Impact of in-situ observations on initial Indian Ocean conditions

Oscar Alves, Yonghong Yin, Li Shi, Harry Hendon and Guomin Wang
(o.alves@bom.gov.au)
Centre for Australian Weather and Climate Research

Abstract

The impact of in-situ obervations on initial conditions of the Indian Ocean using the PEODAS ocean assimilation system is described. PEODAS, the POAMA Ensemble Ocean Data Assimilation System, is a new ocean data assimilation system developed to provide initial conditions for the POAMA dynamical seasonal prediction system. PEODAS is best described as a poor-man's ensemble Kalman filter system. PEODAS is evaluated by comparing it with dependent observations and independent observations such as altimeter and OSCAR surface currents. The PEODAS re-analysis is also compared with the ECMWF re-analysis, a non model based re-analysis from the UK Met Office, the old POAMA-1 re-analysis and a control PEODAS simulation without any observations.

Results show that PEODAS is a signficant improvment over POAMA-1 and is better than the control. PEODAS is also comparable with the latest ECMWF system-3 system. PEODAS, ECMWF and EN3 re-analysis produce similar evolution of the temperatures associated with the Indian Ocean Dipole. However, salinity variations are much weaker in EN3 compared to PEODAS and ECMWF, likely due to the lack of salinity data before Argo. PEODAS and ECMWF indicated that salinity anomalies are likely to be comparable to temperature anomalies during the evolution of the IOD.

1. Introduction

Seasonal prediction is motivated to support a variety of community and industrial groups including agriculture (McIntosh et al. 2007), monitoring and management of coral bleaching events (Spillman and Alves 2008), anticipation and management of vector-borne diseases (Thomson et al. 2006), and for long-range forecasts of climate indicators (e.g., Zhao and Hendon 2009). Seasonal prediction usually involves the initialisation and integration of a coupled ocean-atmosphere model – and often involves ensemble forecasting. An important aspect of a seasonal prediction system is the ocean initialisation, where oceanic observations are combined with a model-generated background field, providing ocean initial conditions for a forecast, or an ensemble of forecasts.

A new ocean analysis system has been developed for operational implementation at the Bureau of Meteorology. The ocean analysis system for the first generation seasonal prediction system in Australia, called the Predictive Ocean Atmosphere Model for Australia (POAMA; Alves et al. 2003; http://poama.bom.gov.au/), is based on a univariate optimal interpolation (OI) system (Smith et al. 1991) that assimilates only in

situ temperature observations. The new analysis system is called the POAMA Ensemble Ocean Data Assimilation System (PEODAS). PEODAS is a variation of the Ensemble Kalman Filter (EnKF; Evensen 1994; Houtekamer and Mitchel 1998; Burgers et al. 1998). PEODAS can be regarded as a poor man's EnKF. Full details of PEODAS and its general performance globally can be found in Yin et all (2009).

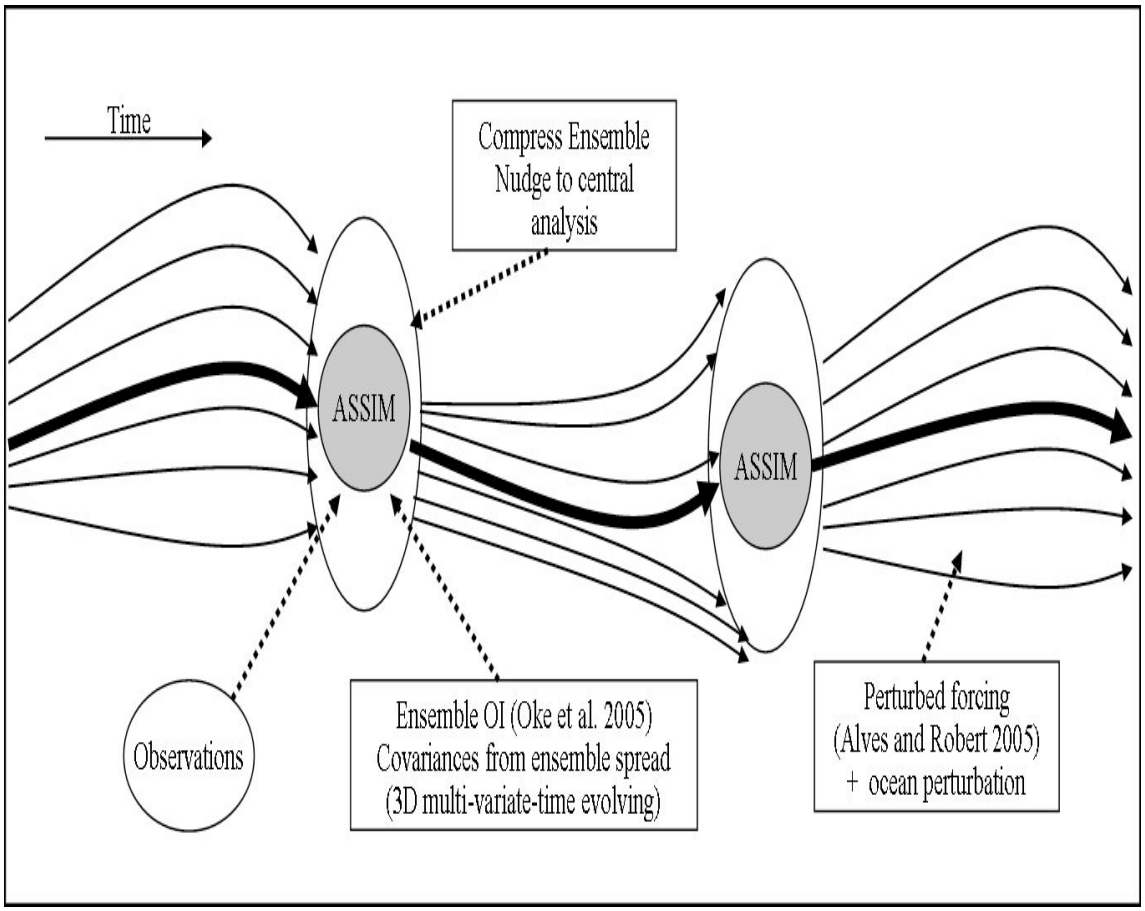

Figure 1. Schematic of the assimilation system, showing the consecutive integration of an ensemble of forecasts, with perturbed forcing (Alves and Robert 2005), the combination of observations with the ensemble mean using an EnOI approach (Oke et al. 2005), and the compression of the ensemble towards the central analysis.

The purpose of this paper is to document the performance of PEODAS in the Indian Ocean. The Indian Ocean is a region of particular interest because, as will be documented below, there has been until recently a paucity of observational data upon which to produce the analyses. Hence, the analysis in the Indian Ocean will depend critically on the performance of the assimilation system in the presence of limited observations. Furthermore, variations of the Indian Ocean are tightly linked to variations of Australian climate, so accurate state estimate of the Indian Ocean is critical for improved climate prediction in the Australian region. Based on hindcasts for the period 1980-2006, the level of skill for predicting seasonal variations of the upper Indian Ocean with the Bureau's seasonal prediction model POAMA is much lower than for the tropical Pacific.

This may reflect model error in the Indian Ocean, a true lack of inherent predictability in the Indian Ocean, or poor quality initial conditions that are provided by the ocean analysis. It it likely that reduced skill in the Indian Ocean compared to the Pacific stems from all three of these issues, and so it is important to assess the relative contributions. Hence, we focus here on the quality of the initial oceans conditions as provided by the ocean analysis, focusing both on the impact of in-situ observations and on improvements in the assimilation system that are provided by PEODAS. Subsequent work will explore the impact of the improved assimilation from PEODAS on seasonal forecast skill.

The performance of PEODAS is assessed using a 27-year reanalysis. The re-analysis is assessed by comparing with in situ and other independent data, against other international re-analsyis, a re-analysis using the old POAMA assimilation system and a re-analysis of the POAMA system with no observations. This paper is organised as follows: a brief description of PEODAS is given in section 2. Section 3 shows the results for PEODAS, including an evaluation against dependent and independent observations, and a comparison with a re-analysis with the old POAMA-1 assimilation and a re-analysis with the ocean model only with no assimilation. Section 4. Compares PEODAS with other international re-analysis. Section 5 investigates the ability of PEODAS to represent Leewin current variability. Section 6 provides a summary and conclusions.

2. Details of PEODAS and the re-analysis

a. PEODAS

The new assimilation system that is introduced here, PEODAS, is a variation of an EnKF system. PEODAS includes the routine generation of an ensemble of forecasts of the upper ocean, and a state-dependent estimate of the background error covariance. PEODAS differs from traditional EnKF systems in that only a single analysis is computed for a central forecast. This analysis is calculated using a modified version of the Bluelink Ocean Data Assimilation System (BODAS; Oke et al. 2008). A schematic of a typical integration of PEODAS is presented in Figure 1. Briefly, a central model run is integrated, along with 11 perturbed ensemble members. These 11 ensemble members are augmented by the 11 ensemble members from each of the previous 9 assimilation cycles that were performed for the preceding month. Together, this gives an ensemble of 110 members. At each assimilation cycle, the 110-member ensemble is used to estimate the background error covariances according to equations (3). Using these background error covariance estimates, the model background field that is derived from the central model run is updated by assimilating a range of different observations using the analysis equations (1-2). PEODAS computes analyses and analysis increments for temperature, salinity, and velocity. The analysis increments for velocity are based on the ensemble-based covariance only, and do not involve any explicit assumption of geostrophy. Full details are given in Yin et all (2009).

$$\mathbf{w}^a = \mathbf{w}^b + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{w}^b),$$

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1}$$

PEODAS has several appealing aspects to it that meet what we regard as essential requirements for a seasonal prediction system. Firstly, it yields an ensemble of forecasts/analyses that are intended to span the actual uncertainty of each forecast. Such an ensemble is useful for generating probabilistic forecasts. Secondly, the background error covariances that are used to assimilate observations are state-dependent and are generated based on our expectations of the dominant sources of error in our system. Thirdly, PEODAS is computationally affordable, and can be scaled according to the computational resources available. We would prefer to integrate a new ensemble of 110-members for each cycle, but this is prohibitively expensive. As computational resources increase, we can readily reconfigure PEODAS to integrate more ensemble members at each cycle, and retain ensemble members from fewer previous cycles (e.g., we may integrate 30 ensemble members and retain ensemble members from the last 4 cycles, yielding a 120-member ensemble).

b. PEODAS re-analysis

The ocean model that underpins both POAMA-1 and PEODAS is version 2 of the Australian Community Ocean Model (ACOM2; Schiller et al. 2002), a global configuration of version 2 of the Modular Ocean Model (MOM2; Pacanowski 1995). The model configuration is described in detail by Schiller and Godfrey (2003). Briefly, the model has constant zonal resolution of 2° and enhanced meridional resolution of 0.5° within 8° latitude of the equator that gradually increases to 1.5° toward the poles. There are 25 vertical levels with 7 levels in the top 100 m. This version of the model includes the hybrid mixed layer model described by Chen et al. (1994).

To assess the relative performance of POAMA-1 and PEODAS, we perform 3 27-year integrations, including a control run with no data assimilation, and two assimilation runs using different ocean analysis systems, one using POAMA-1 and one using PEODAS. Each run is integrated for 27-years, between January 1980 and December 2006. For the assimilating runs analyses are performed every 3 days using observations in a time window of 3 days, centered on the analysis time. For PEODAS, we assimilate observations of in situ temperature and salinity from conductivity-temperature-depth, expendable bathythermograph (temperature only), and Argo profiles, sourced from the ENACT database (Ingleby and Huddleston, 2007). For POAMA-1, only temperature profiles in the upper 500 m are assimilated. Once the analyses are computed, the ocean model is updated in a single time-step by simply replacing the model background fields with analyses. For POAMA-1, the temperature and velocity fields are updated. For PEODAS, the temperature, salinity, and velocity fields are updated.

