Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3

Peter R. Oke[a,*], Pavel Sakov[b], Madeleine L. Cahill[a], Jeff R. Dunn[a], Russell Fiedler[a], David A. Griffin[a], Jim V. Mansbridge[a], Ken R. Ridgway[a], Andreas Schiller[a]

[a] *CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program, Hobart, Tasmania, Australia*

[b] *Bureau of Meteorology Research Centre, Melbourne, Victoria, Australia*

Abstract

The generation and evolution of eddies in the ocean are largely due to instabilities that are unpredictable, even on short time-scales. As a result, eddy-resolving ocean reanalyses typically use data assimilation to regularly adjust the model state. In this study, we present results from a second-generation eddy-resolving ocean reanalysis that is shown to match both assimilated and with-held observations more closely than its predecessor; but involves much smaller adjustments to the model state at each assimilation. We compare version 2 and 3 of the Bluelink ReANalysis (BRAN) in the Australian region. Overall, the misfits between the model fields in BRAN3 and observations are 5-28% smaller than the misfits for BRAN2. Specifically, we show that for BRAN3 (BRAN2) the sea-level, upper ocean temperature, upper-ocean salinity, and near-surface velocity match observations to within 7.7 cm (9.7 cm), 0.68°C (0.95°C), 0.16 psu (0.18 psu), and 20.2 cm/s (21.3 cm/s) respec-

*Corresponding Author: Dr. Peter R. Oke
Email: peter.oke@csiro.au; Tel: +61 3 6232 5387; Fax: +61 3 6232 5123

tively. We also show that the increments applied to BRAN3 - the artificial adjustments applied at each assimilation step - are typically 20-50% smaller than the equivalent adjustments in BRAN2. This leads us to conclude that the performance of BRAN3 is more dynamically consistent than BRAN2, rendering it more suitable for a range of applications, including analysis of ocean variability, extreme events, and process studies.

Keywords: Ocean Reanalysis, Data Assimilation, Ensemble Optimal Interpolation, GODAE, Operational Oceanography

---

1. Introduction

The mesoscale ocean circulation is dominated by the generation, evolution, interaction, and decay of eddies. Eddies typically develop as a result of instabilities associated with either the horizontal shear of the circulation - barotropic instabilities; or vertical shears - baroclinic instabilities (e.g., Lee et al., 1991; Marchesiello et al., 2003; Feng et al., 2005). These instabilities are unpredictable, even on short time-scales (e.g., O'Kane et al., 2011). Data assimilation is therefore a necessary tool for initialising and constraining an ocean model to realistically reproduce the mesoscale ocean circulation in either eddy-resolving or eddy-permitting models (e.g., Carton et al., 2000; Oke et al., 2005; Ferry et al., 2007; Carton and Giese, 2008). A free running model, without data assimilation, can produce realistic mesoscale variability - but without data assimilation, a model will not reliably reproduce particular "eddy events", with eddies in the correct place and time, with the correct intensity and characteristics. Most applications of data assimilation involve the sequential adjustment of the model state to keep it aligned with observations (e.g., Dombrowsky et al., 2009; Zhang et al., 2010; Moore et al., 2011). These updates inevitably interfere with the dynamic balance of the model (e.g., Balmaseda and Anderson, 2009; Oke and Griffin, 2011). The adjustments act as a source of momentum, heat and freshwater that is not easily associated with any specific dynamical process. This makes the use of a data assimilating model for understanding processes somewhat problematic. It is therefore a common goal of a data assimilating model to reduce the magnitude of the adjustments, without compromising the fit to observations. Some data assimilating studies have modified the forcing fields and model parameters, rather than the model state (e.g., Stammer et al., 2002; Koehl et al., 2007; Di Lorenzo et al., 2007; Moore et al., 2009). However, the efficacy of these approaches for eddy-resolving applications, where instabilities are prevalent, is unclear. As a result most data assimilating eddy-resolving models, even those based on variational methods, use a sequential approach involving explicit updates to the model state (e.g., Kurapov et al., 2009; Cummings et al., 2009; Zhang et al., 2010; Moore et al., 2011; Kurapov et al., 2011; Yu et al., 2012). In this study, we present an evaluation of a second generation reanalysis system that is shown to match observations more closely than the first generation system, even though the adjustments during the assimilation step are smaller. This development is a continuation of the Bluelink effort (Schiller et al., 2009a), that was founded under GODAE (Smith, 2000), and continues under GODAE OceanView (www.godae-oceanview.org).

More specifically, we compare the performance of the two most recent versions of the Bluelink ReANalysis (BRAN) - versions 2p1 and 3p5 - hereafter BRAN2 and BRAN3. BRAN is a multi-year integration of the Bluelink ocean model, called the Ocean Forecasting Australian Model (OFAM); and the Bluelink Ocean Data Assimilation System (BODAS; Oke et al., 2008). OFAM and BODAS are combined by sequentially running the model for several days, then combining a model field with observations of sea-level anomaly (SLA), sea-surface temperature (SST), and in situ temperature and salinity from a range of sources. BRAN can be thought of as an observation-based estimate of the ocean circulation, where the model is being used to interpolate between observations that are sparse in time and space, while also extrapolating the observations to provide estimates of unobserved variables. Analogous analyses of ocean observations exist for single variables (e.g., Le Traon et al., 1998) that have no constraint to dynamics, and multiple variables (e.g., Guinehut et al., 2004, 2006; Ridgway and Dunn, 2010) that attempt to respect the ocean water mass properties and linear dynamics (e.g., geostrophy). By contrast, the type of reanalyses presented here (e.g., Ferry et al., 2007; Oke et al., 2008; Schiller et al., 2008; Balmaseda et al., 2012) use primitive equation dynamics to fit data. The risk of this approach is that the penalty for over-fitting the data is potentially much greater (e.g., numerical instability). We therefore monitor this closely by analysing the model mis-match to unassimilated data; and the size of the shocks during each assimilation cycle.

Results from the first BRAN experiment (BRAN1p0; Oke et al., 2005), a 12-year reanalysis, showed that the Bluelink system could produce three-dimensional, time-varying fields that are qualitatively consistent with the real ocean. The configuration of BRAN1p0 was quite immature, and as a result, the model was poorly constrained by observations. The system was refined for BRAN1p5, spanning only 2003-2006, with the addition of the assimilation of SST and other minor changes, resulting in a reanalysis that was closer to observations, but was still poorly constrained (Oke et al., 2008). One of the limitations of BRAN1p5 was the initialisation. BRAN1p5 used a simple Newtonian nudging to initialise the model after each assimilation. This was a conservative approach that succeeded in eliminating much of the "noise" (model-shock) generated after each assimilation, associated with the dynamic imbalance introduced during the update step, but resulting in observations being under-fitted. Version 2p1 of BRAN (Schiller et al., 2008, BRAN2), covered the period 1993-2006, and was largely based on BRAN1p5, but included a few moderate changes to the background error estimates, the initialisation (but still used nudging), and some changes to the model. Like BRAN1p5, BRAN2 under-fitted observations and showed a tendency for the eddies to be somewhat discontinuous in time - a characteristic that is clearly related to the dynamical imbalance introduced after each assimilation. The latest version of BRAN - version 3p5 that is first described here, includes changes to the initialisation (Sandery et al., 2011), localisation method, the assimilation algorithm, and pre-processing of observations and improvements to their error estimates.

In this paper, the model is described in section 2, and the important aspects of the data assimilation system, including the differences between the BRAN2 and BRAN3 configurations, are described in section 3. An overview of the assimilated observations is presented in section 4, followed by a series of comparisons between both assimilated and withheld observations with model fields from BRAN2 and BRAN3 in section 5. An analysis of the increments, or data assimilation adjustments, in section 6, then the conclusions in section 7. The technical details of the assimilation and data-processing are described in Appendix A.

2. Model

The Bluelink ocean model, called the Ocean Forecasting Australia Model (OFAM), has been developed over many years. The first and second versions of OFAM (OFAM1 and OFAM2) are eddy-resolving in the 90°-sector centred on Australia and south of about 20°N. In this study we present results from BRAN2, using OFAM1 - spanning January 1993 to December 2006; and BRAN3, using OFAM2 - spanning January 1993 to September 2012. The key differences between the model used for BRAN2 and BRAN3 are listed in Table 1.

OFAM1 and OFAM2 are configurations of the GFDL Modular Ocean Model (Griffies et al., 2004, OFAM1 uses MOM40d; OFAM2 uses MOM4p1). To date, all versions of OFAM have been developed for analysis and prediction of the upper ocean circulation, so OFAM2 (OFAM1) has 5 m (10 m) vertical grid spacings at the ocean surface and graduated to 10 m vertical grid spacings over the top 200 m. The horizontal grid spacings are 1/10° between 90-180°E and south of about 20°N; 1 ° across the rest of the Indian Ocean and the Pacific to 60°N; and 2° in the Atlantic and far north Pacific Ocean. The horizontal grid spacing changes gradually over 1° between each transition region. To accommodate the inhomogeneous resolution, the horizontal viscosity is resolution and state-dependent, based on the Smagorinsky scheme (Griffies and Hallberg, 2000). The bottom topography for OFAM2 is based on Smith and Sandwell (1997); and OFAM1 is a blend of DBDB2 and GEBCO topography (www7320.nrlssc.navy.mil/DBDB2WWW; www.ngdc.noaa.gov/mgg/gebco/). The turbulence closure model used by OFAM is a version of the hybrid mixed-layer scheme (Chen et al., 1994). OFAM2 also uses an implicit tidal mixing scheme to represent the mixing associated with tides (Lee et al., 2006). Note that OFAM2 does not include explicit tidal forcing - it merely includes a parameterisation that represents the mixing effects of tides.

For both BRAN2 and BRAN3, OFAM is forced with surface fluxes of momentum, heat, and freshwater. BRAN2 uses 2.5°-resolution, 6-hourly fluxes from ERA-40 (Kallberg et al., 2004) between 1993 and 2002, and fields from the European Centre for Medium-Range Weather Forecasting (ECMWF) operational forecasts (http://data.ecmwf.int/data/d/era40 daily) between 2003 and 2006. BRAN3 uses 1.5°-resolution, 3-hourly fluxes from ERA-Interim (Dee and Uppala, 2009). For BRAN2, the above-mentioned fluxes are applied to OFAM1 unaltered. We found that this resulted in a trend in global averaged MSL due to an imbalance between the precipitation and evaporation (and river) fields (recall that MOM is volume conserving). This resulted in a negative bias in BRAN1.5 and BRAN2 that negatively impacted the assimilation (Oke et al., 2008). For BRAN3, we adjust the surface fluxes in advance to ensure that the freshwater fluxes are globally balanced. This is achieved by adding a small amount of precipitation everywhere - a “drizzle”. The magnitude of the drizzle is smaller than all other components of the freshwater budget and changes annually to ensure that the model’s global- and annual-averaged MSL remains constant for the duration of the run. We also scale the long wave flux so that the averaged net heat flux is 1.3 W m^{-2} following Trenberth et al. (2009).

OFAM has been used for many studies, including ocean reanalyses (Oke et al., 2005, 2008; Schiller et al., 2008), observing system experiments (Oke and Schiller, 2007), an investigation of a series of coral bleaching events in the Great Barrier Reef (Schiller et al., 2009b), an analysis of eddy dynamics in the Tasman Sea (Oke and Griffin, 2011), an analysis of fronts in the Southern Ocean (Langlais et al., 2010), an investigation of the seasonality of Chlorophyl a in anti-cyclonic eddies off Western Australia (Dietze et al., 2009), and climate downscaling (Sun et al., 2012). An operational version of OFAM2 is run at the Bureau of Meteorology and is described by Brassington et al. (2007). The most recent version of OFAM, OFAM3, has been integrated for an 18-year run and evaluated by Oke et al. (2013), but a data-assimilating run of OFAM3 has not yet been conducted.

3. Data Assimilation

The data assimilation system used for all BRAN experiments is called BODAS (Oke et al., 2008). An overview of the changes to BODAS for BRAN are summarised below and in Table 2. There are many differences between the version of BODAS used for BRAN2 and BRAN3. The version used for BRAN3 includes many technical changes that were motivated to improve the scalability and robustness of the system, to make it computationally more efficient, and to enable the extraction of additional diagnostics. The improvements to the scalability mean that more observations can be assimilated directly for the same cost, yielding analyses that have a better fit to the data. The other major change for BRAN3 relates to the initialisation - the step in the assimilation when the model state is updated. Other changes relate to the specific implementation of BODAS and the parameters used in its application, including changes to the error estimates of the background field and the observations, and improved pre-processing of observations. The details of the differences in the assimilation system are presented in Appendix A.

The data assimilation method used here is Ensemble Optimal Interpolation (EnOI; Oke et al., 2002; Evensen, 2003). EnOI is based on the Ensemble Kalman Filter (EnKF; Evensen, 1997, 2003), but it uses a time-invariant ensemble to approximate the system's background error covariance. EnOI is inexpensive and robust, and has been tested and shown to be effective for a range of ocean applications (e.g., Oke et al., 2005, 2007, 2008, 2009, 2010; Counillon et al., 2009; Fu et al., 2009; Counillon and Bertino, 2009; Wan et al., 2010; Xie and Zhu, 2010; Srinivasan et al., 2011).

For an EnKF, the ensemble mean and the ensemble perturbations are updated during every assimilation cycle. This yields a time-varying estimate of the system's background error covariance. By contrast, for EnOI, the ensemble is time-invariant, so only the background field (analogous to the ensemble mean in the EnKF) is updated. Here, we apply efficient methods developed for the EnKF to EnOI.

Specifically, the covariance localisation method, used in the previous version of BODAS (Oke et al., 2008), is replaced by a local analysis. Both methods are known to be fundamentally similar (Sakov and Bertino, 2011), except that the local analysis has some significant practical advantages. The main advantage is that while covariance localisation requires an inversion in observation-space, local analysis makes it possible to perform inversions in ensemble-space. For most practical applications, the ensemble size is many orders of magnitude less than the number of observations assimilated. An inversion in ensemble-space is therefore much more efficient than an inversion in observation-space. The details of the assimilation algorithm and localisation are described in Appendix A.1.

The initialisation of the ocean model is as important as the accurate calculation of each analysis. A poor initialisation scheme (e.g., direct insertion), results in a poor forecast. The analyses produced by EnOI are dynamically unbalanced. That is - the analyses are not precisely a solution to the model's equations. However, the analysis increments that are introduced after each analysis can be shown (see Appendix A.1) to be comprised of a linear combination of model anomalies. Since the model produced these anomalies during a free model run, the resulting increments are consistent with the model equations and the model configuration. That is, the ensemble only contains scales and features that the model can generate. Although this doesn't yield analyses that are in perfect dynamic balance, owing to model non-linearities, the consistency between analyses and the model is regarded as a strength of ensemble data assimilation.

