# Impact of improved assimilation of temperature and salinity for coupled model seasonal forecasts

**Mei Zhao · Harry H. Hendon · Oscar Alves · Yonghong Yin**

**Abstract** We assess the impact of improved ocean initial conditions for predicting El Niño-Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) using the Bureau of Meteorology's Predictive Ocean Atmosphere Model for Australia (POAMA) coupled seasonal prediction model for the period 1982–2006. The new ocean initial conditions are provided by an ensemble-based analysis system that assimilates subsurface temperatures and salinity and which is a clear improvement over the previous optimal interpolation system which used static error covariances and was univariate (temperature only). Hindcasts using the new ocean initial conditions have better skill at predicting sea surface temperature (SST) variations associated with ENSO than do the hindcasts initialized with the old ocean analyses. The improvement derives from better prediction of subsurface temperatures and the largest improvements come during ENSO–IOD neutral years. We show that improved prediction of the Niño3.4 SST index derives from improved initial depiction of the thermocline and halocline in the equatorial Pacific but as lead time increases the improved depiction of the initial salinity field in the western Pacific become more important. Improved ocean initial conditions do not translate into improved skill for predicting the IOD but we do see an improvement in the prediction of subsurface temperatures in the Indian Ocean (IO). This result reflects that the coupling between subsurface and surface temperature variations is weaker in the IO than in the Pacific, but coupled model errors may also be limiting predictive skill in the IO.

**Keywords** Seasonal forecast · Prediction skill of ENSO · Data assimilation

M. Zhao (✉) · H. H. Hendon · O. Alves · Y. Yin
Centre for Australian Weather and Climate Research (CAWCR), GPO Box 1289, Melbourne, VIC 3001, Australia
e-mail: m.zhao@bom.gov.au

## 1 Introduction

The capability to make global seasonal climate forecasts with coupled climate models derives primarily from the ability to predict tropical sea surface temperature (SST) variations associated with the El Niño-Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) (Cane et al. 1986; Latif et al. 1993; Kirtman et al. 1997; Ji et al. 1998; Wang et al. 2002; Saha et al. 2006; Luo et al. 2008; Zhao and Hendon 2009; Stockdale et al. 2011). The skill and ultimate utility of the seasonal forecasts is limited by various components of the forecast system, such as the quality of the initial condition, errors in the coupled model, the method of generating the forecast ensemble and the calibration strategy. Here we consider the impact of improved ocean initial conditions, which is the primary source of long lead predictability.

Ocean initial conditions are typically generated by driving the ocean component of the coupled forecast model with atmospheric fluxes of heat, momentum and fresh water in order to provide the first guess to estimate the ocean state. The first guess will have substantial error resulting from errors in the model and forcing fields. This error can be reduced by combining the first guess with oceanic observations of subsurface temperatures, currents and salinity via a data assimilation procedure. Although the tropical oceans historically have not been well observed, the Tropical Ocean-Global Atmosphere (TOGA) Program that was

initiated in the 1980 resulted in the Tropical Atmosphere-Ocean (TAO) array of moored buoys in the equatorial Pacific, a surface drifting buoy program, an island and coastal tide gauge network, and a volunteer observing ship network of expendable bathythermograph measurements (McPhaden et al. 1998), thereby providing good coverage of the subsurface at least in the tropical Pacific which is the heart of the most predictable component of the climate system (ENSO). Globally, the situation improved markedly in the 2000's with the beginning of the deployment of a global array of approximately 3,000 free-drifting profiling floats, known as the ARGO ocean profiling network (http://www.argodatamgt.org/). The ocean observation from ARGO provide real-time measurements of surface winds, SST, subsurface temperature and salinity profiles in the upper 2,000 m ocean, sea level, and ocean velocity. The new data sets play a core role in helping to better understand ENSO and IOD events, to develop improved coupled models, and to provide improved ocean initial conditions via data assimilation. However, to best use these new observations to produce improved ocean initial conditions requires improved methods of data assimilation.

Data assimilation can significantly reduce the model analysis errors and improve both the depiction of the mean state and the analysis of anomalies of the first guess, especially noticeable in the tropical oceans (Zhang et al. 2007; Behringer 2007; Balmaseda et al. 2008; Yin et al. 2011). Assimilated initial states can have a favourable impact on the prediction skill of seasonal forecasts of SST (Ji and Leetmaa 1997; Ji et al. 1998; Schneider et al. 1999; Wang et al. 2002; Alves et al. 2003; Balmaseda et al. 2009; Balmaseda and Anderson 2009; Stockdale et al. 2011). A wide variety of assimilation techniques are used by forecast centres that are making routine coupled model seasonal forecasts including univariate (temperature only) or multivariate (e.g., temperature and salinity) optimal interpolation (OI) systems, three or four-dimension variational (3Dvar or 4Dvar) methods, and the use of variants of the Ensemble Kalman Filter (EnKF); for a review of these systems see Balmaseda et al. (2009).

The first generation data assimilation system that provided initial conditions for coupled model seasonal climate forecast mainly focused on assimilating temperature because subsurface temperature was believed to play the dominant role for providing predictive skill of ENSO. Salinity was typically ignored or simply constrained by climatological relationships. This was motivated in part by the paucity of salinity observations prior to the advent of ARGO. However, even in the absence of abundant salinity data, a proper treatment of salinity in the assimilation process is crucial to properly depict the density field.

In 2002, the Australian Bureau of Meteorology implemented an ocean data analysis system for the Predictive

Ocean Atmosphere Model for Australia (POAMA) seasonal forecast model. This ocean analysis system was based on a univariate (temperature only) optimum interpolation scheme (Alves et al. 2003; http://poama.bom.gov.au/). We refer to it as POAMA Optimum Interpolation (POI). In POI, surface salinity was constrained by the climatological surface flux of fresh water (E-P) but subsurface salinity was otherwise not incremented during the assimilation cycle. A new ensemble-based ocean analysis system called the POAMA Ensemble Ocean Data Assimilation System (PEODAS) has been developed and was implemented in 2010. This second generation data assimilation systems use more sophisticated error covariances (e.g., flow dependent) and includes the assimilation of salinity that is more dynamically and thermodynamically in balance with temperature through the use of covariances of temperature and salinity (Yin et al. 2011).

In a earlier study (Zhao et al. 2013), we highlighted some profound impacts of the changes in the depiction of the mean state T (temperature) and S (salinity) fields using PEODAS compared to POI for ENSO behaviour in forecasts with the POAMA model. We especially highlighted the impacts on the predicted ENSO characteristics due to the erroneous depiction of mean salinity in the Pacific halocline in POI and that those impacts developed quickly (1–2 months) and were long lived because the initial mean-state salinity errors set off a coupled response. This current paper is a continuation of that study but will focus on the impacts on forecast skill of ENSO and the IOD due to the improved depiction of initial anomalies for both temperature and salinity.

A brief description of the assimilation systems and experimental set up is provided in Sect. 2, and a comparison between the old and new ocean analyses is presented in Sect. 3. In Sect. 4, we will compare prediction skills for ENSO/IOD and also subsurface temperatures and salinity in the two hindcast sets that are initialized with the old and new ocean initial conditions. We will argue in Sect. 5 that the improvement in tropical SST skill for ENSO and the IOD derives from the improvement in predicting the subsurface, which in turns stems from an improved depiction of the initial temperature and salinity fields in the upper ocean. Discussion and conclusion follow in Sect. 6.

## 2 The POAMA coupled models and experimental design

### 2.1 Ensemble generation

For this study we explore hindcast skill for seasonal forecasts from the first of each month during January 1982 through December 2006 from two versions of POAMA. POAMA is based on coupled ocean–atmosphere general

**Table 1** Details of ocean assimilation and ensemble perturbations from the two hindcast experiments

| Experiment name | Ocean initial condition | Assimilated observations | Updated variables | Perturbation of initial condition | |
|---|---|---|---|---|---|
| | | | | Ocean | Atmos |
| V1_POI | POI | Temperature | Temperature and currents (geostrophic adjustment) | No | Yes |
| V1_PEO | PEODAS | Temperature and salinity | Temperature, salinity and currents using ensemble cross covariances | Yes | No |

circulation models. The atmospheric model is based on BAM3 (Zhong et al. 2001) and is run at T47L17 resolution (Alves et al. 2003). The ocean model component is version 2 of the Australia Community Ocean Model (ACOM2; Schiller et al. 2002), which is a global configuration of version 2 of the Modular Ocean Model (MOM2; Paca-nowski 1995). The ocean model has 2° zonal resolution and 1° latitudinal resolution that increases to 0.5° in the Tropics. It has 25 vertical levels (0–5,000 m). More detail about the model configuration can be found in Alves et al. (2003), Zhong et al. (2005), Zhao and Hendon (2009) and Wang et al. (2011).

The first set of hindcasts are from POAMA version 1.5b (called V1_POI hereafter), which was the operational seasonal forecast system at BoM from 2002 to 2010. V1_POI was initialized using the POI ocean assimilation. Atmospheric/land surface initial conditions are provided by the Atmosphere–Land Initialization (ALI) scheme (Hudson and Alves 2007; Hudson et al. 2011). The ensemble of forecasts was generated by perturbing the atmospheric initial conditions by successively picking the analysis from a 6 h earlier period (i.e., the tenth member was initialized 2.5 days earlier than the first member). A single ocean initial condition was provided by the POI system for the 1st of each month.

The second version of POAMA is version P2.4c (here-after V1_PEO), which is one of the three versions of the models that forms part of the new operational version of P24. The atmosphere and ocean models are the same version as used for V1_POI but the initialization of the fore-casts is different. The ocean initial conditions are provided by the new PEODAS assimilation system. And, in contrast to V1_POI, the ensemble of initial conditions was gener-ated by using 9 additional ocean initial conditions as pro-vided by PEODAS, while only a single atmospheric initial condition on the 1st of each month, as provided by ALI, was used to initialize the atmosphere. Table 1 lists the major different set up details about two experiments.

All results are based on a 10-member ensemble inte-grated out to 9 months. We adopt the terminology that a lead time of 1 month means a hindcast initialized on, for instance, 1 January that is valid for the month of January. Hindcast anomalies are formed relative to the hindcast model climatology, which is a function of start month and lead time and is unique to each model version. In this fashion, the mean bias from the hindcasts is removed.

## 2.2 Climate indices and verification data

Our focus is on performance of predicting the coupled climate in the Indo-Pacific. We define the region of the tropical Pacific Ocean (PO) as 120°E–90°W, 10°N–10°S, and the region of the tropical Indian Ocean (IO) as 50°E–110°E, 10°N–10°S. Results are also analysed using indices of SST relevant to the ENSO and IOD. ENSO in the Pacific is monitored with the Niño3.4 index (SST anomalies averaged over 170°W–120°W, 5°N–5°S). For the IOD, we use the SST anomaly over the western pole (WIO, 50°–70°E, 10°N–10°S) or eastern pole (EIO, 90°–110°E, 0°–10°S) of the tropical IO (Saji et al. 1999). For verification of SST, we use the NOAA Optimum Interpolation (OI) SST V2 data (Reynolds et al. 2002) for the period 1982–2006. For the subsurface, we verify using the anal-yses from PEODAS, POI or the ENACT (EN3) analysis, which is based on a quality-controlled database (Ingleby and Huddleston 2007).

## 3 Ocean analysis systems

### 3.1 The POI and PEODAS analysis systems

The same ocean model component of POAMA is used for both the coupled model forecasts and the ocean assimila-tion. More details about the ocean model and reanalysis configuration can be found in Yin et al. (2011).