For the runs performed here, we do not use the full coupled system that is used in the operational seasonal forecast system at the Bureau. Here, we simply use the ocean model of this system and prescribe the surface fluxes. We use this approach here to keep the analysis of the relative performance as simple as possible. The unperturbed surface fluxes for momentum, heat, and freshwater are derived from ERA-40 (Uppala et al., 2004,

Troccoli and Kallberg, 2004) for the period 1980-2001; and from NCEP2 (Kanamitsu et al., 2002) for the period 2002-2006. Observation errors are assumed to be correlated in space for POAMA-1, as documented above, and to be uncorrelated in space for PEODAS. For both POAMA-1 and PEODAS, the observation errors are scaled by the assumed, or modelled, background error variance, so that the ratio of the background to observation error is 0.47. We recognise that this negates one of the benefits of the ensemble in PEODAS, but it facilitates a direct inter-comparison with POAMA-1. Such an inter-comparison is one of the main motivations for this study.

c. Other re-analysis

POAMA-1 is a first generation seasonal prediction system that is run operationally at the Bureau of Meteorology (Alves et al. 2003; http://poama.bom.gov.au/). The data assimilation system for POAMA-1 is a variation of the univariate OI scheme that is described by Smith et al. (1991). The details of POAMA-1 are presented by Alves (2003). The ratio of the observation errors to background field errors in POAMA-1 are set to 0.47 and the observation errors are assumed to be correlated in space, with an *e*-folding decorrelation scale of 150 km. The background error covariances $\mathbf{P}^b$, are modelled as Gaussian functions. Within 10^o of the equator, the covariances are assumed to be anisotropic, with *e*-folding zonal and meridional length-scales of 1500 km and 300 km respectively. At mid- and high-latitudes (outside of $\pm 25^o$ latitudes) the covariances are assumed to isotropic, with *e*-folding length-scales of 500 km. Between 10^o and 25^o from the equator the length-scales are simply linearly interpolated.

POAMA-1 assimilates observations of in situ temperature from the top 500 m of the ocean, producing two-dimensional maps of temperature on model levels. Salinity is not updated by POAMA-1. Based on the increments computed for temperature, geostrophic increments are computed for velocity, using a method that is similar to that described by Burgers et al. (2002).

3. PEODAS simulation in the Indian Ocean

a. Obs distribution

Figure 2 shows the number of temperature and salinity profiles in the tropical Indian Ocean each year throughout the re-analysis period. Throughout the 1980's and the first half of the 1990's the number of observations were steady, around 3000 temperature observations and 300-400 salinity observations per year. Between 1996 and 2001 the number of observations decreased. However, from 2001 onwards there has been a dramatic increase in both temperature and salinity profiles, due to Argo. The increase is more dramatic for salinity, since the number of salinity observations pre 2001 was very small. In the most recent years there was approximately 11,000 temperature profiles and

10,000 salinity profiles.

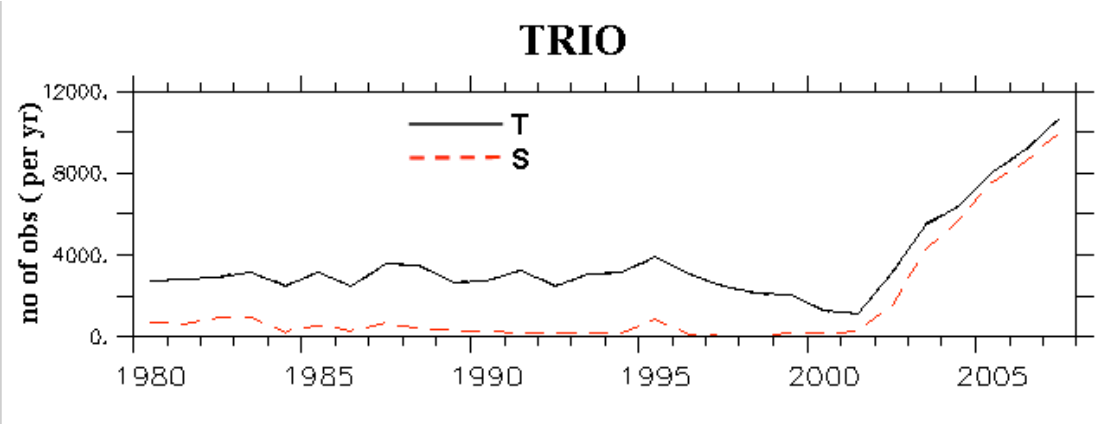

Figure 2. Number of observations per year for temperature and salinity profiles in the tropical Indian Ocean (40°E-120°E,20°S-20°N)

Figure 3 illustrates the spatial distribution of temperature and salinity observations during different periods, expressed as the number of model observations in a year for each model gridpoint. In 1987 temperature observations were available along major shipping tracks (figs 3a). There were also a number of observations which which were not associated with the major shipping lanes, particularly in the western equatorial Indian Ocean and the northern Indian Ocean and off the coast of northwestern Australia. The number of observations per grid point is typically, between 2 and 10, indicating that on a monthly or higher frequency the observations distribution would be much less reduced. During 1987 there was a concentration of gridpoints with up to 4 salinity observations, mainly in parts of the northern Indian Ocean and off the north west coast of Australia.

As mentioned above (see fig 2) there was a minimum in both temperature and salinity observations around 1999-2001. The distribution of observations for temperature and salinity during this period is shown in figs 3b. These show that during this period both temperature and salinity observations were mainly limited to major shipping routes, with widespread gaps particularly in the western half of the Indian Ocean and particularly for salinity.

Argo increased the number of observations significantly from around 2002 onwards. The distribution in 2007 is shown in figs 3c. The impact is dramatic.The distribution is uniform and around 90% of all grid points have at least 2 observations. While there are some gaps reaching some 5-10 gridpoints, there are no large scale gaps. Over at least half of the gripoints there are at least 12 observations, i.e. on average one per month, and for some areas there are gridpoints with over 24 observations per year. The distribution for salinity is similar to that for temperature, as most Argo floats measure both quantities.

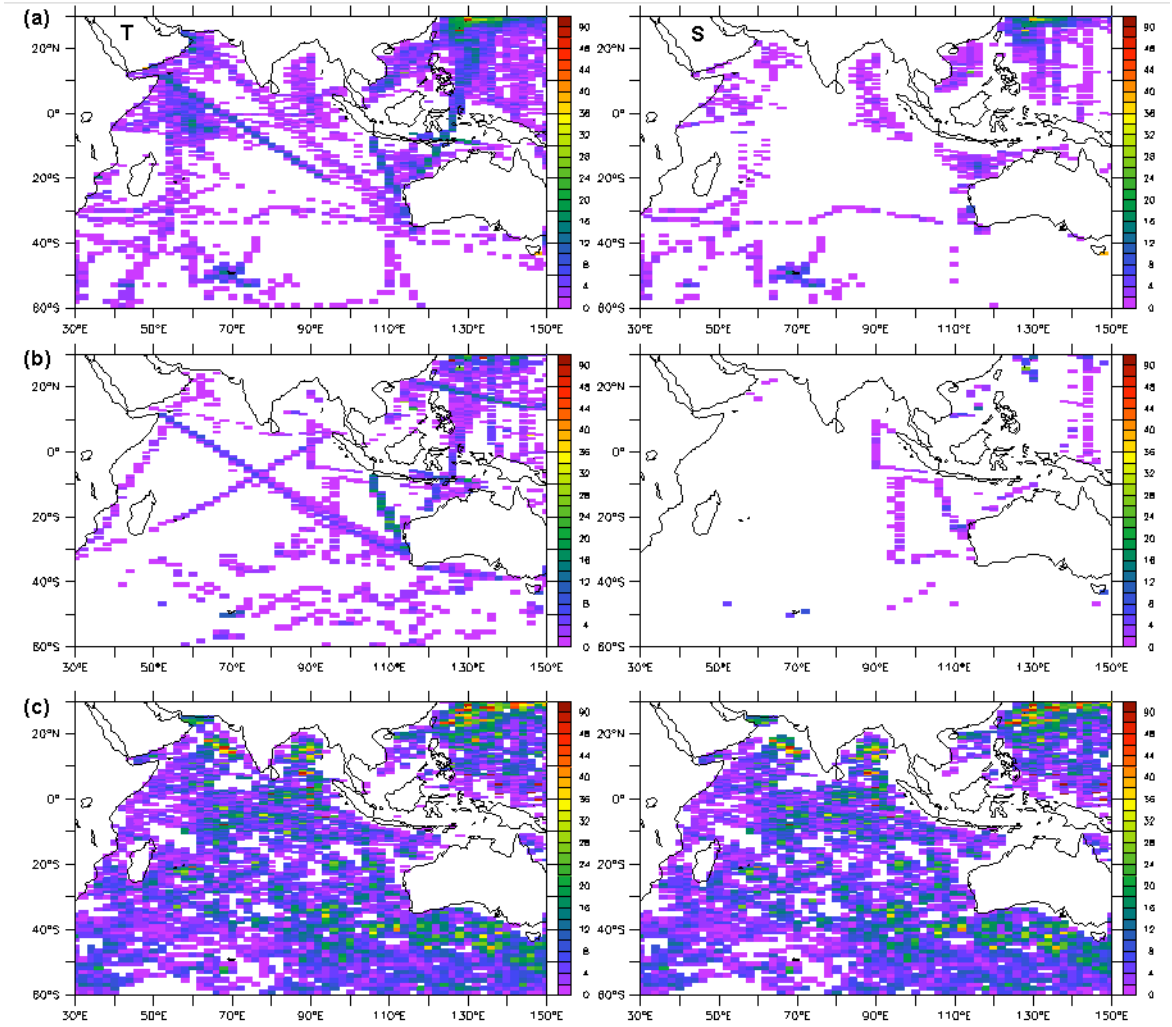

Figure 3. The number of the in-situ observations per model grid point for temperature (left) and salinity (right) profiles assimilated during the one year period of (a) 1987, (b) 2000, and (c) 2007.

b. Assimilation statistics

The PEODAS system makes correction to the model background using each observation. For each observation location a correction is made to all variables and applied over a spatial region that is influenced by that observation. The nature of the correction is determined by the background error covariance matrix, which is calculated by using the ensemble of forecasts about the central analysis. The error covariance structure is simply determined from the ensemble covariance structure. Figure 4 illustrates how a temperature observation in the eastern Indian Ocean at different depths would impact the surrounding model gridpoints. Figure 4a(left) shows how an observation increment at the surface is distributed horizontally. Each observations has an impact over a relatively large spatial area, determined by the covariance structures in the assimilation ensemble. At the

surface the structure has an almost Gaussian shaped elongated along the equator. At depth (fig 4a right) the structure is more complex and shows how errors in the eastern Indian Ocean are related to errors along the coast away from the equator, a reflection that errors are mainly associated with surface wind stress errors and therefore manifest themselves in patterns associated with equatorial Kelvin/Rossby waves and how these waves evolve once they hit the eastern/western boundary - e.g. reflected Rossby waves and coastal trapped Kelvin waves.

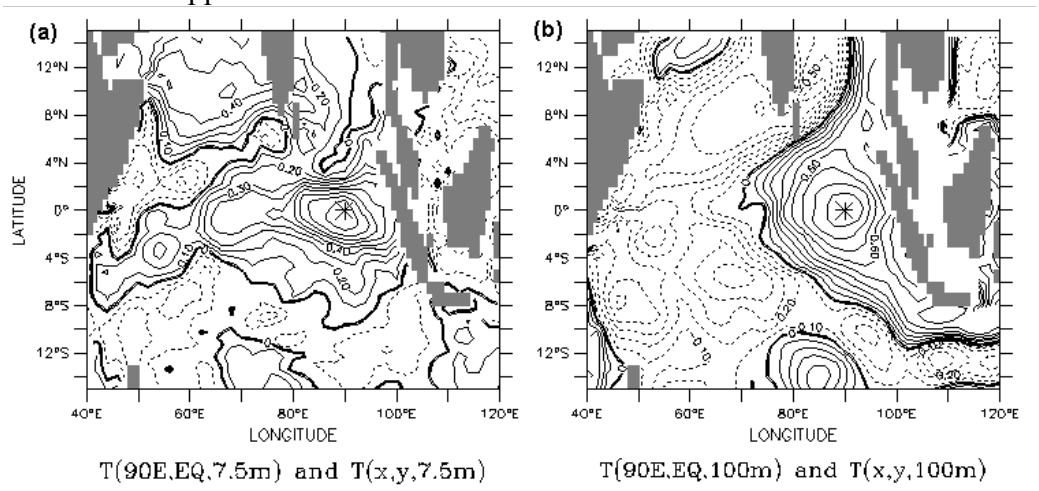

Figure 4. Example of the ensemble-based covariance structure showing cross-correlation between temperature at a reference location (90°E, EQ) in different depth, denoted by the star, and the temperature in the surrounding region

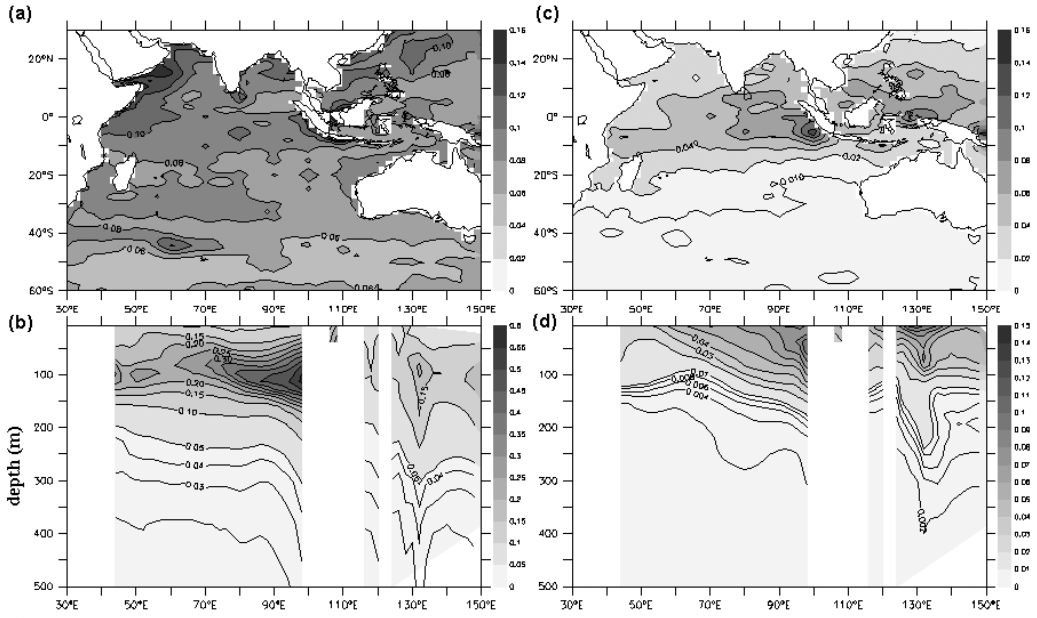

Figure 5. Spread of the ensemble (before assimilation) over the re-analysis period showing fields of (a) SST (Deg C), (b) Temperature section along the equator (Deg C), (c) Sea surface salinity (psu) and (d) salinity section along the equator (psu).