For BRAN2, we calculate an analysis every 7 days, then update the model using nudging over one day with a 12-24 hour nudging time-scale, with shorter time-scales at higher latitudes (see Schiller et al., 2008, for details). For BRAN3, we calculate an analysis every 4-days, then update the model using adaptive initialisation (Sandery et al., 2011). Adaptive initialisation is a more sophisticated form of nudging, where the model is nudged towards an analysis (that was constructed by combining the model state with observations) using a time-scale that changes with time and space. Where and when the difference between the model state and the analysis is large, the nudging time-scale becomes short. As the model state approaches the analysis, the nudging time-scale increases. Using this approach, the discontinuity in the model forcing at the end of the nudging period is greatly reduced, so that the model smoothly transitions back to a free-running, dynamically consistent integration.

Another difference in the initialisation of BRAN2 and BRAN3 is the variables that are explicitly updated. For BRAN2, only the sea-level, temperature and salinity were updated. The velocity field was left to adjust during and after the nudging period. For BRAN3, only the temperature, salinity, and velocity fields are updated. During the initialisation period, the model sea-level is left to adjust according to the model physics without explicit adjustments applied to sea-level directly. In short trial runs of the systems, we found little difference between experiments with and without explicit adjustments to sea-level. However, in some circumstances, it was found that explicitly adjusting sea-level resulted in the generation of barotropic waves that degraded the solution.

The most important differences in the data assimilation system applied to BRAN2 and BRAN3 are described above. Other differences (Table 2) include the ensemble size - BRAN2 and BRAN3 use a 72-member and 144-member ensemble, respectively. The ensemble for BRAN3 (BRAN2) is constructed using fields from a model run called Spinup6p8 (Spinup4/5), a configuration of OFAM2 (OFAM1). BRAN3 uses shorter localising length-scales than BRAN2, allowing BRAN3 to better fit observations. Several aspects of the observations are also different, for example refined instrument error estimates (see Appendix A.3), particularly with respect to the "age error" of observations; and the assimilation of data from different (better) observational databases (e.g., higher resolution satellite SST). The assimilation of altimeter data has been improved in BRAN3 - with more careful processing of altimeter data to avoid biases arising from the assimilation of observations that include the effects of thermal expansion into a (Boussinesq) model that does not include the effects of thermal expansion.

4. Assimilated observations

Both in situ and satellite observations are assimilated in a single step by BODAS. The time-distribution of the assimilated data is displayed in Figure 1 that includes an indication of the availability of data from different satellite missions, and the number of temperature and salinity profiles assimilated at each assimilation step in BRAN3. Data from all available altimeters are assimilated into BRAN2; while data from GFO are withheld from BRAN3. We with-hold GFO from BRAN3 because it has greater errors than other altimeters, and to keep it in reserve as a with-held data set for comparison. For BRAN3, all altimeter data were obtained from RADS in August 2012. BRAN2 was produced in 2005, using the Geophysical Data Records (GDRs) for all altimeters. Satellite SST data are assimilated from AVHRR throughout BRAN2 and BRAN3, using a composite of data from Pathfinder (and NAVO for BRAN3), using the pre-processing described by

Andreu-Burillo et al. (2010). BRAN2 assimilated a 54-km resolution version of Pathfinder SST, while BRAN3 assimilates a 4-km resolution version of the same database. SST from AMSR-E is assimilated in both BRAN2 and BRAN3.

In situ profiles of temperature and salinity are assimilated from a range of sources. Prior to January 1998 we assimilate hydrographic data from World Ocean Circulation Experiment (WOCE) Hydrographic Program (WHP), World Ocean Database 2005 (WOD05; Boyer et al., 2006), and the Quality controlled Ocean Temperature Archive (QuOTA; Gronell and Wijffels, 2008), which contains all XBT data in the Indian and South-West Pacific. After January 1998 we assimilate the WOCE Upper Ocean Thermal (UOT) database that includes global XBT data, except in the Indian Ocean where we use QuOTA for XBTs. This change in data source explains the increase in the number of assimilated temperature profiles in 1998 in Figure 1. We also assimilate profiles from Argo, and temperature and salinity from the TAO array. The dramatic increase in the number of in situ profiles - particularly for salinity - when the Argo program became established is clearly evident (Figure 1a).

Before each assimilation, all observations that are available for assimilation are pre-processed. Although all assimilated observations are sourced from delayed-mode quality controlled sources; we also apply a simple background check, flagging as bad any data that differs significantly from the model background field. Specifically, if an observation differs by more than five times the intraseasonal standard deviation (computed from the ensemble) than the data is not assimilated. Along-track SLA (atSLA; and SST) observations are combined to form super-observations so that there is no greater than one "super-observation" for every 0.2×0.2° box. The details of this processing are described in Appendix A.5. Similarly, the in situ profiles that are available for assimilation are "thinned" prior to assimilation , so that no greater than one profile of each type (i.e., temperature and salinity) is assimilated for every 0.5×0.5° box. A different nominal resolution of the super-obing for atSLA and SST, and sub-sampling for in situ profiles differs because of the different resolution of the original data-sets, with atSLA spaced about 7-km along altimeter tracks, SST spaced 4-km, and Argo profiles typically spaced 100-300 km. Again, the details of this processing are described in Appendix A.5. The other important aspect of observation pre-processing relates to the conversion of the model sea-level to SLA. There are many subtle aspects to this processing that are described in Appendix A.4.

5. Results

5.1. Comparison with altimetry

We compare daily-averaged SLA fields from BRAN2 and BRAN3 with atSLA from Topex/Poseidon (T/P) for the entire T/P mission (1993-2005) in Figure 2. We show time-series of the root-mean-squared difference (RMSD) and the correlation between the observed and modelled fields for different regions that are defined in Table 3 and in Figure 3. Although T/P data are assimilated into both BRAN2 and BRAN3, we regard this comparison as important, because it demonstrates the degree to which the reanalyses match those observations - a necessary, but not sufficient criterion for validating the reanalyses. We interpret the RMSDs and correlations presented here as an indication of the errors in the reanalyses. However, we note the observations are imperfect. Indeed, during the assimilation step, we assume that the T/P data has an error ranging from 3-20 cm, depending on the estimated representation and age errors of the data (see Appendix A.3).

The RMSD and correlation statistics in Figure 2 are produced by first removing a reference MSL from the model's sea-level, and then interpolating the daily-averaged model SLA to each observation location for each day. The resulting model-observation comparisons differ from the comparisons performed during the assimilation step in constructing the innovations (observed fields minus the interpolated background fields; see Appendix A.1 and equation (A.1)). For each analysis in BRAN3 (BRAN2) atSLA altimeter data in a 21-day (11-day) time-window, centred on the analysis time, are assimilated by first differencing them with the model SLA at the analysis time. For assimilation, each observation is weighted by assigning an error variance that includes a significant component due to the relative "age" of each observation (see Appendix A.3 for more details). Further, many observations are combined, forming"super-observations", as described above. By contrast, for the comparisons presented here, we compare the model SLA for each day with the atSLA observations for just the same day. The T/P atSLA observations used here for evaluation are from the RADS database (accessed in August 2012). This is the same data assimilated into BRAN3, but recall that BRAN2 used altimeter data from the GDRs.

The RMSD between the observations and the BRAN3 SLA are less than the BRAN2 SLA for 99.8% of the time (see Figure 2). Similarly, the BRAN3 correlations of SLA with T/P are almost always greater than the BRAN2 correlation. Within the Australian region, the time-averaged (plus or minus the standard deviation) BRAN3 RMSD from T/P atSLA is 7.7$\pm$0.5 cm and the BRAN2 RMSD is 9.7$\pm$0.8 cm.

Maps of the RMSD between BRAN SLA and T/P atSLA are shown in Figure 4, along with a map of the standard deviation of the T/P atSLA. This latter field provides a comparison between the model-data differences and the observed signal. The statistics in Figure 4 are produced by comparing time-series of the modelled and observed fields in $2\times2^\circ$ bins. We find that in all regions around Australia, the RMSD for BRAN3 is less than the standard deviation of the observed signal - indicating that the signal to noise ratio for SLA is greater than one everywhere (and much greater than one in many locations). By contrast, the RMSD for BRAN2 exceeds the observed standard deviation in some locations, including the west Tasman Sea and some eddy-rich regions along the path of the Antarctic Circumpolar Current (ACC). Between the latitudes of 20°S to 20°N, the average RMSD for BRAN3 is about 4 cm, which is comparable to the instrument error of T/P atSLA that is assumed in the assimilation step and estimated by Ponte et al. (2007). This indicates that in those regions, the model is fitting the T/P observations to an optimal degree - fitting any closer would be over-fitting. The RMSD in the west Tasman Sea, where the eddy field associated with the East Australian Current (EAC) is very energetic, shows a local maximum in both BRAN2 and BRAN3. In that region the RMSD for BRAN3 is about half the RMSD in BRAN2. At some locations along the path of the ACC, the RMSD in BRAN2 and BRAN3 are comparable, but at many locations the results for BRAN3 are clearly better than BRAN2.

Maps of the correlation between SLA from the BRAN experiments and T/P atSLA are shown in Figure 5. This comparison demonstrates that the SLA in BRAN3 is better correlated with observations everywhere, compared to BRAN2. BRAN3 shows high correlations between about 30°S and 20°N, with particularly high correlations in the Indian Ocean at about 10-12°S, where seasonal Rossby waves are prevalent (e.g., Schouten et al., 2002; Rao and Behera, 2005). The correlations in BRAN3 are also significantly greater than BRAN2 in the west Tasman Sea and in the region east of the Philippines. This indicates than BRAN3 is more realistically reproducing the variability - and particularly the eddies - in these energetic western boundary current regions. West of New Zealand, and south of the Great Australian Bight (GAB), the correlations for both BRAN2 and BRAN3 are relatively low. However, we note that the magnitude of the SLA signal in those regions is very small (Figure 4c), so the signal to noise ratio in the observations is low rendering the use of T/P observations for model-evaluation in those regions somewhat problematic.

In both BRAN experiments the RMSD increases, and the correlations decrease, at higher latitudes to the south. We suspect that this is due to a combination of shorter length-scales of baroclinic features in the ocean at higher latitudes (due to greater rotation, and weaker stratification), and the increased relative importance of transient, rapidly propagating barotropic signals driven by the strong winds (e.g., Vivier et al., 2005). These barotropic signals are under-sampled by the altimetry in both time and space, and probably not well represented individually in the model because of the limited accuracy of their representation in the 6-hourly archive of analysed wind fields. Recall that for BRAN2, the model SLA was perturbed to correct the misfit to altimetric SLA at each assimilation step, while for BRAN3 we perturb the velocity fields instead. BODAS uses EnOI so an altimeter track that samples a large-scale barotropic signal will be mapped onto the model state via small perturbations of the velocity field, rather than SLA perturbations of similar size to the observation. It appears, from Figures 2-5, that the BRAN3 approach is a better way of using altimetry to constrain the baroclinic features of the ocean.

5.2. Comparison with in situ profiles

We compare modelled and observed profiles of temperature and salinity in Figures 6 and 7, respectively. Specifically, we show the RMSD and the mean bias (computed as the observed minus modelled mean) for different regions around Australia (Table 3 and Figure 3). The statistics presented are based on comparisons with all available profiles for the entire BRAN2 period (1993-2006). Like altimetry, for each analysis, in situ profiles within a centred time-window of 11 (7) days for BRAN3 (BRAN2), are considered for assimilation. However, unlike altimetry, profiles are not combined to form super-observations - and not all profiles are assimilated. Instead, profiles are "thinned", retaining no more than one profile of each variable for every $0.5 \times 0.5°$ box. When more than one profile is present within the given time-window, a single profile is selected by identifying the profile that was measured closest to the analysis time. As for the comparisons with altimetry, much of the data used for this evaluation was assimilated. Despite this, we regard this comparison as a necessary step in the evaluation of each reanalysis.

The comparisons in Figure 6 show that the temperature errors in BRAN3 are almost everywhere less than the errors in BRAN2. BRAN2 has smaller errors than BRAN3 only at about 1250 m depth in the EAC and GAB regions, and around 1000 m in the South-West (SW) region. In each of these isolated regions, the increased RMSD is due to a warm bias in BRAN3. Above 300 m depth, the errors for temperature in BRAN3 are often much less than for BRAN2. Indeed, for the upper 300 m, for the Australian region the area-averaged RMSD for BRAN3 (0.68°C) is 28% less than the BRAN2 (0.95°C) temperature error. The temperature bias in BRAN3 is almost everywhere less than the temperature bias in BRAN2. This difference is most evident in the North-West (NW) region, where the strong negative temperature bias between about 100 and 200 m depth in BRAN2 is virtually eliminated in BRAN3.

The comparisons in Figure 7 indicate that the salinity errors in BRAN2 and BRAN3 are comparable in most regions. In some regions (e.g., Coral Sea, NW region) BRAN3 salinity is significantly better than BRAN2 in the upper ocean, with improvements of 0.05-0.15 psu. But in other regions (e.g., SW and GAB region) BRAN3 salinity is up to 0.05 psu worse than BRAN2. For the upper 300 m, the area-averaged RMSD for BRAN3 (0.155 psu) is 7% less than the BRAN2 (0.167 psu) salinity error. This indicates that overall, in the upper ocean, BRAN3 salinity is about 7% better than BRAN2 salinity. Notably, in several regions (e.g., EAC, SW and GAB regions), BRAN3 salinity has a greater RMSD between about 500 m and 1500 m, owing to a significant negative bias of about 0.1 psu.

The statistics of bias in Figures 6 and 7 indicate that BRAN3 is saltier and warmer than observations between 500 and 1000 m depth. This indicates that the properties of the intermediate water masses are imprecise in BRAN3, and perhaps indicates that BRAN3 either produces too little Intermediate Water, or that the properties of BRAN3's intermediate water are unrealistic. This is an aspect of BRAN that will be further considered in future development. However, aside from the deep ocean comparisons, we find that in the upper ocean temperature and salinity in BRAN3 is more realistic than BRAN2, with reductions of the RMSD with observed profiles of 28% and 7% for temperature and salinity, respectively.

5.3. Comparison with XBT data

Data from several eXpendable BathyThermograph (XBT) transects across the Tasman Sea - including PX34, running between Sydney, Australia, and Wellington, New Zealand - were withheld from both BRAN2 and BRAN3 for some time periods (September 2003 - December 2006). Data from these XBT transects were assimilated for other time periods (February 1993 - July 2003). The PX34 section is occupied 3-4 times each year with high-density sampling. The with-held XBT data along this transect are ideal for independent evaluation - both because they were not used in either reanalysis, and because they traverse a very energetic region of the ocean, with strong seasonality (e.g., Ridgway, 2007) and strong eddies (e.g., Everett et al., 2012). Here, we compare BRAN2 and BRAN3 to data along PX34 for periods when the data are assimilated - to show how tightly each BRAN is constrained to those data; and for periods when the data are with-held - to show how each BRAN matches independent data.

