

The INGV-CMCC Seasonal Prediction System: improved ocean initial conditions

ANDREA ALESSANDRI * AND ANDREA BORRELLI

Centro Euro-Mediterraneo per i Cambiamenti Climatici, Bologna, Italy

SIMONA MASINA, ANNALISA CHERCHI, SILVIO GUALDI AND ANTONIO NAVARRA

Centro EuroMediterraneo per i Cambiamenti Climatici

and Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

PIERLUIGI DI PIETRO

Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

ANDREA F. CARRIL

Centro de Investigaciones del Mar y la Atmósfera (CIMA-CONICET), Ciudad Autónoma de Buenos Aires, Argentina

* *Corresponding author address:* Andrea Alessandri, Centro Euro-Mediterraneo per i Cambiamenti Climatici, Via Aldo Moro 44, 40127 Bologna, Italy.

E-mail: alessandri@bo.ingv.it

ABSTRACT

The development of the INGV (Istituto Nazionale di Geofisica e Vulcanologia)-CMCC (Centro Euro-Mediterraneo per i Cambiamenti Climatici) Seasonal Prediction System (SPS) is documented. In this SPS the ocean initial conditions estimation includes a Reduced Order Optimal Interpolation procedure for the assimilation of temperature and salinity profiles at the global scale. Nine member ensemble forecasts have been produced for the period 1991-2003 for two starting dates per year in order to assess the impact of the subsurface assimilation in the ocean for initialization.

Comparing the results with control simulations (i.e.: without assimilation of subsurface profiles during ocean initialization), we showed that the improved ocean initialization increases the skill in the prediction of tropical Pacific SSTs in our system for boreal winter forecasts. Considering the forecast of the El Niño 1997-1998, the data assimilation in the ocean initial conditions leads to a considerable improvement in the representation of its onset and development.

Our results indicate a better prediction of global scale surface climate anomalies for the forecasts started in November, probably due to the improvement in the tropical Pacific. For boreal winter, in both tropics and extra tropics, we show significant increases in the capability of the system to discriminate above normal and below normal temperature anomalies.

1. Introduction

The scientific basis for seasonal predictions lies in the interaction of the atmosphere with slowly varying components of the climate system such as the ocean (e.g: Navarra 2002; Shukla

and Kinter 2006). Early studies showed that El Niño can be predicted seasons in advance using numerical models of the coupled ocean-atmosphere covering the tropical Pacific (Cane et al. 1986; Zebiak and Cane 1987). Since then, there have been many developments in forecasting Sea Surface Temperature (SST) anomalies in the tropical Pacific (e.g: Latif et al. 1998; Palmer 2006; Balmaseda et al. 2007).

The ability of models to predict ENSO is critically important, as the most significant climate variability on the interannual timescale is related to this phenomenon (Ji and Leetmaa 1995; Trenberth et al. 1998; Wallace et al. 1998) and as the SST in the tropical Pacific have a global impact on atmospheric circulation (e.g: Shukla and Wallace 1983; Trenberth et al. 1998). However, models skill in predicting tropical Pacific SSTs is still limited. For instance, most of the seasonal prediction systems underestimate or do not predict the onset of the exceptional 1997/98 El Niño (e.g: McPhaden 1999; Vitart et al. 2003).

The importance of the oceanic subsurface memory—as expressed by slow variations in the equatorial Pacific upper ocean heat content—for the evolution of ENSO has been shown in many observational and modeling studies (e.g: Chen et al. 1995; McPhaden et al. 1998; Latif et al. 1998; McPhaden 1999; Navarra et al. 2008). Subsurface data assimilation can contribute in obtaining skillful seasonal forecasts and beneficial effects on predictability have been also reported increasing the space-time coverage of the observational network (Rosati et al. 1997; Alves et al. 2004; Ji and Leetmaa 1995; Wang et al. 2002; Balmaseda et al. 2007; Vidard et al. 2006). These studies have often showed reductions of the model errors. However, the reduced model errors not always corresponds to significant increases in prediction skill and the results are substantially depending on the model, the geographical region, the year and the season under consideration (e.g.: Ji and Leetmaa 1995; Vidard et al. 2006;

Balmaseda et al. 2007).

In order to be useful for decision-making, seasonal climate predictions need to be probabilistic and the capability of probability forecasts to provide valuable informations needs to be assessed (e.g: Richardson 2006). Specifically, it would be desirable to evaluate how well a set of probability forecasts is able to discriminate among the occurrence of Mutually Exclusive and Collectively Exhaustive (MECE) climate events, with the simplest possible situation represented by dichotomous yes/no cases (e.g: temperature above normal or not). Considering predictions of dichotomous events, the joint distribution of the observed predictands and of the respective probability forecasts can be conveniently analyzed through the likelihood-base rate factorization (Wilks 2006). This kind of analysis can be used for a direct quantification of the forecasts ability to discriminate among the occurrence of one or the other of a pair of dichotomous events.

This work documents the development of the INGV-CMCC Seasonal Prediction System (SPS), which includes an assimilation of in situ vertical profile observations in the oceanic model in order to produce Initial Conditions (ICs). The ocean data assimilation system has been developed at CMCC-INGV (Pietro and Masina 2009; Bellucci et al. 2007) and it has been used in order to assimilate observed profiles of temperature and salinity through the water column. The main focus of the paper is on the assessment of the impact of the assimilated initial conditions on the forecast skill.

The paper is organized as follows. Section 2 describes the seasonal prediction system, the experiments performed and the data used for validation. After a description of systematic errors and skill performance of the latest release of the system, Section 3 and Section 4 contain the comparison with control forecasts (i.e: without ocean assimilation) to analyze

the effects of the improved ocean ICs on SST bias and on prediction skill, respectively. The effects of the assimilation on the skill of global scale probability forecasts of dichotomous predictands is the topic addressed in Section 5. Finally, Section 6 contains the discussion of the main results and the summary of the conclusions of the study.

2. The Seasonal Prediction System

The SPS documented in the present study represents the evolution of the system described in Gualdi et al. (2004) and developed in the framework of the EU project DEMETER (Palmer et al. 2004). The ICs for the ocean-atmosphere system are prepared separately for the atmosphere and for the ocean. For each start date the atmospheric ICs are obtained from prescribed SST simulations. Differently, the ocean component is obtained from the ocean data assimilation or simply from a flux-forced ocean simulation for the control forecasts. Figure 1 summarizes the hindcasts generation strategy of the latest versions of our seasonal prediction system. The details of the set-up and integrations performed in this study are described in Subsection c.

a. The coupled model

The coupled model included in the System is the ocean-atmosphere CGCM Scale Interaction Experiment-Frontier (SINTEX-F; Gualdi et al. 2003b,a; Luo et al. 2005). SINTEX-F is an evolution of SINTEX (Gualdi et al. 2003b), where both oceanic and atmospheric components have been improved. The model components are ORCA2 for the ocean and ECHAM-4

for the atmosphere.

ORCA2 is the global implementation of the OPA 8.2 ocean modeling system developed by the Laboratoire d’Oceanographie Dynamique et de Climatologie (LODYC) in Paris (Madec et al., 1998; see full documentation at <http://www.lodyc.jussieu.fr/opa/>). It is a finite difference oceanic GCM and solves the primitive equations with a non-linear equation of state on an Arakawa C-grid. The horizontal mesh is orthogonal and curvilinear on the sphere, and its space resolution is roughly equivalent to a geographical mesh of $2^\circ \times 2^\circ$ (with a meridional resolution of 0.5 near the equator). 31 vertical levels are used with 10 levels in the top 100 m. In the configuration used there is no interactive model for the dynamics of sea-ice, whose area coverage is relaxed towards observed monthly climatology.

ECHAM-4 (Roeckner et al. 1996) is the fourth generation of the ECHAM atmospheric general circulation model developed at the Max-Planck-Institut Fur Meteorologie in Hamburg. The model equations are solved on 19 hybrid vertical levels (top at 10 hPa) by using the spectral transform method. In these simulations, ECHAM-4 is used with a triangular truncation T106, corresponding to an associated Gaussian grid of approximately $1.1^\circ \times 1.1^\circ$. As shown in Gualdi et al. (2004), this relatively high resolution improves considerably the prediction skill compared to a coarser atmosphere (T42) and gives a better representation of the delayed oscillator mechanism (Navarra et al. 2008). An exhaustive description of the dynamical and physical structure, and of the simulated climatology of ECHAM-4 are given by Roeckner et al. (1996).

Atmospheric and oceanic components are coupled through OASIS2.4 (Valcke et al. 2000). No flux adjustment or restoring were used in the simulations. Air-sea fluxes and SST between atmosphere and ocean were exchanged every 2 h. The features of the SINTEX-F climatology

and variability have been widely described in the past (e.g.: Gualdi et al. 2003b,a).

b. Ocean assimilation

The latest version of the CMCC-INGV Global Ocean Data Assimilation System (CIGODAS, Pietro and Masina 2009; Bellucci et al. 2007) has been used in order to assimilate observed profiles of temperature and salinity through the water column of the global configuration of the OPA8.2 ocean model. The assimilation scheme used in CIGODAS is based on the System for Ocean Forecasting and Analysis (SOFA; De Mey and Benkiran, 2002) which is a Reduced Order multivariate Optimal interpolation (ROOI) scheme. As described in Bellucci et al. (2007), CIGODAS considerably corrects the subsurface thermal structure of the oceanic model. In particular, tropical Pacific and western boundary currents regions show a beneficial impact from the assimilation. Details of the CIGODAS and of the effects on the ocean model simulated climatology and variability are found in Pietro and Masina (2009) and in Bellucci et al. (2007).

The temperature and salinity profiles used for this study are taken from the EN3 package (an assembling of WOD05, GTSP and ARGO databases, as summarized in Table 1; more informations available at <http://hadobs.metoffice.com/en3/>). Only the profiles which passed all the quality checks described in Ingleby and Huddleston (2007) have been retained for assimilation. The temperature profiles spatial coverage over different latitudes and regions of the globe are reported in Table 2 for the period 1991-2003.

c. Experiments and Data

We performed an experiment composed of 5 months seasonal forecasts for the period 1991-2003. In order to consider the possible impact of the seasonal cycle on the forecasts, the simulations are started from two different dates of the year: 1st of May and 1st of November. Two sets of nine member ensemble forecasts have been produced taking the same ICs for all the coupled model components but the ocean. In the first set the ocean initial states were estimated through the use of the data assimilation system described in Subsection b (hereinafter DAS) while in the second no observed in situ data were assimilated (hereinafter NODAS). For both DAS and NODAS, the ocean model was forced starting from 1955 with momentum, heat and fresh water flux data from the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr reanalysis (ERA-40; Uppala and co authors 2005) before 2002 (ERA40 only covers up to August 2002) and from the ECMWF operational analysis after 2002. Furthermore, in order to keep the simulated SST close to observations, the model field was damped with a timescale of seven days towards the Reynolds SSTs (Reynolds and Smith 1994) from 1982 onward and the ERA40 SSTs before. In practice, the NODAS experiment produced ICs are simply forced from atmospheric fluxes and relaxation to surface SST, whereas the DAS experiment further included assimilation of in situ ocean profiles.

The scientific basis for seasonal predictions lies in the interaction of the atmosphere with slowly varying components of the climate system. As such, seasonal climate predictions are believed to be firstly an initial value problem for the slow ocean component (Palmer 2006; Shukla and Kinter 2006), while the solution for the atmosphere can be conceived

as a boundary value problem (Navarra 2002; Palmer 2006). Consistently, the atmospheric initial conditions were obtained through an Amip-type simulation, that is by prescribing observed SST boundary forcing to the atmospheric model. It is argued that, compared to the method based on atmospheric data assimilation (as widely developed for Numerical Weather Predictions), with the Amip-type approach atmosphere and oceanic conditions (i.e: SST) are more in balance and this could minimize the initial coupling shock (e.g.: Tribbia and Troccoli 2008). The AMIP-type run was performed by using the observed SSTs from the Met Office Hadley Centre's data set (HadISST1.1 Global Sea-ice and SST; Rayner et al. 2003) for the period 1985-2003. In order to represent the uncertainties in the initial state of the system, an ensemble of nine atmospheric ICs has been produced by taking lagged days as initial states. For each starting date we consider the reference date, for instance May 1st (or November 1st) but also the four days before and after (see Fig. 1).

In summary, for each starting date an ensemble of nine atmospheric initial states were created. Starting from these ICs, for both DAS and NODAS (i.e: with and without oceanic assimilation of in situ data), the coupled model has been integrated for five months, producing two sets of nine-member ensemble forecasts covering the period 1991-2003.

1) OBSERVED AND REANALYSIS DATASETS

The predictive skill of the model is assessed comparing the forecasts with analyses and observational products. The ERA-Interim reanalyses (Uppala and co authors 2005; Berrisford et al. 2009) are used for verification of the forecasts. For precipitation, we use the Climate Prediction Center (CPC) Merged Analysis of Precipitation (Xie and Arkin 1997)

dataset. The model and the observed (reanalysis) anomalies are defined as the deviations from the respective climatology for the period 1991-2003.

3. Improved Ocean IC effect on SST Bias

An important source of inaccuracy of the forecasts performed using CGCMs is represented by the model systematic errors (e.g: Gualdi et al. 2004). This problem is especially important for the prediction of the SST anomalies. The magnitude of the model SST bias, in fact, can be as large as the amplitude of the observed SST anomalies that should be predicted. In this section, we present a description and discussion of the effect of oceanic subsurface assimilation on the bias of the SPS forecasts. Before analyzing the details of the impact of the improved ocean IC, we discuss briefly the main features of the bias in the DAS experiment, with the main focus in the tropical Pacific.

a. Bias of the System

The DAS SST systematic error for both May (panel a) and November (panel b) start dates are shown in Fig. 2. The biases are defined as the difference between the forecast ensemble means and ERA-Interim SST climatologies of the period 1991-2003. The months from 2 to 4 of the forecast period are considered in the averages which means that for the May 1st start dates we used the average of monthly means for June, July and August while monthly means for December, January and February are used for the November 1st forecasts. The results shown in Fig. 2 indicate that the systematic error of the model in predicting the

SST field is moderate in most of the tropics. Over a large portion of the tropical belt, in fact, the model exhibits an averaged cold bias smaller than 1 °C. The error is remarkably small in the tropical Indian Ocean. In the south-eastern tropical Pacific and in the upwelling regions off the American coast and in the Gulf of Guinea the model is too warm and the averaged error is larger than 1 °C. The equatorial cold tongue is too pronounced and penetrates too far into the west Pacific, producing SST patterns too symmetrical around the equator. It follows a tendency to produce a double intertropical convergence zone (ITCZ) in the tropical Pacific, consistently to what described in Gualdi et al. (2004).

Figure 2 shows a seasonal dependency of the systematic errors on the date of the ICs. For example, the SST warm bias in the tropical south-eastern Pacific and Atlantic oceans and Southern Hemisphere middle latitudes appears to be more pronounced for the forecasts with start date in November (Fig. 2b), whereas a warm bias in the tropical north-eastern Pacific and Atlantic is found in the forecasts with start dates in May (Fig. 2a). The error in the equatorial Pacific cold tongue, on the other hand, is more evident for the forecasts starting in November.

b. Sensitivity to ocean assimilation

Figure 2 compares the systematic error in the DAS SST (panels a and b), averaged over the 2-to-4 forecast months, with the one suffered by NODAS (panels c and d). Panels e and f report the SST bias difference between DAS and NODAS, with the shading evidencing the areas of significant (10% level, bootstrap method) systematic error increase (light) and decrease (dark) in DAS. The assimilated Ocean IC estimate in DAS leads to a reduced mean

bias over the tropical belt. For both the May (panels a, c and e) and November (panels b, d and f) start date, the warm bias in the upwelling regions off the American Coasts and in the Gulf of Guinea is significantly reduced over large areas. Similarly, the systematic error in the subtropical south-central Pacific is reduced. The SST bias in the regions influenced by the western boundary currents in the extra tropical Pacific and Atlantic appears also to be affected by the subsurface profile assimilation. During boreal summer, the temperature bias over the regions influenced by the Kuroshio current is considerably reduced over large areas. Similarly, the forecasts started in November display a reduction of the SST systematic error over the Gulf current regions in the Atlantic sector poleward of 45°N . In contrast, in the Atlantic southward of 45°N the SST bias in some regions bounded by the Gulf current appears to increase in DAS.