The POI assimilation system used to create ocean initial conditions for the V1_POI forecasts is based on a univar-iate optimal interpolation system and only assimilates in situ temperatures from the top 500 m of the ocean (Smith et al. 1991); velocity fields are updated using the geostrophic relation similar to Burgers et al. (2002). Sub-surface salinity is allowed to evolve during the forecast but is not updated or constrained during the assimilation, resulting in a dynamically unbalanced state after the tem-perature is updated in the assimilation. An accumulated effect of this imbalance is the development of an unreal-istic deep overturning circulations and a systematic bias in

the halocline of the central-western Pacific (Yin et al. 2011). In POI, surface salinity is constrained in the assimilation cycle by the imposed surface flux of freshwater and with an additional relaxation to the climatology from World Ocean Atlas 2001 (WOA2001, Stephens et al. 2002; Boyer et al. 2002) during the assimilation. However, the relaxation is strong (3 days) and so limits depiction of interannual variability at the surface (this impact will become apparent in Sect. 5).

The PEODAS ocean assimilation system, which is used to initialize V1_PEO forecasts, is an ensemble-based data assimilation system that is computationally affordable and easier to implement than a four-dimensional variational method. It also yields an ensemble of ocean initial conditions from which to generate an ensemble of seasonal forecasts. We expect the ensemble of ocean states to span the actual uncertainty in the estimate of the initial conditions (Yin et al. 2011). In contrast to POI, both temperature and salinity profiles are assimilated and, importantly, temperature, salinity, and velocity fields are all updated at all model levels using flow dependent, 3-dimensional error cross-covariances. In addition, surface salinity is forced by imposed observed freshwater fluxes with a slow relaxation to climatology (1-year relaxation time). Therefore, PEODAS provides more realistic depiction of salinity (both mean and variability), and with better dynamical balance between temperature and salinity and much larger (more realistic) interannual variability of surface salinity than POI. More details about those two systems can be found in Yin et al. (2011) and Zhao et al. (2013).

### 3.2 Comparison of the ocean analyses

The improvement in the depiction of the upper ocean salinity, currents and temperature variations from PEODAS compared to POI is summarized below. In Fig. 1, we plot the temporal anomaly correlation coefficient (ACC) of the vertically averaged temperature and salinity in the upper 300 m (T300 and S300) between the ENACT analysis based on the EN3 quality-controlled database (Ingleby and Huddleston 2007) and the PEODAS and POI analyses respectively. Although the EN3 data used in the ENACT analysis also went into both the POI and PEODAS analyses, we compare to it here as "independent" because the ENACT analysis of EN3 is not based on a model.

For T300 (Fig. 1a), PEODAS has a broad band (10°N–10°S) of high correlation along the entire tropical Pacific, whereas the POI has only relatively narrow bands of high correlation confined to the western boundary (i.e., where equatorial Rossby waves make a large contribution to the variability) and in the eastern Pacific cold tongue (i.e., where equatorial Kelvin waves make a large contribution to the variability). Higher correlation seen for the PEODAS

analysis is also seen in the eastern equatorial IO, where Rossby waves stemming from reflected Kelvin waves associated with IOD variability are prominent, and in the south west tropical IO, where Rossby wave associated with both the IOD and ENSO contribute to the variability. The overall correlation in the IO is not as high as for the Pacific, which suggests that the predictability of climate in the IO region will be lower than for the Pacific. This lower agreement of the assimilated state with the independent analysis of the observations in the IO partly stems from the paucity of in situ data especially prior to the ARGO-ERA. Interestingly both systems have relatively high correlation along the west coast of Australia, which reflects strong variability driven by the oceanic teleconnection of ENSO through the Indonesian throughflow and which is the basis for extended range prediction of the Leeuwin current (Hendon and Wang 2009).

Not surprisingly, PEODAS everywhere has higher correlation for vertically integrated salinity to 300 m (S300) than does POI since there is no update for salinity in the POI (Fig. 1d, e). In the Pacific, there are especially large improvements in the western and central Pacific from 0°S to 10°S, which coincides with where the largest increment to temperature is made (Yin et al. 2011) so that neglecting the compensating salinity increments for salinity in POI leads to large errors in the analysis of salinity there (Zhao et al. 2013). The maximum mean state differences of salinity between two analysis systems is also in this region (not shown), but which can be seen in Fig. 5a from Yin et al. (2011) and Fig. 2b from Zhao et al. (2013). The improved depiction of subsurface salinity variations in this region using PEODAS reflects that salinity largely co-varies with temperature, so that vertical displacement of the thermocline results in large variations of salinity in the halocline, which are explicitly accounted for by PEODAS. Regions of large improvement of PEODAS over POI are also seen in the north and southeast IO. Along the equator, the improvement in the IO is not nearly as dramatic as we see in the Pacific. This is due to a much weaker halocline in the tropical IO compared to the western Pacific so that incrementing temperature without adjusting salinity in POI has less impact on salinity there. As for temperature, the overall correlation of S300 in the IO is smaller than in the Pacific.

Yin et al. (2011) argued that PEODAS significantly improved the analysis of zonal surface velocity and sea-level (see their Figs. 13, 14), especially over the tropical oceans where the main modes of climate variability on interannual time scale (e.g., ENSO and the IOD) have large impact. Because POI did not assimilate subsurface salinity, a density imbalance in the subsurface developed as the salinity field was not dynamically constrained when the temperature was incremented, which ultimately degraded the analysis of ocean currents and sea-level anomaly. Yin
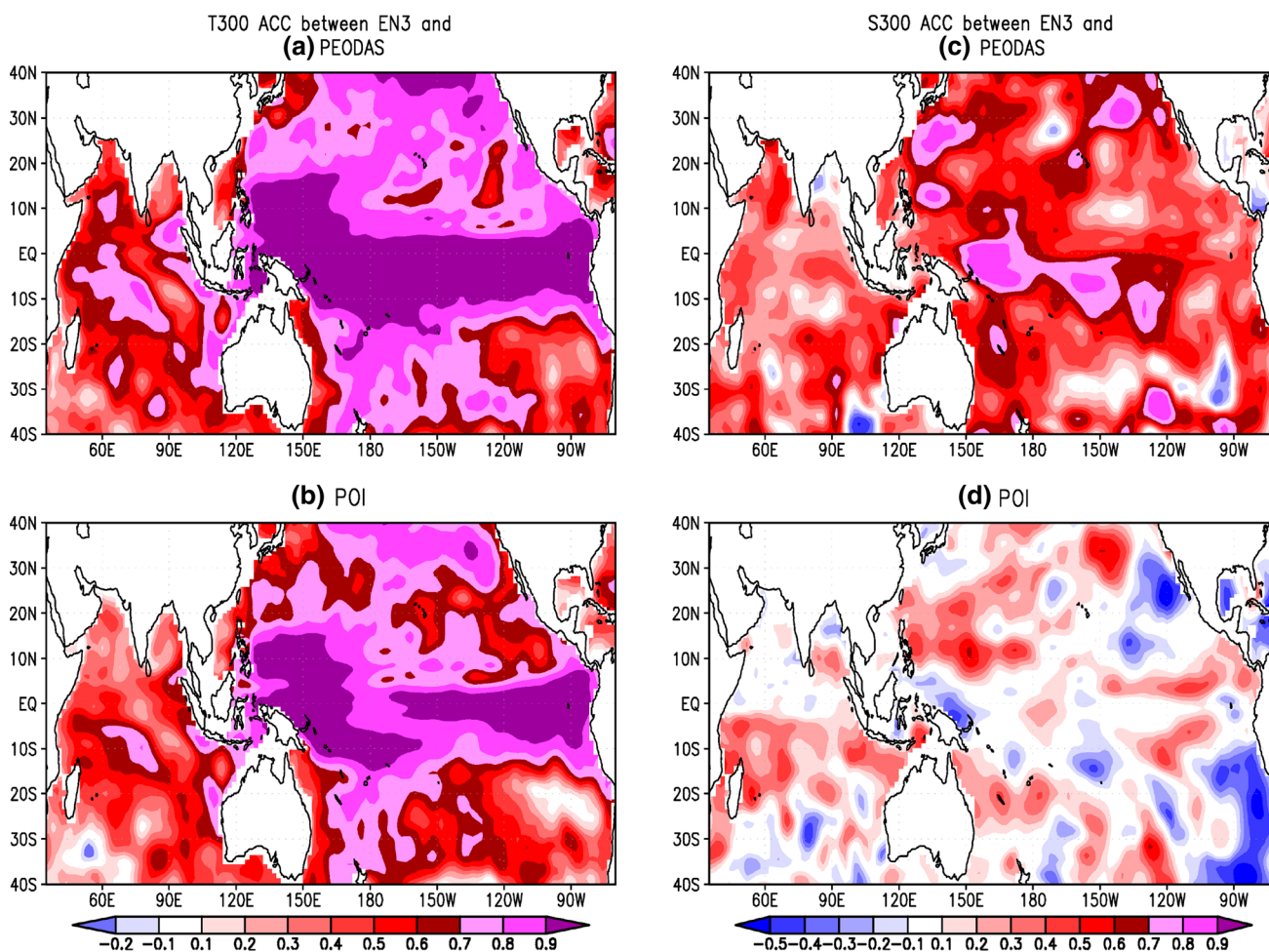
**Fig. 1** Temporal correlation coefficient for anomalies of T300 from EN3 analyses with analyses from **a** PEODAS, and **b** POI. **c**, **d** For S300. Correlation is based on analyses on the 1st of each month for the period 1982–2006

et al. (2011) showed that the analysis of surface currents and sea level provided by POI was even worse than a control assimilation of PEODAS that used no in situ data at all (but was forced by the same observed surface momentum, heat and freshwater fluxes). Note that in the PEODAS control assimilation, the covariance of salinity with temperature was naturally enforced by the model's variability but this co-variation is corrupted in POI when temperature is incremented but salinity is not.

In summary, there is a marked improvement in the analysis of both temperature and salinity throughout the tropical Pacific and IO from PEODAS compared to POI. The improvement is both in the depiction of the variability, for which ENSO/IOD prediction will depend, and in the depiction of the mean state, which affects the simulated coupled climate variability in the model (e.g., Zhao et al. 2013). The focus of the rest of the paper is on assessment of forecast improvement that derives from the improved depiction of temperature and salinity variations in the initial conditions.
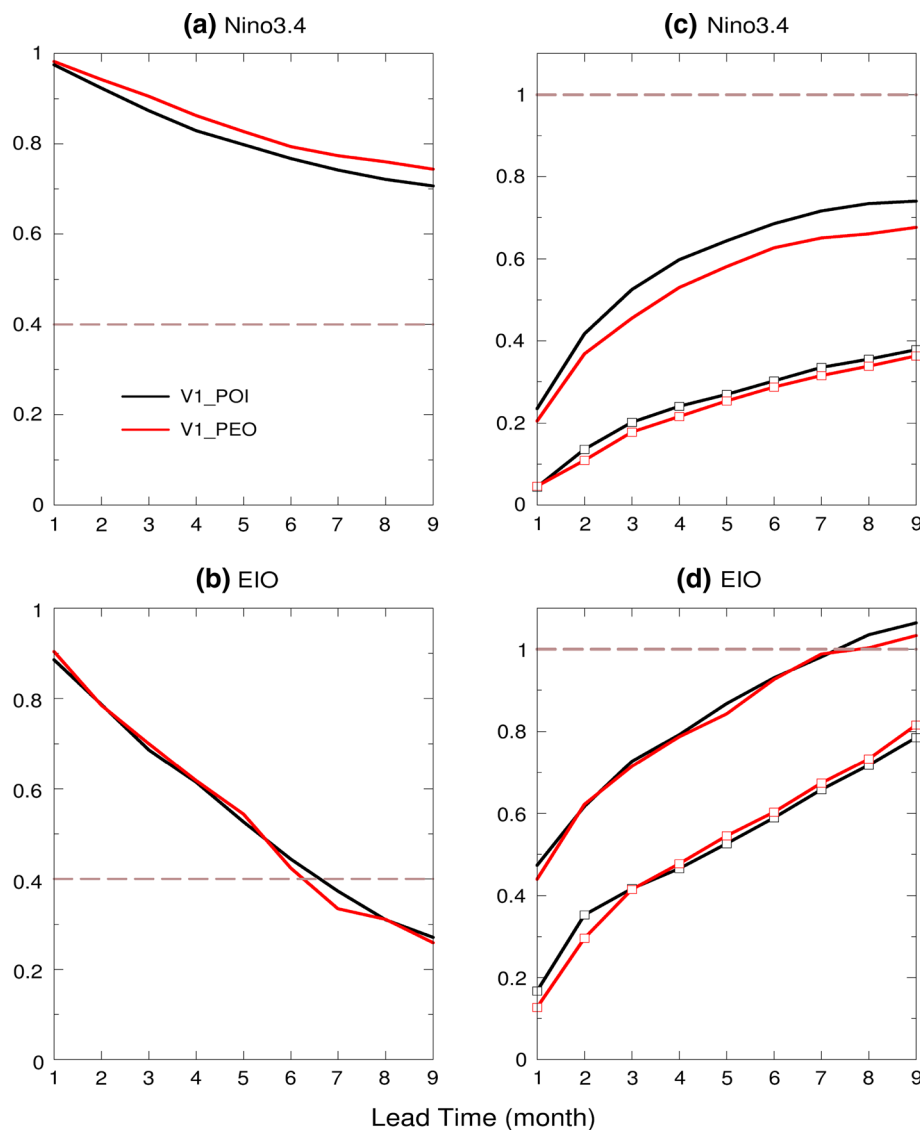
## 4 Comparison of seasonal forecast skills