Figure 8 shows sections of objectively analysed temperature, based on observations that are assimilated into both reanalyses, and temperature from BRAN2 and BRAN3. Each XBT section takes 3-4 days to traverse - however, we simply sample the model on the central day of each section. Overall, both BRAN2 and BRAN3 realistically reproduce the observed features along PX34, but the BRAN3 fields are clearly in better agreement with the assimilated observations. This indicates that the assimilation used for BRAN3 is better at fitting the assimilated observations. Figure 9 shows sections of withheld XBT observations along PX34 and temperature from BRAN2 and BRAN3. Overall, both BRAN2 and BRAN3 realistically reproduce the independently observed features along PX34 in Figure 9. In most cases both reanalyses reproduce almost all of the observed features that are associated with mesoscale variability.

To enable a more quantitive comparison, we compute the depth of the 15°C isotherm (D15) along PX34 for the entire BRAN2 period and compare time series of the D15 anomaly in Figure 10. The temporal sampling of PX34 is insufficient to resolve all of the mesoscale variability there - but the evolution of several large-amplitude, long-lived events is evident. All of the large-amplitude events evident in the observations are also evident in BRAN3; however, they are not all evident in BRAN2. Examples include the large negative anomaly at around 154°E between the start of 1995 and the end of 1996; the positive anomaly at the same longitude during 1997; and a positive, westward-propagating anomaly originating around 158°E between the start of 1999 and the end of 2000. There is excellent correspondence between these events in the observations and BRAN3, which is much better than the BRAN2 estimates. Other short-lived, but large-amplitude anomalies are also clearly evident in both the observations and BRAN3 - but less evident in BRAN2. During the period when the PX34 data are with-held, BRAN3 again appears to be in better quantitive agreement with observations than BRAN2.

The evaluation of D15 along PX34 is quantified in Figure 11, showing the RMSD, bias, and correlation between the observed D15 and the D15 in BRAN2 and BRAN3. The standard deviation of the observed D15 is also shown in Figure 11. A comparison between the RMSDs and the observed standard deviation shows that in the western part of PX34, the signal to noise ratio for D15 is quite good, with errors that are typically 30% less than the observed signal. At some points along PX34, the RMSD for BRAN2 exceeds the size of the observed signal. This is not the case for BRAN3. In the region west of 158°E, the bias in BRAN3 is much smaller than in BRAN2, and the RMSD is reduced by as much as 30%. In the same region, the correlations for BRAN3 exceed the correlations for BRAN2 by 0.1-0.2. In the middle part of PX34, between 158-166°E, there is little difference between the D15 fields in BRAN2 and BRAN3. But along the eastern part, east of 166°E, the BRAN3 correlations again exceed the BRAN2 correlations at several points, and the RMSDs and the bias are less.

The bias of D15 along PX34 in BRAN2 is quite significant (in excess of ± 40 m), particularly west of 158°E. We think that this is an indication that the location of the mean EAC jet and/or its horizontal and vertical shears are wrong in BRAN2. On average, D15 slopes upwards to the east along PX34. The positive-negative shape of the bias in BRAN2, centred around 154°E, means that the upward slope of D15 to the east in BRAN2 is less than observed. As a result, the geostrophic flow associated with the EAC there is less vertically sheared (i.e., more barotropic) in BRAN2. This result is consistent with the analyses of Chiswell and Rickard (2008), who assessed BRAN2 velocities against observed velocities inferred from surface drifting buoys and Argo floats, and concluded that BRAN2 velocities in the deep ocean are too strong (i.e., BRAN2 appeared to be too barotropic in the Tasman Sea). This characteristic appears to be significantly improved in BRAN3, based on the smaller bias evident along PX34.

5.4. Comparison with independent surface drifting buoys

Data from surface drifting buoys are not assimilated into BRAN2 or BRAN3, and are therefore ideal for independent evaluation (e.g., Oke et al., 2012; Blockley et al., 2012). Using daily-averaged velocities derived from krigged drifter positions (obtained from NOAA AOML; www.aoml.noaa.gov/phod/dac/dacdata.php), we compare the model velocities at 12 m depth (the approximate depth of the drifter sea-anchors), with the drifter-derived velocities.

We first present a qualitative comparison of modelled velocities and drifter-derived velocities for a short period in the EAC region (Figure 12). We show only comparisons from a short period - but we note that the results presented here are representative of other periods. Also shown in Figure 12 are observed and reanalysed SST anomalies, and geostrophic velocities derived from a Gridded SLA product (GSLA; see http://oceancurrent.imos.org.au). The observed SSTA fields shown in Figure 12 are 6-day composite AVHRR SST fields, processed at CSIRO under the Australian Integrated Marine Observing System (IMOS; http://oceancurrent.imos.org.au).

The drifter trajectories in Figure 12 are for an 8-day period preceding the day for which the model fields are shown, so precise agreement between the observed and reanalysed trajectories is not expected due to the change of the flow field with time. However, the comparisons show that there is generally good agreement between the model velocities and the drifter-derived velocities. Close inspection suggests that some of the mesoscale fields are slightly mis-placed in BRAN - though it is unclear whether the mis-placement is real, or due to the aliasing referred to above. In some cases shown in Figure 12, where there is good agreement between the BRAN3 fields and observations, there is poor agreement between the BRAN2 fields and observations. The fields on 18 and 26 January 2012 are good examples of this - with good correspondence between the drifter trajectories and the BRAN3 and GSLA-derived velocities, but poor correspondence for BRAN2 fields.

A quantitative comparison of the drifter-derived velocities and the model velocities is presented in Figure 13, showing the RMSD and correlation between the observed and modelled velocities for the whole Australian region, and the other regions defined in Table 3 and Figure 3. Comparisons are for the period 2003-2006 and only include observations when the observed speed exceeds 3 cm/s (to exclude cases where the drifter may have lost its drogue; though we note that this will not exclude all un-drogued drifters; Rio, 2012). The number of drifter observations in the Australian region totals 35000, and the number of observations within each domain is 2000-4000.

We find little difference between the statistics for the zonal and meridional component of velocity - so we present these together in Figure 13. The correlations shown in Figure 13 are the amplitudes of the complex, or vector, correlation (Kundu, 1976). The phase angles of the complex correlation (not shown) are small for all regions ($\pm 10°$) except the GAB, where it is -60°. Also shown in Figure 13 is the standard deviation of the observed speed. The RMSDs in Figure 13 indicate that both BRAN2 and BRAN3 have errors that are less than the observed standard deviation, so the signal to noise ratio in BRAN2 and BRAN3 exceeds one. Further, we find that the BRAN3 velocities have smaller RMSDs than BRAN2, with errors that are typically 1-2 cm/s smaller. This represents an overall, albeit small, improvement in velocity of about 5%. The amplitudes of the complex correlations shown in Figure 13 are only moderate, with value of around 0.3 to 0.5. Despite these relatively low correlations, we note that the large number of observations implies that these correlations are statistically significant (even with only 100 degrees of freedom a correlation of 0.2 is statistically significant). The correlations for BRAN3 are typically about 0.1 greater than BRAN2 - suggesting a significant improvement in BRAN3. The only place where BRAN3 is poorer than BRAN2 is in the GAB region and the region around New Zealand. In those regions, the amplitude of the observed velocities are smallest, so the signal to noise ratio in the observations is relatively low.

6. Analysis

A comparison between the RMS of the increments for temperature and salinity at 100 m depth and sea-level are shown in Figure 14, 15 and 16 respectively. Recall that during each assimilation step the model is initialised to match the analysis field. For BRAN2, temperature, salinity and sea-level are simply nudged to the analysis fields for one day, using a nudging time-scale of one day. For BRAN3, temperature, salinity and velocities are adjusted to match the analyses using adaptive initialisation (Sandery et al., 2011). In BRAN3, sea-level is not adjusted explicitly, but during the initialisation of the other model variables, we find that sea-level adjusts to closely match the analyses computed by BODAS. Recall that the increments that are added to each assimilation step do not necessarily have any physical meaning. They are simply compensating for model limitations, including the inability of a model to reproduce instabilities associated with chaotic and unpredictable dynamics. Ideally, the increments should be as small as possible. Figures 14-16 show that the size of the increments for BRAN3 is significantly less than the size of the increments in BRAN2. Indeed, the area-average ratio of the BRAN3 to BRAN2 increments for temperature is 0.65, for salinity is 0.6, and for sea-level is 0.78. The minimum ratio in the region shown in Figures 14-16 is 0.06, 0.05, and 0.01 for temperature, salinity and sea-level respectively. This indicates that, on average, the BRAN3 increments are 22-40% less than the BRAN2 increments; and as much as 94, 95, and 99% less at some points for temperature, salinity, and sea-level respectively.

Analysis of the magnitude of the increments for all variables at other depths (i.e., above and below 100 m), indicates that for much of the water column, the increments in BRAN3 are typically 30-50% less than the increments in BRAN2. These results indicate that the assimilation system for BRAN3 is doing substantially “less work” than BRAN2.

There are a few reasons why the increments in BRAN3 are so much smaller than BRAN2. BRAN3 updates more frequently (4-days instead of

7-days); the model that underpins BRAN3 is better - with improved parameterisations (e.g., Lee et al., 2006), improved topography, and improved surface fluxes; the pre-processing of the observations is better, and the initialisation scheme in BRAN3 is better. We think that together, these factors allow the model to more realistically evolve the initialised model fields, requiring less adjustment at each assimilation step, resulting in a more dynamically consistent reanalysis that requires less adjustments to stay aligned with observations.

Figures 14-16 include RMS fields for different time periods: 1994-1996 and 2004-2006. The magnitude and spatial distribution of the increments changes with time, depending on the observing system. We chose the above-mentioned periods to highlight the impact of changes in the observing system, and to highlight the multivariate nature of the assimilation. In Figure 14a,b, the increments in 1994-1996 show a clear signature of the XBT transects (IX1, IX12, IX15) in both BRAN2 and BRAN3. Similarly, the increments associated with several XBT tracks in the Pacific (PX05, PX06, PX31, PX30) are also evident - particularly in BRAN3. Interestingly, the influence of many of these temperature observations is also clearly evident in the salinity increments (Figure 15a,b), particularly IX1. However, the influence of these data are not clearly evident during 2004-2006 (Figure 14a,b and 15a,b). This is because during 2004-2006 the number of observations associated with the Argo program increased - so the sampling is much better allowing more observations to do “less work” - which is preferable than fewer observations doing“more work”.

We also note that the impact of the longer length-scales used for the localisation in BRAN2 tend to "smear" the influence of the XBT data in space during 1994-1996 (Figure 14a,b and 15a,b), with broader influence evident in the BRAN2 fields and narrower influence in BRAN3. The narrow influence of the observations from the TAO array is also evident, with small "bullets" in the increment fields.

Apart from the recognition that the magnitude of the increments in BRAN3 is generally much smaller than the magnitude of the increments in BRAN2, we also note the systematic differences in the structure of the increments. This is particularly clear for salinity and sea-level during 2004-2006 in the Pacific between about 20-5°S, where a band of high increments is evident (Figure 15d and 16d). This feature is not present in the BRAN3 increments. The mean increments in BRAN2 and BRAN3 are relatively small in most regions around Australia (not shown) - but are large for BRAN2 in this tropical part of the South Pacific. This indicates that BRAN2 had a bias in this region, requiring constant adjustment in the "same direction". We attribute this problem to two factors in BRAN2. Firstly, the systematic differences between the reference MSL used for BRAN2 and BRAN3 are a factor (see Appendix A.4). The reference MSL plays a key role in determining the mean circulation in each reanalysis - and the MSL used for BRAN3 is superior to the field used for BRAN2 owing to improvements in the model and model forcing. Secondly, the pre-processing of the altimeter observations is a key factor. For BRAN2, we assimilated the atSLA data that included the signal of sea-level rise. We recognise that this is inconsistent, because the (Boussinesq) model does not include the effects of thermal expansion - a key contributor to sea-level rise (Church and White, 2006). The solution to this incompatibility is to use a non-Bousinesq model - but, short of that, we have improved the compatibility by eliminating the global means from the model and the atSLA data prior to assimilation. Although this approach is not perfect - it is an improvement on previous methods used for ocean reanalyses. We attribute the improvements in BRAN3 in this region of the South Pacific to this more careful pre-processing of the altimeter observations.

7. Conclusions

One of the main goals of the Bluelink effort that began in 2001 is the generation of an eddy-resolving ocean reanalysis for the circulation around Australia, that can be used to understand upper-ocean dynamics, teleconnections, and variability. To achieve this goal, a dynamically consistent reanalysis system is the "holy grail". However, the generation and evolution of eddies in the ocean are largely due to instabilities that are unpredictable, even on short time-scales. This means that an eddy-resolving model requires frequent adjustments to keep it aligned with observations. We use an EnOI system to constrain the model to observations by updating the model state regularly. With such a data assimilation approach adopted, the goal then becomes the generation of a reanalysis that matches both assimilated and with-held observations, and involves increments that are as small as possible. Given the chaotic nature of the eddy-scales in the ocean, as discussed above, there is a lower limit to which the size of the increments can be reduced, and yet still keep the model aligned with observations. This lower limit will depend on the observation errors, the length of the assimilation update cycle, and the growth-rate of instabilities. With BRAN3, we are approaching this limit in many parts of the domain of interest here. For example, we show that the reanalysed SLA in BRAN3, is within the error-bars of the observations at low latitudes. Further, we show that the model temperatures and salinity in the upper ocean have errors that are 7-28% less than the previous version of BRAN. Specifically, we show that for BRAN3 (BRAN2) the sea-level, upper ocean temperature, upper-ocean salinity, and near-surface velocity match observations to within 7.7±0.5 cm (9.7±0.8 cm), 0.68±0.08°C (0.95±0.18°C), 0.16±0.02 psu (0.18±0.02 psu), and 20.2 cm/s (21.3 cm/s) respectively.

Somewhat counter-intuitively, we also show while BRAN3 produces re-analyses that more closely match observations, the increments applied to BRAN3 are 20-50% smaller than the equivalent adjustments in BRAN2. This means that the data assimilation system in BRAN3 is doing less work than in BRAN2 - but achieving better results. We attribute these improvements to a few major changes: including the initialisation scheme, the employment of the local analysis approach to localisation, improvements to data processing and improvements to the model configuration; higher frequency of assimilation; and several minor changes, relating to the error estimates and the technical configuration of the assimilation system.

The analyses presented in this study have identified one outstanding issue in BRAN3, namely the quality of the temperature and salinity fields at intermediate depths. We find that the BRAN3 fields are warmer and saltier than observations - and in some places are poorer than the predecessor, BRAN2. This aspect of the BRAN effort will be the focus of future developments.

To summarise, we have shown that BRAN3 produces observations that more closely match observations than BRAN2, and yet requires less adjustment via assimilation. These factors indicate that BRAN3 is more dynamically consistent than BRAN2 - with more realistic reanalyses and less non-dynamical interference. This leads us to conclude that BRAN3 is more suitable for a range of applications, including analysis of ocean variability, extreme events, and process studies.