As the tropical Pacific cold tongue region was shown to be the region with the strongest bias for the forecasts started in both May and November, we focused on the mean systematic errors averaged over the Niño3.4 region (5°S - 5°N ; 190° - 240°E) for DAS (dashed) and for NODAS (dash-dots) experiments (Fig. 3). The forecasts started in May have relatively moderate drifts, reaching about 1 K after 5 months in both NODAS and DAS (Fig. 3). Stronger drifts, exceeding 1.5 K after 5 months, are found for the forecasts starting in November. The cold bias of about 0.5 K already present in the first month (Fig. 3a) indicates rapid adjustments going on due to quite a prominent "coupling shock". Noteworthy, this initialization related adjustments appears to be more effective for the November start date.

The DAS forecasts have a smaller drift than NODAS during the first part of the predictions. In particular the first month averaged bias for the November start date is reduced by about 0.25 K ($\simeq 40\%$) in DAS (Fig. 3a). However, the bias tends to converge to similar

values in the latter stages of the forecasts (from month 4 on), as the forecasts tend toward the modeled climatology (e.g: Alves et al. 2004; Jin and Kinter 2008). Our results indicate that this tendency is reduced for the first 3 months of the forecasts by the use of consistent subsurface temperature and salinity informations in order to initialize the ocean component. Please note that the uncertainty associated with analyzed SST data (based on satellites and in situ measures) is on the order of 0.2 K (e.g: Rayner et al. 2005) and this could limit the significance of the systematic error differences reported above.

4. Improved Ocean IC effect on predictability

This section reports the contribution of the oceanic subsurface assimilation on the skill of the SPS. Subsection a describes the main characteristics of the system skill as obtained for DAS. The comparison between DAS and NODAS is then discussed in Subsection b.

a. Skill of the System

Figure 4 (panels a and b) shows the point by point correlations between predicted (DAS) and observed (ERA-Interim) surface air temperature anomalies (hereinafter TAIRA). Time correlations are computed retaining for each year all the monthly means from the one month lead-time seasonal predictions (forecast months from 2 to 4: June, July and August for May 1st start dates and December, January and February for November 1st start dates). Higher correlation values are found for both the May start date (panel a) and the November start date (panel b) over the tropical Pacific. Positive significant correlations are also found over

most of the tropical Indian and Atlantic oceans with relatively higher values and significant area coverages for the forecasts started in November. The high correlations over tropical Pacific display the skill of our coupled model in predicting ENSO (see the Niño3.4 index in Fig. 5). As summarized in Table 3 the SPS performs particularly well in predicting the Niño3.4 index both for the seasons with one month lead-time (months from 2 to 4; hereinafter lead 1 seasons) and for the seasons with two months lead-time (months from 3 to 5; hereinafter lead 2 seasons). The correlation between predicted and observed monthly Niño3.4 index always exceed 0.9 for both May (0.94 and 0.91 for lead 1 and lead 2 seasons, respectively) and November (0.97 at lead 1 and 0.95 at lead 2) start dates. Correspondingly, the RMSE is moderate with values below 0.4 for November. In the cases with start date May RMSE is 0.44 and 0.53 respectively for lead 1 and lead 2 seasons (Table 3).

From the tropical Pacific, the positive significant correlations tend to irradiate towards the whole tropical belt and towards extra tropics (Fig. 4a,b). Some positive correlations are also found in land regions strongly influenced by ENSO teleconnections (Shukla and Wallace 1983; Trenberth et al. 1998). During boreal winter, such regions are identified over Central and Southern Africa, the Amazon basin and South Eastern South America, the Middle latitude Western Pacific Coasts of Asia and North-Western as well as North-Eastern North America. Boreal summer positive correlations are found in Northern Australia and Indonesian archipelago, Gulf of Mexico and Central America and over south-eastern Asia. Remarkably, significant correlations between forecasts and observed temperatures are found in the Euro-Mediterranean region as well as Middle East during boreal summer, suggesting some predictability over these regions for the forecasts started in May. However, most areas evidencing significant correlations are found over the oceans (Fig. 4a,b). This indicates a

lower predictability over lands and it may follow at least in part from the absence, in the experiments we performed, of any kind of assimilation in order to suitably initialize the long persisting land variables (e.g: Koster and coauthors 2004, 2006; Ferranti and Viterbo 2003; Alessandri and Navarra 2008).

Although the Niño3.4 index shown in Fig. 5 is frequently used to characterize the state of the ENSO and to quantify the quality of simulations and predictions of the oscillation, it describes the averaged SST over only a small portion of the Pacific ocean. The ocean anomalies associated with ENSO, on the other hand, affect the whole tropical basin with the development of wide scale anomaly patterns. It is therefore of interest to check the skill of the forecasting system to predict the evolution of the SST pattern anomalies over the entire tropical Pacific. To this aim, spatial Anomaly Correlation Coefficients (ACCs) and spatial Root Mean Squared Errors (RMSEs) for the forecast SST anomalies in the tropical Pacific (defined as 25°S-25°N, 140°-280°E) have been computed. Figure 6 reports the RMSE (panels a and b) and the ACCs (panels c and d) for the forecasted SST anomalies in the tropical Pacific together with the results obtained for persistence forecasts (dashed line). Thick solid lines (and filled circle marks) are the ensemble means while thin lines stand for each ensemble member of the DAS experiment. Filled triangles also reports the results for the NODAS ensemble mean forecasts—see the next subsection for a comparison between DAS and NODAS. The ACCs and the RMSEs have been computed relative to the ERA-Interim SST and persistence forecasts are made by persisting ERA-Interim monthly anomaly from the month prior to the start date of the model forecasts. For example, SST persistence forecasts for the period May-September 1991 have been made by persisting the SSTA found for the observed April 1991. The results in Fig. 6 indicate that the ensemble mean forecasts

have better skill than any ensemble member. This is in agreement with the results found with other coupled model forecast systems, and can be explained, at least in part, by the fact that the ensemble average reduces the internal dynamics noise present in the individual forecasts (Kirtman and Shukla 2002), thus potentially increasing the correlations and reducing the RMSE.

Overall, DAS displays a good skill in reproducing the observed SST over tropical Pacific, with the ensemble mean forecasts (hereafter, simply 'forecasts') performing usually better also than persistence forecasts (dashed lines), especially on lead times greater than one month. However, both the anomaly correlations and the RMSEs show some seasonal dependence, with some evidence of the so-called 'spring predictability barrier' in our system. Higher capability to beat persistence forecasts, in fact, are found for the forecasts that start in May compared to the forecasts starting in November. These results are consistent with several previous studies (e.g: Gualdi et al. 2004; Schneider et al. 2003).

b. Sensitivity to ocean assimilation

Figure 4 compares the correlations between observed TAIRA and respectively DAS (panels a and b) and NODAS (panels c and d) 2-to-4 months forecasts. Panels e and f report the DAS minus NODAS difference in correlations, with the shaded areas evidencing grid points with significant (10% level, bootstrap method) increase (red) and decrease (blue) in DAS. To better illustrate the skill difference as shown between panels a/b vs c/d of Fig. 4, we report in Table 4 for both DAS and NODAS the fractional areas with significant correlations as well as with correlations exceeding given thresholds. Overall, the global area fraction with

significant correlations between November forecasts and observations is 0.60 in DAS and only 0.55 in NODAS (Table 4). For the forecasts started in May the difference in global fractional areas reduces at 0.03 (0.61 in DAS vs. 0.58 in NODAS). Considering the areas with correlations above 0.8, DAS still outcompete NODAS with 0.08 vs. 0.05 of the global area exceeding this threshold in the forecasts started in November (Table 4). In contrast, the boreal summer forecasts do not display any difference in fractional area with correlations above 0.8 (0.05 in both DAS and NODAS). The areas evidencing significant correlation difference between DAS and NODAS (Fig. 4e, f) are mostly placed over subtropical and middle latitude oceans. For the forecasts started in November (Fig. 4f), DAS displays areas with significant increase over equatorial central Pacific, subtropical central South Pacific, subtropical Indian and Atlantic oceans as well as Northern Hemisphere western boundary currents regions. DAS evidences some significant correlation increase also over continents, and in particular the coastal areas adjacent to North Pacific and Atlantic oceans. The May start date forecasts has fewer grid points displaying significant correlation improvements in DAS (Fig. 4e), which appears to be outperformed by NODAS over large areas surrounding the Kuroshio current. On the other hand DAS is significantly better than NODAS over the North Eastern Pacific.

1) TROPICAL PACIFIC

The prediction of SST over the tropical Pacific appears to be affected by the assimilation of temperature and salinity with stronger improvement for the November start date. From Fig. 6 (panels b and d) it is clear that DAS (filled circles) improves noticeably compared to

NODAS (filled triangles), particularly in terms of ACCs from the third month on. Less clear is the impact for the May start date (panels a and c) where both the ACC and RMSE values appear to be only slightly affected. While DAS improves to some extent in the forecast months 1 and 4, the results are uncertain for month 5 and the skill is almost identical for months 2 and 3.

The significance of the above results can be more clearly evaluated by means of scatterplots of the ACCs and RMSEs computed on the one month lead-time seasonal mean predictions (averages of forecast months from 2 to 4: June, July and August for May 1st and December, January and February for November 1st start dates). Figure 7 compares the forecasts performed with assimilated ICs (DAS) with the control (NODAS) in the tropical Pacific. ACCs (upper panels) and RMSEs (lower panels) for each forecast year (diamonds) as well as the average of the values over all the 13 forecast years (cross marks) are displayed. In some years, an increase of the ACC for DAS in both May and November is visible (Fig. 7). Using a montecarlo bootstrap procedure we checked the significance of the difference in the 13 year means (table 5). We found that the 5% level of significance is verified only for the November case. The ACCs for the lead 2 seasons (averages of forecast months from 3 to 5) display similar results (table 5), evidencing a significant improvement in DAS for the November start dates. For what concern the impact of the assimilation of temperature and salinity profiles on the SST RMSE, it appears to be smaller than that for ACCs, with the 13 year averages which do not pass the significance test for the difference at the 5% level. For completeness, we checked the results also over the Niño3.4 region (Table 6). Compared to the tropical Pacific the ACCs decrease considerably, reflecting the fact that this smaller region has a strong SST signal and may develop anomaly patterns characterized by steep

gradients. Nevertheless, for the lead 1 season the effect of subsurface assimilation is similar to what evidenced by the comparison over the whole tropical Pacific basin. In contrast for the lead 2 season, our analysis do not evidence any significant difference at the 5% level between DAS and NODAS over the Niño3.4 region (Table 6).

A synthesis of the time-space SST variability over the tropical Pacific is reported in table 7. For each forecast month, the standard deviations for the ensemble mean SST anomalies in DAS (1st column), NODAS (2nd column) and for the observations (3rd column) are computed retaining both the space and interannual time variabilities. Compared to NODAS, in both May and November start dates the DAS ensemble mean predictions variability is closer to the observed value, in particular for the forecast months from 3 to 5. This is due to the fact that DAS do not display the marked progressive weakening of the predicted ensemble mean anomalous signal which characterizes NODAS. This result indicates an increased signal-to-noise ratio of the forecasts performed in DAS from month 3 to 5, driven by the improved subsurface initialization of the ocean. As discussed above, this appears to produce an increased skill only for the forecasts started in November, while for May the enhanced signal appears not to correspond to a better fit to the observed anomalies.

2) EL NIÑO 1997-1998

The Niño3.4 index for all start dates in the forecasts performed with NODAS (black) and DAS (red) is reported as box plot time series in Fig. 8. The distribution of predicted monthly mean anomalies is represented by boxes (25th-75th percentiles) and the median is represented by the inside box mark. The shaded band includes the interannual standard

deviation in the observations while the dashed lines refer to the forecasts. DAS appears to better represent the observed (green filled circles) anomalies in the Niño3.4 index compared to NODAS (correlation coefficient 0.94 vs 0.91; see also Table 8 for the comparison of the skill in the lead 1 and lead 2 seasons for each start date). The evolution of the two major El Niño events (1991-1992 and 1997-1998) appear to be better represented in DAS. Particularly, DAS considerably improves the onset amplitude of El Niño 1997-1998. The strength and length of the El Niño 1991-1992, staying well above one standard deviation till May 1992 in the observations, is also more seemly represented in DAS. However, DAS and NODAS appear to anticipate the onset of El Niño 1991-1992 both providing a false alarm in the forecast started May 1991. The considered forecasting period also includes the 1994-1995 and 2002-2003 moderate intensity El Niño events. They appear to be quite well predicted by both DAS and NODAS with the 1994 El Niño onset better captured in NODAS. On the other hand the 2002 onset appears to be slightly better represented in DAS.

The observed El Niño 1997/1998 onset amplitude is very well reproduced by DAS, with an observed averaged anomaly of about 2.3K in September 1997. In contrast NODAS reaches an anomaly of only 1.5K at that time. The Hovmöller diagrams of the evolutions of the heat content anomaly (shaded) and zonal wind stress anomaly (contours) for NODAS (panel a), DAS (panel c) and the ocean analysis (panel b) is reported in Fig 9. Note that for the model ensemble mean anomalies are reported, thus averaging out the "non signal variability" intrinsically working at the shorter time and smaller space scales (Kirtman and Shukla 2002). As shown in Fig. 9, DAS well represents the initial positive heat content anomaly along the equatorial Pacific and off the Peru coast. Similarly, the Kelvin wave train travelling eastward through the tropical Pacific as well as the associated heat content anomaly propagation

appears to be captured correctly by DAS. In contrast, the NODAS heat content anomaly is much weaker than observed and the Kelvin wave signal appears to be less clear and somewhat delayed compared with observations. This results in a reduced development and west-to-east propagation of the heat anomaly and in the consequent weak simulation of the onset of the El Niño 1997-1998 in NODAS. It is important to note here that this weak initial development of the El Niño 1997-98 has been pointed out to be a major problem for most of the dynamical as well as statistical ENSO forecast models (e.g: McPhaden 1999; Vitart et al. 2003).

Interestingly, the second Kelvin wave train do not appear to be captured by DAS, probably because of uninitialized intra-seasonal wind bursts (McPhaden 1999). Nevertheless the initial heat content anomaly to the east of the dateline is correctly propagated eastward in this experiment. The resulting heat content prediction is still close to observed at the end of the DAS forecast.

5. Probabilistic forecasts of dichotomous predictands

In the previous section we have evaluated the impact of the improvement of the oceanic ICs estimation on the prediction of the tropical Pacific SST, which represents the main source of climate predictability at the seasonal time scale (e.g: Ji and Leetmaa 1995; Trenberth et al. 1998; Wallace et al. 1998). In the following we will give an estimate of the associated global scale impact on the performance of probabilistic forecasts of dichotomous observed predictands. In particular, we will concentrate on the forecasting of below normal (i.e: below lower tercile of the sample distribution) and above normal (i.e: above upper tercile of the

sample distribution) observed surface air temperature (hereinafter TAIRA) and observed precipitation (hereinafter PRECA).

The joint distribution of forecasts and observed dichotomous predictands can be conveniently analyzed and displayed graphically through the likelihood-base rate factorization (Wilks 2006):

$$\begin{aligned}
 p(y_i, o_j) &= p(y_i|o_j) \cdot p(o_j) & (1) \\
 i &= 1, \dots, I \\
 j &= 0, 1
 \end{aligned}$$

where the conditional distributions $p(y_i|o_j)$ express the likelihoods that each of the allowable forecast value y_i would have been issued in advance of each of the observed dichotomous event o_j (occurrence $j = 1$; no occurrence $j = 0$). Together with the associated sample climatological probabilities, $p(o_j)$, it completely represents the information of the full joint distribution. Specifically, the conditional likelihood distributions, $p(y_i|o_j)$, are directly indicative of how well a set of forecasts are able to discriminate among the events o_j . Graphically, this can be appreciated through diagrams consisting of superimposed plots of the two likelihood distributions as a function of the forecast probability, y_i (hereinafter discrimination diagrams). As pointed out by Wilks (2006), the above mentioned characteristics of the two likelihood distributions could be used effectively to recalibrate the probability forecasts by calculating posterior probabilities for the two events given each of the possible forecast probabilities.