By comparing metrics of forecast skill from the hindcasts of V1_POI and V1_PEO, we can gauge the impact of the new ocean analyses on the forecast skill. We do need to keep in mind, however, that the ensemble generation strategies are different from the two systems and so it too can affect forecast performance, which we will highlight as appropriate.

### 4.1 Niño3.4 and EIO SST skill

The two key questions to be addressed are "Do improved ocean initial conditions lead to improved skill in predicting ENSO and IOD?", and, if so, "Which aspects of the improved ocean initial conditions matter most?" We begin by looking at the traditional forecast metrics of ENSO prediction: correlation and root mean square error (RMSE) of the Niño3.4 SST index using the ensemble mean forecast. Using all start months, the temporal anomaly

correlation (ACC, Fig. 2a) and normalized root mean square error (NRMSE, i.e., the root mean square error normalized by the SD of the observed index; Fig. 2c), clearly shows that the skill from V1_PEO has improved over V1_POI. For lead times longer than 2 months, the improvement in skill amounts to about 1 month increase in lead time for the equivalent correlation or NRMSE in V1_POI. We will show in Sect. 4.2 that this increase in SST skill is accompanied by (or stems from) an increased in forecast skill for subsurface temperature, which is the heart of ENSO predictability.

We also assess forecast skill for the IO using the SST index in the EIO region (Fig. 2b, d). Somewhat disappointingly, there is no skill improvement from using the V1_PEO with the improved PEODAS ocean initial conditions compared to V1_POI using the older POI ocean initial conditions. The skill over the WIO region is very

similar in the two models as well, but skill in the WIO drops less sharply with lead time than for EIO (figure not shown). For instance, the correlation skill for the WIO remains above 0.4 for all lead times to 9 months (as does the skill of the Niño3.4 index), but the skill for the EIO drops quickly and is below 0.4 by 6 months. Zhao and Hendon (2009) have interpreted this more rapid drop off in skill in EIO compared to WIO to derive from the stronger relationship of WIO SST to ENSO and the fact that a lot of the variability of EIO SST is driven by local air-sea fluxes (Hendon 2003) that are not a source of long lead predictability. The overall lack of improvement in skill in the IO probably is a reflection of the large model errors and biases (e.g., Zhao and Hendon 2009), which prevent the coupled model from exploiting any improvement provided by the new ocean initial conditions, and, as we will discuss below, the relatively weak relationship between variations in

subsurface and surface temperatures as compared to the Pacific so that any profound improvement in the initial depiction of the subsurface variations in the IO are not manifested by improvements in predictions for the surface.

The improvement in skill seen for Niño3.4 may also be contributed to by different ensemble generation strategies between the two systems, which might change the spread-error relationship in the forecast. The forecast spread (normalized by the observed SD of the indices, NSPREAD) is shown in Fig. 2c (for Niño3.4) and Fig. 2d (for EIO). The difference in spread between the old and new systems is negligible, indicating that the reduced NRMSE and increased correlation for the Niño3.4 index is not a result of an improved spread-error relationship. We note, however, that both forecast systems are under-dispersive (i.e., the NRMSE is greater than the NSPREAD) and so the POAMA prediction system could further benefit from an improved ensemble generation strategy that increases forecast spread while decreasing forecast error.

## 4.2 Subsurface skill and relationship with surface skill

We now look at improvements in prediction skill of subsurface temperatures and salinity with the intent to determine if this is the source of improved prediction skill of Niño3.4 index seen in Fig. 2. We firstly present spatial correlation (SCOR) of T300 anomalies over the PO regions (Fig. 3a). Here, the spatial correlation of forecast anomalies is computed over each region (i.e., we get a single score for each forecast) and then average this correlation over all forecasts. For this analysis, we verify against the anomalies from the PEODAS analysis. We also tried verifying against the analyses from POI and from EN3 because there is no common default "observation" data for the subsurface as there is for SST. Although the subsurface skill is dependant on which verification data we used, we show the verification here using the PEODAS analysis, noting that the results are very similar if we use the EN3 analysis. We feel some what justified in using PEODAS because verification against the initial conditions is a generally adopted method by other prediction studies and the PEODAS analyses are regarded as good quality (Xue et al. 2011).

As for Niño3.4 index, the forecast skill of subsurface temperature using spatial correlation over the PO region is much improved in V1_PEO over V1_POI (Fig. 3a). An even more dramatic improvement in forecast skill is seen for spatial correlation of subsurface salinity in the PO region (Fig. 3c), which is not surprising given the huge improvement in the initial condition of salinity as provided by PEODAS over that from POI. We also compute the spatial correlation for SST anomalies over the PO and IO regions (Fig. 3e, f), which is complimentary to the temporal correlation for Niño3.4 that we displayed in Fig. 2a.

In both approaches, V1_PEO is better than V1_POI for lead times beyond 2 months.

Subsurface forecast skill of V1_PEO over V1_POI in the IO region (Fig. 3b) is also increased, but, which we noted above for correlation for EIO, is not reflected in improved SST forecasts in the IO (Fig. 3f). Subsurface forecast skill, as for surface skill, drops off much more rapidly in the IO than in the Pacific, which might also help explain why the improved prediction of the subsurface is not evidenced by improved prediction at the surface: forecast skill in the subsurface is not high enough to impact prediction at the surface. This rapid drop off in skill compared to the Pacific also reflects that predictability of the coupled climate is much lower for the tropical IO than for the Pacific (e.g., Zhao and Hendon 2009) but could also reflect greater model error in the IO compared to the Pacific. Predictive skill of subsurface salinity in the IO is also much higher from V1_PEO than from V1_POI (Fig. 3d), but, interestingly, the predictive skill of subsurface salinity in V1_POI is higher in the IO than in the Pacific. We think this reflects a poorer representation of salinity using POI in the Pacific than in the IO (Fig. 1d), which comes about for the reasons we discussed previously: POI develops a more spurious depiction of salinity in the Pacific because more increments are made to temperature in the Pacific, which due to the lack of a corresponding increment on salinity in POI in the presence of a sharper halocline, results in a more degraded depiction of salinity there compared to the IO.

## 4.3 Relationship with ENSO strength

One possible source of the skill improvement using V1_PEO for prediction of Niño3.4 SST index is from improved prediction of the large ENSO events. For instance, Jin et al. (2008), using the DETEMER hindcast set, found that strong ENSO's were better predicted than neutral-ENSOs, which presumably results from increased signal to noise when ENSO variability is large. Because we expect that the subsurface has a longer and stronger impact on the surface during strong ENSO events, we may then anticipate that improved subsurface initial conditions might provide the most impact for prediction of the strong ENSO events.

To address the relationship of forecast skill to ENSO amplitude, we present the time series of anomaly spatial correlation for predicted SST in the PO domain from the V1_PEO forecasts (Fig. 4a) along with the time series of observed Niño3.4 index so that the magnitude of the ENSO variability can be assessed. High and long lasting skill in predicting SST over the tropical Pacific domain occurs during strong ENSO events (El Niño and La Niña),
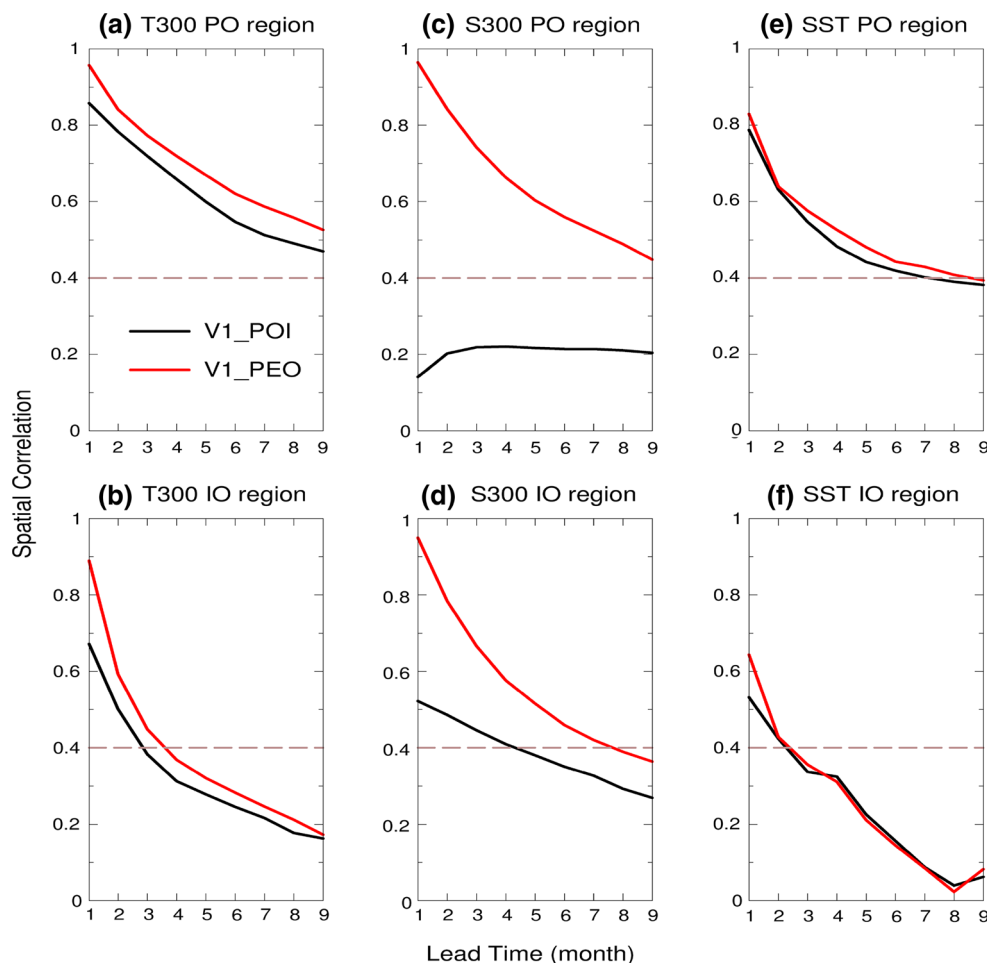
confirming the results of Jin et al. (2008). However, the improvement in skill compared to V1_POI, which we show in Fig. 4b as the difference in the anomaly spatial correlation between V1_PEO and V1_POI, is seen to mainly occur during periods of low-amplitude ENSO when the skill for SST is relatively low (Fig. 4a).