Acknowledgments. Financial support for this research is provided by CSIRO, the Bureau of Meteorology and the Royal Australian Navy as part of the Bluelink project, and the US Office of Naval Research. The authors also acknowledge the contributions of the Bluelink Science Team. Satellite altimetry is provided by NASA, NOAA and CNES. Drifter data are provided by NOAA-AOML and SST observations are provided by NOAA and Remote Sensing Systems. Coastal sea-level around Australia is provided by the National Tidal Centre and state authorities.

References

Andreu-Burillo, I., Brassington, G. B., Oke, P. R., Beggs, H., 2010. Including a new data stream in the Bluelink ocean data assimilation system. Aust Meteorol Ocean 59, 77–86.

Balmaseda, M. A., Anderson, D., 2009. Impact of initialization strategies and observations on seasonal forecast skill. Geophysical Research Letters 36 (1), L01701.

Balmaseda, M. A., Morgensen, K., Weaver, A. T., 2012. Evaluation of the ecmwf ocean reanalysis system oras4. Quarterly Journal of the Royal Meteorological Society.

Blockley, E. W., Martin, M. J., Hyder, P., 2012. Validation of FOAM near-surface ocean current forecasts using Lagrangian drifting buoys. Ocean Science 8, 551–565.

Boyer, T. P., Antonov, J. I., Garcia, H. E., Johnson, D. R., Locarnini, R. A., Mishonov, A. V., Pitcher, M. T., Baranova, O., Smolyar, I., Levitus, S., 2006. World Ocean Database 2005. NOAA Atlas NESDIS 60, 190pp.

Brassington, G. B., Pugh, T. F., Spillman, C., Schulz, E., Beggs, H., Schiller, A., Oke, P. R., 2007. Bluelink development of operational oceanography and servicing in Australia. J Research Pract Inf Tech 39 (2), 151–164.

Carton, J., Chepurin, G., Cao, X., 2000. A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part II: Results. Journal of Physical Oceanography 30 (2), 311–326.

Carton, J. A., Giese, B. S., 2008. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). Monthly Weather Review 136, 2999–3017.

Chen, D., Rothstein, L., Busalacchi, A., 1994. A hybrid vertical mixing scheme and its application to tropical ocean models. Journal of Physical Oceanography 24 (10), 2156–2179.

Chiswell, S. M., Rickard, G. J., 2008. Eulerian and Lagrangian statistics in the Bluelink numerical model and AVISO altimetry: Validation of model eddy kinetics. Journal of Geophysical Research 113 (C10), C10024.

Church, J. A., White, N. J., 2006. A 20th century acceleration in global sea-level rise. Geophysical Research Letters 33, L01602.

Counillon, F., Bertino, L., 2009. Ensemble optimal interpolation: multivariate properties in the Gulf of Mexico. Tellus A 61 (2), 296–308.

Counillon, F., Sakov, P., Bertino, L., 2009. Application of a hybrid EnKF-OI to ocean forecasting. Ocean Science 5 (4), 389–401.

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

Dee, D. P., Uppala, S., 2009. Variational bias correction of satellite radiance data in the ERA-interim reanalysis. Quarterly Journal of the Royal Meteorological Society 135 (644), 1830–1841.

Di Lorenzo, E., Moore, A. M., Arango, H. G., Cornuelle, B. D., Miller, A. J., Powell, B. S., Chua, B. S., Bennett, A., 2007. Weak and strong constraint data assimilation in the inverse Regional Ocean Modeling System (ROMS): Development and application for a baroclinic coastal upwelling system. Ocean Modelling 16 (3-4), 160–187.

Dietze, H., Matear, R., Moore, T., 2009. Nutrient supply to anticyclonic meso-scale eddies off Western Australia estimated with artificial tracers released in a circulation model. Deep Sea Research 56, 1440 – 1448.

Dombrowsky, E., Bertino, L., Brassington, G. B., Chassignet, E. P., Davidson, F. J. M., Hurlburt, H. E., Kamachi, M., Lee, T., Martin, M. J., Mei, S., Tonani, M., 2009. GODAE systems in operation. Oceanography 22 (3), 80–95.

Evensen, G., 1997. Advanced data assimilation for strongly nonlinear dynamics. Monthly Weather Review 125 (6), 1342–1354.

Evensen, G., 2003. The ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dynamics 53 (4), 343–367.

Everett, J. D., Baird, M. E., Oke, P. R., Suthers, I. M., 2012. An avenue of eddies: Quantifying the biophysical properties of mesoscale eddies in the tasman sea. Geophysical Research Letters 39 (L16608).

Feng, M., Wijffels, S. E., Godfrey, J. S., Meyers, G. A., 2005. Do eddies play a role in the momentum balance of the Leeuwin Current? Journal of Physical Oceanography 35 (6), 964–975.

Ferry, N., Remy, E., Brasseur, P., Maes, C., 2007. The Mercator global ocean operational analysis system: Assessment and validation of an 11-year reanalysis. Journal of Marine Systems 65 (1-4), 540–560.

Fu, W., Zhu, J., Yan, C., 2009. A comparison between 3DVAR and EnOI techniques for satellite altimetry data assimilation. Ocean Modelling 26 (3-4), 206–216.

Griffies, S. M., Hallberg, R. W., 2000. Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. Monthly Weather Review 128 (8), 2935–2946.

Griffies, S. M., Pacanowski, R. C., Rosati, A., 2004. A technical guide to MOM4. GFDL ocean group technical report no. 5. NOAA/Geophysical Fluid Dynamics Laboratory 5.

Gronell, A., Wijffels, S. E., 2008. Semiautomated approach for quality controlling large historical ocean temperature archives. Journal of Atmospheric and Oceanic Technology 25, 990–1003.

Guinehut, S., Traon, P.-Y. L., Larnicol, G., 2006. What can we learn from global altimetry/hydrography comparisons? Geophysical Research Letters 33 (10), L10604.

Guinehut, S., Traon, P.-Y. L., Larnicol, G., Philipps, S., 2004. Combining Argo and remote-sensing data to estimate the ocean three-dimensional temperature fields - a first approach based on simulated observations. Journal of Marine Systems 46 (1-4), 85–98.

Hunt, B. R., Kostelich, E. J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman Filter. Physica D 230 (1-2), 112–126.

Kallberg, P., Simmons, A., Uppala, S., Fuentes, M., 2004. The ERA-40 archive, European Centre for Medium-range Weather Forecasts (ECMWF), ECMWF Re-Analysis Project (ERA). ERA-40 Project Report Series 17, 131pp.

Koehl, A., Stammer, D., Cornuelle, B. D., 2007. Interannual to decadal changes in the ECCO global synthesis. Journal of Physical Oceanography 37 (2), 313–337.

Kundu, P., 1976. Ekman veering observed near ocean bottom. Journal of Physical Oceanography 6 (2), 238–242.

Kurapov, A., Egbert, G., Allen, J., Miller, R., 2009. Representer-based analyses in the coastal upwelling system. Dynamics of Atmospheres and Oceans 48 (13), 198 – 218, ¡ce:title¿Modeling and Data Assimilation in Support of Coastal Ocean Observing Systems¡/ce:title¿.

Kurapov, A. L., Foley, D., Strub, P. T., Egbert, G. D., Allen, J. S., 2011. Variational assimilation of satellite observations in a coastal ocean model off Oregon. Journal of Geophysical Research 116, C05006.

Langlais, C., Schiller, A., Oke, P. R., 2010. Southern ocean fronts in the Bluelink reanalysis. Mercator Quarterly Newsletter 36, 50–57.

Le Traon, P.-Y., Nadal, F., Ducet, N., 1998. An improved mapping method of multisatellite altimeter data. Journal of Atmospheric and Oceanic Technology 15 (2), 522–534.

Lee, H.-C., Rosati, A., Spelman, M., 2006. Barotropic tidal mixing effects in a coupled climate model: ocean conditions in the northern Atlantic. Ocean Modelling 11, 464–470.

Lee, T., Yoder, J. A., Atkinson, L. P., 1991. Gulf-stream frontal eddy influence on productivity of the southeast United-States continental-shelf. Journal of Geophysical Research 96 (C12), 22191–22205.

Marchesiello, P., McWilliams, J. C., Shchepetkin, A., 2003. Equilibrium structure and dynamics of the California Current System. Journal of Physical Oceanography 33 (4), 753–783.

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., Robinson, P., 2011. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part II performance and application to the California current system. Progress in Oceanography 91, 50–73.

Moore, A. M., Arango, H. G., Lorenzo, E. D., Miller, A. J., Cornuelle, B. D., 2009. An adjoint sensitivity analysis of the southern California current circulation and ecosystem. Journal of Physical Oceanography 39 (3), 702–720.

Nerger, L., Janjic, T., Schroter, J., Hiller, W., 2011. A regulated locatization scheme for ensemble-based Kalman filters. Quarterly Journal of the Royal Meteorological Society 138, 802–812.

O'Kane, T. J., Oke, P. R., Sandery, P. A., 2011. Predicting the East Australian Current. Ocean Modelling 38, 251–266.

Oke, P. R., Allen, J. S., Miller, R. N., Egbert, G. D., Kosro, P. M., 2002. Assimilation of surface velocity data into a primitive equation coastal ocean model. Journal of Geophysical Research 107 (C9), 3122.

Oke, P. R., Brassington, G. B., Cummings, J., Martin, M., Hernandez, F., 2012. GODAE inter-comparisons in the Tasman and Coral Seas. Journal of Operational Oceanography 5 (2), 11–24.

Oke, P. R., Brassington, G. B., Griffin, D. A., Schiller, A., 2008. The Bluelink Ocean Data Assimilation System (BODAS). Ocean Modelling 21 (1-2), 46–70.

Oke, P. R., Brassington, G. B., Griffin, D. A., Schiller, A., 2010. Ocean data assimilation: a case for ensemble optimal interpolation. Aust Meteorol Ocean 59, 67–76.

Oke, P. R., Griffin, D. A., 2011. The cold-core eddy and strong upwelling off the coast of New South Wales in early 2007. Deep Sea Research 58 (5), 574–591.

Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J. V., Lenton, A., Cahill, M., Chamberlain, M. A., Ridgway, K., 2013. Evaluation of a near-global eddy-resolving ocean model. Geoscientific model development gmd-2012-113.

Oke, P. R., Sakov, P., 2008. Representation error of oceanic observations for data assimilation. Journal of Atmospheric and Oceanic Technology 25 (6), 1004–1017.

Oke, P. R., Sakov, P., Corney, S. P., 2007. Impacts of localisation in the EnKF and enoi: experiments with a small model. Ocean Dynamics 57 (1), 32–45.

Oke, P. R., Sakov, P., Schulz, E., 2009. A comparison of shelf observation platforms for assimilation in an eddy-resolving ocean model. Dynamics of Atmospheres and Oceans 48 (1-3), 121 – 142.

Oke, P. R., Schiller, A., 2007. Impact of Argo, SST, and altimeter data on an eddy-resolving ocean reanalysis. Geophysical Research Letters 34 (19), L19601.

Oke, P. R., Schiller, A., Griffin, D. A., Brassington, G. B., 2005. Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. Quarterly Journal of the Royal Meteorological Society 131 (613), 3301–3311.

Ponte, R. M., Wunsch, C., Stammer, D., 2007. Spatial mapping of time-variable errors in jason-1 and TOPEX/poseidon sea surface height measurements. Journal of Atmospheric and Oceanic Technology 24 (6), 1078–1085.

Rao, S., Behera, S., 2005. Subsurface influence on sst in the tropical Indian Ocean: structure and interannual variability. Dynamics of Atmospheres and Oceans 39 (1-2), 103–135.

Ridgway, K. R., 2007. Long-term trend and decadal variability of the southward penetration of the East Australian Current. Geophysical Research Letters.

Ridgway, K. R., Dunn, J. R., 2010. Using satellite altimetry to correct mean temperature and salinity fields derived from Argo floats in the ocean regions around Australia. Deep Sea Research 57, 1137–1151.

Rio, M. H., 2012. Use of altimeter and wind data to detect the anomalous loss of svp-type drifter's drogue. Journal of Atmospheric and Oceanic Technology 29, 1663–1674.

Sakov, P., Bertino, L., 2011. Relation between two common localisation methods for the EnKF. Computational Geoscience 15, 225–237.

Sandery, P. A., Brassington, G. B., Freeman, J., 2011. Adaptive nonlinear dynamical initialization. Journal of Geophysical Research 116 (C01021).

Schiller, A., Meyers, G., Smoth, N., 2009a. Taming Australia's last frontier. Bulletin of Americal Meteorological Society 90, 437–440.

Schiller, A., Oke, P. R., Brassington, G. B., Entel, M., Fiedler, R., Griffin, D. A., Mansbridge, J. V., 2008. Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort. Progress in Oceanography 76 (3), 334–365.

Schiller, A., Ridgway, K. R., Steinberg, C. R., Oke, P. R., 2009b. Dynamics of three anomalous SST events in the Coral Sea. Geophysical Research Letters 36, L06606.

Schouten, M., de Ruijter, W., van Leeuwen, P. J., Dijkstra, H., 2002. An oceanic teleconnection between the equatorial and southern Indian Ocean. Geophysical Research Letters 29 (16), 1812.

Smith, N. R., 2000. The global ocean data assimilation experiment. Advanced in Space Research 25 (5), 1089–1098.

Smith, W., Sandwell, D., 1997. Global sea floor topography from satellite altimetry and ship depth soundings. Science 277 (5334), 1956–1962.

Srinivasan, A., Chassignet, E. P., Bertino, L., Brankart, J. M., Brasseur, P., Chin, T. M., Counillon, F., Cummings, J. A., Mariano, A. J., Smedstad, O. M., Thacker, W. C., 2011. A comparison of sequential assimilation schemes for ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM): Twin experiments with static forecast error covariances. Ocean Modelling 37, 85–111.

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., Adcroft, A., Hill, C. N., Marshall, J., 2002. Global ocean circulation during 1992-1997, estimated from ocean observations and a general circulation model. Journal of Geophysical Research 107 (C9), 3118.

Sun, C., Feng, M., Matear, R., Chamberlain, M., Craig, P., Ridgway, K., Schiller, A., 2012. Marine downscaling of a future climate scenario for Australian boundary currents. Journal of Climate 25, 2947 – 2962.

Taylor, J. R., 1997. An introduction to error analysis. University of Science Books, Sausalito, California.

Trenberth, K. E., Fasullo, J. T., Kiehl, J., 2009. Earth's global energy budget. Bulletin American Meteorlogical Society March, 311–323.

Vivier, F., Kelly, K. A., Harismendy, M., 2005. Causes of large-scale sea level variations in the southern ocean: Analyses of sea level and a barotropic model. Journal of Geophysical Research 110.C9.