Differently from previous sections which were mostly focused on climatology and pre-

dictability of seasonal means, here we focus on the probabilities of monthly means to exceed tercile thresholds and on the capability of the forecasts to discriminate the corresponding observed outcomes. To this aim we retain all the forecast months from 2 to 5 in the analysis that follows. Figure 10 shows a comparison between DAS and NODAS of the discrimination diagrams evaluated over the tropics (0-360E; 25S-25N) and considering all the monthly means from month 2 to 5 of the forecasts started in November (panels a and b; December, January, February and March monthly means considered) and in May (panels c and d; June, July, August and September monthly means considered). The left panels (a and c) relates to the forecast probabilities of the event of TAIRA being below the lower tercile of the climatological sample distribution (E_T^-). In contrast, panels b and d refer to the forecast probabilities to exceed the upper tercile (E_T^+). In Fig. 10 the dashed lines represent the likelihood distribution given the no occurrence of the event [$p(y_i|o_0)$], while solid lines are the likelihood distributions verified o_1 [$p(y_i|o_1)$]. Both NODAS in red and DAS in blue are displayed in the same diagram. In the forecasts started in November (panels a and b), DAS displays larger $p(y_i|o_0)$ values (dashed lines) for the smaller forecast probabilities for both E_T^- and E_T^+ . Similarly, conditional probabilities given o_1 (solid lines) are higher in DAS compared to NODAS for the larger probability forecast outcomes. This determines an increase in DAS of the separation between the two respective likelihood distributions, indicating an improved ability of the forecasts to discriminate warm and cold events.

The separations of the two likelihood distributions are also plotted in the same figure (bottom horizontal bars) in the form of discrimination distances (d), a scalar attribute defined as the difference between the means of the two likelihood distributions (μ) following Wilks

(2006):

$$d = | \mu_{y|o_1} - \mu_{y|o_0} | \quad (2)$$

In Fig. 10, star marks are placed in correspondence of the discrimination values to indicate that they are significantly (at the 5% level, using a montecarlo bootstrap method) higher than the other experiment. DAS increases significantly d compared to NODAS for both $E_{T,Nov}^-$ and $E_{T,Nov}^+$ (Fig. 10a, b). This result indicates that, for the forecast started in November, the assimilation of temperature and salinity profiles improves the ability of the model to discriminate between the occurrence of below normal or above normal events over the tropics. Differently from the November start date case, the forecasts started in May (Fig. 10c, d) do not display any significant (5% level) difference in d . In fact, considering both $E_{T,May}^-$ and $E_{T,May}^+$, the two likelihood distributions, compared between DAS and NODAS, are respectively very close to each other, resulting in discrimination distances almost identical.

The panel insets in the discrimination diagrams in Fig. 10 also report the refinement distributions, $p(y_i)$. The dispersion of $p(y_i)$ reflects the overall confidence of the forecasts, so that forecasts that deviate rarely from their average value exhibit little confidence (Wilks 2006). Interestingly, all the refinement distributions in Fig. 10 evidence an increased confidence in DAS. This is verified even when the discrimination distance is not affected considerably, as it is the case for the May start date. This result indicates that the addition of subsurface informations to the ocean IC tends to increase the signal-to-noise ratio of the predictions and consequently drives the predictions to fall preferentially outside normal conditions. However, consistently with the results in Section 4, this produces increased discriminations only for the November start date, while for May the enhanced signal does

not appear to coincide with better correspondence to observations. Table 9 summarizes, for both November and May, the discrimination distances for $E_{T,P}^-$ and $E_{T,P}^+$ (where T and P indicates Temperature and Precipitation, respectively) in the tropics, northern extra tropics and southern extra tropics. Considering precipitation over tropics, DAS show an enhanced $d_{P,Nov}^+$ (5% significance level) and, interestingly, also the forecasts started in May display an improvement due to the initialization of the subsurface ocean. In fact $d_{P,May}^-$ increases significantly compared to NODAS.

From table 9, it is shown that discrimination distances outside the tropics tend to decrease to values close or below 0.1 in both DAS and NODAS. Nevertheless, extra tropics discrimination appears to be affected by the assimilation of temperature and salinity as well. Considering winter northern extra tropics TAIRA, $E_{T,Nov}^-$ and $E_{T,Nov}^+$ are improved considerably in DAS compared to NODAS (5% level significance verified). Differently, the PRECA field discrimination is affected during boreal summer in the northern extra tropics, with $d_{P,May}^+$ which increases significantly in DAS. Similarly to the Northern Hemisphere case, the comparison of the d values for the southern extra tropics display improvements for TAIRA during boreal winter. In this case $d_{T,Nov}^-$ is increased in DAS compared to NODAS. Interestingly, $d_{T,May}^+$ displays an higher value in NODAS (not significant however). This is probably due to the fact that only a very small number of profiles is available in the southern extra tropics during austral winter and spring for initialization purposes (see table 2).

Table 10 reports the comparison of the discrimination distances between DAS and NODAS for the three tropical ocean basin sectors. For the tropical Pacific, results emphasize the analysis performed in Section 4. Compared to NODAS, the DAS discrimination distances over this ocean basin increase significantly (5% level) in the November start date forecasts for

all the cases considered ($E_{T,Nov}^+$, $E_{T,Nov}^-$, $E_{P,Nov}^-$ and $E_{P,Nov}^+$). Differently, for the May start dates the discriminations appear to be little affected, with no significant differences between DAS and NODAS over tropical Pacific. On the other hand, the assimilation of subsurface observations leads to the enhancement of the May prediction performance in the Indian Ocean. A significantly increased discrimination has been verified for $E_{T,May}^+$, $E_{P,May}^-$ and $E_{P,May}^+$ (Table 10).

Tropical Atlantic behaves differently from the other tropical oceans. Considering the May start date, in this basin the assimilation of subsurface temperature and salinity leads to a clear worsening of the predictions in terms of discrimination distance. In fact, the discriminations for $E_{T,May}^-$ and $E_{P,May}^+$ increase significantly in NODAS compared with DAS (Table 10). From Fig. 11a it is shown a clear reduction of the overlapping between $p(y_i|o_0)$ and $p(y_i|o_1)$ in NODAS compared with DAS and this ends up in a significantly (5% level) higher $d_{T,May}^-$ value. Fig. 11d documents the increased NODAS discrimination for $E_{P,May}^+$. Noteworthy, in this case the refinement distribution (inset histogram) show much less confidence than for the temperature cases, and correspondingly the discrimination values appear to be reduced. Nevertheless, there is a considerable improvement of NODAS compared to DAS. The forecasts over tropical Atlantic started in November (Fig. 12) display improved temperature discriminations in DAS. In fact, $E_{T,Nov}^-$ and $E_{T,Nov}^+$ are significantly better discriminated by DAS (Fig. 12a, b). In order to understand the opposite behavior of boreal winter and boreal summer forecasts over tropical Atlantic, we compared the subsurface thermal climatology of the ICs taken from the ocean analysis with a long free simulation of the coupled model (radiative boundary conditions of the 1991-2003 period were used). We found that the equatorial Atlantic subsurface thermal structure is very badly represented in the

coupled model during boreal spring and early summer (not shown). During these seasons the coupled model shows an opposite slope of the thermocline with respect to observations. This result suggests that the subsurface correction, due to the data assimilation during spring, drives the coupled model too far from and not "in balance" with the state it would have and thus leading to a negative impact on the forecasts started in May.

Table 11 compares the DAS and NODAS discriminations over the Pacific North American (150E-300E; 40N-65N; PNA) and the Euro-Atlantic (80W-40E; 35N-65N) regions. For the PNA region, the results indicate increased discriminations in the forecasts started in November for the TAIRA prediction. Differently, the performances for the precipitation are in general very little affected. Even if discriminations appear to decrease considerably compared to tropical regions, the refinement distributions show good confidences for both November (Fig. 13a, b) and May (Fig. 13c, d) start dates. In this context DAS exhibits a considerably higher sharpness of the forecast distribution. This result appears to drive the increased discrimination in DAS, which is verified for significance (5% level) for $E_{T,Nov}^+$ and $E_{T,Nov}^-$.

Figure 14 is the same as Fig. 13 but considering the Euro-Atlantic region. Compared to Fig. 13, it is noted that discrimination performances over this area is considerably lower than for PNA (see also table 11). However, for the forecasts starting in November, DAS improves the TAIRA results if compared with NODAS. In particular, for the E_T^+ case there is a significant (5% level) increase (2.5 times) of the discrimination from the very low value of 0.011 to 0.026, leading to a noticeable increase in the performance. This is an important result, as it leads to increased predictability in the region. However, increased discrimination is achieved by reducing the overlapping in the "not conventional" direction: in fact DAS

displays lower $p(y_i|o_0)$ values for the smaller forecasts probabilities and the $p(y_i|o_1)$ are smaller for the larger forecasts probabilities. From a physical point of view, this is of course deprecable, as it means that the model tends to represent the opposite than reality in this region. Nevertheless, as soon as it is systematic, it is as well useful for predictions.

6. Conclusions

The forecasts performed using the assimilation of subsurface temperature and salinity profiles during ocean initialization (DAS experiment) has evidenced an enhanced signal to noise ratio of the predicted surface temperature anomalies. In particular, in the tropical Pacific the magnitude of the time-space variability of the ensemble-mean SST anomalies predicted by DAS appears to be very well simulated when compared with observations. It is improving over NODAS (i.e: without subsurface assimilation during initialization), which underestimate considerably the predicted anomalous SST in the tropical Pacific from the third forecast month on. In terms of ACCs and RMSEs, subsurface initialization improves the November start date forecasts over tropical Pacific. In fact, the averaged (over the thirteen forecast years) ACCs computed on the seasonal mean anomalies increase significantly (5% level) in the tropical Pacific for the forecasts started in November (considering the one month lead-time predictions it increases by 8% with respect to NODAS). The impact on the SST RMSE is smaller (5% reduction for the one month lead-time predictions) and the significance in this case can only be verified at the 10% level. In contrast, for the forecasts started in May ACC and RMSE values are only slightly affected by the subsurface oceanic initialization and the averages over all forecast years are not significantly modified.

NODAS shows the tendency to underpredict the ENSO anomalies. This has been reported as a major problem also for other dynamical as well as statistical ENSO forecast models (e.g: McPhaden 1999; Vitart et al. 2003). In particular, when no subsurface data are assimilated, our system appears to considerably underestimate the development of El Niño 1997/1998. In contrast, the assimilation of temperature and salinity profiles in DAS leads to a considerable improvement of the representation of the initial positive heat content anomaly along the equatorial Pacific and off the Peru coast. Similarly, the observed Kelvin wave train travelling eastward through the tropical Pacific as well as the associated heat anomaly propagation appears to be captured correctly by DAS. These results further highlight the importance of subsurface data assimilation for the improvement in forecasting the development and evolution of El Niño events .

The assimilated ocean initial conditions, and the resulting enhanced tropical Pacific prediction, has been shown to improve the probabilistic predictions of dichotomous events globally. Considering the tropical belt, DAS displays significantly (5% level) enhanced capability to discriminate both warm (above upper tercile of the sample distribution, E_T^+) and cold (below lower tercile, E_T^-) surface air temperature events in the forecasts started in November. For boreal winter, the discrimination distance of anomalous temperature forecasts increases by 7% for E_T^- and by 9% for E_T^+ over the tropics. In some cases, the enhancement in the discrimination distance has been shown to be significant (5% level) not only for temperature but also for precipitation. In particular, wet events appear to be significantly better discriminated during boreal winter (12% increase with respect to NODAS), while dry events improve in boreal summer (6% increase with respect to NODAS). Differently from November, the forecasts started in May do not display a significant increase in the discrimination

of anomalous temperature events considering the whole tropics. And this even though the refinement distributions show increased confidence of the forecasts in a way similar to the November case. This result indicates that the addition of subsurface informations to the ocean IC tends to increase the signal-to-noise ratio of the forecasts and consequently drives the predictions to fall preferentially outside normal conditions. However, this improves the skill in discrimination only for the November start dates, while for May the enhanced signal does not appear to coincide with significantly better correspondence to observations.

The Tropical Atlantic is found to behave differently from the other tropical basins, with a marked and significant worsening in DAS of the discrimination distance for the prediction of cold and wet events during boreal summer. By contrast, the November forecasts improve the prediction of the anomalous temperature quite similarly to the other tropical basins. The results in the Atlantic are in agreement with previous studies evidencing the tendency to a degradation of the prediction skill over tropical Atlantic when subsurface data assimilation is used (e.g: Vidard et al. 2006; Balmaseda et al. 2007). Our analysis suggests that the subsurface correction due to the assimilation of observed profiles in spring leads to a change too drastic in the tropical Atlantic and drives the system too far from the state that the coupled model would have there—in this season the slope of the equatorial Atlantic thermocline in free coupled-model long runs is opposite than observed—thus leading to a negative impact on the forecasts started in May.

The assimilation of subsurface data in the ocean IC impacts the extra tropics as well in the forecasts started in November. For the boreal winter forecasts, the discrimination distances appear to increase significantly (5% level) for anomalous temperature events in both northern middle latitudes (compared to NODAS, the discrimination of both warm and

cold events increase respectively by 24% and 38%) and southern middle latitudes (cold events discrimination increases by 10%). The Pacific-North American sector displays the strongest improvement in the Northern Hemisphere middle latitudes. In this region, a significant increase in the discrimination for both warm (46% increase with respect to NODAS) and cold (39% increase with respect to NODAS) events for the forecasts started in November is evidenced. Remarkably, DAS increases the capability to discriminate warm dichotomous events during boreal winters even in the Euro-Atlantic region. However, this is achieved through smaller conditional likelihoods for the smaller forecast probabilities, given the non-occurrence of the event. Concurrently, when the event occurs, smaller conditional likelihoods correspond to the higher forecast probabilities.

Dynamical Weather and Climate Prediction is challenging: progress did not occur in the last 30 years because of drastic breakthroughs, but because of slow incremental progresses and through a great deal of hard work (Shukla and Kinter 2006). This study has evidenced beneficial effects on the boreal winter forecasts deriving from the subsurface initialization of the ocean model in our prediction system. However, the impact of the ocean data assimilation is small and mostly negligible for the forecasts started in May. Large SST biases in the central tropical Pacific and quite a prominent initial "coupling shock" have been evidenced in our system. With this regard, more efforts are needed on the initialization of the coupled model and on reducing SST biases as they are probably limiting the positive impact of subsurface assimilation in our system.

In this study we considered the climate predictability using our SPS at scales up to the first 5 months and mostly focusing on the one month lead-time seasonal (months 2-4) forecasts. The predictability of seasonal climate at longer lead times is also an important

issue that requires a large effort and it will be considered in future works. Furthermore, given the relatively short prediction scales considered in this work, the atmospheric initial conditions may have an impact on the skill of the system. Recently, some dynamical forecasts of Intraseasonal Oscillations produced rather credible simulations of the Madden Julian Oscillation, with evidence of some prediction skill out to a lead-time of about 2 weeks (Kim et al. 2007; Vitart et al. 2007). Whether or not this kind of atmosphere initialization can affect the prediction of ENSO and of climate at seasonal time scale still needs to be evaluated.

Acknowledgments.

Many thanks for the generous support from several CMCC staff. We especially thank Alessio Bellucci, Fabrizio Massari, Loredana Amato, Alberto Troccoli, Antje Weisheimer and Francisco D. Reyes for the help and for the inspiring discussions.

