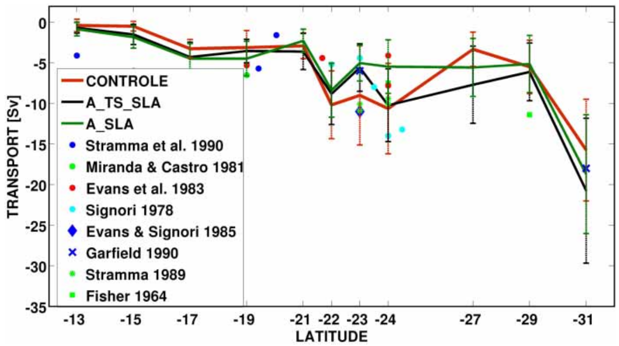

Figure 9. Mean southward BC transport from 1 April 2011 to 31 March 2012 along multiple latitudes. The vertical bars represent one standard deviation. The literature values were taken from Silveira et al. (2000).

and the analyses were projected back to the model hybrid coordinate structure. The impact of the assimilation on the 24 h forecasts was evaluated against a control run without assimilation.

The A_SLA run significantly improved the prediction of the SLA, reducing the RMSD by 27.1%. It also improved the currents, especially the BiSEC, which migrated from 16°S, in control run, to 11°S, in much better agreement with the literature. Assimilation of SLA produced, in general, positive impacts in the subsurface temperature and salinity, reducing RMSD by 3.6% and 3.1% respectively, but it was insufficient to effectively constrain the vertical thermohaline structure. For SST, the A_SLA run presented positive impacts in some regions, reducing the RMSD by more than 1.0°C, and negative impacts in others, increasing the RMSD by 1.0°C. This indicates that, for some regions of the METAREA V, the parameterization of the background error covariance employed here for SLA was not able to produce dynamically consistent corrections on SST (Cummings et al. 2009). On the other hand, assimilation of both Argo and SLA data produced positive and significant impacts in all the evaluated fields with respect to the control run. The RMSD was reduced by 21.4%, 11.5%, 28.1%, and 15.8%, for SLA, SST and subsurface temperature and salinity, respectively, with respect to the control run. The BiSEC position was not sensitive to the assimilation of Argo data at 11°S, but significant improvements in the mean velocity and vertical extension of the BC were verified at 24°S. This latitude also showed the largest difference on BC transport between the two assimilation runs with the A_TS_SLA run presenting the best prediction in agreement with literature. Therefore, the assimilation of both Argo and SLA with the present data-assimilation system seems to be crucial for an accurate 24 h forecast of the BC and to constrain at the same time the surface and subsurface circulation and the thermohaline structure in general (Moore et al. 2011a; Oke et al. 2008). Also, an increase in the Argo profilers in sub-regions of METAREA V seems to be necessary to produce better analyses and forecasts in the region.

This work has demonstrated that the assimilation system developed here could improve the three-dimensional ocean state and variability in HYCOM 24 h forecasts. The employed schemes have a very low computational cost and can be easily implemented for operational purposes. Assimilation of both Argo and SLA data, took only 20 min in a single 2 GHz intel processor. Nevertheless, in order to increase HYCOM short-term predictability, a higher model resolution, particularly in the vertical direction, and a fully multivariate data-assimilation system should be pursued.

Acknowledgements

The authors would like to thank M.Sc. Leonardo Lima and M.Sc. Davi Mignac Carneiro by REMO for their help in the construction of the data-assimilation system used here. The authors also thank commander João Bosco Rodrigues Alvarenga by the CHM, the HYCOM Consortium and the Mercator Ocean for making their analysis available for comparison.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by PETROBRAS and Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, Brazil, via the Oceanographic Modeling and Observation Network [0050.0071946.11.9]. The first author acknowledges support from the FAPESB Foundation, State of Bahia, Brazil [Proc. 1225/2011], and the second author acknowledges support of the CAPES Foundation, Ministry of Education of Brazil [Proc. BEX 3957/13-6].

References

Bleck R. 2002. An oceanic general circulation model framed in hybrid isopynic-Cartesian coordinates. Ocean Model. 37:55–88.

COADs. 2007. Comprehensive ocean data set. Available at ftp://hycom.rsmas.miami.edu/awall/hycom/

Cooper M, Haines K. 1996. Altimetric assimilation with water property conservation. J Geophys Res. 101:1059–1077.

Chassignet EP, Hulburt HE, Smedstad OM, Halliwell GR, Wallcraft AJ, Metzger EJ, Blanton BO, Lozano C, Rao DB, Hogan PJ, Srinivasan A. 2006. Generalized vertical coordinates for eddy-resolving global and coastal ocean forecasts. Oceanogr. 19(1):118–129.

Chassignet EP, Hulburt HE, Smedstad OM, Halliwell GR, Hogan PJ, Wallcraft AJ, Baraille R, Bleck R. 2007. The HYCOM (HYbrid coordinate ocean model) data assimilative system. J Mar Sys. 65:60–83.