We quantify this tendency for greater increase in skill to occur for lower amplitude ENSO events in Table 2, where we tabulate the number of occurrences when the increase in SCOR exceeds 0.1 from Fig. 4b and bin according to the amplitude of the observed Niño3.4 index. The distribution of points is similar for other thresholds of skill improvement. From Table 2 we see that at all lead times the greatest number of points where an improvement is achieved occurs for low amplitude ENSO episodes. We seldom see any forecast improvement for ENSO events (i.e. when Niño3.4 amplitude exceeds 1 °C). We also see that the total number of improved forecasts for low amplitude ENSO episodes increases with lead time. Although overall skill is highest during large ENSO events, the greatest improvement in skill especially at longer lead times comes about from improved predictions during low-amplitude ENSO episodes.

From an examination of the difference in skill of T300 (Fig. 4c), we see that the times of high improvement in SST skill (Fig. 4b) correspond to times of high improvement in subsurface skill. This is quantified by computing the correlation of the difference in skill for SST (Fig. 4b) with the difference in skill for T300 (Fig. 4c), which we show in Fig. 5 (black curve) as a function of forecast lead time. At short lead time, the SST skill difference is not strongly associated with the skill differences in T300, but the correlation becomes strong after about 3 month lead time. This lagged development of the association of skill difference at the surface and in the subsurface is consistent with the notion that the Pacific climate is a coupled system and that the adjustment time of the subsurface equatorial Pacific Ocean is about 3 month (e.g., initial difference in the subsurface in the western Pacific will take 2–3 months to propagate into to the central and eastern Pacific where they impact SST, see Figs. 5a and 6e in Zhao et al. 2013, and we will discuss more in Sect. 5).

We also compute the difference in skill for upper ocean salinity (S300), which is shown in Fig. 4d. Not surprisingly, V1_PEO forecasts of S300 have systematically higher skill relative to V1_POI and this difference is

**Fig. 4 a** Spatial correlation (SCOR) over the PO region from the V1_PEO hindcasts of SST as a function of forecast start time (each month 1982–2006) and lead time in months (*left hand y-axis*). The SCOR is shown at each lead time for a constant verification time (e.g., the SCOR shown for lead time 1 month at 1 January 1982 used the forecast that was initialized on 1 January 1982 and the SCOR shown for lead 9 at 1 January 1982 used the forecast that was initialized on 1 May 1981). **b** The difference in SCOR between predictions from V1_PEO and V1_POI for SST anomalies over the PO region. **c** As in **b** except for the difference in SCOR for T300. **d** As in **b** except for the difference in SCOR for S300. The heavy *black curve* is the time series of the observed Niño3.4 index (units are °C on *right hand y-axis*). The *right hand y-axis* extends back to the initial time, which for **a** signifies that the SCOR was computed using the PEODAS analyses of SST against the Reynolds OI v2 analyses. In **b**–**d** the SCOR at the initial time is computed using the initial anomalies from PEODAS and POI analyses
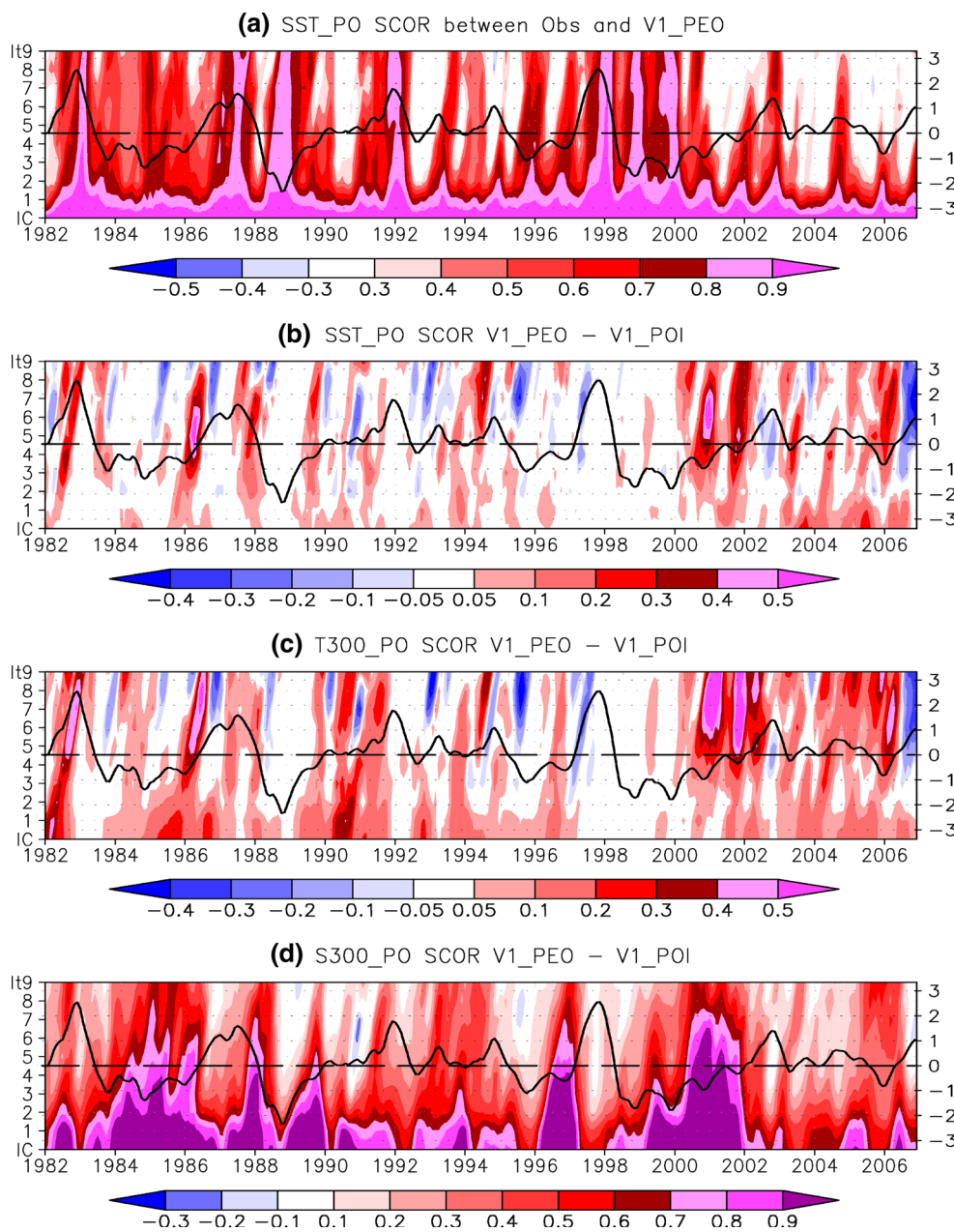


**Table 2** The number of forecasts when the difference of SCOR (V1_PEO minus V1_POI) for SST in PO exceeds 0.1, binned according to the amplitude of the observed Niño3.4 index at each lead time

| Niño3.4 | Lt1 | Lt2 | Lt3 | Lt4 | Lt5 | Lt6 | Lt7 | Lt8 | Lt9 |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0–0.5   | 32  | 20  | 31  | 39  | 38  | 35  | 36  | 46  | 44  |
| 0.5–1   | 11  | 13  | 24  | 23  | 33  | 24  | 21  | 19  | 22  |
| 1–1.5   | 0   | 2   | 4   | 7   | 6   | 8   | 5   | 8   | 5   |
| 1.5–2   | 0   | 0   | 2   | 1   | 1   | 2   | 3   | 3   | 2   |
| 2–2.5   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 1   |
| >2.5    | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 1   |

immediately apparent at the shortest lead time. However, the largest skill improvement at longer lead time also appears to occur when ENSO amplitude is weaker. Furthermore, there is a strong relationship between the variation of improved skill in SST and improved skill S300 (Fig. 5), that also does not appear until lead time 2–3 months.

Although the largest improvement in skill as measured using correlation comes from the low-amplitude ENSO periods, this does not mean that the forecast errors as measured by RMSE are most improved then as well. To address this, we show a similar set of plots as in Fig. 4a, b

but for RMSE (Fig. 6). Clearly, the RMSE of the SST forecast increases with the lead time (Fig. 6a) but it also increases with ENSO amplitude (i.e., the largest absolute errors occur with the largest amplitude events).
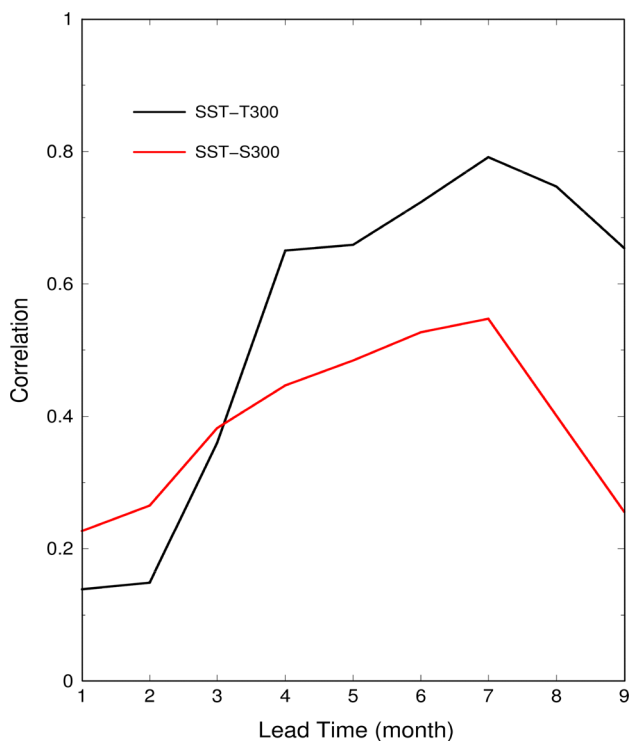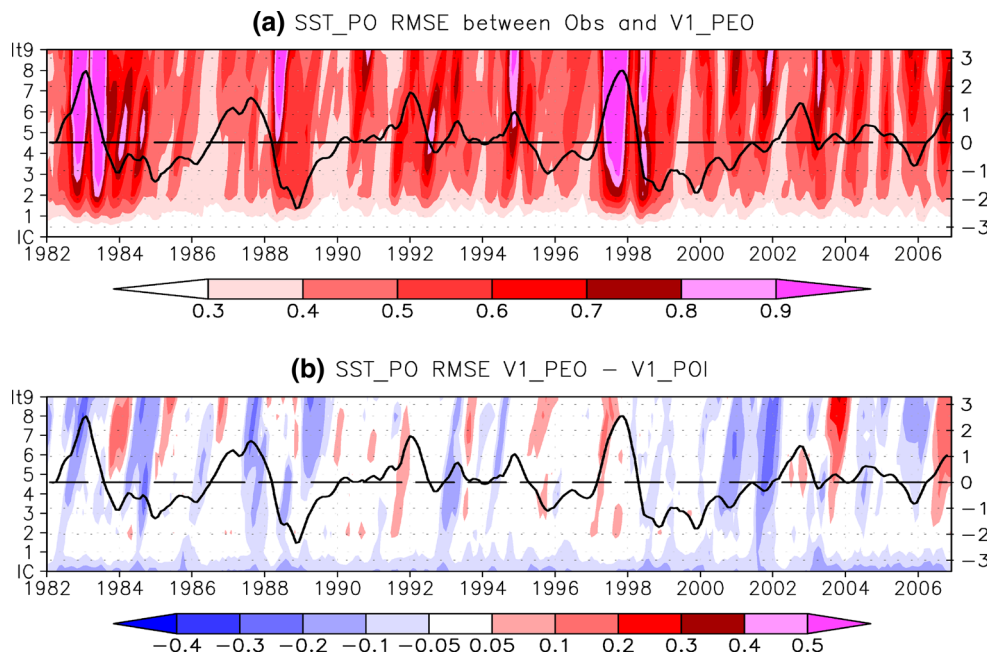


**Fig. 5** Correlation of the differences in SCOR for each forecast from V1_PEO and V1_POI from SST and T300 (*black curve*) and from SST and S300 (*red curve*) as a function of forecast lead time. The SCOR was computed over the tropical Pacific region

Nonetheless, the greatest reduction in SST error (blue colours in Fig. 6b) also occurs for times of weak ENSO amplitude. A similar decrease in error is also seen for T300 (not shown). For S300 the reduction in RMSE for V1_PEO compared to V1_POI is not as long lasting as for correlation (Fig. 4d) and clearly is maximum at the initial time and decays with forecast lead time (not shown). This implies that reduced error in the initial salinity analysis impacts forecast more through the phasing of the predicted S300 anomalies (i.e., as reflected in improved spatial correlation score) than through amplitude (i.e., reduced RMSE).