REFERENCES

- Alessandri, A. and A. Navarra, 2008: On the coupling between vegetation and rainfall inter-annual anomalies: Possible contributions to seasonal rainfall predictability over land areas. *GRL*, **35**, L02 718, doi:10.1029/2007GL032415.
- Alves, O., M. Balmaseda, D. Anderson, and T. Stockdale, 2004: Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *QJRMS*, **130**, 647–667.
- Balmaseda, M., D. Anderson, and A. Vidard, 2007: Impact of argo on analyses of the global ocean. *GRL*, **34**, L16 605, doi:10.1029/2007GL0304452.
- Bellucci, A., S. Masina, P. D. Pietro, and A. Navarra, 2007: Using temperature-salinity relations in a global ocean implementation of a multivariate data assimilation scheme. *MWR*, **135**, 3785–3807.
- Berrisford, P., D. Dee, K. Fielding, M. Fuentes, P. Kallberg, S. Kobayashi, and S. Uppala, 2009: The era-interim archive. Technical report ERA Report Series No.1, ECMWF, 16 pp.
- Cane, M., S. Zebiak, and S. Dolan, 1986: Experimental forecasts of el niño. *NAT*, **321**, 827–832.
- Chen, D., S. Zebiak, and A. Busalacchi, 1995: An improved procedure for el niño forecasting: Implications for predictability. *SCI*, **269**, 1699–1702.

- Ferranti, L. and P. Viterbo, 2003: The european summer of 2003: Sensitivity to soil water initial conditions. *JCLI*, **19**, 3659–3680.
- Gualdi, S., A. Alessandri, and A. Navarra, 2004: Impact of atmospheric horizontal resolution on el niño southern oscillation forecasts. *TELLUS*, **57**, 357–374.
- Gualdi, S., E. Guilyardi, A. Navarra, S. Masina, and P. Delecluse, 2003a: The interannual variability in the tropical indian ocean as simulated by a cgcm. *CDYN*, **20**, 567–582.
- Gualdi, S., A. Navarra, E. Guilyardi, and P. Delecluse, 2003b: Assessment of the tropical indo-pacific climate in the sintex cgcm. *Ann. Geophys.*, **46**, 1–26.
- Ingleby, B. and M. Huddleston, 2007: Quality control of ocean profiles: historical and real-time data. *JMS*, **65**, 158–175.
- Ji, M. and A. Leetmaa, 1995: Impact of data assimilation on ocean initialization and el niño prediction. *MWR*, **125**, 742–753.
- Jin, E. and J. Kinter, 2008: Characteristics of tropical pacific sst predictability in coupled gcm forecasts using the ncep cfs. *CDYN*, **32**, 675–691, doi:10.1007/s00382-008-0418-2.
- Kim, H., I. Kang, B. Wang, and J. Lee, 2007: Interannual variations of the boreal summer intraseasonal variability predicted by ten atmosphere. *CDYN*, **30**, 485–496, doi:10.1007/s00382-007-0292-3.
- Kirtman, B. P. and J. Shukla, 2002: Interactive coupled ensemble: A new coupling strategy for cgcms. *GRL*, **29**, 17–20, doi:10.1029/2002GL014834.

- Koster, R. D. and coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *SCI*, **305**, 1138–1140.
- Koster, R. D. and coauthors, 2006: Glace: The global land-atmosphere coupling experiment. part i: Overview. *JHM*, **7**, 590–610.
- Latif, M., et al., 1998: A review of the predictability and prediction of enso. *JMS*, **103**, 14375–14393.
- Luo, J.-J., S. Masson, E. Roeckner, G. Madec, and T. Yamagata, 2005: A review of the predictability and prediction of enso. *JCLI*, **18**, 2344–2360.
- McPhaden, M. J., 1999: Genesis and evolution of the 1997-98 el niño. *SCI*, **283**, 950–954.
- McPhaden, M. J., et al., 1998: The tropical ocean-global atmosphere observing system: A decade of progress. *JGR*, **103**, 14169–14240.
- Navarra, A., 2002: Ensembles, forecasts and predictability. *Ocean forecasting: conceptual basis and applications*, N. Pinardi and J. Woods, Eds., Springer-Verlag, 131–148.
- Navarra, A., et al., 2008: Atmospheric horizontal resolution affects tropical climate variability in coupled models. *JCLI*, **21**, 730–750.
- Palmer, T., 2006: Predictability of weather and climate: from theory to practice. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 1–29.
- Palmer, T., et al., 2004: Development of a european multi-model ensemble system for seasonal to inter-annual prediction (demeter). *BAMS*, **85**, 853–872.

- Pietro, P. D. and S. Masina, 2009: The cmcc-ingv global ocean data assimilation system (cigodas). Technical Report rp0071, Centro euro-Mediterraneo per i Cambiamenti Climatici, 39 pp.
- Rayner, N. A., P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J. Ansell, and S. F. B. Tett, 2005: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The hadisst2 dataset. *J. Climate*, **19**, 446–469.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analysis of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **18**, doi:10.1029/2002JD002670.
- Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Richardson, D., 2006: Predictability and economic value. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 628–644.
- Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, and co authors, 1996: The atmospheric general circulation model echam-4: model description and simulation of present-day climate. Technical Report 218, Max-Planck-Institut für Meteorologie, 94 pp.
- Rosati, A., R. Gudgel, and K. Miyakoda, 1997: The impact of ocean initial conditions on enso forecasting with a coupled model. *Mon. Wea. Rev.*, **5**, 754–772.

- Schneider, E., D. DeWitt, A. Rosati, B. Kirtman, L. Ji, and co authors, 2003: Retrospective enso forecasts: sensitivity to atmospheric model and ocean resolution. *Mon. Wea. Rev.*, **131**, 3038–3060.
- Shukla, J. and J. C. Kinter, 2006: Predictability of seasonal climate variations: a pedagogical review. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 306–341.
- Shukla, J. and J. M. Wallace, 1983: Numerical simulation of the atmospheric response to equatorial pacific sea surface temperature anomalies. *J. Atmos. Sci.*, **40**, 1613–1630.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N. Lau, and C. Ropelewski, 1998: Progress during toga in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, **103**, 14 291–14 324.
- Tribbia, J. and A. Troccoli, 2008: Getting the coupled model ready at the starting blocks. *Seasonal Climate: Forecasting and Managing Risk*, DLT and M. SJ, Eds., NATO Science Series, Springer Academic Publishers, 91–126.
- Uppala, S. and co authors, 2005: The era-40 reanalysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.
- Valcke, S., L. Terray, and A. Piacentini, 2000: The oasis coupler user guide version 2.4. Technical Report TR/CMGC/00-10, CERFACS, 85 pp.
- Vidard, A., D. Anderson, and M. Balmaseda, 2006: Impact of ocean observation systems on ocean analysis and seasonal forecasts. *Mon. Wea. Rev.*, **135**, 409–429, doi:10.1175/MWR3310.1.

- Vitart, F., M. Balmaseda, L. Ferranti, and D. Anderson, 2003: Westerly wind events and the 1997/98 el niño event in the ecmwf seasonal forecasting system: A case study. *J. Climate*, **16**, 3153–3170.
- Vitart, F., S. Woolnough, M. Balmaseda, and A. Tompkins, 2007: Monthly forecast of the madden-julian oscillation using a coupled gcm. *Mon. Wea. Rev.*, **135**, 2700–2715.
- Wallace, J. M., E. M. Rasmusson, T. P. Mitchell, W. E. Kousky, N. S. Sarachik, and H. von Storch, 1998: On the structure and evolution of enso-related climate variability in the tropical pacific: Lessons from toga. *J. Geophys. Res.*, **103**, 14 241–14 259.
- Wang, G., R. Kleeman, N. Smith, and F. Tseitkin, 2002: The bmrc coupled general circulation model enso forecast system. *Mon. Wea. Rev.*, **130**, 975–991.
- Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences Second Edition*. Academic Press, 630 pp.
- Xie, P. and P. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.
- Zebiak, S. and M. Cane, 1987: A model of the el niño-southern oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278.

List of Tables

- 1 Data sources in the EN3_v1c dataset, which is an assembling of World Ocean Database '05 (WOD05), Global Temperature-Salinity Program (GTSP) and Argo database. The left column list the instrument codes as reported in the respective database documentation. We also reported start and end year of availability of each data source for the assimilation we performed to produce Ocean initial conditions. 41
- 2 Fractional space coverage of the quality checked EN3 Temperature profiles used for this work. The annual cycle is divided into the May-to-October and the November-to-April time periods. The temperature profiles have been distributed to a $0.5^{\circ} \times 0.3^{\circ}$ global map of monthly means before computation of the averaged fraction for each region. We considered as covered the grid points with at least one measure available in the first 100 mt ocean depth in period 1991-2003. 42
- 3 Skill in the prediction of the monthly Niño3.4 index (averaged temperature anomaly over 5S-5N; 190E-280E) for the DAS experiment. Monthly time correlations and RMS Errors are computed against ERA Interim for both one month lead-time (upper rows) and two months lead-time (lower rows) predicted seasons. All the forecast months from 2 to 4 (one month lead-time) and months from 3 to 5 (two months lead-time) have been used in the computation of the skill. 43
- 4 DAS vs. NODAS global area fractions with significant correlations (r) between modeled and observed TAIRA (first row; see also Fig. 4). Rows 2 to 4 also displays the fractional areas with correlations significant and above 0.5, 0.6 and 0.8, respectively. 44

- 5 Comparison between DAS and NODAS of the 13 years mean ACCs and RMSEs computed against ERA Interim over the tropical Pacific (25S-25N; 140E-280E) for one month lead-time (months 2-to-4; upper rows) and two months lead-time (months 3-to-5; lower rows) seasonal mean predictions. Using a montecarlo bootstrap procedure we checked the significance of the differences in the 13 year averages. Two asterisks (one asterisk) indicate the 5% (10%) significance level. 45
- 6 Same as table 5 but for the Niño3.4 region. 46
- 7 Standard deviations of the SST anomalies over the tropical Pacific (25°S-25°N; 140-280°E) for each forecast month: 1st and 2nd columns are the DAS and NODAS ensemble mean forecasts, respectively; observations are reported in the 3rd column. Both the space and interannual time variabilities are retained in the computations. 47
- 8 Comparison between DAS and NODAS of the prediction skill for the monthly Niño3.4 index (averaged temperature anomaly over 5S-5N; 190E-280E). Monthly time correlations and RMS Errors are computed against ERA Interim for both the one month lead-time (upper rows) and the two months lead-time (lower rows) seasons. All forecast months from 2 to 4 (one month lead-time) and months from 3 to 5 (two months lead-time) have been used in the computation of the skill. Using a montecarlo bootstrap procedure we checked the significance of the DAS vs NODAS differences. Two asterisks (one asterisk) indicate the 5% (10%) significance level. 48

9	<p>NODAS vs DAS discrimination distances for the tropics, northern extra tropics and southern extra tropics. The probability forecasts for above upper tercile ($E_{T,P}^+$) and below lower tercile ($E_{T,P}^-$) of the sample climatological distributions are reported for both temperature (left columns) and precipitation (right columns). Underlined bold discriminations stand for values which are significantly increased compared to the other experiment at the 5% significance level.</p>	49
10	<p>Same as table 9 but for tropical Pacific, tropical Indian and tropical Atlantic Oceans.</p>	50
11	<p>Same as table 9 but for Pacific North American (PNA) and Euro-Atlantic regions.</p>	51

Data sources in EN3_v1c dataset; Ingleby and Huddleston, 2007			
http://hadobs.metoffice.com/en3			
WOD05 World Ocean Database '05 (WOD05; Boyer et al. 2006).			
Data availability and documentation: http://www.nodc.noaa.gov/OC5/WOD05/pr_wod05.html			
Instrument Code	Description	Starting Year	Ending Year
OSD	Ocean Station Data	1958	2003
CTD	Conductivity-Temperature-Depth (Pressure) Data	1961	2003
XBT	Expendable Bathythermograph Data	1966	2003
PFL	Profiling Floats Data	1994	2003
MBT	Mechanical Bathythermograph Data	1958	2003
MRB	Moored Buoy Data	1980	2003
DRB	Drifting Buoy Data	1985	2003
UOR	Undulating Ocean Recorders Data	1992	2000
APB	Autonomous Pinniped Bathythermograph Data	1997	1999
GTSP Global Temperature-Salinity Profile Program			
(Data availability and documentation: http://www.nodc.noaa.gov/GTSP/)			
Instrument Code	Description	Starting Year	Ending Year
BA	BATHY radio message (temperature with depth)	1990	2003
BF	Undulating Oceanographic Recorder (e.g. Batfish CTD)	1996	2003
BO	Bottle temperature data	1990	2003
CD	Conductivity-Temperature-Depth (Pressure) down trace	1990	2003
CT	Conductivity-Temperature-Depth (Pressure) data, up or down	1990	1999
CU	Conductivity-Temperature-Depth (Pressure) up trace	1991	2003
DT	Digital Bathythermograph	1990	2003
MB	Mechanical Bathythermograph	1992	1993
PF	Profiling float	1992	2003
TE	Temperature Salinity Current (TESAC) radio message	1990	2003
TO	Towed Conductivity-Temperature-Depth (Pressure)	1993	1993
TR	Thermistor chain	1990	2003
XB	Expendable Bathythermograph	1990	2003
XC	Expendable Conductivity-Temperature-Depth (Pressure)	1999	1999
Argo (Data availability and documentation: http://www.argo.net)			
Instrument Code	Description	Starting Year	Ending Year
-	Argo profiling float	1999	2003

TABLE 1. Data sources in the EN3_v1c dataset, which is an assembling of World Ocean Database '05 (WOD05), Global Temperature-Salinity Program (GTSP) and Argo database. The left column list the instrument codes as reported in the respective database documentation. We also reported start and end year of availability of each data source for the assimilation we performed to produce Ocean initial conditions.

Profiles spatial availability		
Region	<i>November – to – April</i>	<i>May – to – October</i>
Tropics lat(25N-25S) - lon(0-360)	0.17	0.17
NorthernExtraTropics lat(30N-90N) - lon(0-360)	0.19	0.24
SouthernExtraTropics lat(90S-30S) - lon(0-360)	0.05	0.02
TropicalPacific lat(25N-25S) - lon(190E-240E)	0.17	0.17
IndianOcean lat(20N-25S) - lon(50E-110E)	0.1	0.11
TropicalAtlantic lat(40N-35S) - lon(60W-10E)	0.17	0.16
PacificNorthAmerican lat(40N-65N) - lon(150E-300E)	0.22	0.32
Euro – Atlantic lat(35N-65N) - lon(80W-40E)	0.78	0.83

TABLE 2. Fractional space coverage of the quality checked EN3 Temperature profiles used for this work. The annual cycle is divided into the May-to-October and the November-to-April time periods. The temperature profiles have been distributed to a $0.5^\circ \times 0.3^\circ$ global map of monthly means before computation of the averaged fraction for each region. We considered as covered the grid points with at least one measure available in the first 100 mt ocean depth in period 1991-2003.

Niño3.4 index prediction skill		
	May	November
	Lead 1 - months 2-4	
Correlation	0.94	0.97
RMSE	0.44	0.34
	Lead 2 - months 3-5	
Correlation	0.91	0.95
RMSE	0.53	0.38

TABLE 3. Skill in the prediction of the monthly Niño3.4 index (averaged temperature anomaly over 5S-5N; 190E-280E) for the DAS experiment. Monthly time correlations and RMS Errors are computed against ERA Interim for both one month lead-time (upper rows) and two months lead-time (lower rows) predicted seasons. All the forecast months from 2 to 4 (one month lead-time) and months from 3 to 5 (two months lead-time) have been used in the computation of the skill.

	Start date 1 May		Start date 1 November	
	DAS	NODAS	DAS	NODAS
significant r	0.61	0.58	0.60	0.55
significant with $r > 0.5$	0.29	0.27	0.30	0.25
significant with $r > 0.6$	0.19	0.17	0.22	0.16
significant with $r > 0.8$	0.05	0.05	0.08	0.05

TABLE 4. DAS vs. NODAS global area fractions with significant correlations (r) between modeled and observed TAIRA (first row; see also Fig. 4). Rows 2 to 4 also displays the fractional areas with correlations significant and above 0.5, 0.6 and 0.8, respectively.