Wan, L., Bertino, L., Zhu, J., 2010. Assimilating altimetry data into a HYCOM model of the Pacific: Ensemble optimal interpolation versus ensemble Kalman filter. Journal of Atmospheric and Oceanic Technology 27 (4), 753–765.

Xie, J., Zhu, J., 2010. Ensemble optimal interpolation schemes for assimilating Argo profiles into a hybrid coordinate ocean model. Ocean Modelling 33 (3-4), 283–298.

Yu, P., Kurapov, A. L., Egbert, G. D., Allen, J. S., Kosro, P. M., 2012. Variational assimilation of HF radar surface currents in a coastal ocean model off Oregon. Ocean Modelling 49-50, 86–104.

Zhang, W. G., Wilkin, J. L., Arango, H. G., 2010. Towards an integrated observation and modeling system in the New York Bight using variational methods. Part I: 4DVAR data assimilation. Ocean Modelling 35 (3), 119–133.

Table 1: Summary of the key differences between models used for BRAN2 and BRAN3. The term "globally balanced" refers to the freshwater fluxes that have been adjusted so that the annual averaged, global average freshwater fluxes are zero; and the global average of applied net heat flux is adjusted to the observed global average.

	BRAN2	BRAN3
Model	OFAM1	OFAM2
MOM version	MOM40d	MOM4p1
Period	1/1993-12/2006	1/1993-9/2012
Vertical resolution	10-m surface	5-m surface
Vertical mixing	Chen	Chen + Lee
Topography	DBDB2+GEBCO	Smith&Sandwell
Forcing	ERA-40+ECMWF	ERA-Interim
	6-hourly	3-hourly
	Unaltered	Globally balanced

Table 2: Summary of the key differences between data assimilation system used for BRAN2 and BRAN3.

	BRAN2	BRAN3
BODAS version	BODAS5p0	BODAS8p2
Update cycle	7-days	4-days
Ensemble size	72	144
Ensemble run	Spinup4/5	Spinup6p8
Localisation method	Covariance	Local analysis
Localising length-scale	8°	250 km
Initialisation	Weak nudging	Adaptive initialisation
Updated variables	T, S, sea-level	T, S, U, V
Age error	RMS of ensemble	RMS of time-difference
Altimetry data	All	GFO with-held
Altimetry window	11 days	21 days
Altimeter processing	Unaltered	Volume conserving
Altimeter mask	200 m depth	200 m depth
Reference MSL	Spinup4/5	OFAM3
SST window	1 day	5 days
AVHRR SST data	54-km Pathfinder	4-km Pathfinder
SST mask	100 km from coast	20 m depth
T/S window	7 days	11 days

Table 3: Summary of the different regions around Australia for which statistics are computed throughout this study. See also Figure 3.

Region	Longitudes	Latitudes
Australian region (Aust)	90-180°E	60°S - equator
East Australian Current (EAC)	147-165°E	50-25°S
Coral Sea	143-165°E	25-5°S
North-Western Australia (NW)	100-143°E	20-5°S
South-Western Australia (SW)	100-116°E	35-20°S
Great Australian Bight (GAB)	116-147°E	50-30°S
Antarctic Circumpolar Current (ACC)	90-180°E	70-50°S
New Zealand (NZ)	165-180°E	50-25°S

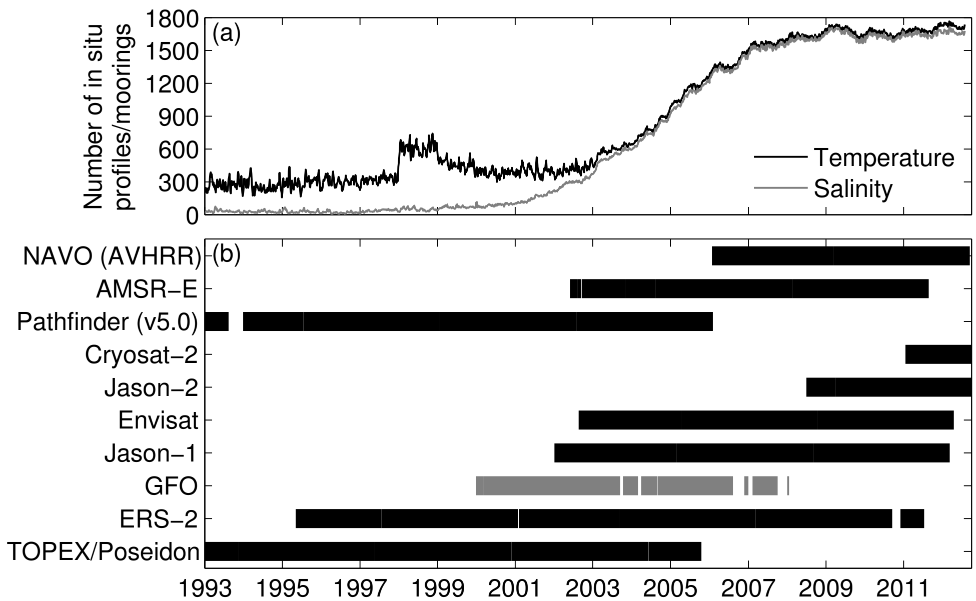

Figure 1: Time series of (a) the number of in situ temperature and salinity profiles or moorings (including the TAO array) assimilated at each analysis step in BRAN3 (note that many profiles are used in multiple consecutive assimilation cycles); and a schematic showing data availability from each SST database and altimeter mission. The grey bar for GFO indicates that these data were with-held from the assimilation. Altimeter data were accessed from RADS in August 2012 (and updated for 2012-data in October 2012).

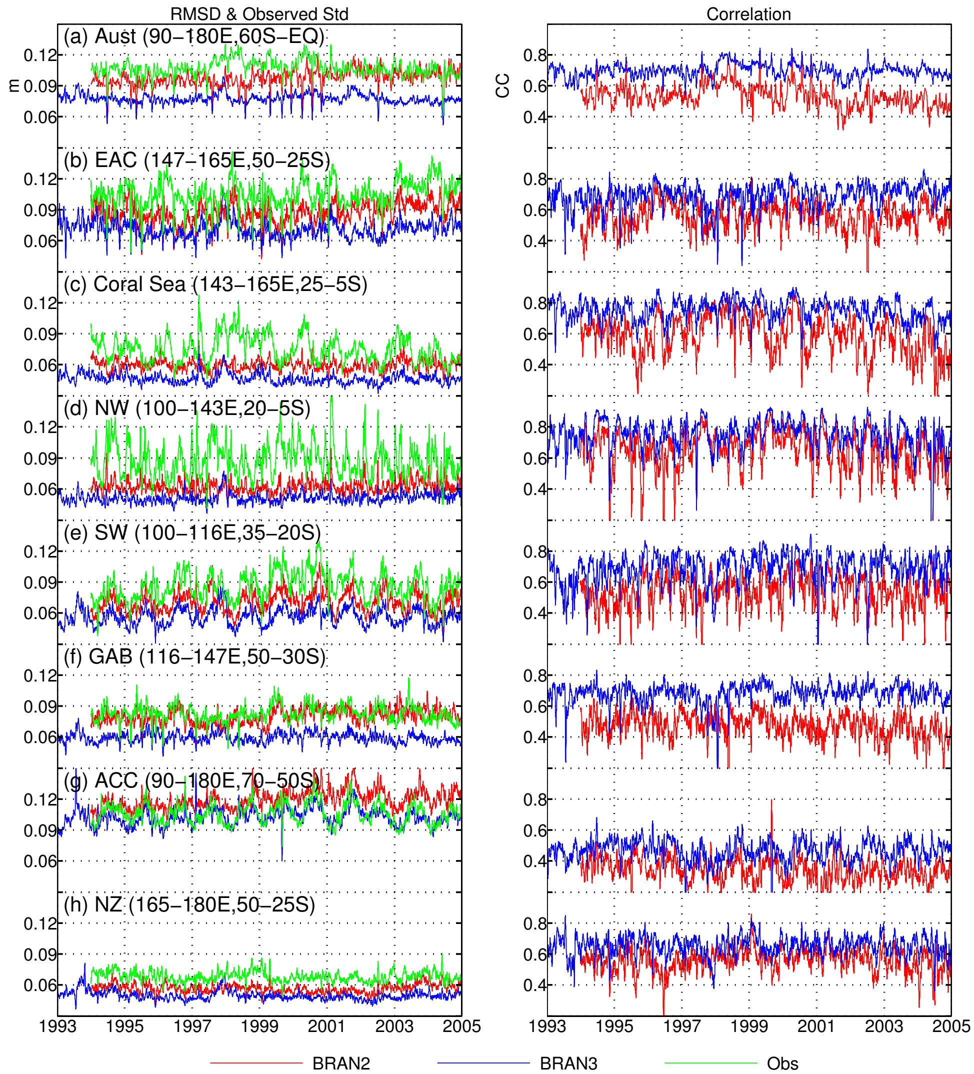

Figure 2: Time series of the RMSD (left) and anomaly correlation (right) between T/P atSLA BRAN2 (red) and BRAN3 (blue) SLA for different regions (see Table 3). The observed standard deviation (green) is also shown.

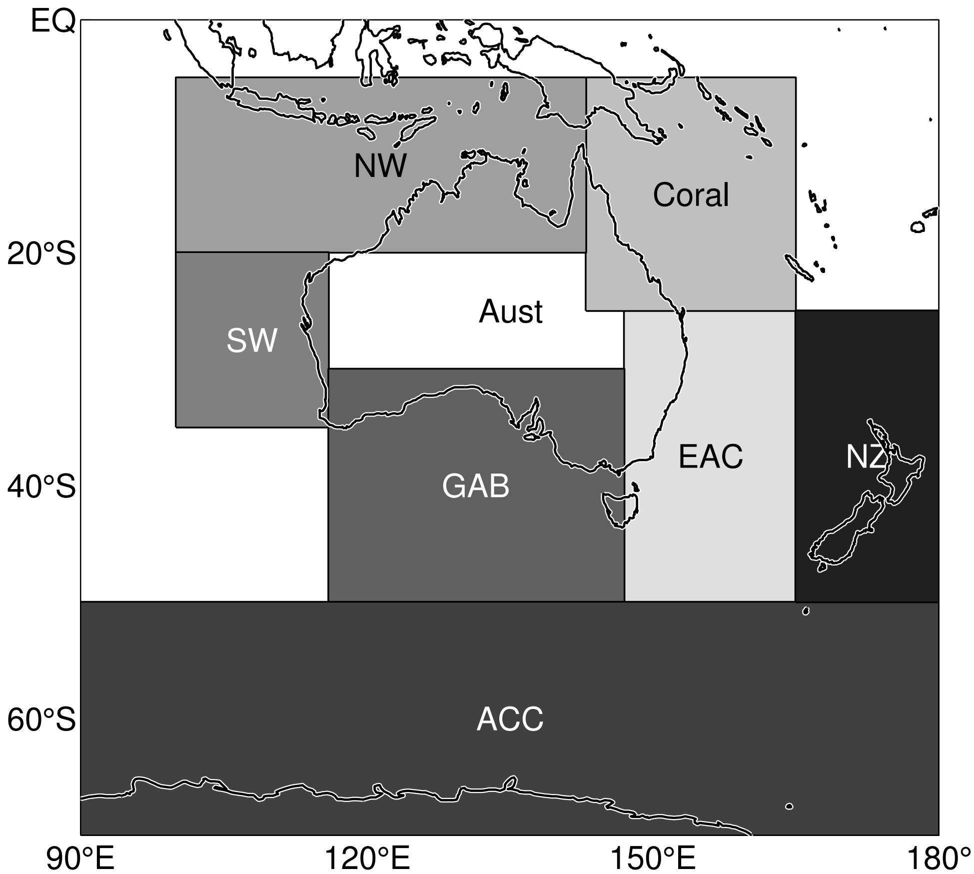

Figure 3: Map of the region of interest showing the different regions around Australia for which statics are computed throughout this study (see Table 3).

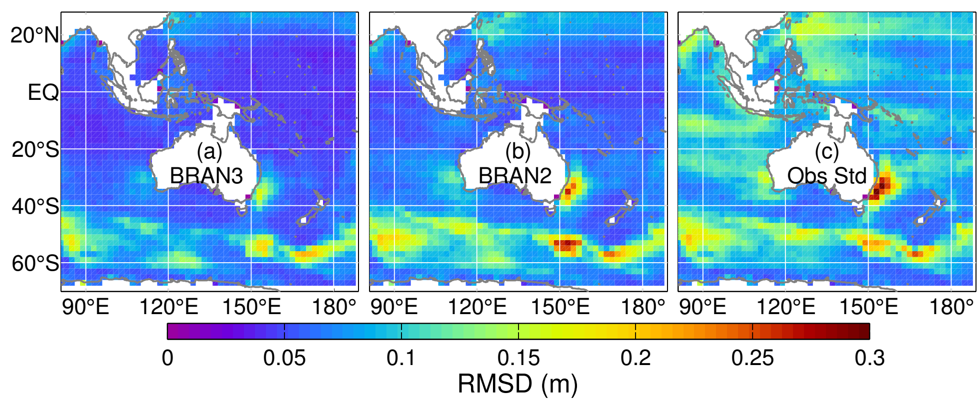

Figure 4: Map of the RMSD between T/P atSLA and (a) BRAN3 and (b) BRAN2 SLA; and (c) the standard deviation of the T/P atSLA observations. Data have been analysed in 2x2° bins and processed for the period 1/1993-12/2004.

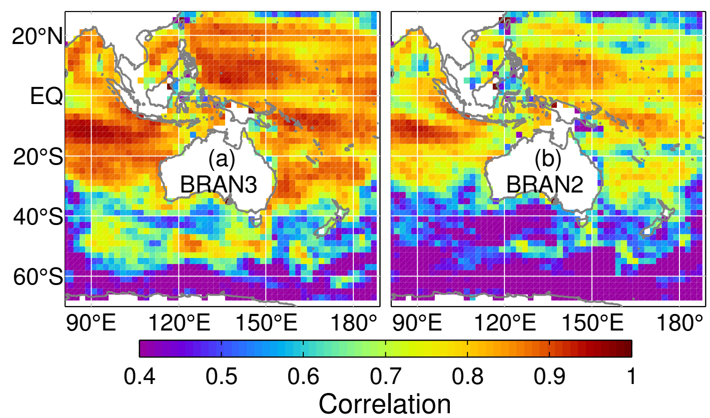

Figure 5: Map of the correlation between T/P atSLA and SLA from (a) BRAN3 and (b) BRAN2. Data have been binned over 2x2° bins and processed for the period 1/1993-12/2004.