### 4.4 Source of improved skill

The above results are suggestive that improved forecast skill in V1_PEO compared to V1_POI for prediction of ENSO primarily stems from improved prediction of subsurface temperatures and salinity. Maps of forecast skill for V1_PEO and V1_POI of T300 and S300, as measured by temporal correlation, are displayed in Fig. 7 for forecast lead time 6 months. For reference, we also include the maps of SST skill. Here we use the PEODAS analysis for subsurface verification, and Reynolds et al. (2002) for SST verification. Clearly forecast skill is higher in V1_PEO for T300, especially in the regions where ocean memory has a strong influence on the coupled climate (i.e., the equatorial Pacific). We note that this improvement in skill for V1_PEO relative to V1_POI is independent of the choice of subsurface temperature verification data (e.g., a similar improvement is seen if we verify with EN3; not shown). There is also a large increase in T300 skill where ENSO is
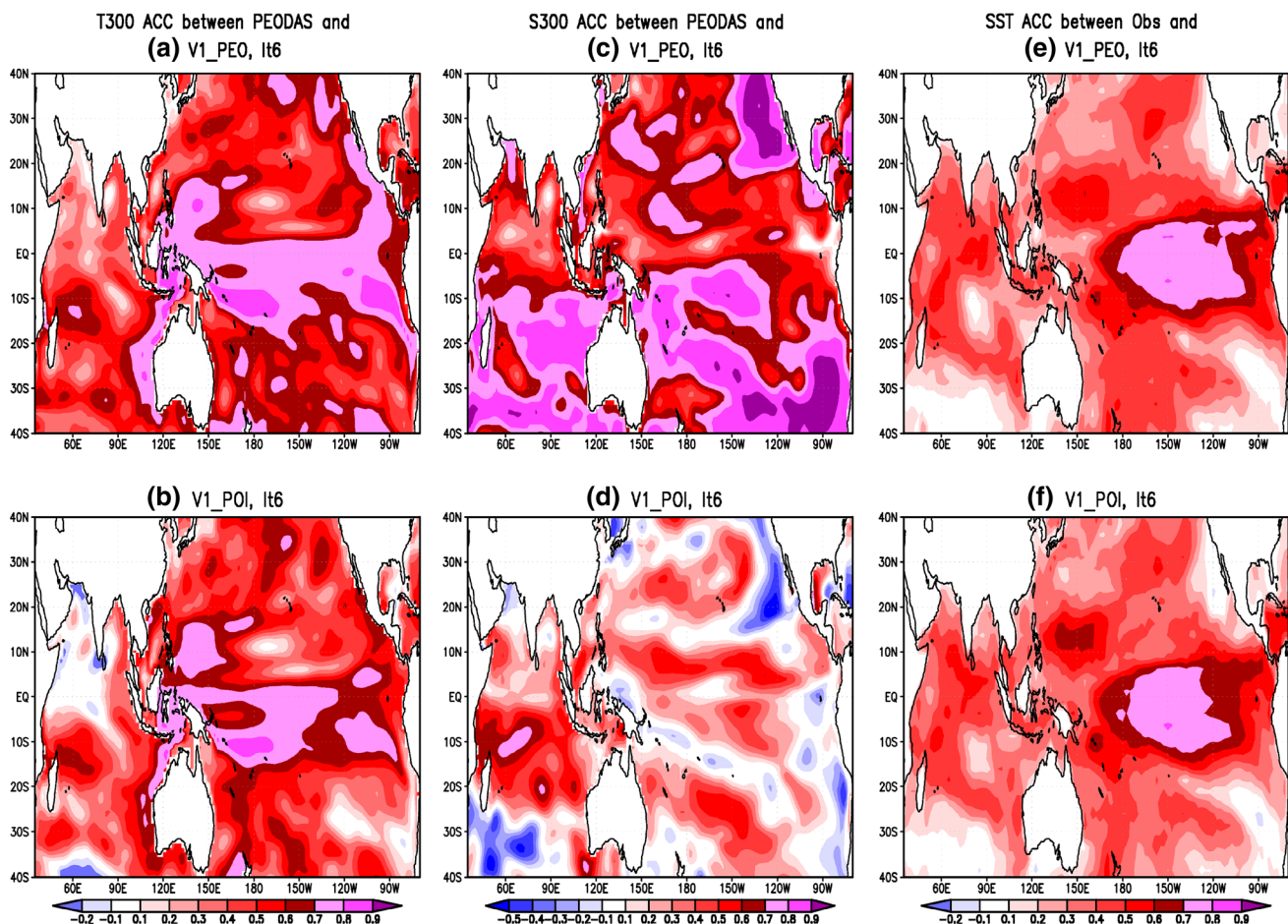
**Fig. 6** The same as Fig. 4a, b, but for root mean square error (RMSE)

**Fig. 7** Correlation of predicted anomalies of T300 at lead time 6 months from **a** V1_PEO and **b** V1_POI. Verification is PEODAS analysis. **c, d** Predictions of S300. **e–f** Predictions of SST and are verified with Reynolds OI v2 observational analyses

known to have large remote impacts, for instance on the west coast of Australia, which is associated with the oceanic teleconnection through the Indonesian throughflow (e.g. Hendon and Wang 2009), and on the north and south American coasts where coastally trapped Kelvin waves carry the signal of ENSO from the equator to high latitudes. Notably skill is modestly improved in the equatorial IO, including over the WIO and west of the Java–Sumatra, presumably reflecting improved prediction of the subsurface evolution of the IOD in V1_PEO.

The improvement in predicting subsurface salinity (S300) using V1_PEO compared to V1_POI is even moved prominent (Fig. 7c, d). The regions of high skill for salinity are located primarily south of the equator in both basins, but the high skill regions for temperature are more centered on the equator, where coupled interactions associated with ENSO occur. Improvement in the prediction of salinity by V1_PEO in the regions south of the equator in the Pacific would not be expected to be expressed as an improvement in forecasts of the coupled climate at least not until longer lead time. This is because although subsurface salinity will

directly affect density, these density changes off of the equator would have to get into the equatorial region to be important for ENSO dynamics, and the time scale for off equatorial density anomalies to get into the equatorial region is long (many months; cf e.g., Zhao et al. 2013). However, the near equatorial salinity improvements seen in the central Pacific (Fig. 7c compared to 7d) can have a more rapid impact on ENSO variability (Zhao et al. 2013).

The region of improved skill for SST in the Pacific roughly coincides with the region of improved skill for T300 (Fig. 7a, e). In contrast, although V1_PEO has better skill than V1_PO for T300 in much of the IO, there is little improvement for SST there. This highlights that coupling between the subsurface and surface temperatures in the IO is generally much weaker compared to in the equatorial Pacific and where the coupling is strongest (south west tropical IO and west of Java-Sumatra), the improvement in prediction of T300 is relatively small (Fig. 7a, b).

To better highlight the key role of equatorial subsurface anomalies of T and S for providing improved long lead forecast skill of surface climate, we look at the forecast
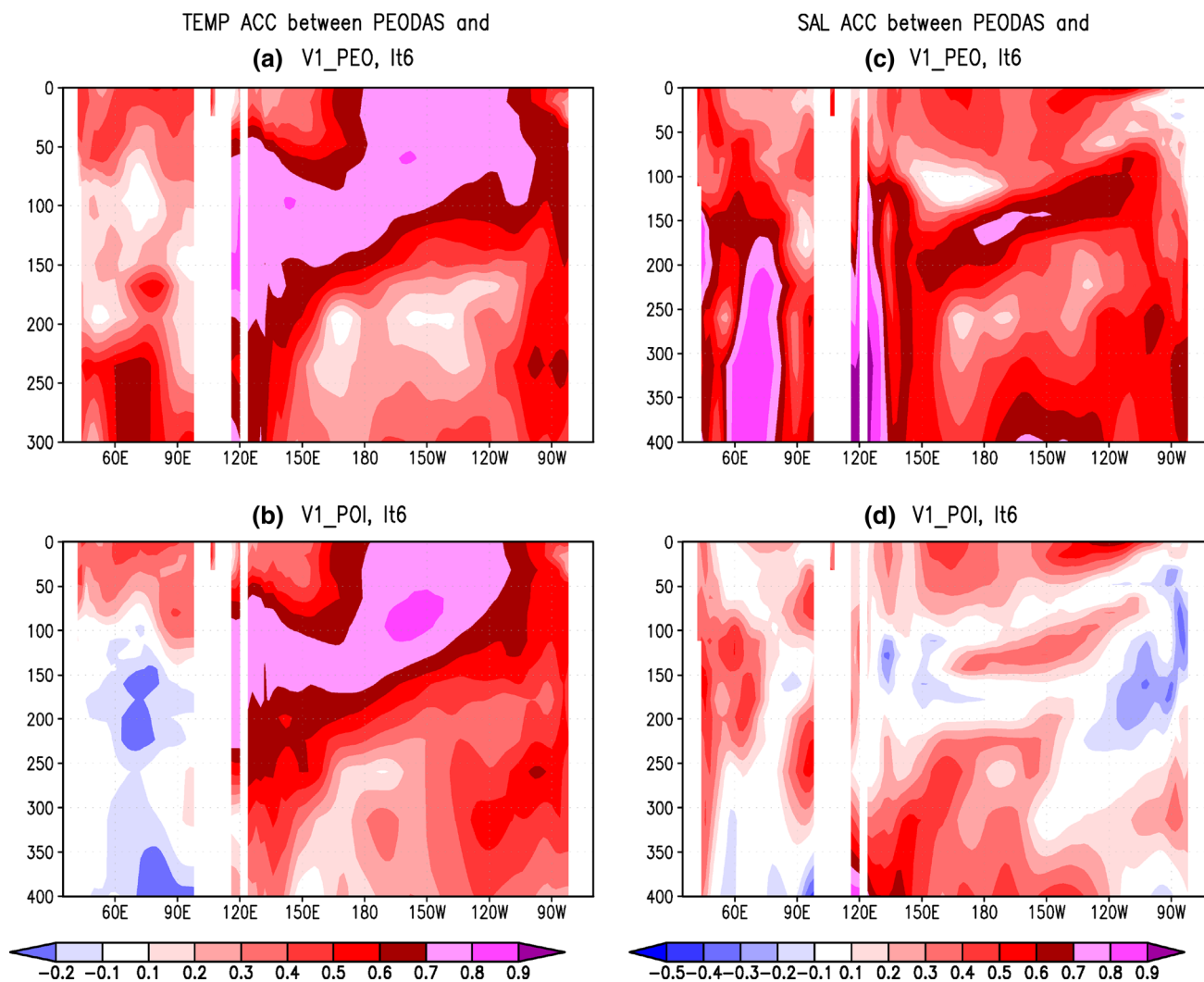
## TEMP ACC between PEODAS and
### (a) V1_PEO, lt6

## SAL ACC between PEODAS and
### (c) V1_PEO, lt6

### (b) V1_POI, lt6

### (d) V1_POI, lt6

−0.2  −0.1  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9

−0.5 −0.4 −0.3 −0.2 −0.1  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9

**Fig. 8** Correlation of predicted anomalies of temperature along equator (5°N–5°S) as function of depth (0–400 m) from **a** V1_PEO and **b** V1_POI at the lead time 6 months. **c, d** For salinity. Verification is PEODAS analysis

skill as a function of depth along the equator for temperature and salinity at lead time 6 months (Fig. 8). V1_PEO shows higher skill for both temperature and salinity along the equatorial Pacific thermocline, which is the region that is actively involved in ENSO variability. V1_POI shows similar forecast skill as V1_PEO for near surface salinity but, as expected, little skill for subsurface salinity. Even though there is a poor depiction of near-surface salinity variations in the initial state of V1_POI, similar forecast skills for V1_POI and V1_PEO at 6 month lead time results because of the fast time scale of near surface salinity to respond to the model's surface fluxes and induced near-surface currents even if the near-surface salinity is initialized to climatology (as is the case for V1_POI).