Skill comparison in the Tropical Pacific				
	May		November	
	DAS	NODAS	DAS	NODAS
Lead 1 - months 2-4				
ACC	0.63	0.61	0.68**	0.63**
RMSE	0.36	0.36	0.39*	0.41*
Lead 2 - months 3-5				
ACC	0.58	0.58	0.61**	0.58**
RMSE	0.41	0.41	0.43	0.44

TABLE 5. Comparison between DAS and NODAS of the 13 years mean ACCs and RMSEs computed against ERA Interim over the tropical Pacific (25S-25N; 140E-280E) for one month lead-time (months 2-to-4; upper rows) and two months lead-time (months 3-to-5; lower rows) seasonal mean predictions. Using a montecarlo bootstrap procedure we checked the significance of the differences in the 13 year averages. Two asterisks (one asterisk) indicate the 5% (10%) significance level.

Skill comparison in the Niño3.4 region				
	May		November	
	DAS	NODAS	DAS	NODAS
Lead 1 - months 2-4				
ACC	0.51	0.50	0.47**	0.43**
RMSE	0.43	0.44	0.34*	0.39*
Lead 2 - months 3-5				
ACC	0.33	0.33	0.35*	0.33*
RMSE	0.53	0.51	0.43*	0.48*

TABLE 6. Same as table 5 but for the Niño3.4 region.

	SSTA Standard Deviation		
	DAS	NODAS	OBS
May	0.60	0.54	0.60
	0.58	0.51	0.56
	0.56	0.49	0.55
	0.57	0.50	0.57
	0.57	0.50	0.58
November	0.70	0.63	0.69
	0.70	0.62	0.72
	0.66	0.57	0.70
	0.63	0.52	0.65
	0.55	0.45	0.59

TABLE 7. Standard deviations of the SST anomalies over the tropical Pacific (25°S-25°N; 140-280°E) for each forecast month: 1st and 2nd columns are the DAS and NODAS ensemble mean forecasts, respectively; observations are reported in the 3rd column. Both the space and interannual time variabilities are retained in the computations.

Skill comparison of the Niño3.4 index				
	May		November	
	DAS	NODAS	DAS	NODAS
Lead 1 - months 2-4				
Correlation	0.94	0.93	0.97**	0.93**
RMSE	0.44	0.45	0.34*	0.39*
Lead 2 - months 3-5				
Correlation	0.91	0.90	0.95*	0.92*
RMSE	0.53	0.52	0.38	0.39

TABLE 8. Comparison between DAS and NODAS of the prediction skill for the monthly Niño3.4 index (averaged temperature anomaly over 5S-5N; 190E-280E). Monthly time correlations and RMS Errors are computed against ERA Interim for both the one month lead-time (upper rows) and the two months lead-time (lower rows) seasons. All forecast months from 2 to 4 (one month lead-time) and months from 3 to 5 (two months lead-time) have been used in the computation of the skill. Using a montecarlo bootstrap procedure we checked the significance of the DAS vs NODAS differences. Two asterisks (one asterisk) indicate the 5% (10%) significance level.

		<i>Temperature</i>		<i>Precipitation</i>	
		E_T^-	E_T^+	E_P^-	E_P^+
Tropics lat(25S-25N)-lon(0-360)					
d_{May}	DAS	0.200	0.192	<u>0.113</u>	0.072
	NODAS	0.201	0.190	0.107	0.067
d_{Nov}	DAS	<u>0.208</u>	<u>0.219</u>	0.109	<u>0.070</u>
	NODAS	0.195	0.200	0.108	0.062
Northern Extra Tropics lat(30-65N)-lon(0-360)					
d_{May}	DAS	0.083	0.077	0.020	<u>0.009</u>
	NODAS	0.084	0.076	0.016	0.003
d_{Nov}	DAS	<u>0.076</u>	<u>0.078</u>	0.016	0.016
	NODAS	0.055	0.063	0.016	0.014
Southern Extra Tropics lat(65-30S)-lon(0-360)					
d_{May}	DAS	0.131	0.121	0.026	0.021
	NODAS	0.126	0.129	0.029	0.022
d_{Nov}	DAS	<u>0.125</u>	0.115	0.021	0.017
	NODAS	0.114	0.109	0.018	0.014

TABLE 9. NODAS vs DAS discrimination distances for the tropics, northern extra tropics and southern extra tropics. The probability forecasts for above upper tercile ($E_{T,P}^+$) and below lower tercile ($E_{T,P}^-$) of the sample climatological distributions are reported for both temperature (left columns) and precipitation (right columns). Underlined bold discriminations stand for values which are significantly increased compared to the other experiment at the 5% significance level.

		<i>Temperature</i>		<i>Precipitation</i>	
		E_T^-	E_T^+	E_P^-	E_P^+
Tropical Pacific lat(25S-25N)-lon(190-240E)					
d_{May}	DAS	0.318	0.262	0.136	0.112
	NODAS	0.320	0.260	0.130	0.109
d_{Nov}	DAS	<u>0.299</u>	<u>0.267</u>	<u>0.137</u>	<u>0.117</u>
	NODAS	0.287	0.251	0.130	0.100
Indian Ocean lat(25S-20N)-lon(50-110E)					
d_{May}	DAS	0.135	<u>0.156</u>	<u>0.070</u>	<u>0.042</u>
	NODAS	0.128	0.137	0.055	0.032
d_{Nov}	DAS	0.147	0.198	0.139	0.063
	NODAS	0.150	0.192	0.140	0.060
Tropical Atlantic lat(40N-35S)-lon(60W-10E)					
d_{May}	DAS	0.105	0.083	0.48	0.012
	NODAS	<u>0.135</u>	0.090	0.045	<u>0.024</u>
d_{Nov}	DAS	<u>0.093</u>	<u>0.136</u>	0.065	0.015
	NODAS	0.069	0.113	0.064	0.010

TABLE 10. Same as table 9 but for tropical Pacific, tropical Indian and tropical Atlantic Oceans.

		<i>Temperature</i>		<i>Precipitation</i>	
		E_T^-	E_T^+	E_P^-	E_P^+
Pacific North American lat(40-65N)-lon(150-300E)					
d_{May}	DAS	0.092	0.079	0.006	0.006
	NODAS	0.084	0.074	0.006	0.003
d_{Nov}	DAS	<u>0.101</u>	<u>0.103</u>	0.008	0.010
	NODAS	0.069	0.074	0.014	0.006
Euro-Atlantic lat(35-65N) - lon(80W-40E)					
d_{May}	DAS	0.026	0.035	0.020	0.009
	NODAS	0.026	0.036	0.019	0.005
d_{Nov}	DAS	0.012	<u>0.026</u>	0.005	0.008
	NODAS	0.011	0.011	0.001	0.002

TABLE 11. Same as table 9 but for Pacific North American (PNA) and Euro-Atlantic regions.

List of Figures

- 1 Scheme summarizing the hindcasts generation strategy adopted in this work. For each start date we obtain the ocean initial conditions from off-line forced ocean analysis. Similarly, the atmospheric ICs are obtained from a simulation performed with prescribed observed SSTs. For each start date, an ensemble of 9 atmospheric initial states is produced by taking lagged days for the atmospheric component. Specifically, we sample the atmospheric IC by taking not only the actual start date but also the four days before and after, respectively. See text for further details. 56

- 2 Systematic error of seasonal mean predicted SST (target months from 2 to 4). DAS forecasts with starting dates in (a) May and (b) November. The error is defined as the difference between the 1991-2003 climatologies obtained respectively from the forecast ensemble means and from the ERA-Interim SST. Dark (light) shading indicates values above 0.5 (below -0.5) K. Contours: solid lines correspond to positive values and dashed lines correspond to negative values. Contour interval is 0.5 K. (c) and (d) are the same as (a) and (b) but for NODAS forecasts. (e) and (f) are the SST bias difference between DAS and NODAS, respectively for May and November (Contour interval 0.3 K). Shaded are the areas of increase (light) and decrease (dark) in DAS which passed a significance test at the 10% level. 57

- 3 Mean drift of the Niño3.4 SST over the period 1991-2003 for both May 1 and November 1 starting dates. (a) difference between modeled and observed climatologies and (b) absolute values. Dashed (dash-dot) lines indicates DAS (NODAS). In (b) the solid line refers to the observations. 58
- 4 Ensemble mean forecasts vs. ERA-Interim surface air temperature anomalies: point by point correlations of months 2-to-4 of the predictions. DAS forecasts with starting dates (a) May 1st and (b) November 1st. The grid points in which correlations are significant at the 10% level (bootstrap method) are shaded. (c) and (d) are the same as (a) and (b) but for NODAS forecasts. (e) and (f) are the DAS minus NODAS difference in correlations, respectively for May and November (Contours interval 0.2). Shaded are the areas of increase (light) and decrease (dark) in DAS which passed a significance test at the 10% level. 59
- 5 Spaghetti plot of the monthly mean Niño3.4 index (SST anomaly averaged in the region 190-240E, 5S-5N). Red thin lines indicate the predicted (DAS) anomalies for each ensemble member and the red square marks stand for the ensemble means. Observations are in black. 60

- 6 RMSEs (upper panels) and ACCs (lower panels) between predicted (DAS) and ERA-Interim SST anomalies over tropical Pacific (25S-25N; 140-280E). The 1991-2003 averaged RMSEs and ACCs for the forecasts with starting dates May 1st (a and c) and November 1st (b and d) are plotted as a function of the forecast target month. Solid thick lines and filled circle marks are for the ensemble means, while thin lines show the results for each ensemble member forecast. The dashed lines stand for the persistence forecasts, obtained persisting the monthly anomaly observed during the month prior to the start date of the model forecasts. The filled triangles indicate RMSEs and ACCs obtained for the ensemble mean forecasts performed with the system without ocean assimilation (NODAS). 61
- 7 NODAS vs. DAS scatter diagram of the Anomaly Correlation Coefficients (ACCs; a and b) and of the Root Mean Squared Error (RMSE; c and d) computed against ERA-Interim for the tropical Pacific (25S-25N; 140-280E) for each forecast year (diamonds). The cross marks also report the 13 year averages. Left (right) panels show the results for the 2 to 4 months averaged forecasts with starting date May 1 (November 1). 62
- 8 Box plot showing the Niño3.4 index in the NODAS (black) and in the DAS (red) forecasts. Green filled circles stand for observations. The distribution of predicted monthly mean anomalies is represented by boxes (25th-75th percentiles), the median is represented by the inside box marks and empty circles indicate outliers. The interannual monthly standard deviation is indicated by the shaded bands for the observations and by the dashed lines for the forecasts. 63

- 9 Onset of the El Niño 1997-1998: hovemoller diagram of the evolution of the heat content anomaly (Shaded; units are $\text{J}/\text{m}^2 \times 10^8$) and the zonal wind stress anomaly (contour; interval is 0.04 Pa). (a) NODAS, (b) the ocean analysis and (c) DAS. Forecast anomalies are ensemble averages. 64
- 10 Discrimination diagrams for the forecasts started in November (a and b) and May (c and d) computed using each grid point and the forecast months from 2 to 5 of the surface air temperature anomalies (TAIRA) over the tropics (0E-360E; 25S-25N) and using ERA-Interim data as reference. Left panels represent the dichotomous event of temperature being below the lower tercile of the sample climatological distribution and right panels are for the case of the temperature exceeding the upper tercile. The inset histogram in the diagrams indicate the refinement distribution. In red is the NODAS experiment and in blue DAS. Discrimination distance values are reported and a star mark is placed to indicate if the value is significantly higher compared to the other experiment at the 5% level (montecarlo method). 65
- 11 As in Fig. 10 but computed over the tropical Atlantic (35S-35N; 60W-10E) for May TAIRA (a and b) and for May Precipitation anomalies (PRECA; c and d) 66
- 12 As in Fig. 11 but for November starting dates. 67
- 13 As in Fig. 10 but for the PNA region (40-65N; 150-300E). 68
- 14 As in Fig. 13 but for the Euro-Atlantic region (35-65N; 80W-40E). 69

SYSTEM INITIALIZATION SCHEME

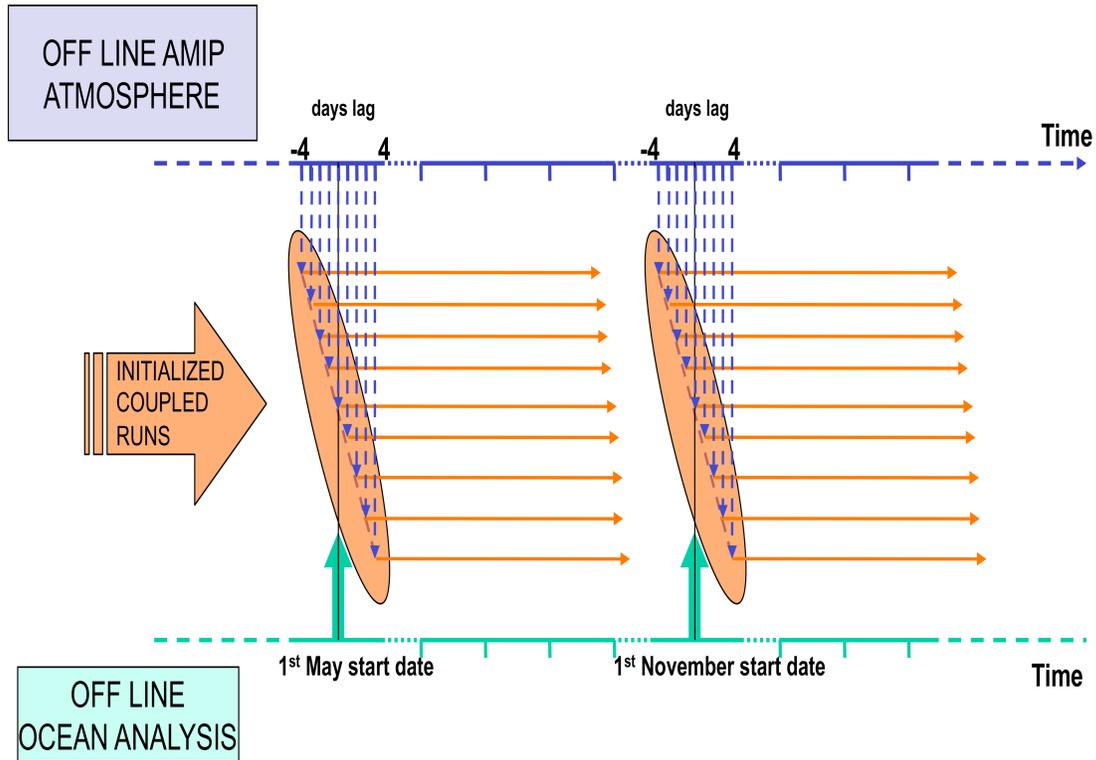
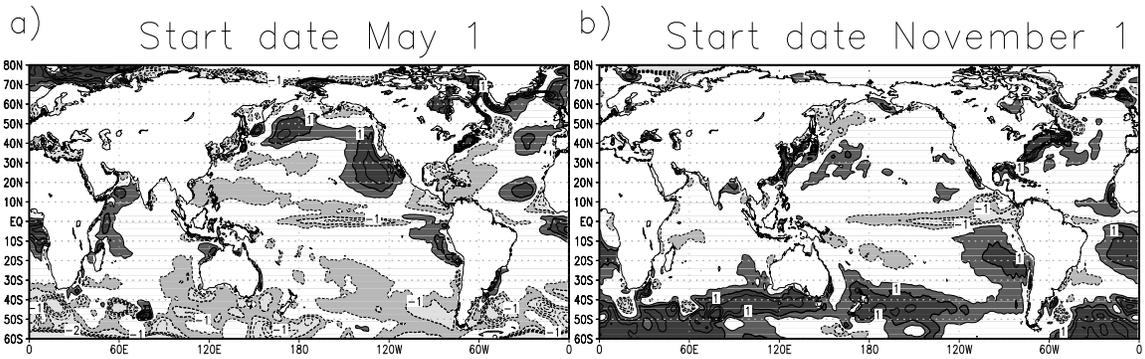


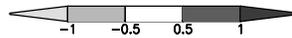
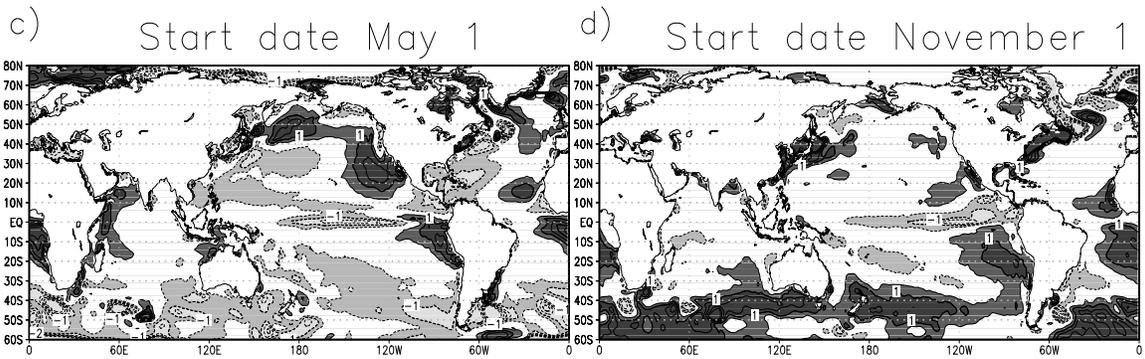
FIG. 1. Scheme summarizing the hindcasts generation strategy adopted in this work. For each start date we obtain the ocean initial conditions from off-line forced ocean analysis. Similarly, the atmospheric ICs are obtained from a simulation performed with prescribed observed SSTs. For each start date, an ensemble of 9 atmospheric initial states is produced by taking lagged days for the atmospheric component. Specifically, we sample the atmospheric IC by taking not only the actual start date but also the four days before and after, respectively. See text for further details.