The time-averaged ensemble spread (ensemble standard deviation) is presented for surface and sub-surface temperature and salinity in Figure 5. The time-averaged ensemble spread (Figure 5) is intended to approximate the time-averaged magnitude of the background field error. This information is not used in PEODAS, although it is normally used in a tradition Ensemble Kalman Filter. However, the ensemble members will be used to perturb the coupled model forecasts. The field presented in Figure 4 demonstrate that the magnitude of the background field error for temperature and salinity has significant spatial variation. The ensemble spread for both SST and SSS identify the regions of high uncertainty in the Indian Ocean. The ensemble spread for SSS is high in the eastern tropical Indian Ocean.

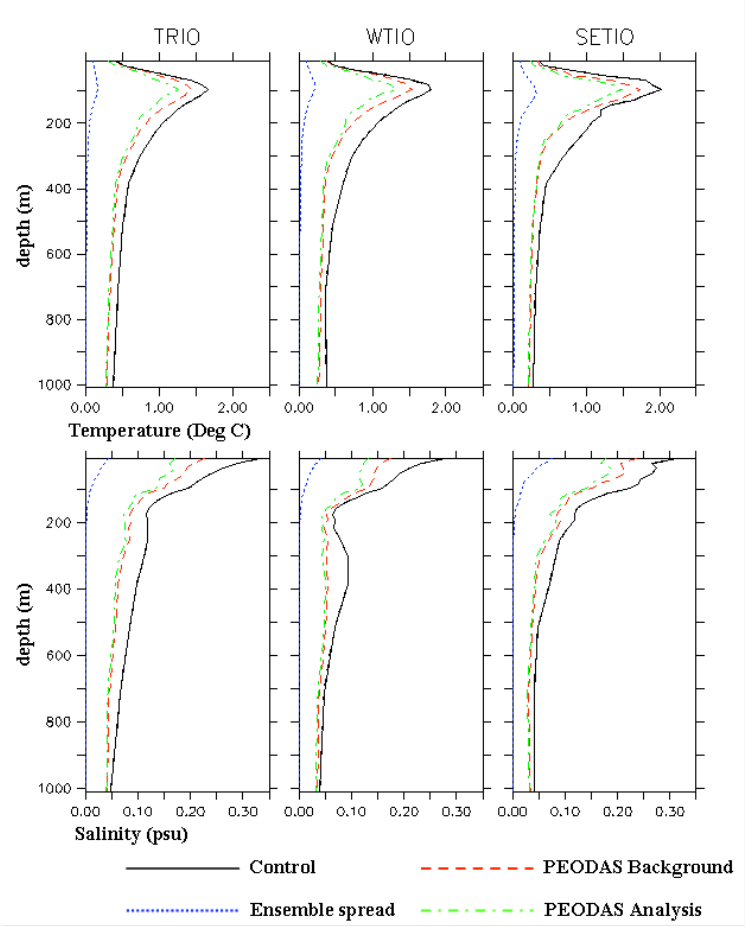

Figure 6. RMS difference between re-analysis and observations, for all observations within each area and over the re-analysis period.

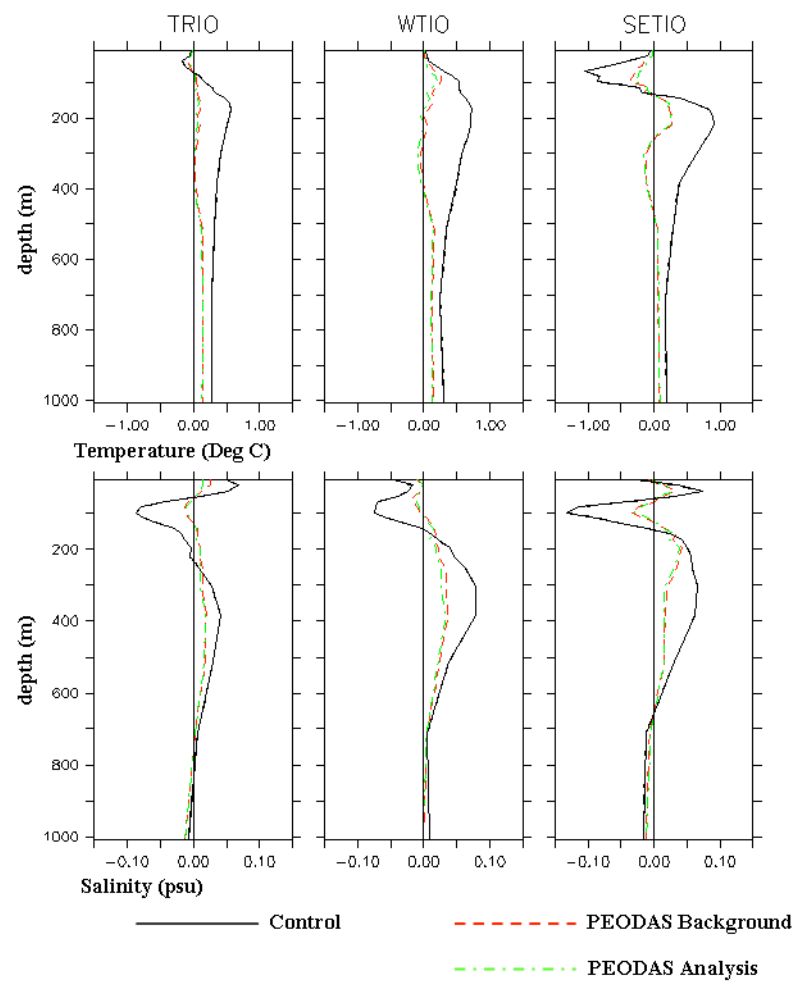

Figure 7. As figure 6 except for mean difference between re-analysis and observations.

In this section we compare observed and modelled temperature and salinity fields for different locations, calculated at each observation location and averaged for each region. The RMS difference between observed and modelled sub-surface temperature and salinity in different regions for the entire 27-year reanalysis period are shown in Figure 6. Specifically, we show the RMS difference (RMSD) between the observed temperature and salinity and the Control run (with no data assimilation), the PEODAS background field (immediately before assimilation) and the PEODAS analysis (immediately after assimilation). We also show the profile of the time-averaged ensemble spread. The ensemble spread is intended to represent the magnitude of the background error. It should therefore have a structure and magnitude that is similar to the RMSD between the observed and background fields. The profiles presented in Figure 6 are produced by interpolating each model product to the observation location in time and space, and then averaged all profiles in different regions. The regions shown include TRIO (40-120^{o}E, 20^{o}S-20^{o}N), WETIO (50-70^{o}E, 10^{o}S-10^{o}N), SETIO (90-110^{o}E, 10^{o}S-EQ).

The RMSD profiles in Figure 6 show that the control run has the largest error for all regions for both temperature and salinity. The background field from PEODAS has the next largest error, and the PEODAS analysis has the smallest error. This is exactly as expected. The difference between the temperature and salinity errors for the Control and for the PEODAS products is significant. This difference is as large as 0.4^{o}C and 0.1 psu at some locations. The variance of the Eulerian temperature is greatest at the depth of the thermocline, due to the vertical excursion of isotherms in response to external forcing and wave propagation. As a result, the profile of error for temperature shows a maximum about the thermocline depth (around 100m depth; Figure 6). The analysis error for temperature about the thermocline depth is around 1.3^{o}C for all regions considered, and the corresponding background error is typically 0.1^{o}C greater than that of the analysis. This indicates that the error in temperature near the thermocline typically grows by about 0.1^{o}C between each analysis. For the control, with no assimilation, the error in the thermocline region is around 1.7^{o}C. Therefore the assimilation of the data has reduced the error from around 1.7^{o}C to 1.3^{o}C in the thermocline region.

The variance of the salinity is greatest at the surface, owing to the impact and uncertainty of the surface fluxes. The analysis error for salinity ranges from about 0.2-0.25 psu at the surface, and the corresponding background field error is about 0.05 psu greater than that of the analysis (Figure 6).In the deeper ocean, below about 300 m, the error for both temperature and salinity decreases (Figure 6), owing to the reduced variability there. For those depths the ensemble spread is very small. Moreover, at those depths there is also little difference between the quality of the analysis and background fields.

Figure 7 shows the mean difference between the control and assimilation fields and the observations. These plots indicated the biases in the control and assimilation fields, and to what extent the assimilation has reduced these biases. Temperature biases (Figs 7a) are small at the surface due to a strong relaxation to the Reynolds SST. In the western part (WTIO) there is a warm bias in the control of around 0.5^{o}C, generally below 150m. This has been significantly reduced by the assimilation. In the eastern node of the dipole (SETIO) there is a negative bias of up to -1^{o}C at around 100m depth, and a warm bias of up to +1^{o}C below 200m. Both these biases have been significantly reduced by assimalation, indicating that the assimilation is able to make significant reductions to model bias.

Figures 7b show the biases in salinity. In general salinity is too fresh around the thermocline and too salty below around 200m. The salinity biases have an opposite impact on density than those of temperature, indicating some compensation. Again assimilation reduces the salinity biases in all regions considered.

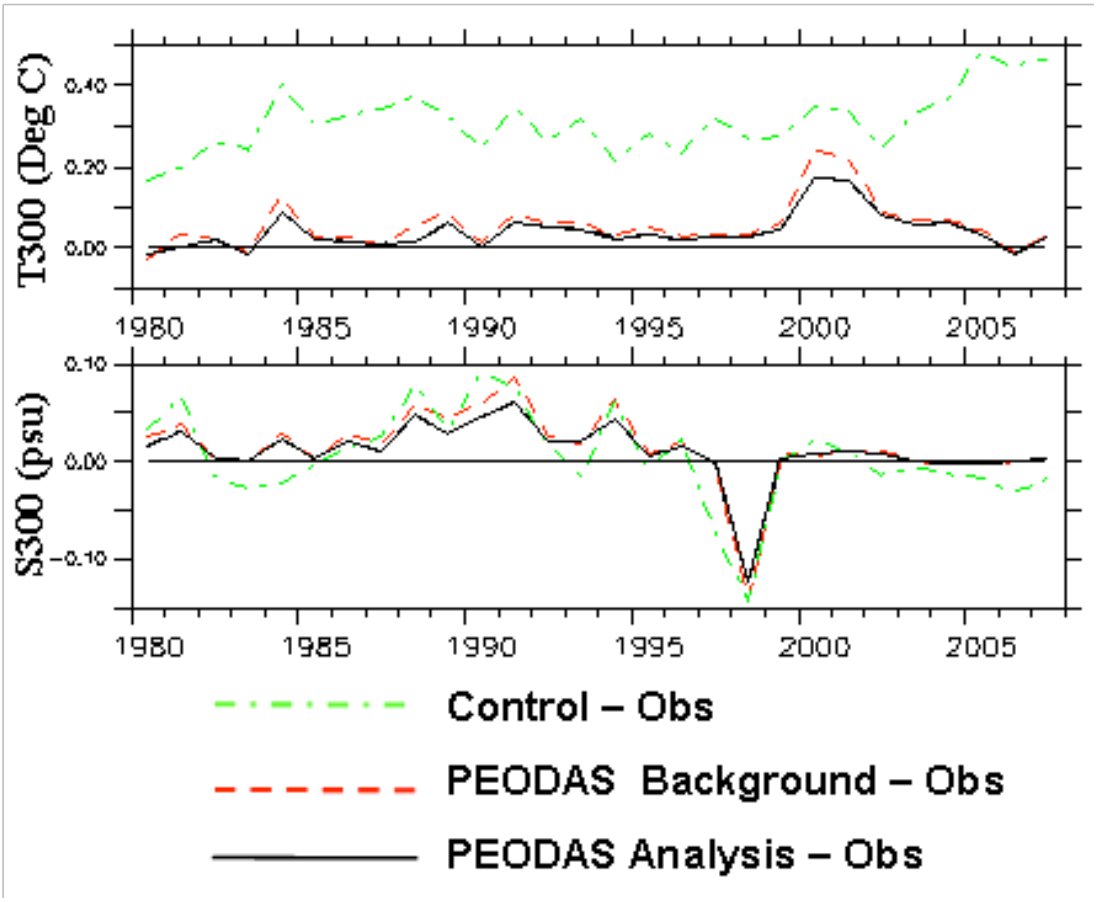

Fig. 8 Annual mean of T300 (°C) and S300 (psu) differences between model and observations for PEODAS analysis (solid), background (or first guess, dashed lines) and Control (dot-dashed lines). The model data are interpolated in observational locations in both time and space, and then averaging all profiles within top 300 meters for the region TRIO.