In summary, subsurface forecast skill has clearly improved in V1_PEO because of improved depiction of the upper ocean at the initial time. Importantly, these improvements in both temperature and salinity occur in regions that provide the effective memory (predictability) of the coupled atmosphere–ocean climate, thus promoting improved predictions of ENSO from V1_PEO over V1_POI at longer lead times.

## 5 Source of forecast differences

The preceding analysis has demonstrated that V1_PEO provides improved predictions of SST variation associated with ENSO, and that these improved predictions are tied to improved prediction of subsurface temperature and salinity variations in the near-equatorial region of the Pacific, where subsurface variations are tightly coupled to subsequent surface climate variations. It is an outstanding question as to which differences in the depiction of the

initial ocean state matter most for the differences in forecasts of surface climate (especially forecasts of equatorial SST) in the two systems. To address this, we relate the differences in prediction of the Niño3.4 SST index between V1_PEO and V1_POI at different lead times to the differences in the depiction of the initial sate of temperature and salinity between the two ocean analyses. We explore this by directly relating the differences in the initial state of temperature and salinity to differences in the prediction of Niño3.4 using multiple linear regression. That is, we use the initial differences in temperature and salinity at each grid point in the equatorial-depth plane as predictors for the subsequent difference in Niño3.4 index at each lead time. In developing the regression, we use multiple linear regressions to account for temporal co-variation of the differences in T and S at each grid point. However, we do not account for any spatial co-variation of the initial temperature and salinity differences. But, this method allows us to get a direct measure of the total variance of the difference in the prediction of Niño3.4 that can be accounted for by initial differences in T and S.

We display the regression coefficients in the equatorial-depth plane as sensitivities of the differences in prediction of the Niño3.4 SST index to the standardized differences in initial temperature (Fig. 9a–d) and salinity (Fig. 9e–h). That is, the units of the contoured quantities in Fig. 9 are °C difference for prediction of Niño3.4 per SD difference of the initial temperature and salinity differences. At short lead time (1 month; Fig. 9a, e), differences in the prediction of the Niño3.4 SST index derive primarily from differences in the depiction of near-surface temperature in the vicinity of the Niño3.4 region. This makes sense because at short lead time, the only differences in the initial state that could matter to variations of Niño3.4 SST will be near-surface temperature in the vicinity of the Niño3.4 region. We do see a modest sensitivity to near surface salinity (fresh) anomaly at the eastern edge of the Pacific warm pool (∼160°W), which would go together with an eastward expanded warm pool as depicted by the temperature differences in Fig. 9a.

As lead time increases, the sensitivity to the initial differences moves down and westward into the halocline/thermocline of the western Pacific and the sensitivity to the initial salinity differences become as important as temperature differences. We also see a modest sensitivity to surface salinity differences at lead times 3 and 6 months, indicative of an initially eastward expanded fresh pool resulting in warmer Niño3.4 index. However, the largest sensitivity is in the halocline/thermocline. This downward and westward shift of the sensitivity with lead time makes sense because the typical evolution of anomalies in the subsurface that subsequently will affect the eastern Pacific SST in the Niño3.4 region is from the western Pacific to the eastern Pacific along the thermocline.

Interestingly, already by lead time 3 months (Fig. 9b, f), differences in the prediction of Niño3.4 SST are more sensitive to initial salinity differences in the halocline in the western Pacific than they are to the initial temperature differences in the thermocline. The sign of this sensitivity makes sense: a positive difference in the prediction of the Niño3.4 SST index will derive from a warmer thermocline difference (negative density perturbation) in the western Pacific 6–9 months earlier. But a negative salinity difference in the same region will also cause a negative density difference that will also be carried to the eastern Pacific in the form of a temperature perturbation (see for example Zhao et al. 2013). Alternatively, a warmer thermocline in the western Pacific can be viewed to be caused by adiabatic downward displacement, and this downward displacement will result in freshening above the halocline (the saltiest water sits in the halocline which lies just above the thermocline).

Zhao et al. (2013) showed that this region around the thermocline in the equatorially western Pacific is where the largest differences in the depiction of the mean state salinity occur between PEODAS and POI. They showed that the mean salinity bias in POI results in a density anomaly that sets off a coupled response in the forecasts akin to that at the onset of El Niño. The mean behavior of El Niño is then affected by the subsequent changes in the mean sate. The present study suggests that differences in depiction of anomalies of salinity in this region are also a source of differences in prediction of individual ENSO events. Although this correlation/regression analysis, which highlights patterns and regions of forecast sensitivity, does not differentiate between improved or degraded forecast skill, it is suggestive that the improved depiction of subsurface temperatures and, importantly and perhaps unexpectedly, salinity in PEODAS in the tropical western Pacific is a source of improved predictions of ENSO by V1_PEO at longer lead times.

We quantify the relative contribution of the temperature and salinity differences at the initial time to the total explained variance of the predicted Niño3.4 differences using analysis of variance. We calculate the total explained variance of the predicted differences in Niño3.4 index by the initial temperature and salinity differences at each grid point and then break this down into the explained variance resulting from the co-variation of the initial temperature and salinity difference, and the explained variances by the temperature differences that are independent of the salinity differences, and by the salinity differences that are independent of temperature. We show this for a lead time of 9 months (Fig. 10), where the contoured quantity is the explained variance of the difference in predicted Niño3.4 index by the difference in the initial temperature and salinity at each grid point. We see that the maximum total
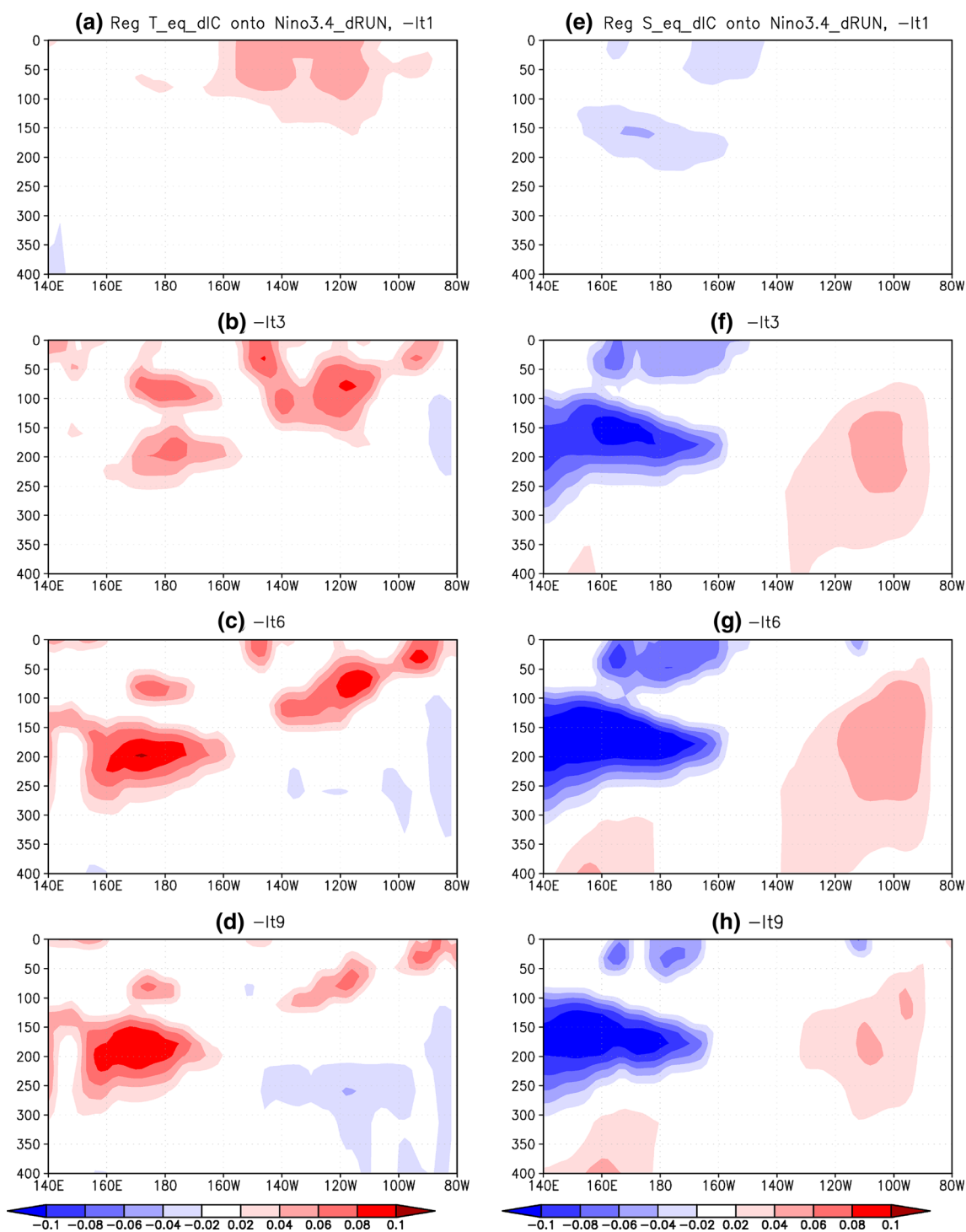
**Fig. 9** Multiple regression of the difference in prediction of the Niño3.4 index at lead times 1 (**a**, **e**), 3 (**b**, **f**), 6 (**c**, **g**), and 9 (**d**, **h**) months onto the initial differences in temperature (*left panel*) and salinity (*right panel*) at each grid point in the equatorial-depth plane (5°N–5°S). The regression coefficient is *shaded* at each grid point and is expressed as a sensitivity of the difference in the predicted Niño3.4 index (°C) to a one SD difference in initial temperature (*right panel*) and salinity (*left panel*)

explained variance ($\sim$20 %) is located in the western Pacific thermocline (Fig. 10a), and that around 8 % comes from the independent temperature difference (Fig. 10c) and 12 % from the independent salinity differences (Fig. 10d).

There is little contribution from co-variance between initial temperature and salinity differences (Fig. 10b).