SST BIAS OF THE SYSTEM

Assimilated Ocean IC



Ocean IC with no assimilation



Assimilated IC vs. no Assimilation

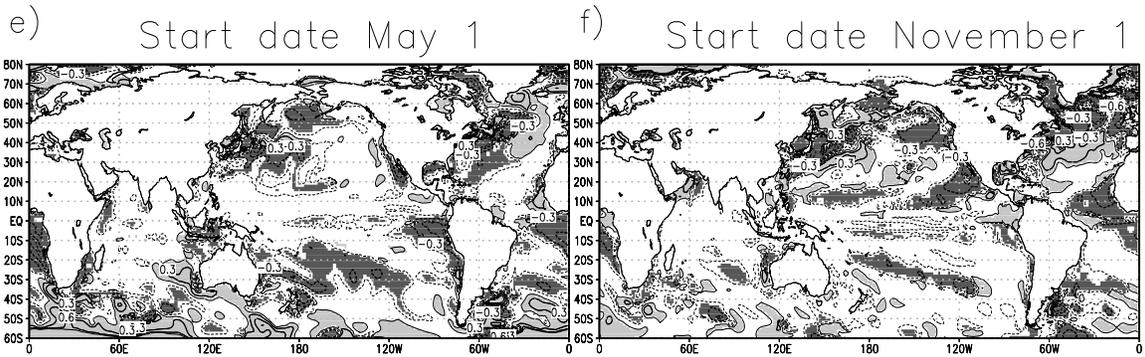


FIG. 2. Systematic error of seasonal mean predicted SST (target months from 2 to 4). DAS forecasts with starting dates in (a) May and (b) November. The error is defined as the difference between the 1991-2003 climatologies obtained respectively from the forecast ensemble means and from the ERA-Interim SST. Dark (light) shading indicates values above 0.5 (below -0.5) K. Contours: solid lines correspond to positive values and dashed lines correspond to negative values. Contour interval is 0.5 K. (c) and (d) are the same as (a) and (b) but for NODAS forecasts. (e) and (f) are the SST bias difference between DAS and NODAS, respectively for May and November (Contour interval 0.3 K). Shaded are the areas of increase (light) and decrease (dark) in DAS which passed a significance test at the 10% level.

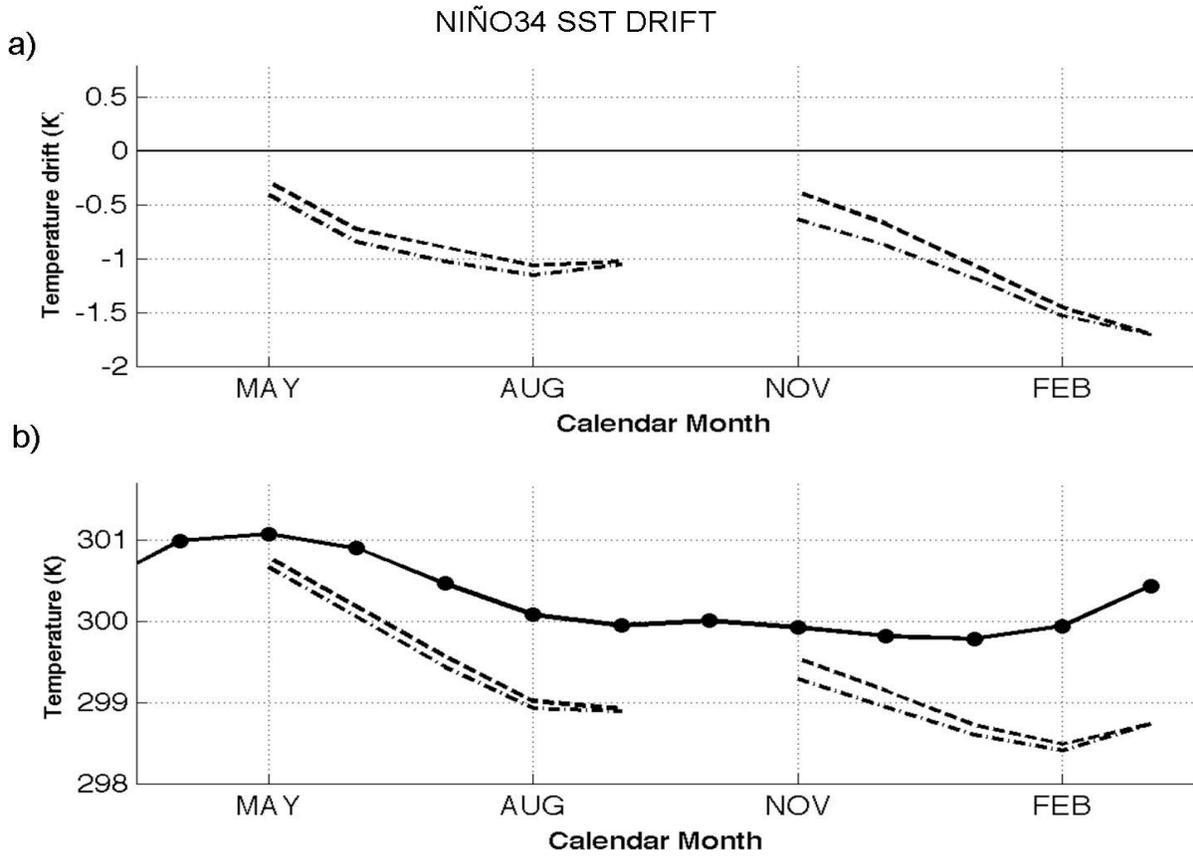
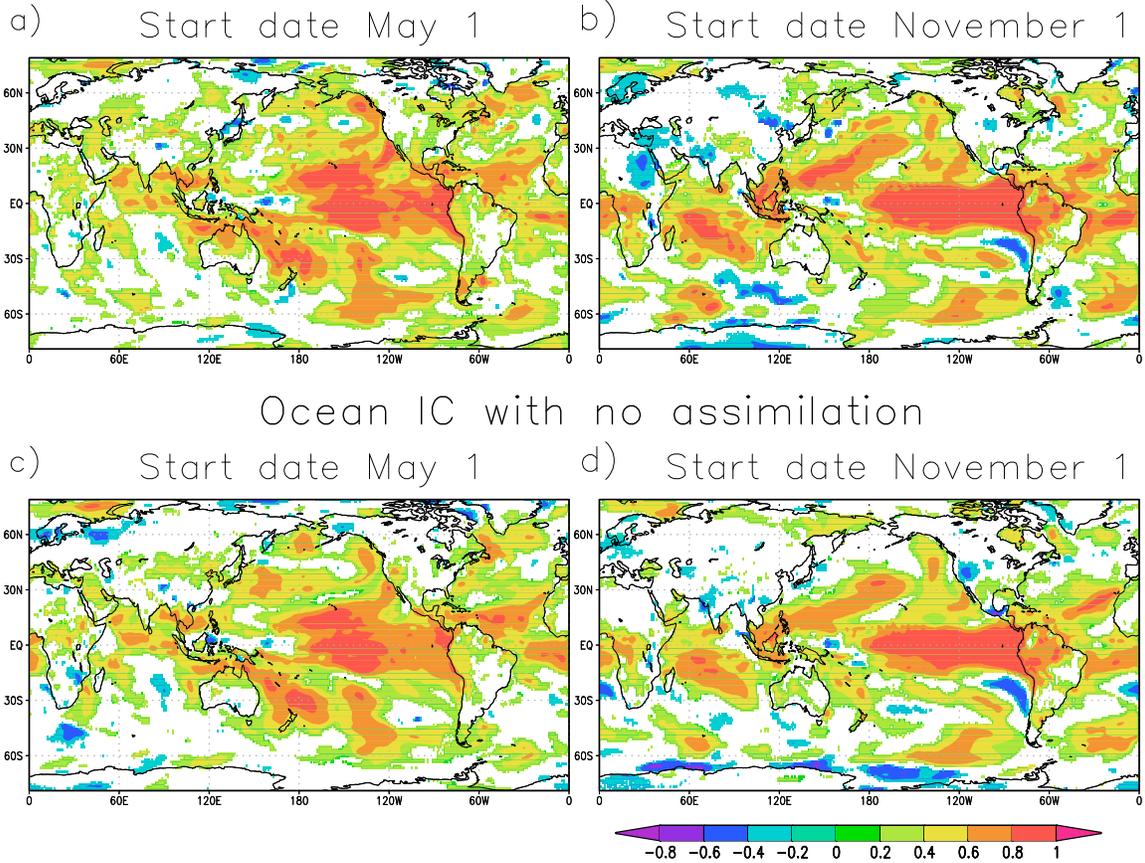


FIG. 3. Mean drift of the Niño3.4 SST over the period 1991-2003 for both May 1 and November 1 starting dates. (a) difference between modeled and observed climatologies and (b) absolute values. Dashed (dash-dot) lines indicates DAS (NODAS). In (b) the solid line refers to the observations.

Assimilated Ocean IC



Assimilated IC vs. no Assimilation

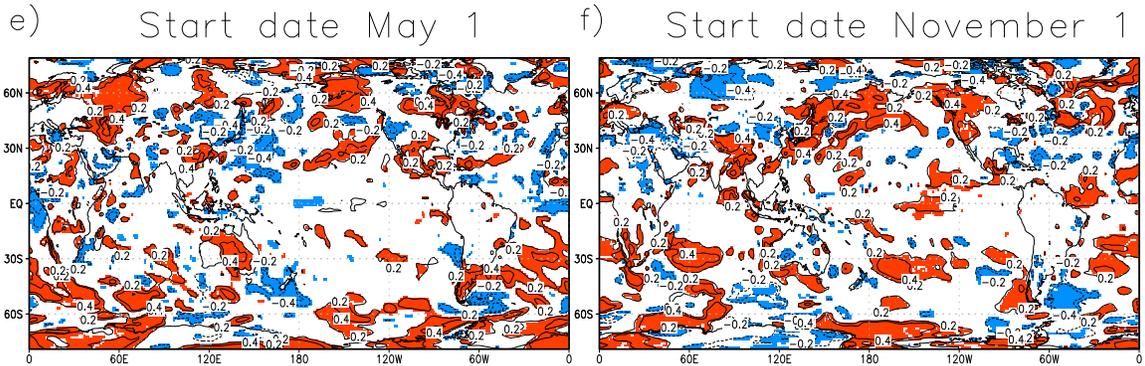


FIG. 4. Ensemble mean forecasts vs. ERA-Interim surface air temperature anomalies: point by point correlations of months 2-to-4 of the predictions. DAS forecasts with starting dates (a) May 1st and (b) November 1st. The grid points in which correlations are significant at the 10% level (bootstrap method) are shaded. (c) and (d) are the same as (a) and (b) but for NODAS forecasts. (e) and (f) are the DAS minus NODAS difference in correlations, respectively for May and November (Contours interval 0.2). Shaded are the areas of increase (light) and decrease (dark) in DAS which passed a significance test at the 10% level.

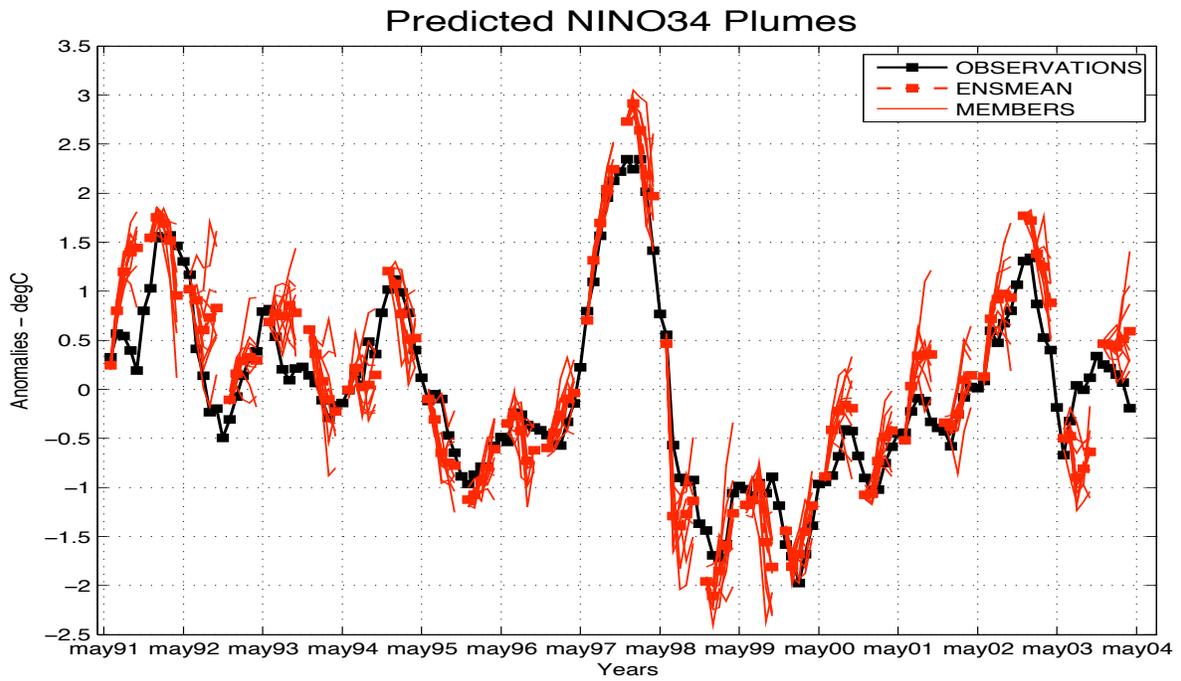


FIG. 5. Spaghetti plot of the monthly mean Niño3.4 index (SST anomaly averaged in the region 190-240E, 5S-5N). Red thin lines indicate the predicted (DAS) anomalies for each ensemble member and the red square marks stand for the ensemble means. Observations are in black.