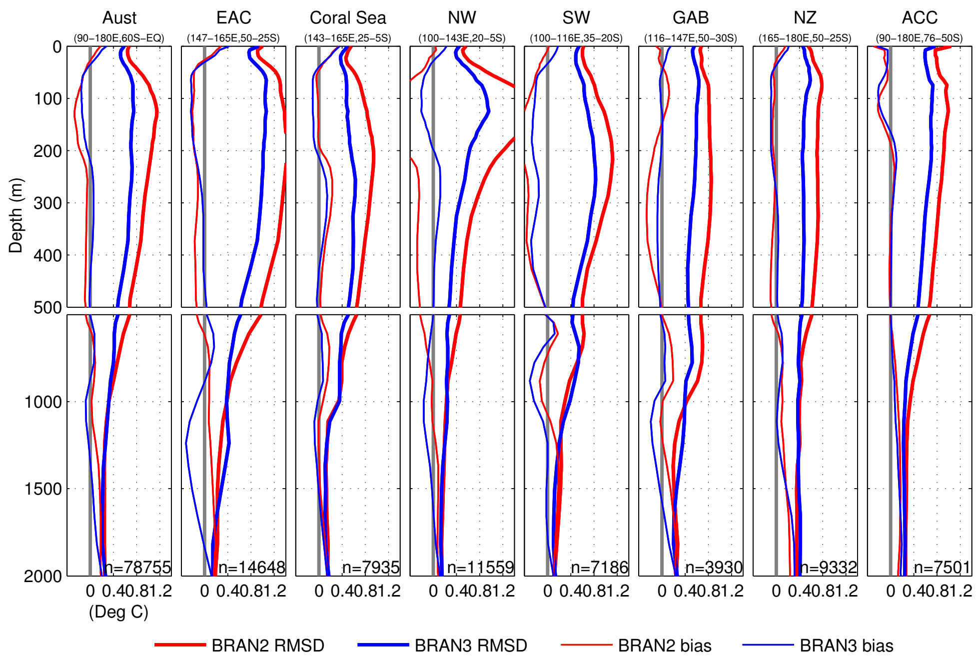

Figure 6: RMSD (bold lines) and bias (thin line; observed minus model) between the temperature from BRAN3 (blue) and BRAN2 (red) for different regions (see Table 3). Comparisons are made for the period January 2003 to December 2006. The average number of observations n, at each depth in the top 700 m is recorded in the bottom of each panel.

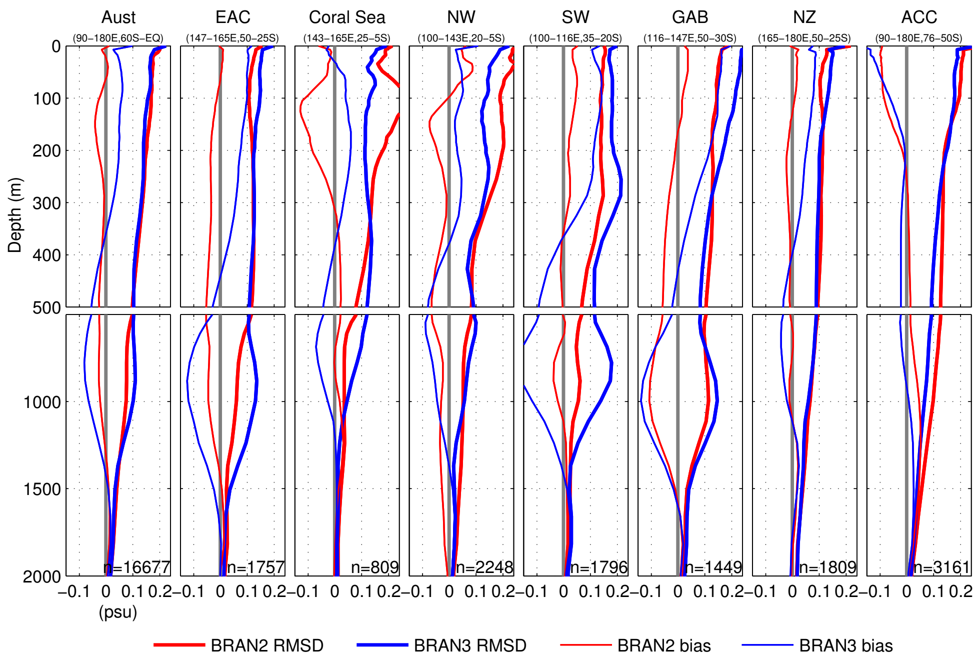

Figure 7: As for Figure 6, except for salinity.

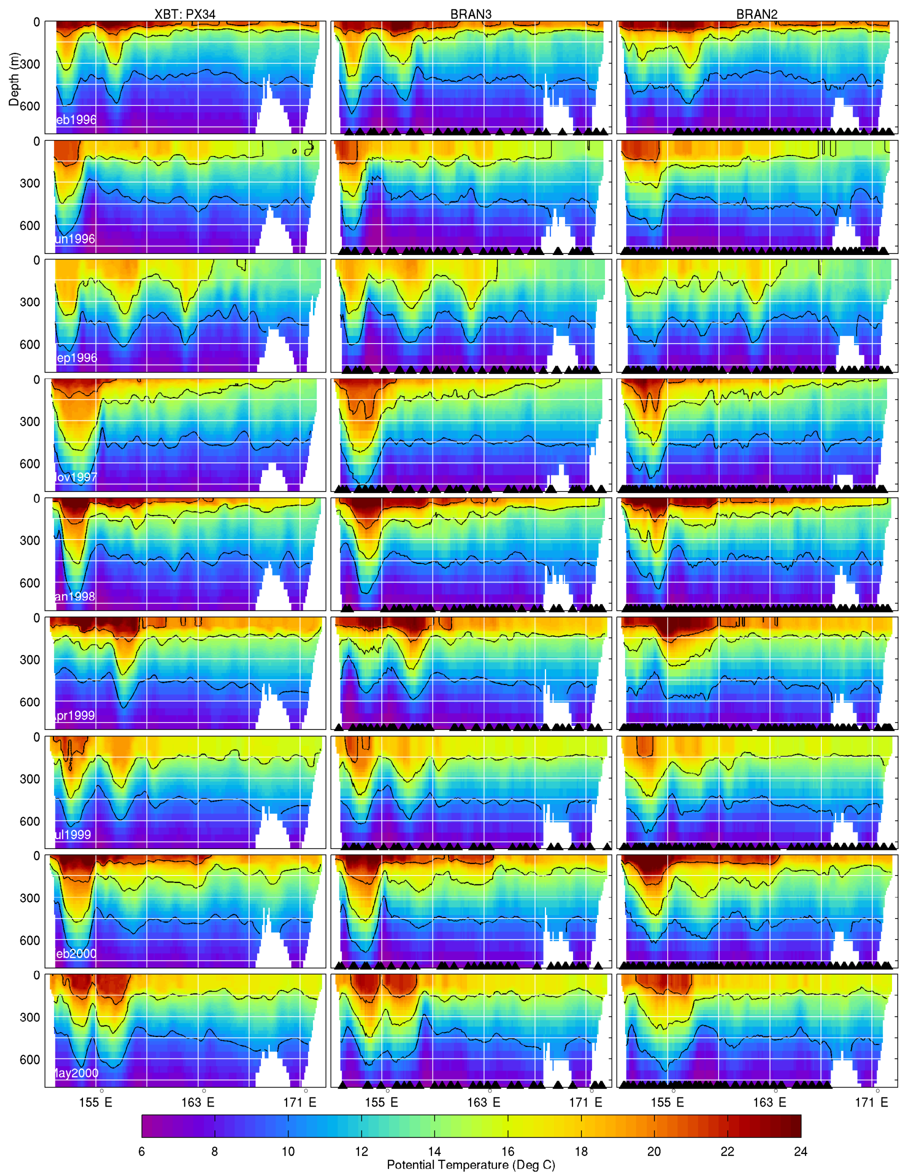

Figure 8: Comparisons between observed and assimilated temperature along the XBT track PX34 (left), and temperature from BRAN3 (middle), and BRAN2 (right), for different times (recorded to the left of each row). The triangles along the bottom of the BRAN3 and BRAN2 panels indicate that the corresponding observed profile is assimilated.

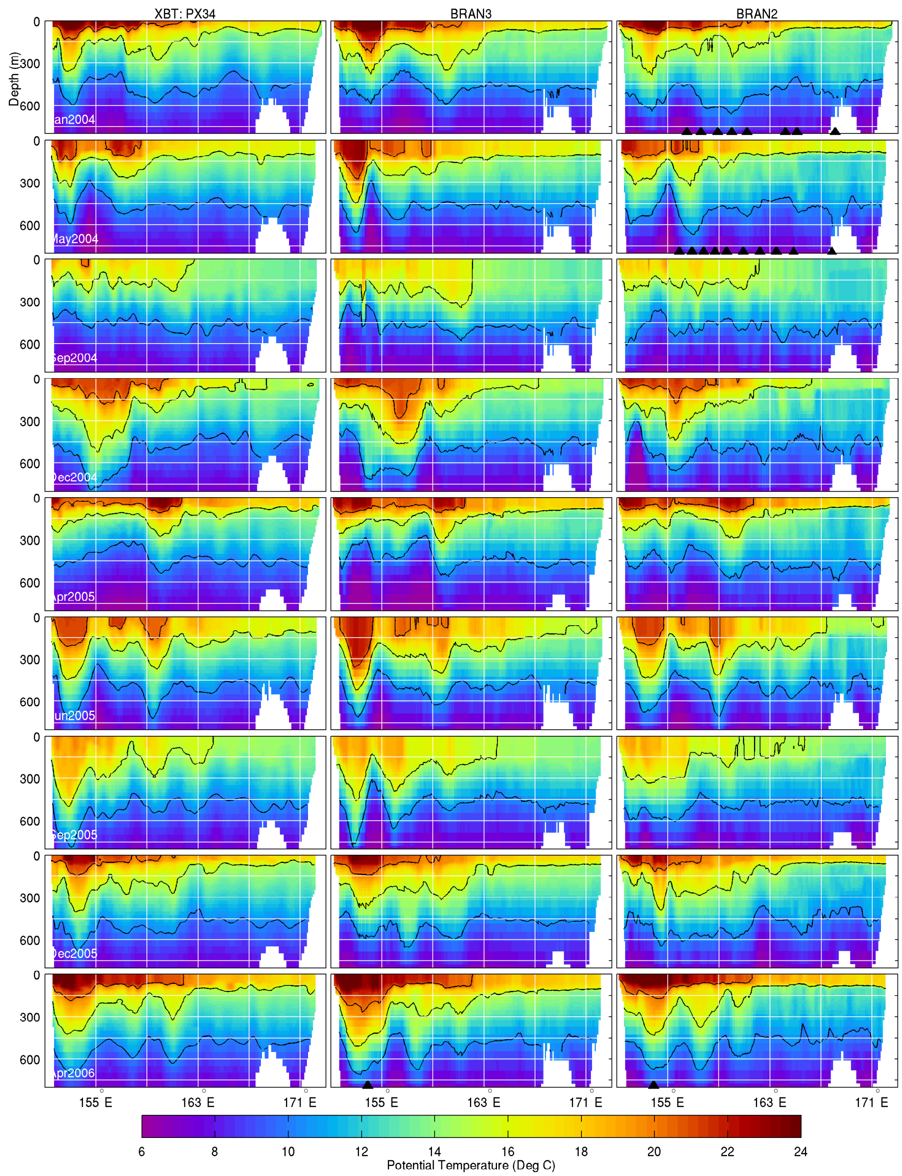

Figure 9: As for Figure 8, except showing comparisons with with-held XBT observations.

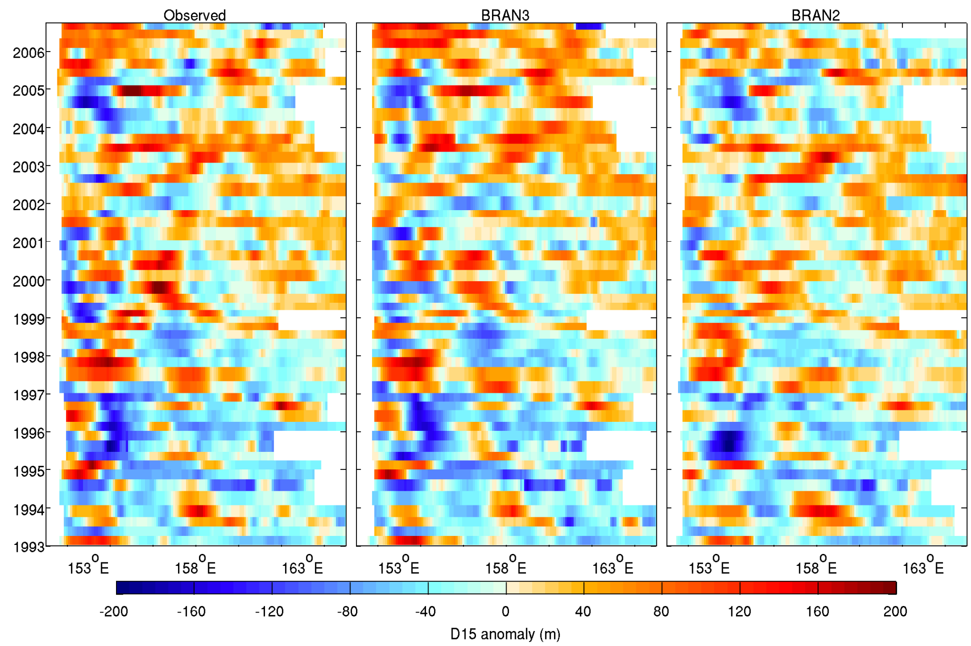

Figure 10: Hovmoller diagram showing the D15 anomaly from XBT observations along the PX34 line (left), and from BRAN3 (middle), and BRAN2 (right). XBT data along PX34 are assimilated before July 2003, and with-held thereafter.

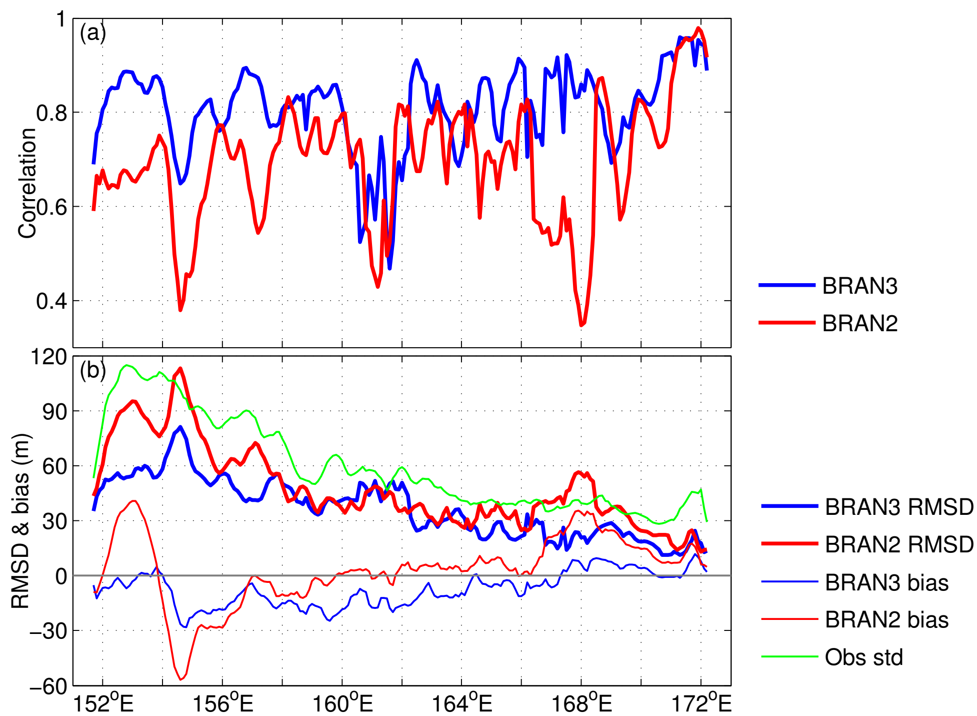

Figure 11: (a) Correlation and (b) RMSD (bold) and bias (thin line; observed minus modelled) between D15 derived from observations and D15 derived from BRAN3 (blue) and BRAN2 (red) along the PX34 XBT line. Also shown on panel (b) is the standard deviation of the observed D15 (green).