Figure 8 shows time series of area-averaged temperature and salinity, averaged over the top 300 m, across the tropical Indian Ocean. These time series are derived by interpolating the model products to observation locations, in both time and space, and then averaging all profiles for each region. The comparisons in Figure 8 show that the control is generally biased warm and salty. The assimilation seems to do a good job of eliminating these biases.

Figure 8 also shows significant multi-annual variability in the errors. For the control the error is around 0.3°C up until 2002, and then starts to increase to around 0.45°C. The reason for this increase is not known, but likely due to at least one of two reasons. Firstly from 2002 onwards NCEP2 rather ERA40 forcing is used and this could be a reflection of greater errors in the NCEP forcing compared to ERA40, but this has not been verified. Secondly the dramatic increase in observations due to Argo could lead to the control being assessed in areas where it was not being assessed before (due to no observations

pre-Argo) and these new areas are such that the control has larger errors in these regions; this has also not been verified. The assimilation fields do however show decreased error during the Argo period, indicating that the increased observation distribution is overcoming the increased error when observations are not assimilated.

The salinty field (fig 8b) shows mixed results, although in general the assimilation is better than the control, especially during the Argo period. During 1998 both assimilation and control show large errors, this is likely due to insufficient observations during this period. In fact, pre Argo there is some question over the validity of these diagnostics due to the small number of observations.

d. Comparison with other data

In this section we compare the PEODAS reanalysis with independent data, i.e. data that has not been assimilated. This provides an independent measure of the quality of the re-analysis. For comparison we also include the control simulation and the re-analysis from POAMA-1. Both altimeter data (Ducet et al., 2000) and sea surface currents from OSCAR (Bonjean and Lagerloef 2002) are used as independent data.

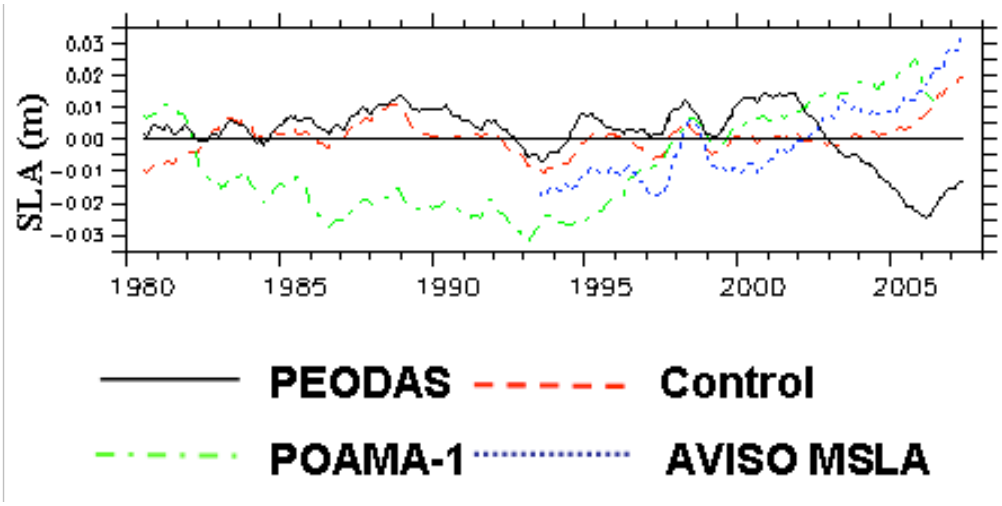

Fig 9. Monthly time series of steric sea level anomalies (m) for PEODAS (solid), Control (dash), dynamical high anomalies for POAMA-1 (dot-dash) and satellite altimeter MSLA (dot) data averaged in the region of TRIO. The monthly mean time series have been smoothed with 13-month running mean average.

Figure 9 shows a timeseris of sea level anomalies for PEODAS, Control, POAMA-1 and Altimeter. Altimeter data is only available from 1993 onwards. PEODAS and control are closest up until 2002, but then diverge. POAMA-1 is an outlier with strong decadal

changes. From 1993 -2002 PEODAS shows some agreement with altimeter, although there is an offset. From 2002 onwards there is considerable disagreement and PEODAS is an outlier showing a decreasing trend, while the other products show an increasing trend. Part of the reason for the discrepancy between PEODAS and the altimeter is because altimeter measures the total sea level, which includes the increasing sea level trend. The model based sea levels, are effectively dynamic height, ie departures from the global mean sea level, since rigid lid models do not represent the total sea level. While this partly explains the difference between PEODAS and the altimeter data, it does not explain the difference between PEODAS and POAMA-1 or Control. This is likely due to the increased Argo salinity data used in PEODAS but not used in POAMA-1 and the Control. Further work is necessary to understand these discrepancies. Balmaseda et al. (2007) indicated that such changes in sea level was associated with the evolution of the observing system, especially when the data assimilation is trying to correct the model systematic error. Indeed, from 2002 onwards the systematic error is warm and fresh (see Figure 8), the increased Argo temperaure and salinity data used in PEODAS le to an analysis which was colder and saltier both of which contribute to a decreased sea level. In order to avoid such discontinuities in the ocean analysis due to changing observing system, Balmaseda et al. (2007) introduced for ECMWF an *a priori* estimate of the bias term and found poorly observed regions, such as the equatorial Indian Ocean, are better represented if the bias term is introduced.

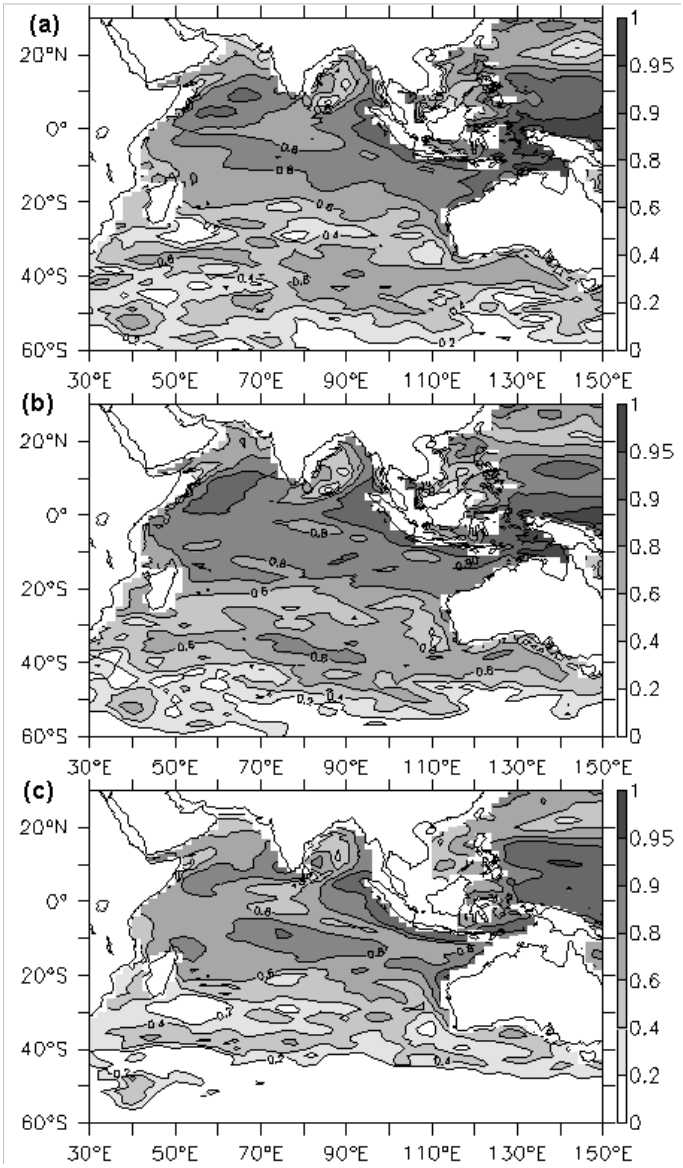

Figure 10. Correlations of SLA from AVISO maps with SLA from (a) PEODAS, (b) Control, and (c) POAMA-1 for the period 1993-2006.

Another way of assessing the re-analysis against sea level data is to calculate the spatial anomaly correlation between the re-analysis and the altimeter data. Figure 10 shows the anomaly correlation for PEODAS, control and POAMA-1 for the period 1993-2006. All three re-analysis show the highest correlation between 20S and 20N, particularly in the eastern half and near the Java-Sumatra coast and off the coast of western Australia. PEODAS has correlations up to 0.8 throughout within 20 degrees of the equator, increasing to over 0.9 off the Jave-Sumatra coast and off the Australian coast. The correlation patterns are very similar for the control run, although the control has slightly

higher correlations throughout the equatorial region and off the coast of Java-Sumatra. However, the correlation off the Western Australia coast only reaches 0.8, compared to 0.9 in PEODAS.

POAMA-1 shows the lowest correlation throught the equatorial region, although it does have higher skill than the control off the coast of Australia. Sea level depends both on salinity and temperature, the lower correlation for POAMA-1 is likely because the salinity field was not well represented in the POAMA-1reanalysis since no salinity corrections were made.

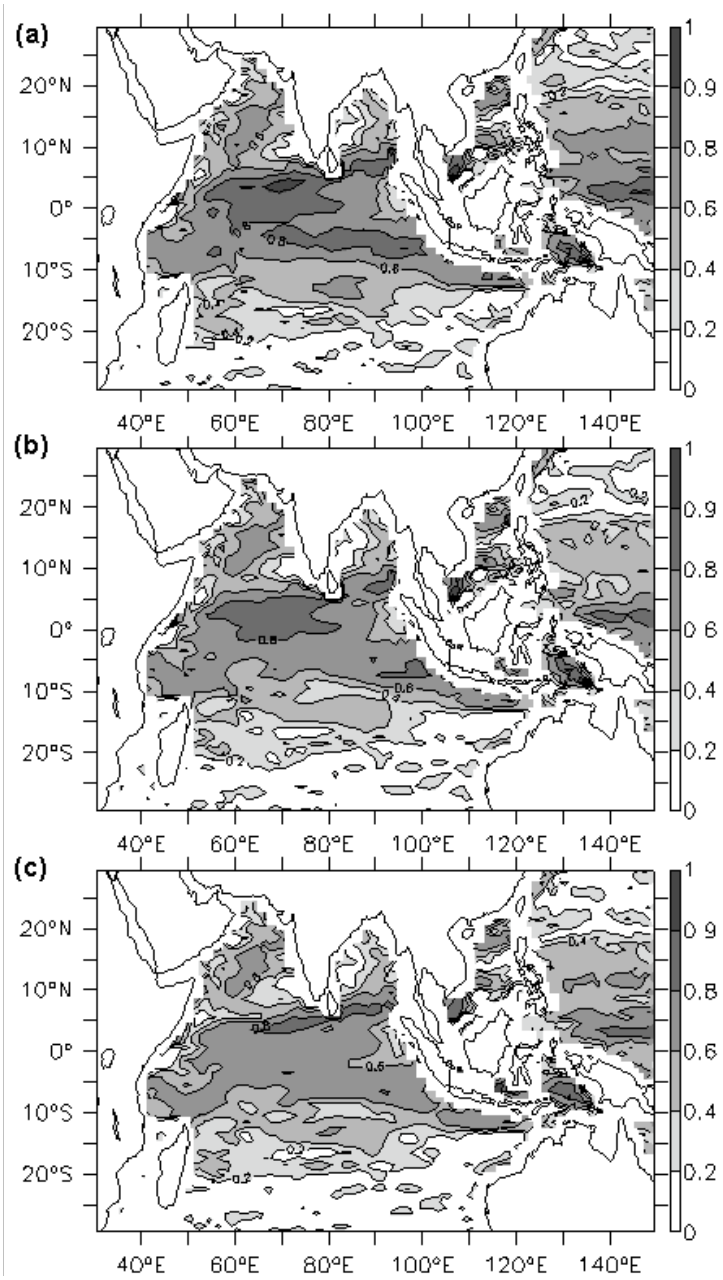

Figure 11. Correlations between the zonal surface velocity from OSCAR and (a) PEODAS, (b) Control, and (c) POAMA-1 for the period 1993-2006.