In attempt to understand why the initial salinity differences have a larger impact on the predicted Niño3.4 index

**Fig. 10** **a** The total explained variance of the predicted difference in Niño3.4 index at lead time 9 months by the multiple regression onto the initial difference in temperature and salinity at each grid point in the equatorial-depth plane. The total explained variance is broken down into **b** the part that is due to the covariation of the initial temperature and salinity differences and the part that is due to **c** the initial temperature variations that are independent of the initial salinity variations, and **d** the initial salinity variations that are independent of the initial temperature variations
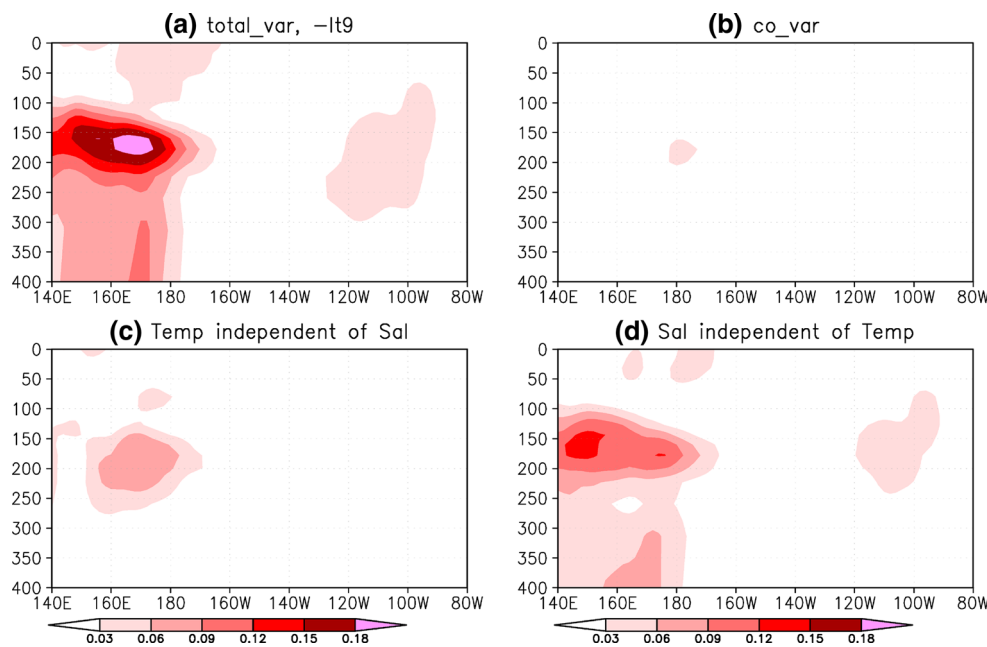


than do the initial temperature difference, we examine the magnitude (SD) of the initial differences in temperature (Fig. 11a) and salinity (Fig. 11b). Not surprisingly, the maximum SD of the initial temperature differences aligns along the thermocline, where the observed temperature variability is greatest. The maximum SD of the difference in salinity (Fig. 11b) is in the near surface in the western Pacific, which too is not surprising because the observed interannual SD of salinity is large there but the surface salinity in the POI analyses is constrained to be climatological (i.e., the neat surface SD of the differences in salinity displayed in Fig. 11b looks nearly identical to the SD directly from the PEO analyses). We do see, however, that there is also a maximum of SD of the differences in subsurface salinity in the vicinity of the thermocline in the central and western Pacific.

The question is raised as to the relative importance of these subsurface salinity differences (Fig. 11b) compared to the temperature differences (Fig. 11a) in the central and western Pacific. One way to answer this question is to estimate the relative contribution of each of these initial differences to the initial density differences. We compute the relative contribution of the temperature and salinity variations to the density using a linearized equation of state (e.g., Gill 1982) and display the relative contribution to the density difference as the ratio of the contribution of the salinity differences to the contribution of the temperature differences (Fig. 11c). Interestingly, we see that salinity differences dominate in the near surface, but in the thermocline region of the western Pacific where the sensitivity of the predicted Niño3.4 index is greatest for salinity differences (e.g., Fig. 10d), we see that the temperature

differences make a stronger contribution to the local density difference. The mean ratio in the western Pacific thermocline is about 0.4–0.8 which indicates that the initial temperature differences are associated with 50–75 % greater density perturbations than are the salinity differences. So why does the predicted difference in Niño3.4 show a greater sensitivity to the initial salinity differences, especially at longer lead times?

The answer to this question stems from the fact that the initial salinity differences, although producing an initial density difference that is roughly 50 % weaker than from the initial temperature differences, results in sustained temperature differences because the adjustment time for salinity in the western Pacific is much longer than for temperature (e.g., Zhao et al. 2013). We demonstrate this by regressing the predicted difference in temperature at each grid point onto the time series of area mean (160°E–170°W, 142–218 m) difference in temperature and salinity at the initial time (Fig. 12). To account for covariation of the box-averaged differences of temperature and salinity at the initial time, we use multiple regression. To be clear, our two predictors are the box-averaged initial difference in temperature and salinity and our predictand is the predicted difference in temperature at every grid point. We display the regression coefficients with units °C of predicted temperature difference per SD of each predictor. For clarity, we show regression at the initial time so that the initial temperature differences are clear.

Concentrating first on the temperature response to an initial temperature difference (left hand columns of Fig. 12), we see that the initial temperature difference is concentrated in the region where the temperature box was
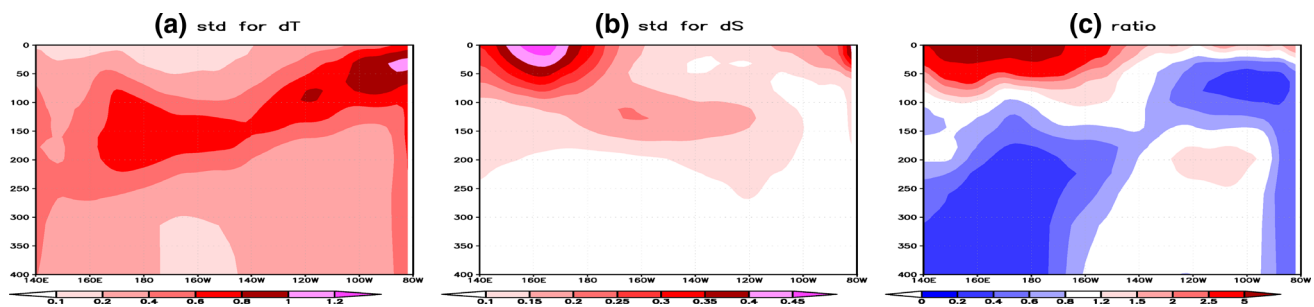
**Fig. 11** The SD of the initial differences in **a** temperature (units °C) and **b** salinity (units psu). **c** The relative contribution to the initial difference in density expressed as a ratio of the salinity contribution to the temperature contribution. Values greater than (less than) 1 mean that the initial salinity differences (temperature differences) account for a larger impact on density

defined (Fig. 12a) and this temperature anomaly then propagates eastward along the thermocline, demonstrating well known characteristics of Kelvin wave propagation. The temperature signal surfaces in the eastern Pacific after about 3 months, and then gradually weakens. In contrast, the initial temperature anomaly associated with the box-mean salinity difference is relatively weak locally, is negatively signed (as explained above, an initial positive salinity difference will cause a positive density difference which is reflected by a negative temperature difference). This initial cold anomaly then propagates eastward and surfaces in the eastern Pacific (most prominently after 3 months but is even evident after 1 month). Locally in the western Pacific, a warm anomaly develops in order to counter the density perturbation cause by the initial salinity difference (c.f., Zhao et al. 2013), and because the initial salinity difference in the west is slow to adjust, the resultant warm anomaly in the west and upward tilt of the thermocline to the east is maintained through to 9 month lead time (Fig. 12j). Hence, the impact of the initial salinity difference on the subsequent surface temperature in the east is greater than that from the initial temperature differences even though the initial density difference resulting from the initial temperature differences are greater than those from the initial salinity differences.

## 6 Conclusions

We have assessed the impact of improved ocean initial conditions for prediction of ENSO and IOD using the Bureau of Meteorology's POAMA forecast systems. Improved ocean analyses are provided by the PEODAS ocean analysis system, which uses an ensemble of analyses, explicit estimates of state-dependent background error covariance, and dynamically consistent, multivariate model updates for both subsurface temperature and salinity. PEODAS gives a clear improvement for depiction of the initial state of key oceanographic variables, such as vertical

mean heat content (T300) and salinity content (S300), sea level anomaly, and ocean currents, over the previous POI system, which used static error covariances and was uni-variate (temperature only), leading to dynamical inconsistencies (Yin et al. 2011).

A pair of hindcast experiments from the POAMA seasonal forecast system was examined for which the only difference was initial conditions as provide by PEODAS and the older POI. Hindcasts using initialized with PEODAS lead to notable increases in the forecast skill after about 2 months for SST and subsurface heat content (as measured by T300) over the Pacific. This result is consistent with previous studies (Kleeman et al. 1995; Ji and Leetmaa 1997; Rosati et al. 1997; Wang et al. 2002; Balmaseda et al. 2009; Stockdale et al. 2011) in which more accurate data assimilation had a significant positive impact in improving ENSO forecast skill. Specifically focusing on the Niño3.4 SST index the hindcasts initialized with PEODAS improved the prediction of Niño3.4 SST index by $\sim 1$ month lead time. We further showed that most of the forecast improvement arose for prediction of weaker ENSO events or neutral years, which makes sense because improved initialization would provide the greatest benefit for times of a smaller signal to noise.

We think that PEODAS provides better ocean initialization because the ocean subsurface is more dynamically balanced and possibly represents more accurately low-frequency variability in the ocean initial conditions which are desirable for ENSO prediction. Zhao et al. (2013) showed that systematic bias in subsurface salinity from POI subsequently affected the simulated mean state temperature and simulated ENSO variability and that this impact developed rapidly but lasted at least 9 months (duration of hindcast). The present results goes one step further and indicates that increased forecast skill of SST anomalies over the Pacific is linked to the increased forecast skill of subsurface temperature and salinity anomalies, which derive from improved depiction of subsurface temperature and salinity variations at the initial time.

**Fig. 12** Multiple-regression of the predicted difference in temperature at each grid point in the equatorial-depth plane (0–400 m averaged 5°N–5°S along 140°E–80°W) onto the initial differences in temperature (*left panel*) and salinity (*right panel*) averaged in the box bounded by (160°E–170°W, 142–218 m). The predicted differences in temperature (units of °C) are *scaled* for one SD anomalies of box-averaged initial differences in temperature and salinity and are shown at the initial time and at lead times 1, 3, 6, and 9 months
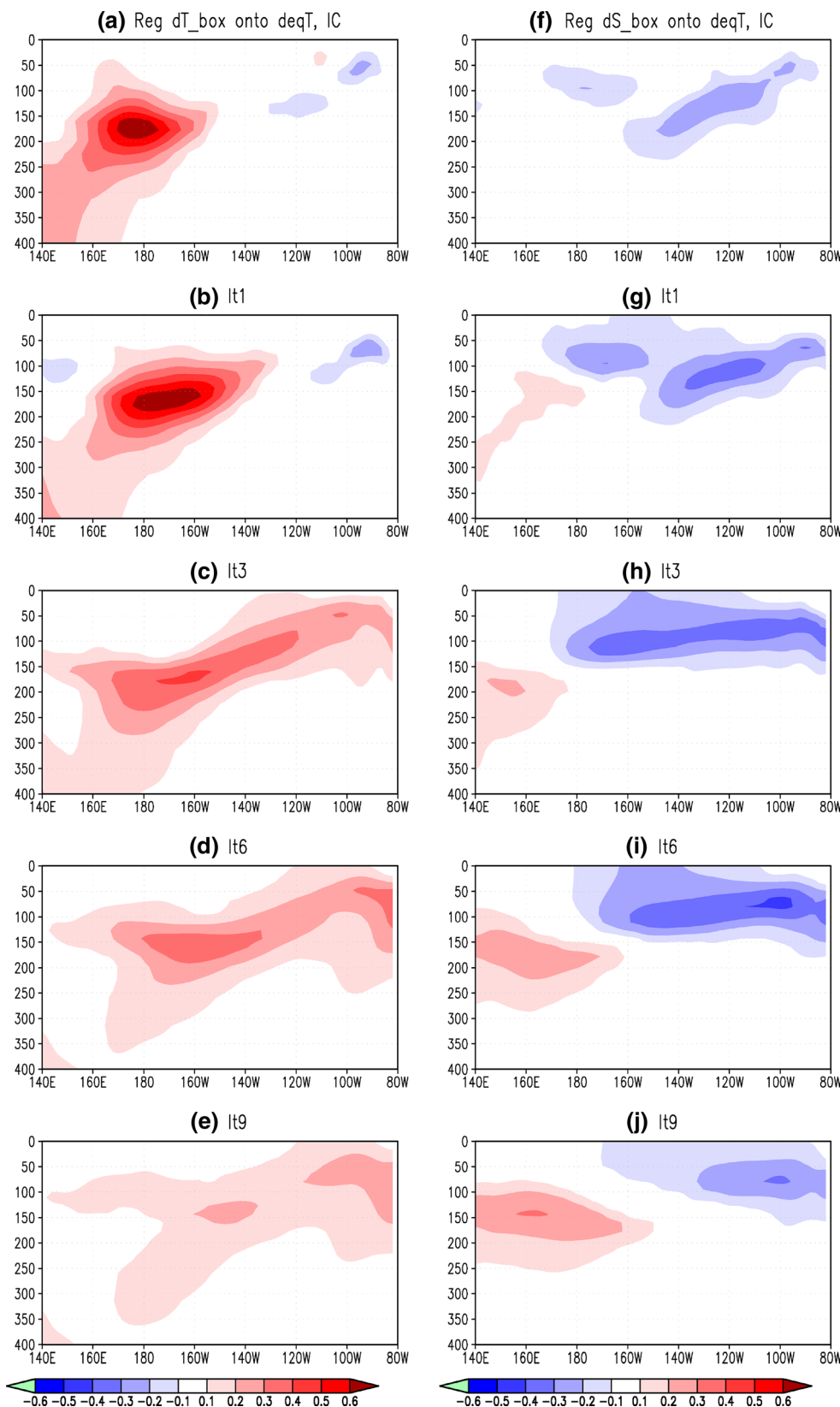