Tropical Pacific (25S-25N; 140-280E) monthly Performance

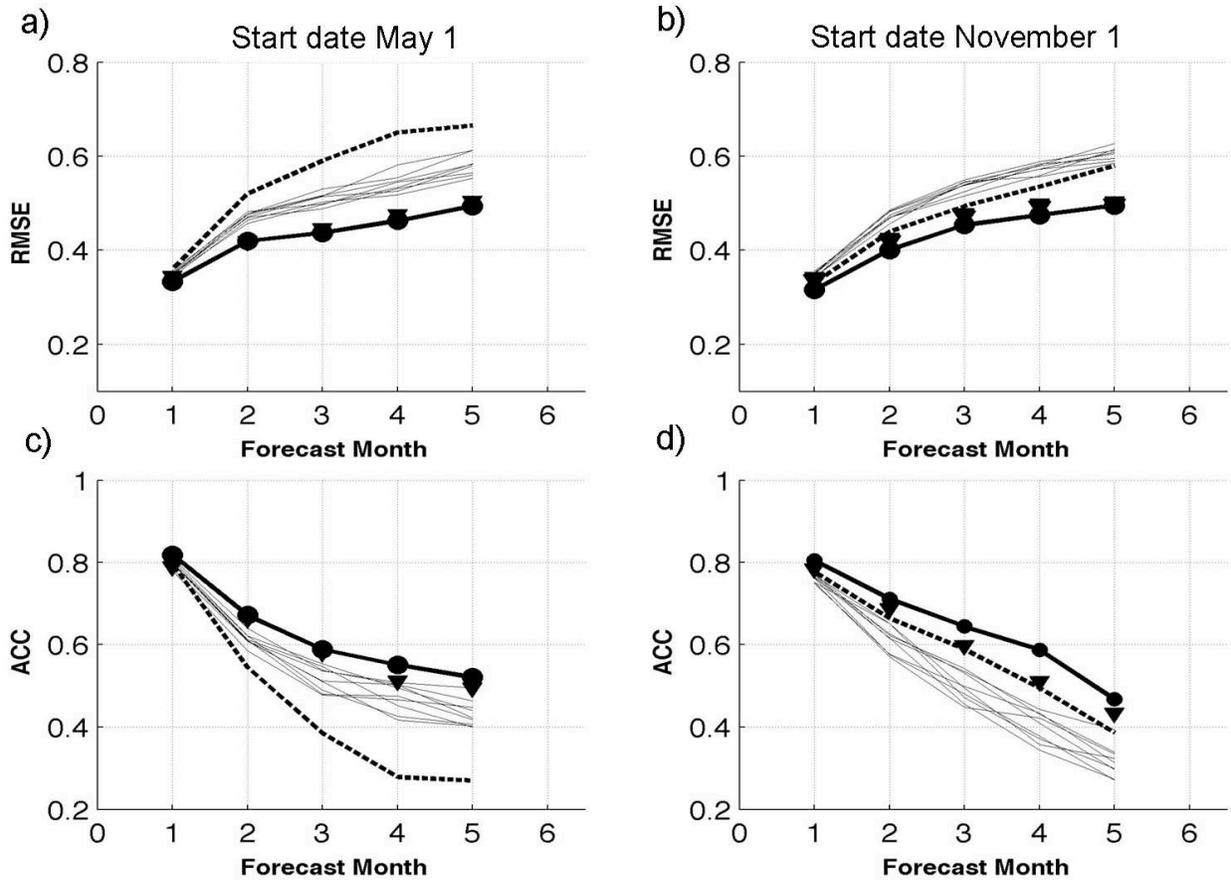


FIG. 6. RMSEs (upper panels) and ACCs (lower panels) between predicted (DAS) and ERA-Interim SST anomalies over tropical Pacific (25S-25N; 140-280E). The 1991-2003 averaged RMSEs and ACCs for the forecasts with starting dates May 1st (a and c) and November 1st (b and d) are plotted as a function of the forecast target month. Solid thick lines and filled circle marks are for the ensemble means, while thin lines show the results for each ensemble member forecast. The dashed lines stand for the persistence forecasts, obtained persisting the monthly anomaly observed during the month prior to the start date of the model forecasts. The filled triangles indicate RMSEs and ACCs obtained for the ensemble mean forecasts performed with the system without ocean assimilation (NODAS).

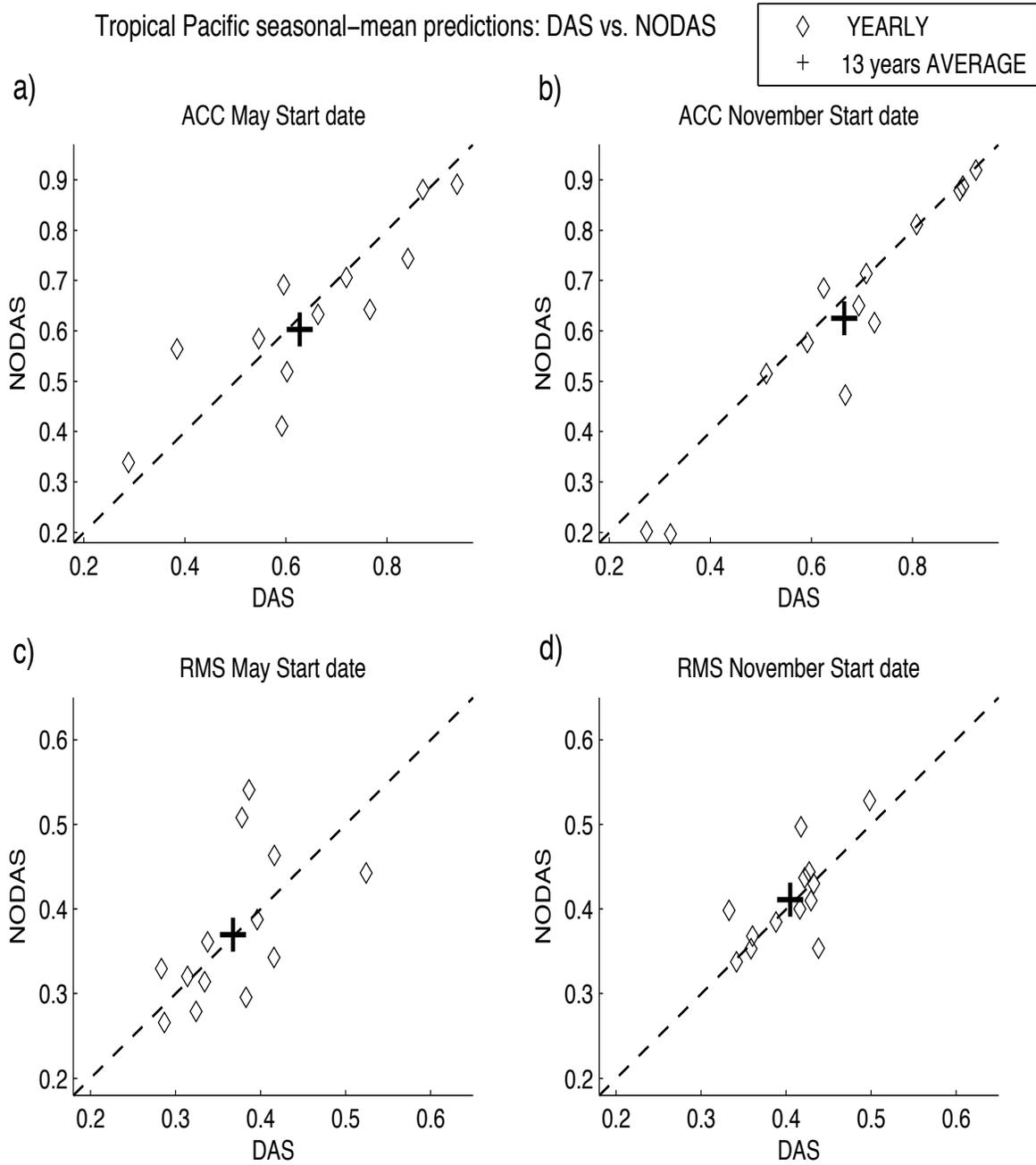


FIG. 7. NODAS vs. DAS scatter diagram of the Anomaly Correlation Coefficients (ACCs; a and b) and of the Root Mean Squared Error (RMSE; c and d) computed against ERA-Interim for the tropical Pacific (25S-25N; 140-280E) for each forecast year (diamonds). The cross marks also report the 13 year averages. Left (right) panels show the results for the 2 to 4 months averaged forecasts with starting date May 1 (November 1).

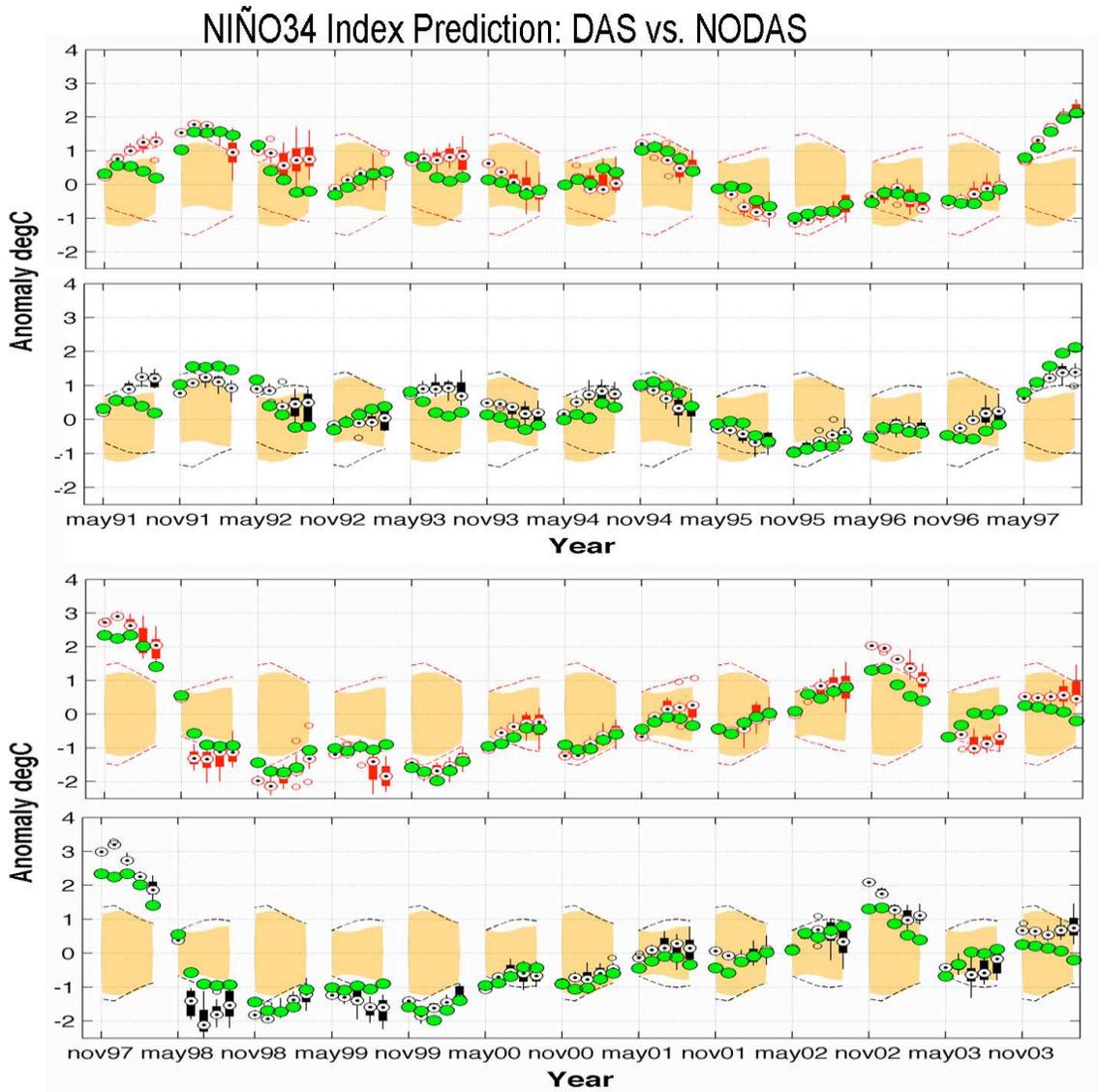


FIG. 8. Box plot showing the Niño3.4 index in the NODAS (black) and in the DAS (red) forecasts. Green filled circles stand for observations. The distribution of predicted monthly mean anomalies is represented by boxes (25th-75th percentiles), the median is represented by the inside box marks and empty circles indicate outliers. The interannual monthly standard deviation is indicated by the shaded bands for the observations and by the dashed lines for the forecasts.

ONSET OF El Niño 1997-1998

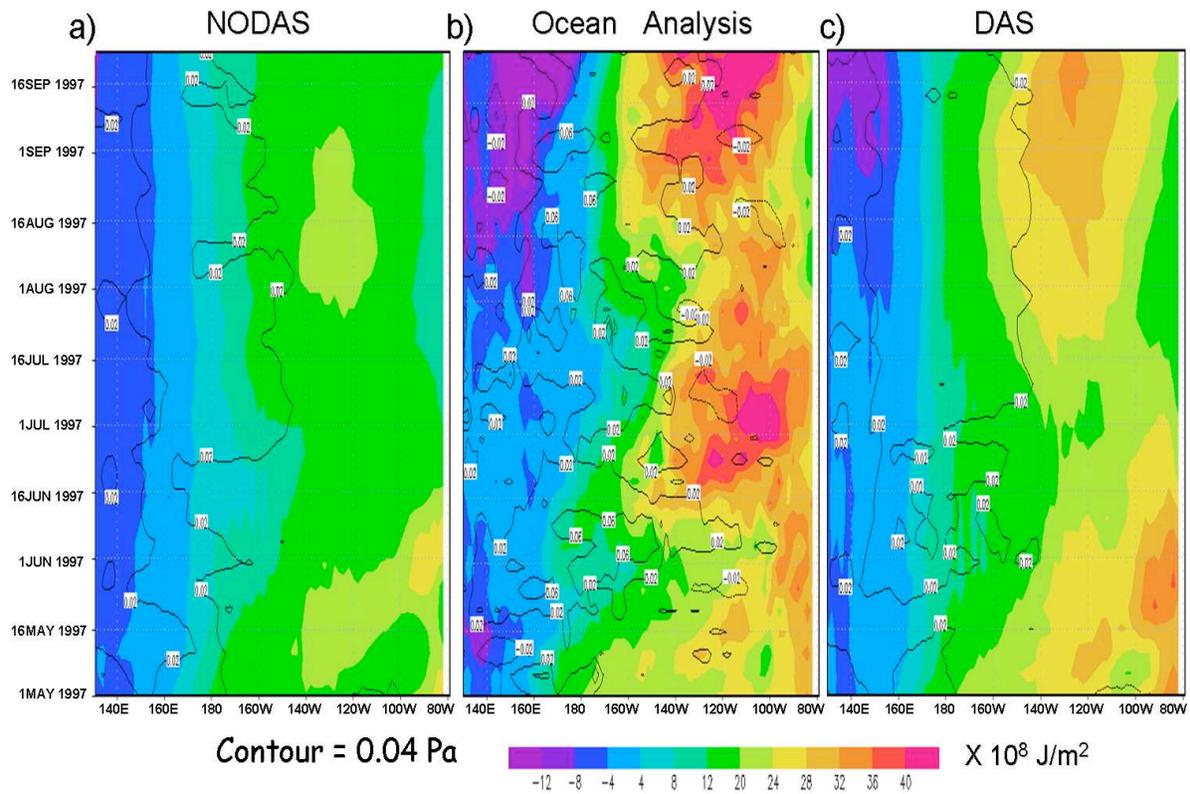


FIG. 9. Onset of the El Niño 1997-1998: hovmöller diagram of the evolution of the heat content anomaly (Shaded; units are $\text{J/m}^2 \times 10^8$) and the zonal wind stress anomaly (contour; interval is 0.04 Pa). (a) NODAS, (b) the ocean analysis and (c) DAS. Forecast anomalies are ensemble averages.