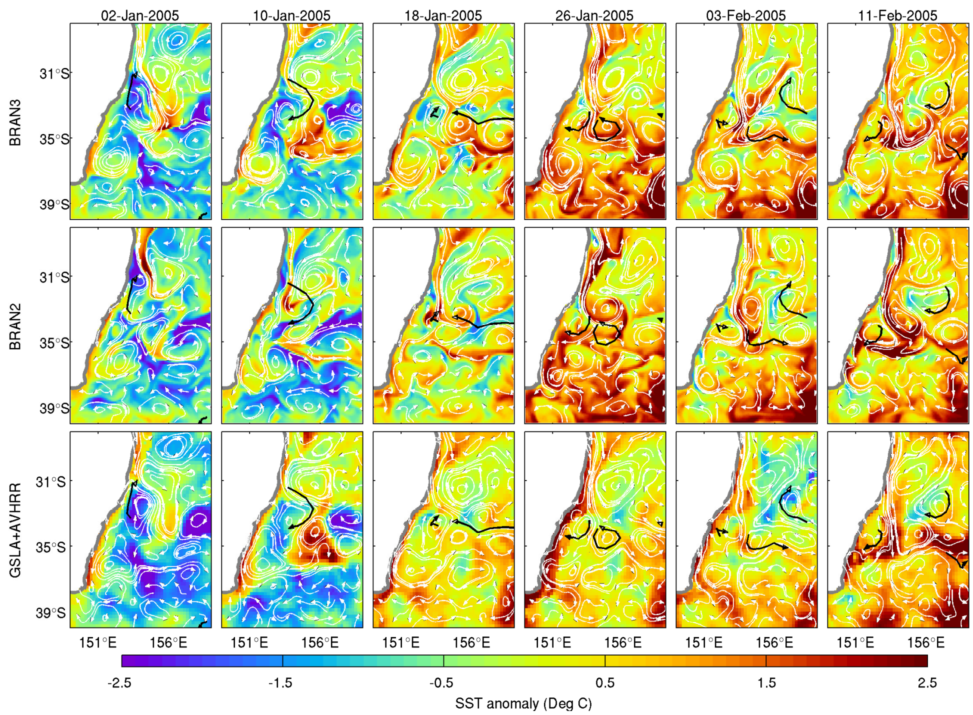

Figure 12: Sequence of daily-averaged SST anomalies and near-surface velocities (white vectors) off south-east Australia in early 2006 from BRAN3 (top), BRAN2 (middle), and observations (bottom), with observed surface drifting buoy trajectories overlaid (black vectors). SST anomalies are with respect to a 15-year seasonal climatology from a spin-up run of OFAM2. Model velocities represent flow over a 5 day period. Drifter trajectories are for a 8-day period preceeding the date of each image.

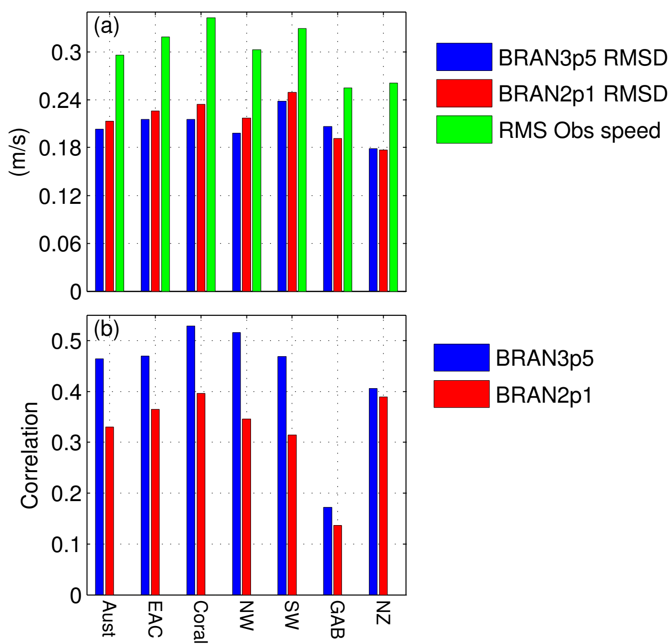

Figure 13: Comparison between BRAN near-surface velocity (12 m depth) and drifter-derived velocity (drogued between 10-15 m), showing (a) the RMSD and the RMS of the observed standard deviation, and (b) the magnitude of the vector correlation for different regions (see Table 3).

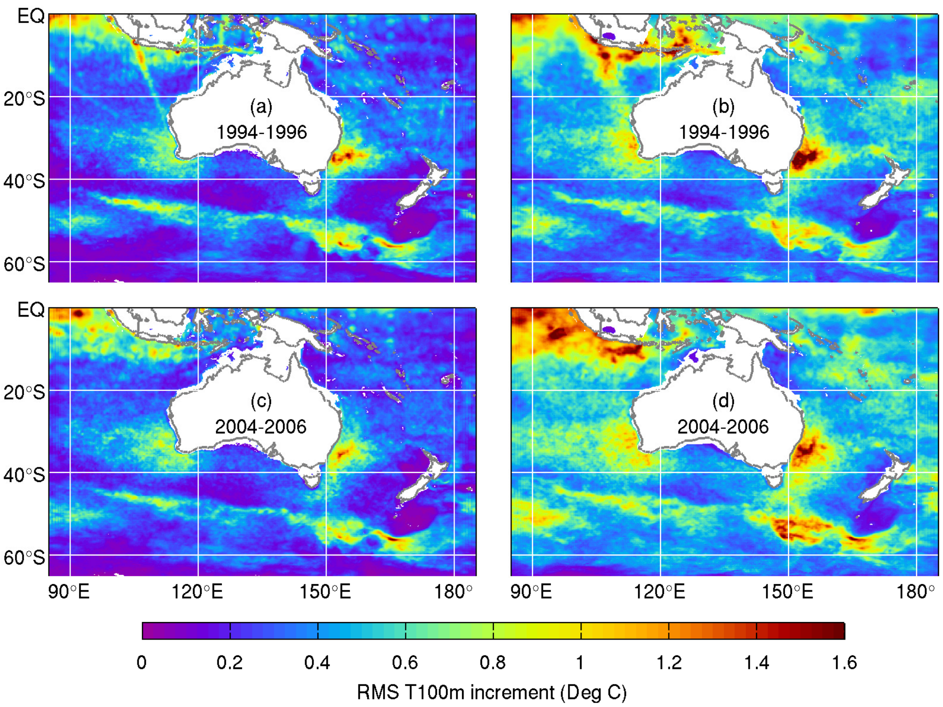

Figure 14: RMS increments for potential temperature at 100 m depth over three different time periods (a-b) 1994-1996 and (c-d) 2004-2006 from (a,c) BRAN3 and (b,d) BRAN2. The area-averaged and minimum ratio of the increments in BRAN3 and BRAN2 are 0.60 and 0.06 respectively.

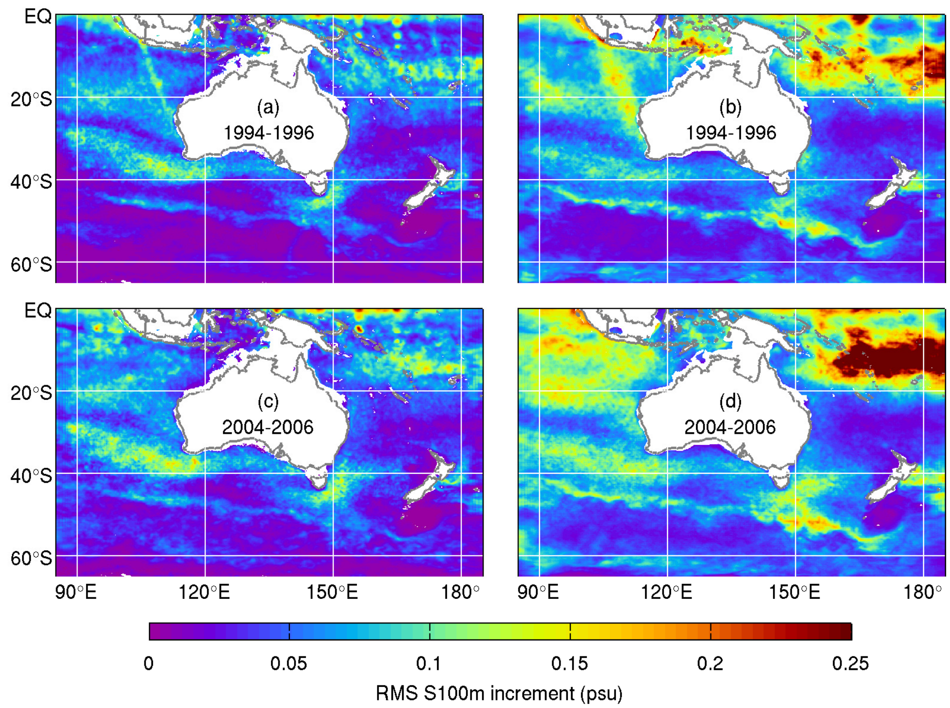

Figure 15: As for Figure 14, except for salinity at 100 m depth. The area-averaged and minimum ratio of the increments in BRAN3 and BRAN2 are 0.66 and 0.05 respectively.

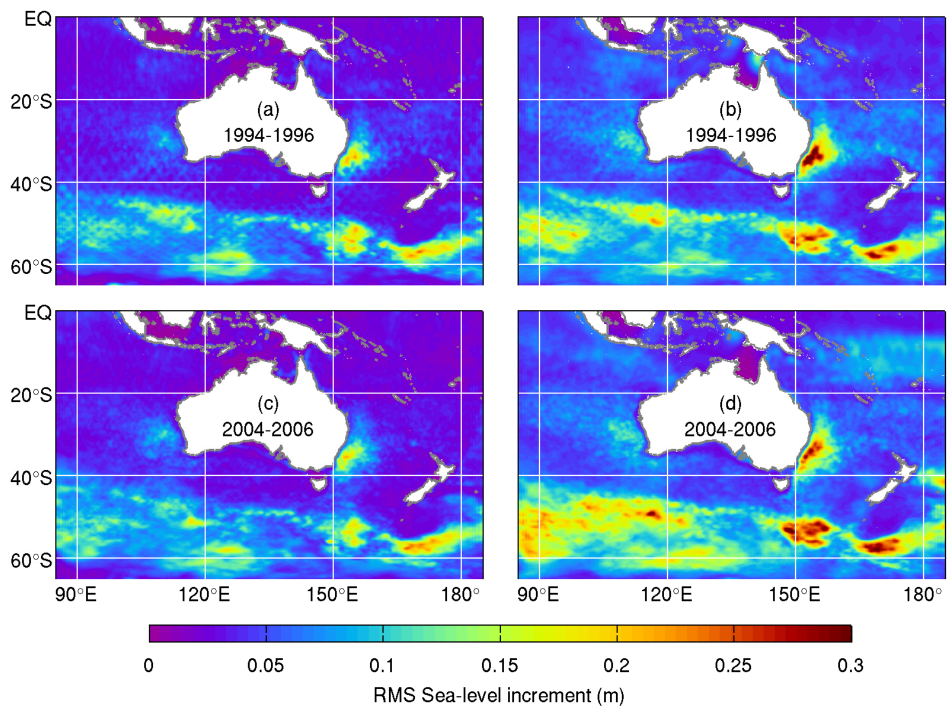

Figure 16: As for Figure 14, except for sea-level. The area-averaged and minimum ratio of the sea-level increments in BRAN3 and BRAN2 are 0.7 and 0.01.

Appendix A. Assimilation details

Appendix A.1. Assimilation algorithm and localisation

Calculation of an analysis using either an EnKF or EnOI is typically performed using either covariance localisation or a local analysis. The version of BODAS used for BRAN2 uses covariance localisation, as described by Oke et al. (2008), while the version used for BRAN3 uses a local analysis that is described below. BODAS uses EnOI, combining an array of observations $\mathbf{y}$ ($p \times 1$, where p is the number of observations) of different types, with a model background field $\mathbf{w}^f$ ($n \times 1$, where n is the dimension of the model state), yielding an analysis $\mathbf{w}^a$ ($n \times 1$), using the standard Kalman filter update equation:

$$\mathbf{w}^a = \mathbf{w}^f + \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{w}^f) \qquad \text{(A.1)}$$

where $\mathbf{P}^f = \frac{1}{m-1}\mathbf{A}\mathbf{A}^T$ ($n \times n$) is the background error covariance matrix, $\mathbf{H}$ ($p \times n$) is the linearised observation operator, and $\mathbf{R}$ ($p \times p$) is the observation error covariance matrix. Equation (A.1) can be re-written in terms of the ensemble transform as follows:

$$\begin{aligned} \mathbf{w}^a &= \mathbf{w}^f + \mathbf{A}\mathbf{b} && \text{(A.2)} \\ &= \mathbf{w}^f + \sum_{i=1}^{m} \mathbf{A}_i \mathbf{b}_i && \text{(A.3)} \end{aligned}$$

where $\mathbf{A}_i$ is the ith ensemble member, and $\mathbf{b}_i$ is the weight of the ith member, where

$$\begin{aligned} \mathbf{b} &= \mathbf{S}^T(\mathbf{I} + \mathbf{S}\mathbf{S}^T)^{-1}\mathbf{s} && \text{(A.4)} \\ &= (\mathbf{I} + \mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{s}, && \text{(A.5)} \end{aligned}$$

and $\mathbf{S}$ and $\mathbf{s}$ are standardised ensemble anomalies and standardised innovations that are given by

$$\begin{aligned} \mathbf{S} &\equiv \mathbf{R}^{-1/2}\mathbf{H}\mathbf{A}/\sqrt{m-1} && \text{(A.6)} \\ \text{and} \qquad \mathbf{s} &\equiv \mathbf{R}^{-1/2}(\mathbf{y}-\mathbf{H}\mathbf{x}^f)/\sqrt{m-1}. && \text{(A.7)} \end{aligned}$$

Equations (A.4) and (A.5) are formally equivalent, but require inversion of matrices of different sizes. Equation (A.4) requires an inversion of a $p \times p$ matrix, where p is the number of observations, while equation (A.5) requires an inversion of a $m \times m$ matrix, where m is the ensemble size.