Another source of independent data is the OSCAR surface currents from http://www.oscar.noaa.gov/. The calculation of the OSCAR currents uses altimeter data. Figure 11 shows the correlation of the re-analysis currents with OSCAR for PEODAS, Control and POAMA-1. PEODAS shows the highest correlation in the equatorial region, with highest correlations peaking near the equator with values up to 0.9. Correlations are also high off the equator in the eastern Indian Ocean at around 5-15 degrees from the equator. In the far eastern equatorial region, near the Java Sumatra coast correlations are much lower, it is not clear if this is a problem in PEODAS or in the OSCAR currents, since the influence of coastal upwelling may not be well represented in OSCAR.

The control shows similar correlations to PEODAS, but slightly smaller, indicating that the assimilation of data as led to improved representation of the surface currents. However, the POAMA-1 correlations are the lowest, showing that the lack of increments to the velocity and/or salinity fields led to dynamical imbalances and to a deterioration of the surface currents.

4. Comparison with other international re-analysis

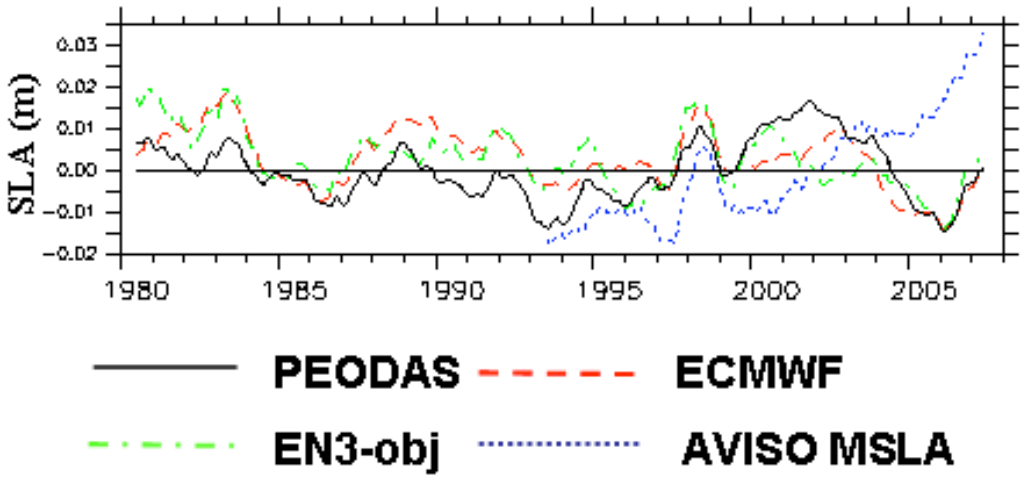

Figure12. Monthly time series of dynamical high anomalies (m) for PEODAS (solid), ECMWF (dash), EN3 objective analysis (dot-dash) and satellite altimeter MSLA (dot) data averaged in the region of TRIO. The monthly mean time series have been smoothed with 13-month running mean average.

Figure 12 compares the timeseries of tropical Indian Ocean sea level anomalies between PEODAS, ECMWF (Balmaseda et al., 2008), EN3 (objective analysis, available on http://hadobs.metoffice.com/en3/data/EN3_v2a/download_EN3_v2a.html , along with EN3 T/S quality-controlled profiles) and the AVISO altimeter sea level. ECMWF detrend their sea level anomaly before they incorporate it in their assimilation. The PEODAS, ECMWF and EN3 products are the dynamic sea levels estimated from the 0-500 m temperature and salinity analyses. Anomalies are calculated by removing their long term mean (1993-2007), respectively. Before the altimeter period (1993) there is reasonable agreement between ECMWF and EN3. There is also good agreement between PEODAS and ECWMF, except for an offset of approx 0.01m. The similar results are seen during altimeter period, and there is especially good agreement between the three re-analysis from 2003 onwards, presumably due to Argo. The biggest disagreement between PEODAS and both ECMWF and EN3 is during 2001 when PEODAS has a significant anomaly reaching up to 0.02m whereas ECMWF and EN03 are near neutral. The reasons for this is explained below. The altimeter values are dominated by a trend, which is not included in the re-analysis.

The disagreement between PEODAS, ECWMFand EN3 are likely caused by differences in the assimilation strategy. For PEODAS, there is no specific bias correction scheme, the model bias is reduced automatically as a result of data assimilation. Therefore bias would remain where there are insufficient observations, as is a case for tropical Indian Ocean before Argo. ECMWF, however, applied a sophisticated bias-correction scheme to correct the model background before assimilation. The bias correction algorithm consists of an *a priori* correction to temperature, salinity and pressure gradient, as well as a time dependent bias estimated on-line, which is applied only to pressure gradient (Balmaseda et al., 2008). The *a priori* bias is prescribed to be time-independent and estimated as the annual mean difference between a climatological model run and the Levitus climatology (Balmaseda et al., 2007). This leads to the ECMWF analysis having a bias correction, even during the period before Argo. As a result, the ECMWF agrees well with EN3, and PEODAS displays a near-constant offset to both of them before 2000 (Figure 12). During the period 2000-2003 when the observation system changed continuously (Figure 2), the biggest disagreement between PEODAS and both ECMWF and EN3 implies that the *a priori* bias correction did not work well for such a situation. From 2004 onwards, all the re-analysis agree because the Argo observation density became large enough that the assimilation strategy had a smaller impact.

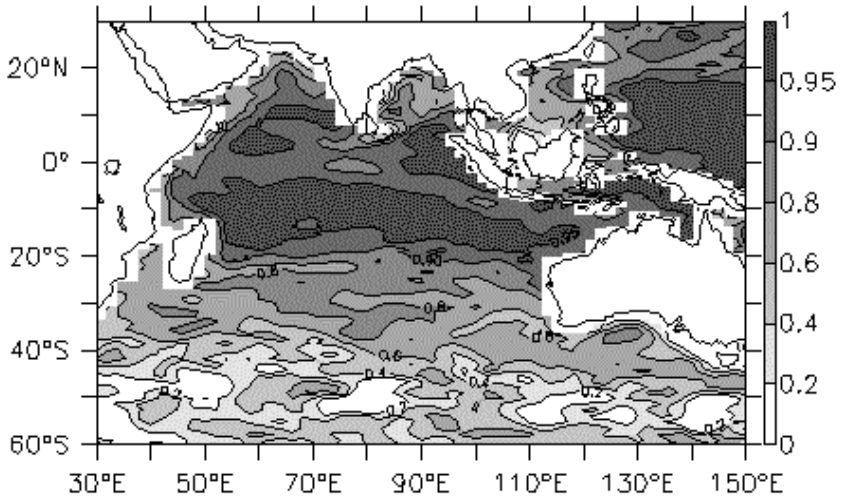

Fig 13. Correlations of SLA from AVISO maps with SLA from ECMWF.

The correlation between ECWMF and AVISO sea level is shown in figure 13. This shows overall stronger correlation that does the PEODAS correlation with AVISO (fig. 10). However, the AVISO sea level was assimilated in ECMWF but not PEODAS, so this result is expected. The correlation between ECWMF zonal current and OSCAR zonal current is shown in figure 14. The correlation is very similar to that obtained using PEODAS, perhaps slightly higher for ECMWF. This is also likely due to the use of AVISO sea level by ECMWF but not PEODAS. The OSCAR currents are also derived from altimeter measurements.

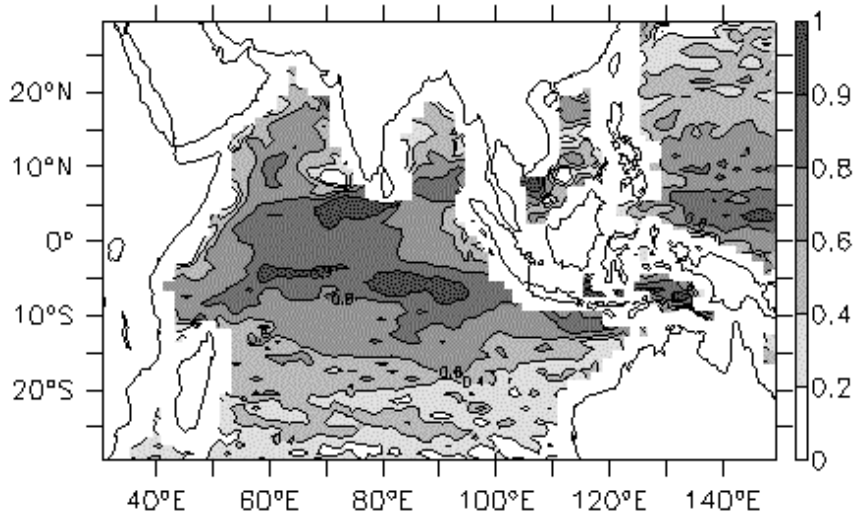

Fig. 14. Correlations between the zonal surface velocity from OSCAR and ECMWF.

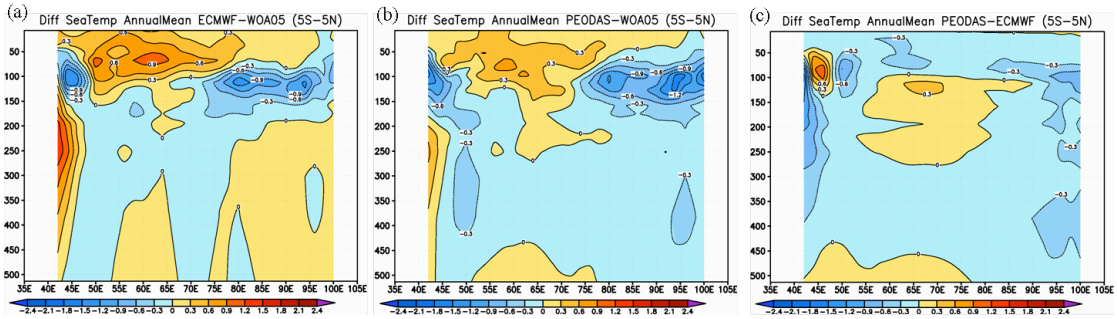

Figure 15. The vertical sections (upper 500 m) of differences between annual mean temparture along the equator (averaged over 5ºS-5ºN) in the Indian Ocean for the period 1982-2006 from ECMWF-WOA05 (a), PEODAS-WOA05 (b) and PEODAS-ECMWF (c). The color bar's interval is 0.3ºC.

The representation of the mean state by the PEODAS and ECWMF re-analysis can be assessed by comparing the re-analysis climatology with the climatology from WOA05. This is shown in figure 15 for a temperature section across the equatorial Indian Ocean. Fig 15a shows that the ECMWF re-analysis is too warm in the upper 150m in the west Indian Ocean by up to 0.9ºC and too cold in the east by up to -1ºC. This error is related to the Indian Ocean thermocline, i.e. it does not slope enough to the east. The PEODAS error shows a similar pattern to ECWMF, but the errors are relatively smaller.

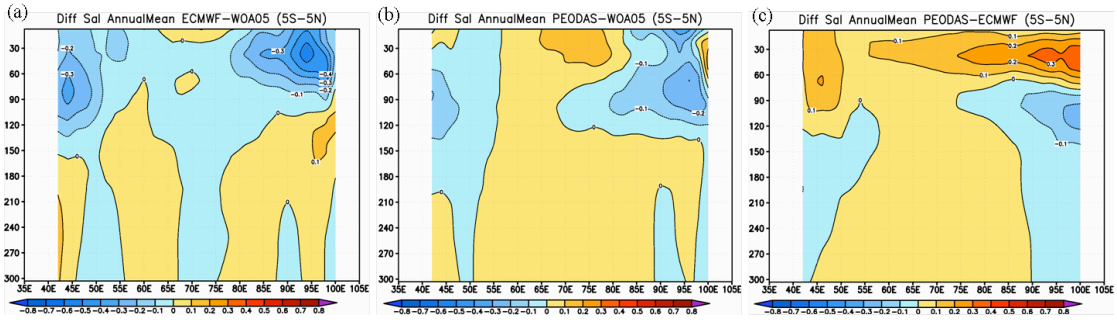

Figure 16. The vertical sections (upper 300 m) of differences between annual mean salinity along the equator (averaged over 5ºS-5ºN) in the In for the period 1982-2006 from ECMWF-WOA05 (a), PEODAS-WOA05 (b) and PEODAS-ECMWF (c). The color bar's interval is 0.1 psu.