Chassignet EP, et al. 2009. US GODAE: global ocean prediction with the HYbrid coordinate ocean model (HYCOM). Oceanogr. 22(2):64–75.

Cummings J, Bertino L, Brasseur P, Fukumori I, Kamachi M, Martin MJ, Mogensen K, Oke P, Testut CE, Verron J, Weaver A. 2009. Ocean data assimilation systems for GODAE. Oceanogr. 22(3):96–109.

Daley R. 1991. Atmospheric data analysis. Cambridge: Cambridge University Press; p. 457.

Dombrowsky E, Bertino L, Brassington GB, Chassignet EP, Davidson F, Hurlburt HE, Kamachi M, Lee T, Martin MJ, Mei S, Tonani M. 2009. GODAE systems in operation. Oceanogr. 22(3):80–95.

Evans D, Signorini SR, Miranda LB. 1983. A note on the transport of the Brazil Current. J Phys Oceanogr. 13:1732–1738.

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

Ezer T, Mellor GL. 1997. Data assimilation experiments in the gulf stream region: how useful are Satellite-Derived surface data for nowcasting the subsurface fields? J Atmos Oceanic Technol. 14:1379–1391.

Kalnay E, et al. 1996. The NCEP/NCAR 40-year reanalysis project. Bulletim Am Meteorol Soc. 77:437–471.

Lima LN, Tanajura CAS. 2013. 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. Revista Brasileira de Geofísica. 31(2):271–288.

Moore AM, Arango HG, Broquet G, Edwards C, Veneziani M, Powell B, Doyle JD, Costa D, Robinson P. 2011a. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. III – Observation impact and observation sensitivity in the California Current system. Prog Oceanogr. 91:74–94.

Moore AM, Arango HG, Broquet G, Weaver AT, Zavala-Garay J. 2011b. The regional ocean modeling system (ROMS) 4-dimensional variational data assimilation systems Part I – System overview and formulation. Prog Oceanogr. 91:34–49.

Oke PR, Brassington GB, Griffin DA, Schiller A. 2008. The Bluelink ocean data assimilation system (BODAS). Ocean Model. 21:46–7.

Oke PR, Schiller A. 2007. Impact of ARGO, SST, and altimiter data on an eddy-resolving ocean reanalysis. Geophys Reasearch Lett. 34:L19601.

Rodrigues RR, Rothstein LM, Wimbush M. 2007. Seasonal variability of the South equatorial Current bifurcation in the Atlantic Ocean: a numerical study. J Phys Oceanogr. 37:16–30.

Sakov P, Evensen G, Bertino L. 2010. Asynchronous data assimilation with the EnKF. Tellus. 62(A):24–29.

Silva M, Araujo M, Servain J, Penven P, Lentini CAD. 2009. High-resolution regional ocean dynamics simulation in the southwestern tropical Atlantic. Ocean Model. 30(4):256–269.

Silveira ICA, Calado L, Castro BM, Cirano M, Lima JAM, Mascarenhas AS. 2004. On the baroclinic structure of the Brazil Current–intermediate western boundary Current system at 22°–23°S. Geophys Res Lett. 31:L14308:1–5; doi:10.1029/2004GL020036.

Silveira ICA, Schmidt ACK, Campos EJD, Godoi SS, Ikeda Y. 2000. A Corrente do Brasil ao largo da Costa Leste Brasileira. Rev Bras Oceanografia. 48:171–183.

Stramma L. 1989. The Brazil Current transport south of 23°S. Deep Sea Res. 36(4A):1875–1886.

Stramma L, England M. 1999. On the water masses and mean circulation of the South Atlantic Ocean. J Geophys Res. 104 (C9):20863–20883.

Stramma L, Ikeda Y, Peterson RG. 1990. Geostrophic transport in the Brazil Current region. Deep Sea Res. 37(1A):1875–1886.

Tanajura CAS, Costa FB, Silva RR, Ruggiero GA, Daher VB. 2013. Assimilation of sea surface height anomalies into HYCOM with an optimal interpolation scheme over the Atlantic Ocean METAREA V. Rev Bras Geofís. 31(2):257–270.

WOA. 2001. World ocean Atlas, national oceanic and Atmospheric administration US Department of Commerce. Prepared by the Ocean Climate Laboratory National Oceanographic Data Center Silver Spring. [online] available at http://www.nodc.noaa.gov/OC5/WOA01/pr_woa01.html.

Xie J, Counillon F, Zhu J, Bertino L. 2011. An eddy resolving tidal-driven model of the South China Sea assimilating along-track SLA data using the EnOI. Ocean Sci. 7:609–627.

Xie J, Zhu J. 2010. Ensemble optimal interpolation schemes for assimilating ARGO profiles into a hybrid coordinate ocean model. Ocean Model. 33:283–298.