The ensemble size is typically many orders of magnitude less than the number of degrees of freedom of the model - so a naive implementation of an EnKF or EnOI by simply solving (A.3) yields are a poor fit to observations because the ensemble is severely rank-deficient. This is one of the main reasons why ensemble data assimilation requires localisation (e.g., Oke et al., 2007). Here, we implement localisation by adopting a local analysis (Evensen, 2003), also called domain localisation (Nerger et al., 2011). This approach involves the calculation of a separate analysis for every horizontal grid point in the model. For each such analysis, only observations within a prescribed distance (here we use 250 km) are used, and the calculated ensemble weights $\mathbf{b}$, from (A.3), are stored for each grid point. The analysis in adjacent grid points uses almost the same observations, so the ensemble weights change smoothly over space. To further ensure this smoothness in space, we reduce the magnitude of the ensemble anomalies as a function of distance from each analysis location. This is equivalent to the approach introduced by Hunt et al. (2007), who increased the observation error variance as a function of distance from the each analysis location. The resulting ensemble weights are then spatially dependent, so (A.3) becomes:

$$\mathbf{w}^a = \mathbf{w}^f + \sum_{i=1}^{m} \mathbf{A}_i \mathbf{b}_i(x, y). \tag{A.8}$$

In practice, the ensemble weights in (A.8) are computed for each horizontal grid point independently on multiple processors (we use 192 processors) and stored for later use. The analysis update, the second term in the right-hand-side of (A.8), is constructed after all calculations to compute $\mathbf{b}_i(x, y)$ are complete.

Appendix A.2. Ensemble

EnOI uses a time-invariant ensemble $\mathbf{A}$, to approximate the system's background error covariance matrix $\mathbf{P}^f$. For both BRAN2 and BRAN3 we use output from a long model run to construct an ensemble of intraseasonal anomalies. These anomalies are generated by calculating the difference between a 3-day mean and a 3-month mean. One ensemble member is computed for each month of a long model run, with the 3-day means computed by averaging fields from the 14-16th of each month; and the 3-month means computed for the 3-month period centred on the 15th of each month. For BRAN2, we use fields from the last 6-years of a 9-year integration of OFAM1 (called spinup4/5), to generate a 72-member ensemble. For BRAN3, we use fields from the last 12-years of an 18-year integration of OFAM2 (called spinup6p8), to generate a 144-member ensemble.

Appendix A.3. Observation error standard deviation estimates

Every observation that is assimilated requires an explicit estimate of the observation error variance. Observation error is here considered to have three components: instrument error, representation error, and age error. Here, we estimate each component of the observation error explicitly, and combine them with a quadrature sum. Instrument error arises simply because observations are imperfect, and prone to measurement noise. A list of the assumed instrument errors for different platforms is presented in Table A.1.

Representation error arises from the fact that observations typically measure a point in time and space, that represents processes on all time- and space-scales; while the model represents only a finite range of time- and space-scales. The mis-match between these scales is called representation error - because the observation "represents" variability that differs from the variability that the model "represents". For both BRAN2 and BRAN3 we use representation error estimates based on the approach described by Oke and Sakov (2008). The representation error is typically large (e.g., up to 10 cm for SLA) - particularly in boundary currents and the ACC, where small-scale variability is prevalent.

The last component of the observation error is the age error. For each assimilation step, we typically assimilate observations from a time-window that is centred around the analysis time. For BRAN3 (BRAN2) we use all observations of SLA, SST, and in situ T/S that were made within 21 (11), 5 (1), and 11 (7) days of the analysis time, respectively. As a result, most assimilated observations were made at a different time to the analysis time. The absolute value of this difference in time is here referred to as the "age" of an observation. An observation with a small age is assigned a smaller age error than an equivalent observation with a large age. For BRAN2, we simply chose a time-scale of 3 days - and assumed that the age error increases as a Gaussian over this time-scale, approaching the background variability (approximated with the RMS of each variable from a spin-up run). For BRAN3, we adopt a more novel approach to the age error. Taking model fields from an 18-year spinup run, we calculate the difference between each variable n days apart at all model grid points. We calculate these differences for every month of the model run, and then calculate the RMS of the resulting differences. This yields estimates of how much each variable changes over n days for each location in the model. Assuming the observations change with similar magnitudes and on similar time-scales to the model , these RMS fields represent the age error of the observations. We estimate the age error for each variable for ages ranging from 1-10 days, and use them for each assimilation step. Examples of the age error for SST and SLA are presented in Figure A.1. This figure shows that in regions of high variability, such as the boundary currents, the ACC, and near the coast, the age error of an observation increases with age, as we expect. The age error for SLA near the coast increases quickly, saturating after 2-3 days. In some regions, where the variability is small (e.g., offshore of the GAB, and west of NZ) the age error remains insignificant for all ages. This indicates that even "old" observations in those regions, are useful for data assimilation.

Appendix A.4. Reference MSL

A reference MSL field is used during each assimilation step to convert the model sea-level into SLA. This allows the model SLA to be compared directly to atSLA from satellite altimetry - the first step in the assimilation process. The reference MSL is critical for the success of the reanalysis because it largely determines the mean circulation. For BRAN2, we use the time-averaged sea-level from a 13-year model run of OFAM1 (Spinup4/5); and for BRAN3, we use the time-mean sea-level from the last 18-years of a 33-year run of OFAM3 (Oke et al., 2013). A comparison of the different MSL products is presented in Figure A.2. This comparison shows that there is generally good agreement between the different reference MSL products. However, there are several significant differences. Most relevant for the comparisons in this study is the difference between the reference MSL used for BRAN2 and BRAN3 (Figure A.2d). This difference field shows several systematic differences, including a band of positive and negative difference along the path of the ACC. This indicates that the mean ACC is at a different latitude in these fields. Also, the large differences south of Papua New Guinea indicate that the strength and structure of the South Papua gyre is different in the different reference field. There is also a broad band of positive difference in the Indian Ocean at low latitudes that extends to the south-western corner of Australia; and a broad band of negative difference south of this. Similarly, in the Pacific Ocean there is a band of negative difference between about 20-5°S, and a narrow band of positive difference along the path of the South Equatorial Current. These large-scale differences in the reference MSL fields used in BRAN2 and BRAN3 indicate that the mean circulation associated with these fields are very different.

The differences between the BRAN reference MSL fields and the CNES-CLS09 MSL (Figure A.2e-f) are also significant. There are broad scale differences that are similar in structure to the differences between the BRAN2 and BRAN3 references. These differences are smaller for BRAN3, but in both BRAN fields they are large. The reason for using a model-based estimate of the reference MSL is to ensure that the mean circulation associated with the reference field is compatible with the model - though we note that other groups have adopted observation-based reference fields in preference to model-based reference fields (Cummings et al., 2009).

Appendix A.5. Observation pre-processing

Note that for BRAN3, we have carefully prepared the surface fluxes so that the annual- and global-average MSL remains constrained for the duration of the model run. Consistent with this, we also calculate and remove the global mean sea-level from the atSLA observations prior to each assimilation step. In this way, we have processed the altimetry so that it is effectively volume-conserving - consistent with the model. This eliminates a small, but significant, source of bias in the reanalysis that was identified by Oke et al. (2008).

Not all available observations are assimilated. For altimeter and SST observations, we combine individual observations into super-observations. Within the $1/10°$-resolution region of the model domain, we prepare super-observations on a nominal $0.2 \times 0.2°$ grid. Where no observations are available, no super-observations are computed. The estimated error of each super-observation is calculated based on the estimated errors of the raw observations using standard error propagation techniques (e.g., Taylor, 1997, , pp45). For in situ measurements, we don't compute super-observations. Instead, we simply "thin" the database to retain no greater than one profile of temperature or salinity for every 0.5 degrees.

Altimeter data tends to have larger errors in shallow water, owing to limitations of global tide models and atmospheric corrections used in their processing. We attempt to eliminate contaminated altimeter observations by only assimilating observations over water depths that exceed 200 m. Similarly, satellite SST observations can become contaminated near the coast. To eliminate contaminated SST observations in BRAN3, we only assimilate SST observations over water depths that exceed 20 m. In BRAN2, we eliminated more SST data in this step - with-holding satellite SST data within 100 km of any coastline - an approach that was too conservative and with-held data unnecessarily from the reanalysis.

Table A.1: Summary of the assumed standard deviation of the instrument errors used in BRAN2 and BRAN3.

	BRAN2	BRAN3
CTD temperature	0.01°C	0.05°C
CTD salinity	0.05 psu	0.02 psu
XBT temperature	0.2°C	0.2°C
AVHRR SST	0.5°C	0.5°C
AMSR-E SST	0.25°C	0.4°C
T/P, J1, J2	3 cm	3 cm
Envisat, Cryosat	5 cm	5 cm

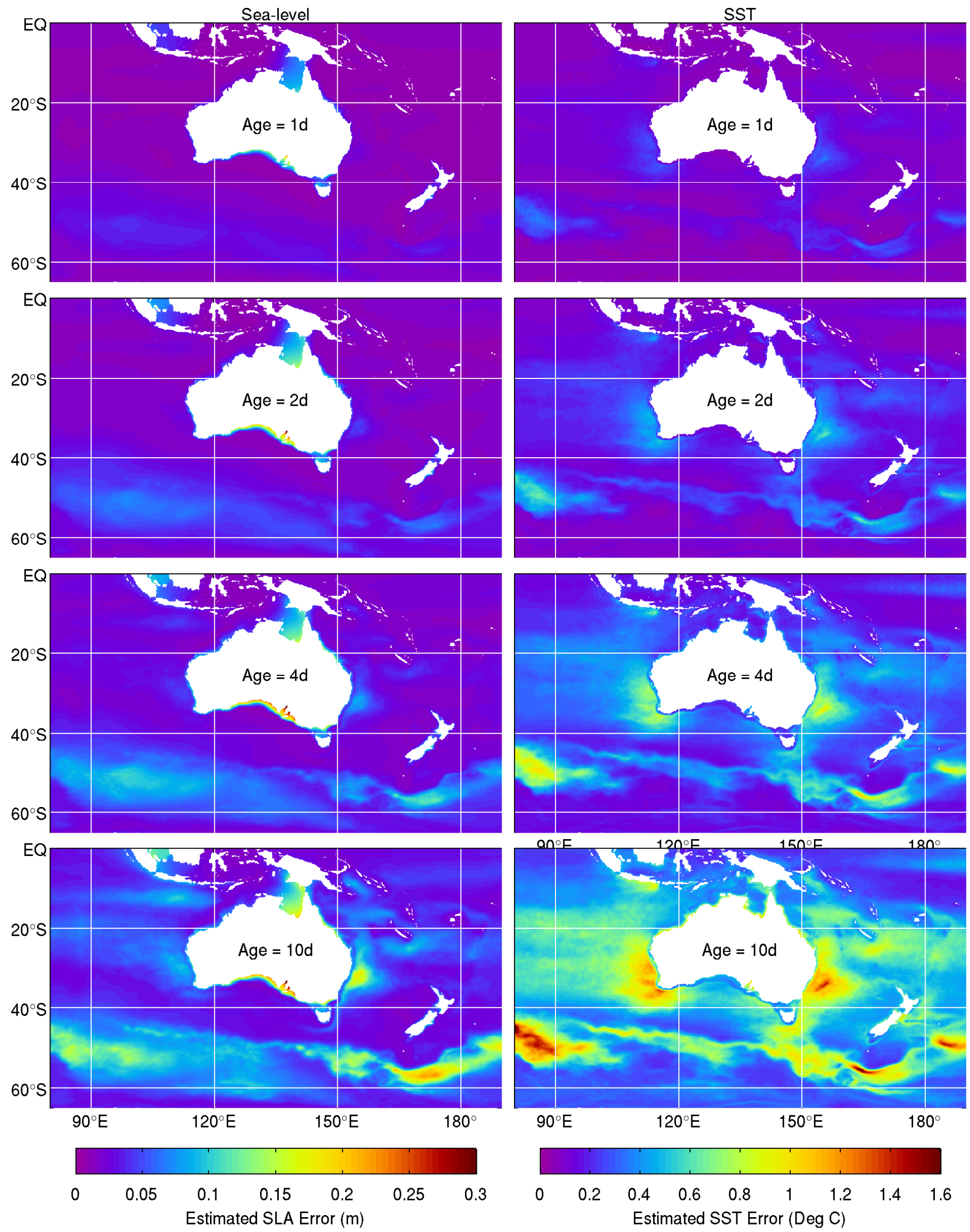

Figure A.1: Age error estimates for sea-level (left) and SST (right) for different ages (1, 2, 4, and 10 days; top-to-bottom).

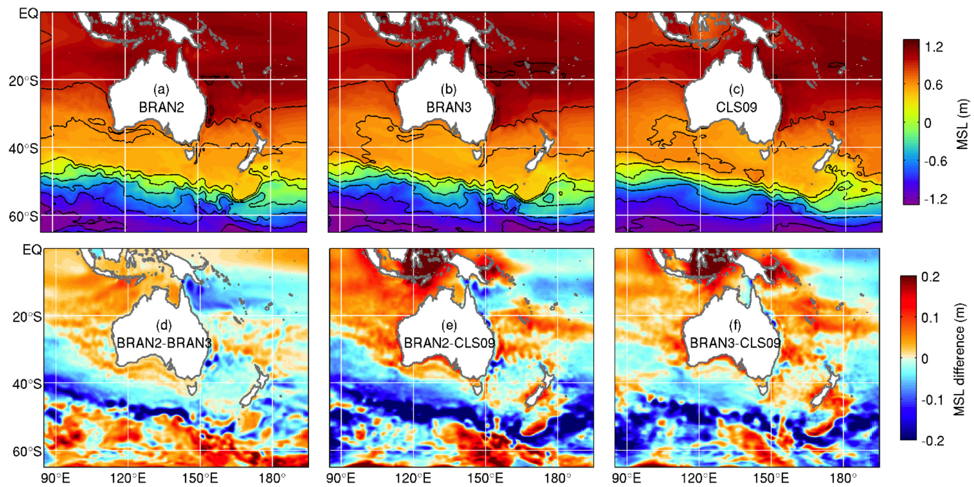

Figure A.2: Comparison of the reference MSL used for (a) BRAN2, (b) BRAN3, and (c) the CNES-CLS09 MSL; and the difference between the (d) BRAN2 and BRAN3 reference MSL, (e) BRAN2 and CLS09 reference MSL, and (f) BRAN3 and CLS09 reference MSL. The contour interval in panels (a-c) is 0.3 m.