The salinity differences relative to climatology are shown in figure 16. ECMWF re-analysis (fig 16a) is too fresh across the Indian Ocean,but particularly so in the east where it is up to 0.5ppt too fresh and in the far west where it is up to 0.3ppt too fresh. PEODAS

re-analysis (fig 16b) is also too fresh in the east and west, but by only about half of the amount of the ECMWF re-analysis. PEODAS is more salty than ECMWF in the upper 70m or so, particularly in the east where it is up to 0.3ppt more salty.

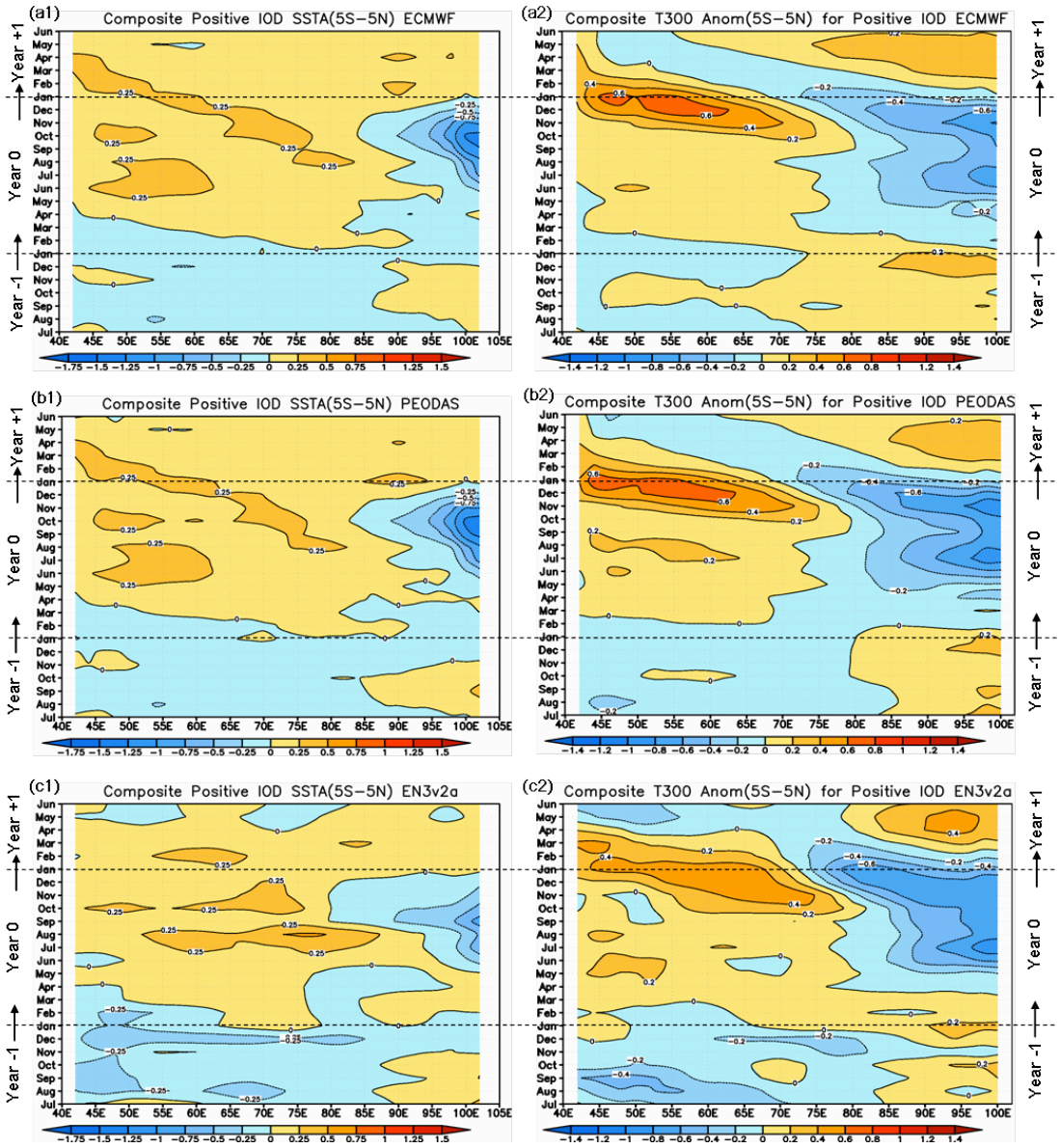

Figure 17: Comparison of the evolutions of SST anomalies (a1; interval 0.25ºC), T300 anomalies (a2; interval 0.2ºC) along the equator (5ºS-5ºN averaged) of the composite Positive IOD event (1980-2007) from Year -1 to Year +1 for the ECMWF. (b1)-(b2) same as (a1)-(a2), except for PEODAS. (c1)-(c2) same as (a1)-(a2), except for EN3. The selection criteria of Positive IOD events for the composite is defined as the monthly IOD indices (the difference of SSTanomaly between the western (50–70ºE, 10ºS–10ºN) and

the eastern (90–110ºE, 10–0ºS) Indian Ocean) meet or exceed +/- its standard deviation (0.48°C) for a period of at least 3 consecutive months. So, during the composite period 1980-2007, there are six positive IOD events selected for the composite, ie. 1982/83, 1983/84, 1994/95, 1997/98, 2003/04 and 2006/07.

The principal mode of variability in the Indian Ocean is the Indian Ocean dipole (IOD; Saji et al., 1999). We investigate how well the re-analysis represent the IOD by creating IOD composites relative to the peak of the IOD signal which occurs in Sep/October. The time evolution of SST and heat content anomalies along the equator for the IOD composite is shown in figure 17 for ECWMF, PEODAS and the EN3 re-analysis. Since both ECWMF and PEODAS are strongly relaxed to Reynolds SST, their IOD composite is almost identical. At the peak of the far east is cool by up to -1.25C and the west is warm by up to 0.25C. EN3 shows a similar pattern to PEODAS and ECMWF, but the cooling in the east is about half the amplitude. Figure 17 also shows the composite heat content of the upper 300m. Again ECMWF and PEODAS are similar, although PEODAS shows slightly more cooling in the east. They show a build up of heat in the west and a decrease of heat in the east. The peak in the east/west gradient occurs later than that in SST. The peak SST gradient occurs around September/October, whereas the peak in heat content occurs in December and then decays rapidly. EN3 shows a similar pattern to ECMWF and PEODAS, but the warming in the west is not as high.

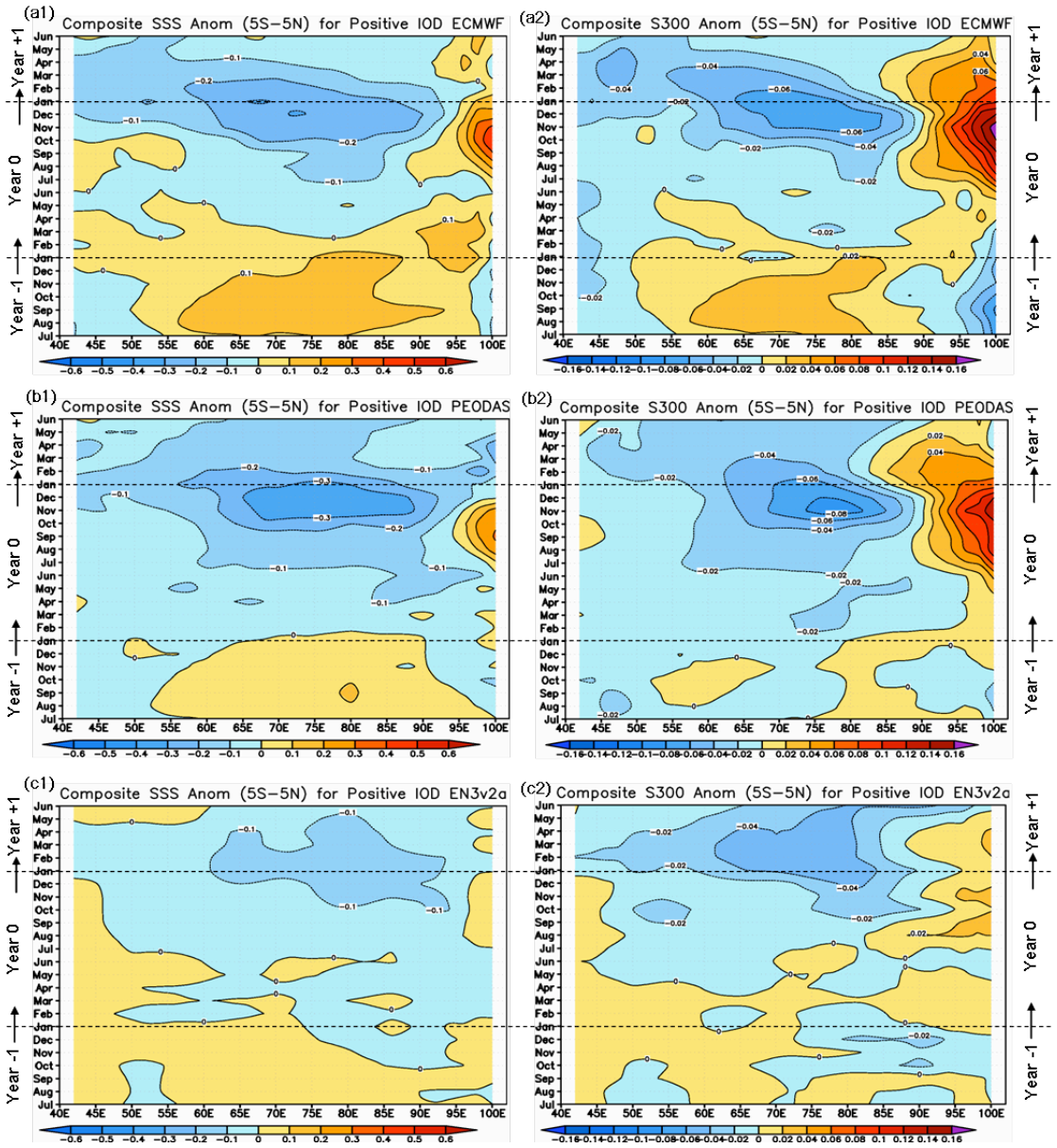

Figure 18: as Figure 17 but for SSS (interval: 0.1 ppt) and S300 (interval: 0.02 ppt) anomalies.

The composite surface salinity (SSS) and top 300m salt content (S300) composite IOD patterns are shown in figure 18. Both ECMWF and PEODAS show similar patterns. In the latter half of the IOD year an east/west gradient in salinity develops similar to that in heat content. The SSS salinity anomaly increases rapidly in the far east, peaking around October and reaches 0.3ppt in PEODAS and 0.4ppt in ECMWF. In the central Indian Ocean the salinity freshens, peaking at the end of the year at just below 0.3ppt in ECMWF and over 0.3ppt in PEODAS. Interestingly the re-analysis show a biennial cycle, with the opposite, although weaker signal the previous year. The EN3 anomalies, while in general of the same sign, are only about 1/3 of the size and not so significant.

The impact of the salinity anomalies on density seen in PEODAS and ECMWF are are comparable to those due to temperature anomalies, and have an impact on density in the same direction as those from temperature. However, the salinity anomalies in EN3 are small relative to the temperature anomalies.

ECMWF and PEODAS show similar salt content composite patterns, with a strong dipole pattern peaking towards the end of the year. However, PEODAS and ECMWF show different patterns the preceeding year. ECMWF shows a reverse signal the preceeding year, indicating a biennial mode, however PEODAS thoes not show this. EN3, like for SSS, has anomalies that are significantly weaker. This is likely due to insufficient salinity obseravations.

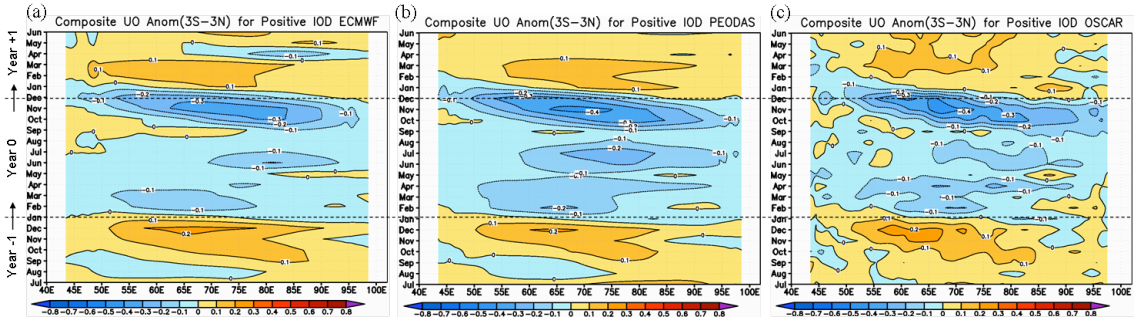

Figure 19: Comparison of the evolutions of UO anomalies (zonal surface velocity at 15m depth, interval: 0.1 m/s) along the equator (3ºS-3ºN averaged) for the composite Positive IOD event (1993-2007) from Year -1 to Year +1 between the ECMWF(a), PEODAS (b) and OSCAR (c), respectively. For the composite of zonal surface current, the exact positive IOD event are 1994/95, 1997/98, 2003/04 and 2006/07.