Although we see a pronounced improvement in prediction of surface and subsurface temperatures in the Pacific associated with ENSO, we do not see the same improvement in prediction of the IOD. We do see a modest improvement in the prediction of the subsurface temperatures in the IO, but this does not translate into improvement

prediction at the surface (i.e., the IOD). This reflects that there is less control of the subsurface temperature variations on the surface temperature variations in the IO than in the Pacific (i.e., the predictability of the IO surface climate due to ocean memory is lower than in the Pacific). Furthermore, atmospheric noise associated with, for instance the MJO, is stronger over the IO than the Pacific, and systematic model error in the IO region (e.g. associated with depiction of surface fluxes heat fluxes) also act to limit the capability to exploit improved ocean initialization.

A sensitivity analysis of the difference in predicted Niño3.4 SST anomaly to the differences in the depiction in the initial state of temperature and salinity shows that sensitivity shifts westward and downward into the Pacific thermocline and halocline with increasing lead time, and the relative sensitivity to the initial salinity differences increases. After about lead time 3 months, the sensitivity per SD of initial salinity difference is greater than per SD of initial temperature difference, which suggests that errors in the depiction of salinity variations are more important than errors in the depiction of initial temperature variations. However, the errors in depiction of salinity in the POI analyses in the western Pacific thermocline region are larger than what can be achieved using a modern assimilation system that accounts for the covariation of salinity and temperature even in the absence of any in situ salinity observations (Zhao et al. 2013) and so the errors in POI and hence the differences between POI and PEODAS would not be expected to be representative of typical analysis errors in a well designed assimilation system (such as PEODAS). However, our results emphasize that salinity perturbations in the halocline/thermocline region of the western equatorial Pacific can rapidly affect surface climate in the eastern Pacific and this impact is maintained for a long time, thus, it should be considered a source of longer lead predictability of ENSO. Thinking about longer lead (i.e., decadal) predictions, a proper depiction of slower variations of salinity in this thermocline region could potentially be critical to make, for example, decadal predictions of ENSO activity.

# References

Alves O, Wang G, Zhong A, Smith N, Tseitkin F, Warren G, Schiller A, Godfrey S, Meyers G (2003) POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. In: Proceedings of the National Drought Forum, Brisbane, pp 49–56. Available from DPI Publications, Department of Primary Industries, GPO Box 46, Brisbane, Qld 4001, Australia

Balmaseda M, Anderson D (2009) Impact of initialisation strategies and observations on seasonal forecast skill. Geophys Res Lett 36:L01701. doi:10.1029/2008GL035561

Balmaseda MA, Vidard A, Anderson D (2008) The ECMWF ocean analysis system: ORA-S3. Mon Weather Rev 136:3018–3034

Balmaseda MA, Alves OJ, Arribas A, Awaji T, Behringer DW, Ferry N, Fujii Y, Lee T, Rienecker M, Rosati T, Stammer D (2009) Ocean initialization for seasonal forecasts. Oceanography 22:154–159

Behringer DW (2007) The global ocean data assimilation system at NCEP. In: 11th symposium on integrated observing and assimilation systems for atmosphere, oceans, and land surface, AMS 87th annual meeting, Henry B. Gonzales Convention Center, San Antonio, Texas, 12 pp

Boyer TP, Stephens C, Antonov JI, Conkright ME,. Locarnini RA, O'Brien TD, Garcia HE (2002) World Ocean Atlas 2001, vol 2: Salinity. In: Levitus S (ed) NOAA Atlas NESDIS 50. U. S. Government Printing Office, Washington, 165 pp., CD-ROMs

Burgers G, Balmaseda MA, Vossepoel FC, van Oldenborgh GJ, van Leeuwen PJ (2002) Balanced ocean-data assimilation near the equator. J Phys Oceanogr 32:2509–2519

Cane MA, Zebiak SE, Dolan SC (1986) Experimental forecasts of El Niño. Nature 321:827–832

Gill AE (1982) Atmosphere–ocean dynamics. Academic Press, London, p 662

Hendon HH (2003) Indonesian rainfall variability: impacts of ENSO and local air–sea interaction. J Clim 16:1775–1790

Hendon HH, Wang G (2009) Seasonal prediction of the Leeuwin Current using the POAMA dynamical seasonal forecast model. Clim Dyn. doi:10.1007/s00382-009-0570-3

Hudson D, Alves O (2007) The impact of land–atmosphere initialisation on dynamical seasonal prediction. CAWCR Research Report No. 133, Bur. Met., Melbourne, Australia, 4 pp

Hudson D, Alves O, Hendon H, Wang G (2011) The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST. Clim Dyn 36:1155–1171. doi:10.1007/s00382-010-0763-9

Ingleby B, Huddleston M (2007) Quality control of ocean temperature and salinity profiles—historical and real-time data. J Mar Syst 65:158–175

Ji M, Leetmaa A (1997) Impact of data assimilation on ocean initialization and El Niño prediction. Mon Weather Rev 125:742–753

Ji M, Behringer DW, Leetmaa A (1998) An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: the coupled model. Mon Weather Rev 126:1022–1034

Jin EK, Kinter JL III, Wang B, Park CK, Kang LS, Kirtman BP, Kug JS, Kumar A, Luo JJ, Schemm J, Shukla J, Yamagata T (2008) Current status of ENSO prediction skill in coupled ocean–atmosphere models. Clim Dyn 31:647–664. doi:10.1007/s00382-008-0397-3

Kirtman BP, Shukla J, Huang B, Zhu Z, Schneider EK (1997) Multiseasonal prediction with a coupled tropical ocean global atmosphere system. Mon Weather Rev 125:89–808

Kleeman R, Moore AM, Smith NR (1995) Assimilation of subsurface thermal data into an intermediate tropical coupled ocean–atmosphere model. Mon Weather Rev 123:3103–3113

Latif M, Sterl A, Maier-Reimer E, Junge MM (1993) Structure and predictability of the El Niño/Southern Oscillation phenomenon in a coupled ocean–atmosphere general circulation model. J Clim 6:700–708

Luo JJ, Masson S, Behera S, Yamagata T (2008) Extended ENSO predictions using a fully coupled ocean-atmosphere model. J Clim 21:84–93

McPhaden MJ, Busalacchi AJ, Cheney R, Donguy J, Gage KS, Halpern D, Ji M, Julian P, Meyers G, Mitchum GT, Niiler PP, Picaut J, Reynolds RW, Smith N, Takeuchi K (1998) The tropical ocean-global atmosphere observing system: a decade of progress. J Geophys Res 103(C7):14169–14240

Pacanowski RC (1995) MOM2 documentation user's guide and reference manual. Version 1.0, GFDL Tech. Rep. 3, 232 pp

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. J Clim 15:1609–1625

Rosati A, Miyakoda K, Gudgel R (1997) The impact of ocean initial conditions on ENSO forecasting with a coupled model. Mon Weather Rev 125:754–772

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, van den Dool HM, Pan HL, Moorthi S, Behringer D, Stokes D, White G, Lord S, Ebisuzaki W, Peng P, Xie P (2006) The NCEP climate forecast system. J Clim 15:3483–3517

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401:360–363

Schiller A, Godfrey JS, McIntosh PC, Smith NR, Alves O, Wang G, and Fiedler R (2002) A new version of the Australian community ocean model for seasonal climate prediction. SCIRO Marine Research Report No. 240

Schneider EK, Huang BH, Zhu ZX, DeWitt DG, Kinter JL III, Kirtman BP, Shukla J (1999) Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM. Mon Weather Rev 127:1187–1207

Smith NR, Blomley JE, Meyers G (1991) A univariate statistical interpolation scheme for subsurface thermal analysis in tropical oceans. Prog Oceanogr 28:219–256

Stephens C, Antonov JI, Boyer TP, Conkright ME, Locarnini RA, O'Brien TD, Garcia HE (2002) World Ocean Atlas 2001, vol 1: Temperature. In: Levitus S (ed) NOAA Atlas NESDIS 49. U.S. Government Printing Office, Wash., D.C., 167 pp, CD-ROMs

Stockdale TN, Anderson D, Balmaseda MA, Doblas-Reyes F, Ferranti L, Mogensen K, Palmer TN, Molteni F, Vitart F (2011) ECMWF seasonal forecast system 3 and its prediction of sea surface temperature. Clim Dyn 37:455–471. doi:10.1007/s00382-010-0947-3

Wang GM, Kleeman R, Smith N, Tseitkin F (2002) The BMRC coupled general circulation model ENSO forecast system. Mon Weather Rev 130:975–991

Wang GM, Hudson D, Yin Y, Alves O, Hendon H, Langford S, Liu G, Tseitkin F (2011) POAMA-2 SST skill assessment and beyond. CAWCR Res Lett 6. Sandery PA, Leeuwenburg T, Wang G, Hollis AJ, Day KA (eds). http://www.cawcr.gov.au/publications/researchletters.php

Xue Y, Balmaseda MA, Boyer T, Ferry N, Good S, Ishikawa I, Rienecker M, Rosati T, Yin Y, Kumar A (2011) Comparative analysis of upper ocean heat content variability from ensemble operational ocean analyses. US CLIVAR Var 9(1):7–10

Yin YH, Alves O, Oke PR (2011) An ensemble ocean data assimilation system for seasonal prediction. Mon Weather Rev 139:786–808. doi:10.1175/2010MWR3419.1

Zhang S, Harrison MJ, Rosati A, Wittenberg AT (2007) System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. Mon Weather Rev 135:3541–3564

Zhao M, Hendon H (2009) Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. Q J R Meteorol Soc 135:337–352. doi:10.1002/qj.370

Zhao M, Hendon HH, Alves O, Yin Y, Anderson D (2013) Impact of salinity constraints on the simulated mean state and variability in a coupled seasonal forecast model. Mon Weather Rev 141(1):388–402

Zhong A, Colman R, Smith N, Naughton M, Rikus L, Puri K, Tseitkin F (2001) Ten-year AMIP1 climatologies from versions of the BMRC atmospheric model. BMRC Res Rep 83, 33 pp

Zhong A, Hendon HH, Alves O (2005) Indian Ocean variability and its association with ENSO in a global coupled model. J Clim 18:3634–3649