The IOD composite surface currents are shown in figure 19. PEODAS and ECWMF produced similar patterns shown to some extent a biennial signal. During the IOD year there are
westward currents peaking in the central Indian Ocean towards the end of the year. The peak value is 0.3 m/s for ECMWF and 0.4 m/s for PEODAS. The enhanced westward currents at the peak of the IOD, around October/ November, decay rapidly to be replaced by approx 0.1 m/s eastward currents by January of the following year. IOD composite of observed currents were obtained from the OSCAR current dataset (fig. 19c). This shows similar patterns to those in the re-analysis, and the magnitudes are well with those produced by PEODAS.

5. Representation of the Leeuwin Current in the re-analysis

The relatively coarse resolution of the analyses would appear to preclude an accurate depiction of the relatively narrow Leeuwin Current. Nonetheless, it is of interest to compare the depiction of the gross structure of the Leewuin Current from the different analyses to both act as a benchmark for comparison with future upgrades to the analyses

that may, for instance, be provided at higher resolution and to assess how well the large-scale driver of the Leeuwin Current is depicted in the available analyses. Figure 20 displays the meridional current along 32S, which is roughly the latitude of Fremantle and where the mean Leeuwin Current is strongest. A semblance of the Leeuwin Current is depicted in all of the analyses, with all of the analyses showing a peak current at the surface that is confined to the coast above about 300 m . The weakest current is depicted in the Control, followed by PEODAS and ECMWF. The strongest current is depicted in POAMA-1 followed by NCEP-GODAS. It is unclear why the current should be analysed weaker in PEODAS as compared to POAMA-1. However, in all of the analyses, the Leeuwin Current is depicted to be much broader and weaker than observed. For instance, Feng et al. (2003) who estimated the Leeuwin Current geostrophically from temperature observations, show the mean current to have maximum velocity of 30-50 cms-1 but to be confined to a narrow jet of less than 1 deg longitude width centered centered on about 115E (Fig. 21). However, the total mass transport, when integrated from the coast to 105E and from the surface to 300 m is realistic in POAMA-1. Hence, the large-scale driver of the Leeuwin Current appears to depicted in the analyses (i.e. the north-south steric gradient that drives the west-east geostrophic flow that converges on the coast and flows southward as the Leeuwn Current), thus resulting in a realistic large-scale north-south transport. But, the coarseness of the analyses prevents a realistic depiction of the narrowness and strength of the actual Leeuwin Current.

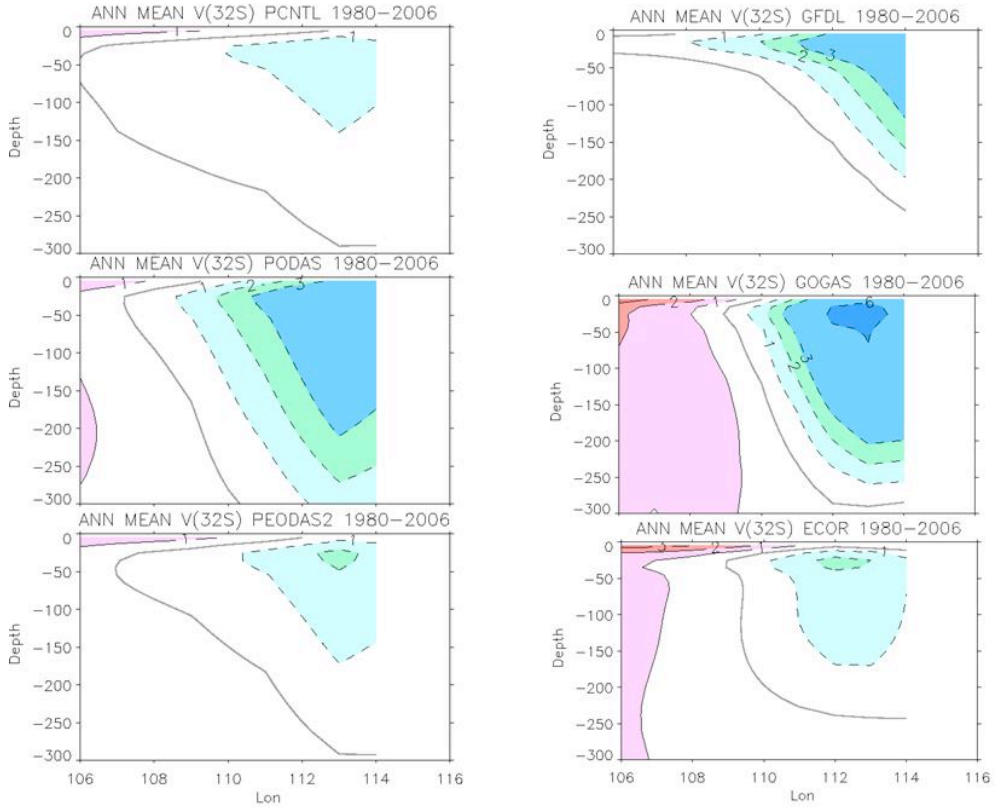

Figure 20. Mean meridional current along 32ºS from a) Control, b) POAMA-1, c) PEODAS, d) the GFDL analysis, e) NCEP GODAS, and f) ECMWF. Contour interval is 1 cm s^{-1}.

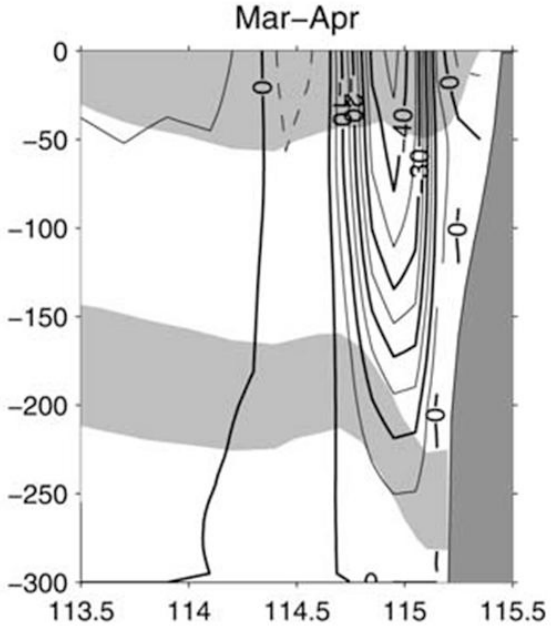

Figure 21 Mean geostrophic current along 32ºS from Feng et al. (2003). Contour interval is 5 cm s^{-1}.

The representation of the large-scale driver of the variation of the Leeuwin Current is demonstrated by correlating the analyzed sea level anomaly at every grid point with the observed times series of sea level at Fremantle. Feng et al (2003) howed that SLA at Fremantle is a good proxy for the volume transport of the Leeuwin Current, with the current varying in phase with sea level. Feng et al. also showed that correlation with the observed altimeter data is strongest on the northwest shelf (indicative of La Nina conditions) and with a strong narrow tongue of correlation extending down the west coast, that wraps around into the Bight. This feature is depicted well in PEODAS (Fig. 22a), with the correlation off of the northwest shelf being stronger than in the control (Fig. 22b). POAMA-1 (Fig. 22c) does not capture the positive correlation that wraps around into the bight. This lack of a signal in the bight in POAMA-1 nicely reflects the benefits of using flow-dependent covariances in PEODAS verus the prescribed static covariances in POAMA-1: there are probably a limited number of near-coastal bservations in the bight but the PEODAS covariance cosntrains the analysis in a realistic fashion. It is interesting to note that EN3 also does not well represent the behavior into

the bight. Finally, PEODAS exhbits stronger correlation on the north west shelf than ECMWF (Fig. 22e), despite the fact that ECMWF assimilates altimeter data..

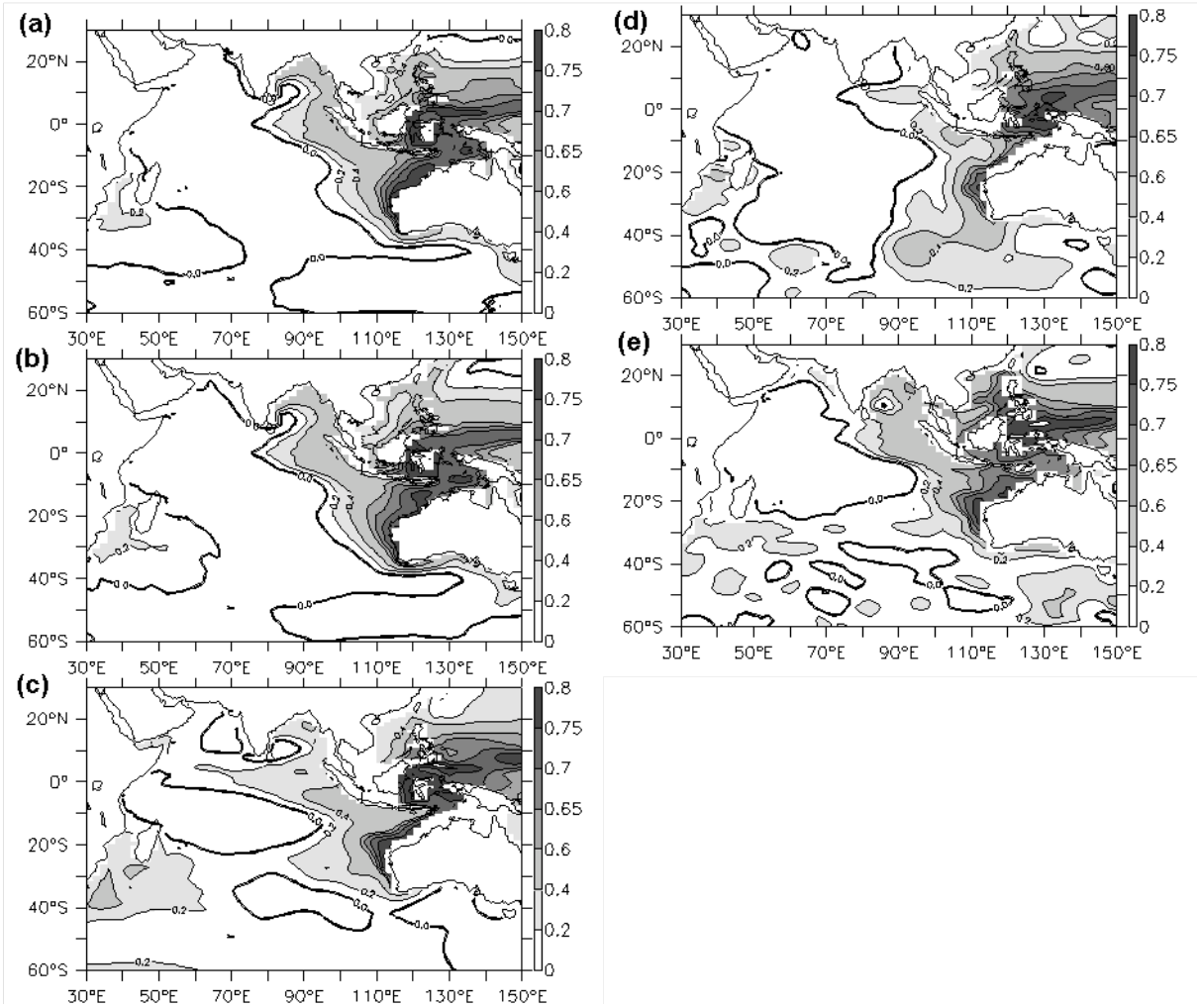

Figure 22 Correlation of analyzed SLA with observed SLA at Fremantle for a) PEODAS, b) Control, c) POAMA-1, d) EN3, and e) ECMWF. Contour interval is 0.1 above 0.4 with additional 0.2 contour shown.

An impact of in-situ observations on the depiction of the Leeuwin Current is estimated by comparing the analyzed current using the available analyses from 1980-2006 and from just 2003-2006 when ample Argo data are available. Figure 23 shows the mean current from the Control, POAMA-1 and PEODAS. As no observations are used in the control, the similarity of the current for the two periods for the Control suggests that there was little interannual variability of the large-scale forcing of the current in the two periods. Hence, the much stronger currents in both PEODAS and POAMA-1 (local increase by a

factor of two) reflects the benefit of increased observations in the late period. However, to fully exploit the Argo observations will require a higher resolution ocean model with which to do the assimilation.

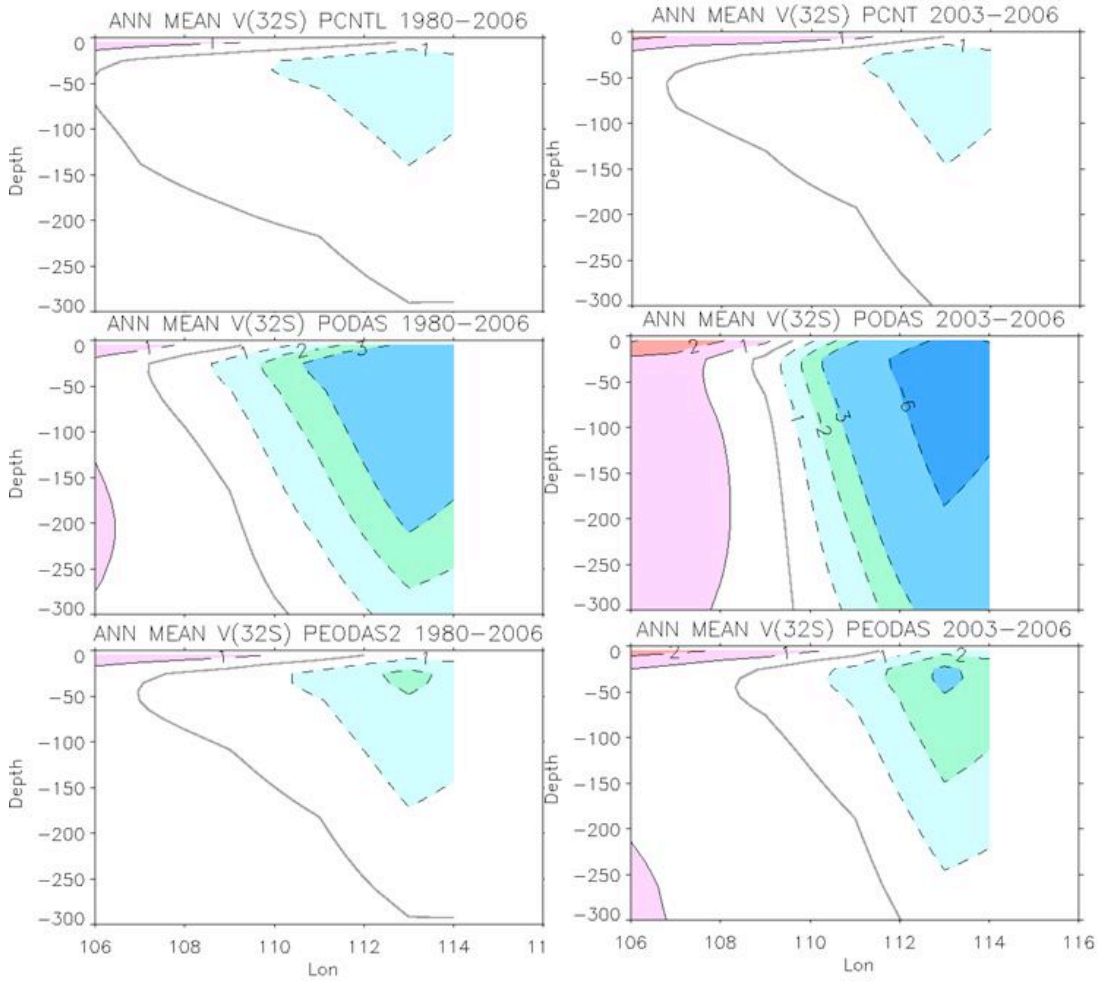

Figure 23 Mean meridional current along 32ºS from a) Control, b) POAMA-1, and c) PEODAS for 1980-2006 (rh panels) and 2003-2006 (lh panels). Contour interval is 1 cm s^{-1}.

7. Summary and conclusions

The a new ocean ocean data assimilation systems based on a simplified Ensemble Kalman Filter approach has been developed for the POAMA dynamical seasonal prediction system. Compared with the old POAMA-1 assimilation system, PEDOAS incoporates several improvements including:

Assimilation salinity in addition to temperature

Three-dimensional multi-variate assimilation with increments for every model variable

Background statistics based on time evolving ensemble based covariances

A re-analysis with PEODAS was carried out from 1980 to 2006. Over this period, and particularly in the Indian Ocean, there have been dramatic changes in density of ocean observations. Before Argo (~2002) there were large gaps in temperature observations and salinty observations were very limited. Argo changed all this, with an almost even distribution of tempertaure and salinity observations, at least once per month, over most of the Indian Ocean.

PEODAS was assessed by comparing it against dependent and independent observations. The results show that when compared against dependent observations (ie observations that have been assimilated) both background and analysis fields show significant improvements compared to the control run with no data assimilation. This is expected for the analysis fields but not guaranteed for the model background fields, which are forecasts from the previous assimilation.

Difficulty was found when comparing PEODAS with altimeter sea level data, mainly because PEODAS and the control produce effective dynamic sea level which does not incorporate the mean sea level height, whereas altimeter is the total sea level and includes mean sea level trends. The introduction of Argo is also believed to have introduce artificial trends, as also found by Balmaseda et al (2007). Pre Argo there were insufficient salinity data to adequately correct the model bias, however with the introduction of Argo there was sufficient data to correct the model bias, introducing a significant change in re-analysis pre and post introduction of Argo.

PEODAS was also compared with OSCAR surface current. For both altimeter and OSCAR surface zonal currents PEODAS outperformed both the control and the old POAMA-1 assimilation. In fact the old POAMA-1 assimilation was worst, probably because it did neither assimilate salinity, nor did it generated salinity corrections based on the temperature corrections. This led to dynamical imbalances in POAMA-1 assimilation which in turn led to a deterioration of the salinity through spurious dynamical processes.

The PEODAS re-analysis was also compared with the ECMWF system-3 re-analysis. ECMWF and PEODAS re-analysis were found to be very similar. Both re-analysis used the same surface forcing and in situ observations, but the assimilation approaches are

very different. ECMWF did produce slightly high correlations with altimeter sea level and OSCAR surface currents, but ECMWF did also assimilate the altimeter data.

The ability of PEODAS, ECMWF and EN3 re-analysis to represent Indian Ocean variability was investigated by comparing composites of the evolution of the Indian Ocean dipole. In general the composites for temperature and heat content were very similar for all three re-analysis. This was not the case for salinity and salt content. While PEODAS and ECWMF did agree reasonable well, the EN3 re-analysis had significantly reduced anomalies due to lack of observations. When there are no observations EN3 essentially resorts back to climatology, which is used as the background field. EN3 shows that there are no significant salinity anomalies associated with IOD. However, both PEODAS and ECMWF show that SSS and salt content anomalies developed associated with the IOD that are comparable with the SST and heat content anomalies. These results suggest that salinity anomalies may be important for accurate simulation of the IOD.

A resemblance of the Leeuwin Current is depicted in all of the analyses, with all of the analyses showing a peak current at the surface that is confined to the coast above about 300 m . The weakest current is depicted in the Control, followed by PEODAS and ECMWF. The strongest current is depicted in POAMA-1 followed by NCEP-GODAS. It is unclear why the current should be analysed weaker in PEODAS as compared to POAMA-1. Much stronger currents in both PEODAS and POAMA-1 (local increase by a factor of two) during the Argo period compared to pre-Argo reflects the benefit of increased observations in the late period. However, to fully exploit the Argo observations will require a higher resolution ocean model with which to do the assimilation.

Acknowledgements

This work was supported by the Western Australia Marine Science Initiative.The observations used in the re-analysis were supplied by the U. K. Met Office via http://hadobs.metoffice.com/en3/. The ECWMF re-analysis data was kindly provided by Dr Magdalena Balmaseda.

References

Alves, O., G. Wang, G. Zhong, N. Smith, F. Tzeitkin, G. Warren, A. Schiller, J. S. Godfrey, G. Meyers, 2003: POAMA: Bureau of Meteorology operational coupled model forecast system. National Drought Forum, Brisbane, 15/16 April.

Alves, O., C. Robert, 2005: Tropical Pacific Ocean model error covariances from Mont Carlo simulations. *Quart. J. Roy. Meteor. Soc.*, 131, 3643-3658.

Balmaseda, M.A., D. Dee, A. Vidard and D. L. T. Anderson, 2007: A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans, Q. J. R. Meteorol. Soc. 133, 167-179.

Balmaseda, M.A., A. Vidard, and D.L.T. Anderson, 2008: The ECMWF Ocean Analysis System: ORA-S3. *Mon. Wea. Rev.*, **136**, 3018–3034.

Bonjean, F., G. S. E. Lagerloef, 2002: Diagnostic model and analysis of surface currents in the tropical Pacific Ocean. *Journal of Physical Oceanography,* **32**, 2938-2954.

Burgers, G., M.A. Balmaseda, F.C. Vossepoel, G.J. van Oldenborgh, P.J. van Leeuwen, 2002: Balanced Ocean-Data Assimilation near the Equator. *J. Phys. Oceanogr.*, **32**, 2509–2519.

Chen, D., A. J. Busalacchi and L. M. Rothstein, 1994: The Roles of Vertical Mixing, Solar-Radiation, and Wind Stress in a Model Simulation of the Sea-Surface Temperature Seasonal Cycle in the Tropical Pacific-Ocean. *J. Geophy. Res.*, **99**(C10): 20345-20359.

Ducet, N., P.-Y. Le Traon, G. Reverdin, 2000: Global highresolutionmapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. *J. Geophys. Res.,* **105,** 19 477–19 498.

Evensen, G.,1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Mont Carlo methods to forecast error statistics. J. Geophys. Res.,99(C5), 10143-10162.

Feng, M., G. Meyers, A. Pearce, and S. Wijffels, 2003: Annual and interannual variations of the Leeuwin Current at 32ºS, J. Geophys. Res., 108(C11), 3355, doi:10.1029/2002JC001763.

Houtekamer, P. L., H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* 126, 796–811.

Ingleby, B., M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles — Historical and real-time data. *Journal of Marine Systems,* **65**, 158-175.

McIntosh, P. C., M. J. Pook, J. S. Risbey, S. N. Lisson, M. Rebbeck, 2007: Seasonal climate forecasts for agriculture: Towards better understanding and value. *Field Crops Research*, **104**, 130-138.

Oke, P. R., A. Schiller, G. A. Griffin, G. B. Brassington, 2005: Ensemble data assimilation for an eddy-resolving ocean model. *Quart. J. Roy. Meteor. Soc.*, 131, 3301-3311.

Oke, P. R., G. B. Brassington, D. A. Griffin, A. Schiller, 2008: The Bluelink Ocean Data Assimilation System (BODAS), *Ocean Modelling, 21*, 46-70, doi:10.1016/j.ocemod. 2007.11.002.

Pacanowski, R. C., 1995: MOM2 documentation user's guide and reference manual. Version 1.0, GFDL Tech. Rep. 3, 232 pp.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, T. Yamagatta, 1999: A dipole mode in the tropical Indian Ocean. *Nature,* **401**, 360-363.

Schiller A., J. S. Godfrey, P. C. McIntosh, N. R. Smith, O. Alves, G. Wang, R. Fiedler, 2002: A new version of the Australian Community Ocean Model for seasonal climate prediction. CSIRO Marine Research Report No. 240.

Schiller, A., and J.S. Godfrey, 2003: Indian Ocean Intraseasonal Variability in an Ocean General Circulation Model. , **16**, 21–39.

Smith, N. R., J. E. Blomley, G. Meyers, 1991: A univariate statistical interpolation scheme for subsurface thermal analysis in tropical oceans. *Progress in Oceanography,* **28**, 219-256.

Spillman, C. M., O. Alves, 2008: Dynamical seasonal prediction of summer sea surface temperatures in the Great Barrier Reef. *Coral Reefs,* **28**, 197-206.

Sprintall, J., A. Gordon, R. Mutugudde, D. Susanto, 2000: A semi-annual Indian Ocean forced Kelvin wave observed in the Indonesian Seas. *J. Geophys. Res.,***105**, 17217-17230.

Troccoli, A. and Kallberg, P., 2004: Precipitation correction in the ERA-40 reanalysis, ERA-40 Project Report Series, 13.

Uppala, S., and coauthors, 2005: The ERA-40 Reanalysis. *Q. J. R. Meteorol. Soc*. 131, 2961-3012.

Yin, Y., O. Alves, P. R. Oke, 2009: An ensemble ocean data assimilation system for seasonal prediction. in preparation for *Mon. Wea. Rev.*

Zhao, M, H. H. Hendon, 2009: Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. *Quarterly Journal of the Royal Meteorological Society,* **135**, 337-352.