DISCRIMINATION DIAGRAMS -Tropics (25S-25N; 0-360E)

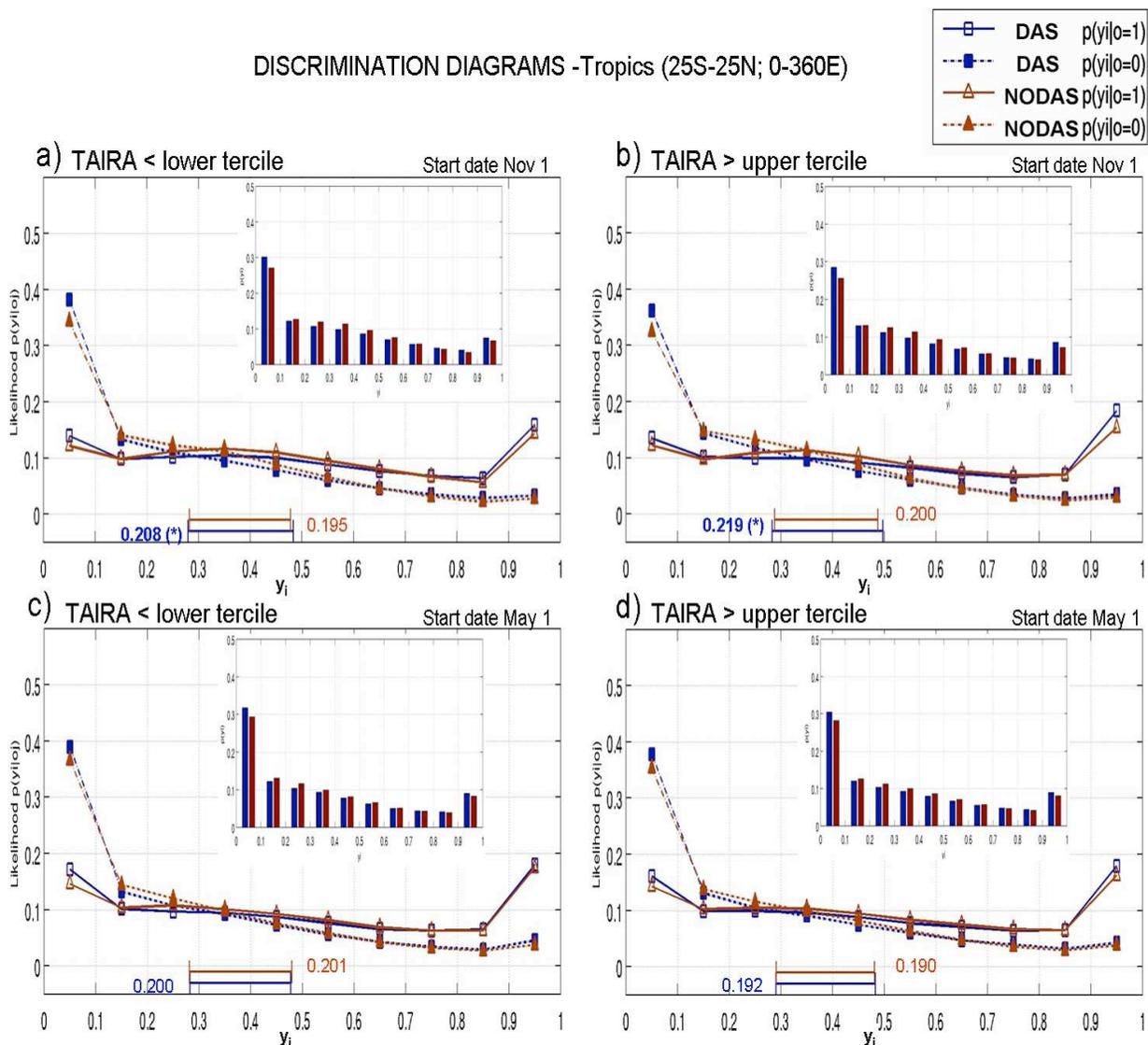


FIG. 10. Discrimination diagrams for the forecasts started in November (a and b) and May (c and d) computed using each grid point and the forecast months from 2 to 5 of the surface air temperature anomalies (TAIRA) over the tropics (0E-360E; 25S-25N) and using ERA-Interim data as reference. Left panels represent the dichotomous event of temperature being below the lower tercile of the sample climatological distribution and right panels are for the case of the temperature exceeding the upper tercile. The inset histogram in the diagrams indicate the refinement distribution. In red is the NODAS experiment and in blue DAS. Discrimination distance values are reported and a star mark is placed to indicate if the value is significantly higher compared to the other experiment at the 5% level (montecarlo method).

DISCRIMINATION DIAGRAMS –Tropical Atlantic (35S-35N; 60W-10E)

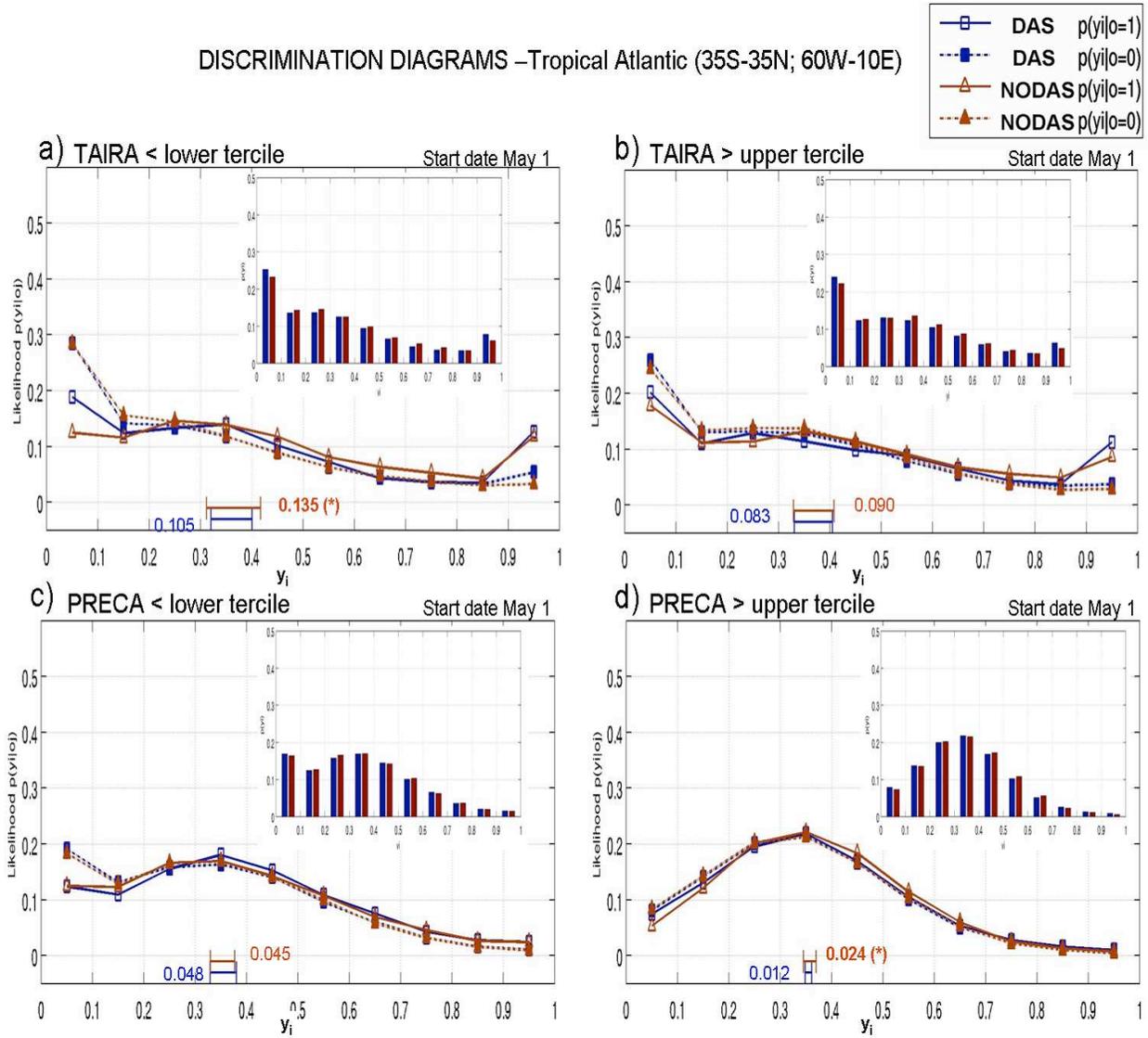


FIG. 11. As in Fig. 10 but computed over the tropical Atlantic (35S-35N; 60W-10E) for May TAIRA (a and b) and for May Precipitation anomalies (PRECA; c and d)

DISCRIMINATION DIAGRAMS –Tropical Atlantic (35S-35N; 60W-10E)

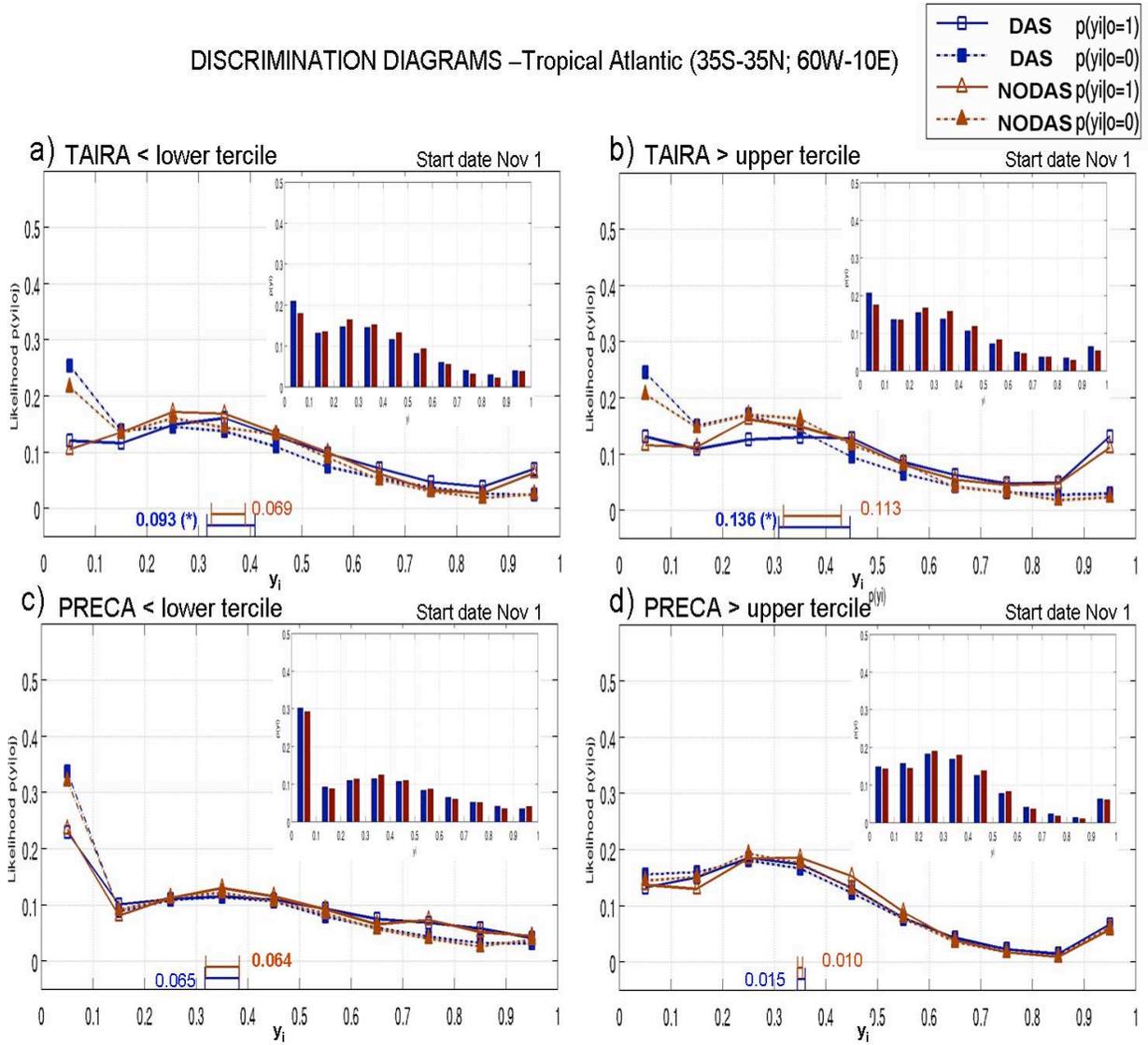


FIG. 12. As in Fig. 11 but for November starting dates.

DISCRIMINATION DIAGRAMS – Pacific North American (40N-65N; 150-300E)

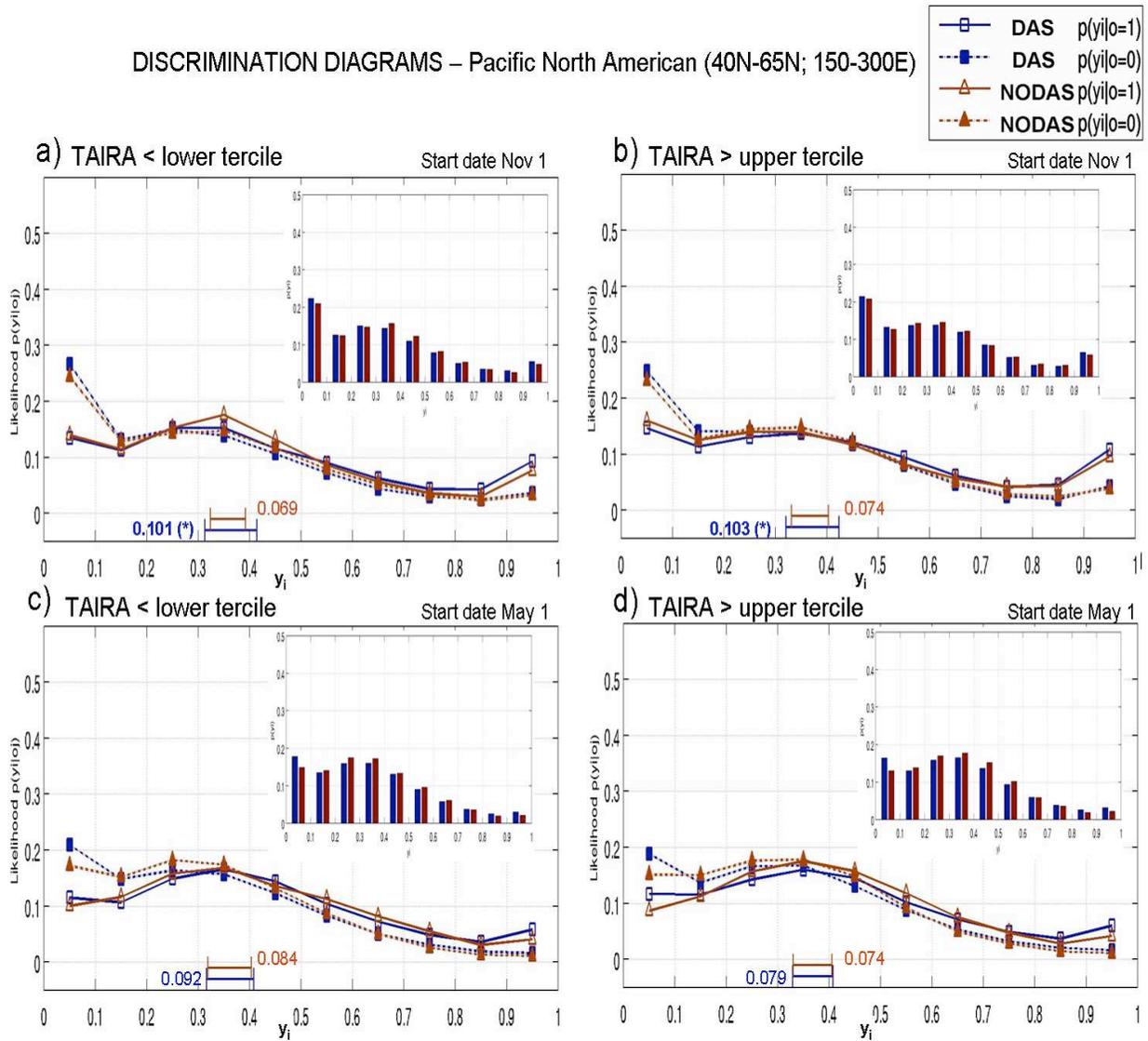


FIG. 13. As in Fig. 10 but for the PNA region (40-65N; 150-300E).

DISCRIMINATION DIAGRAMS – Euro Atlantic (35N-65N; 80W-40E)

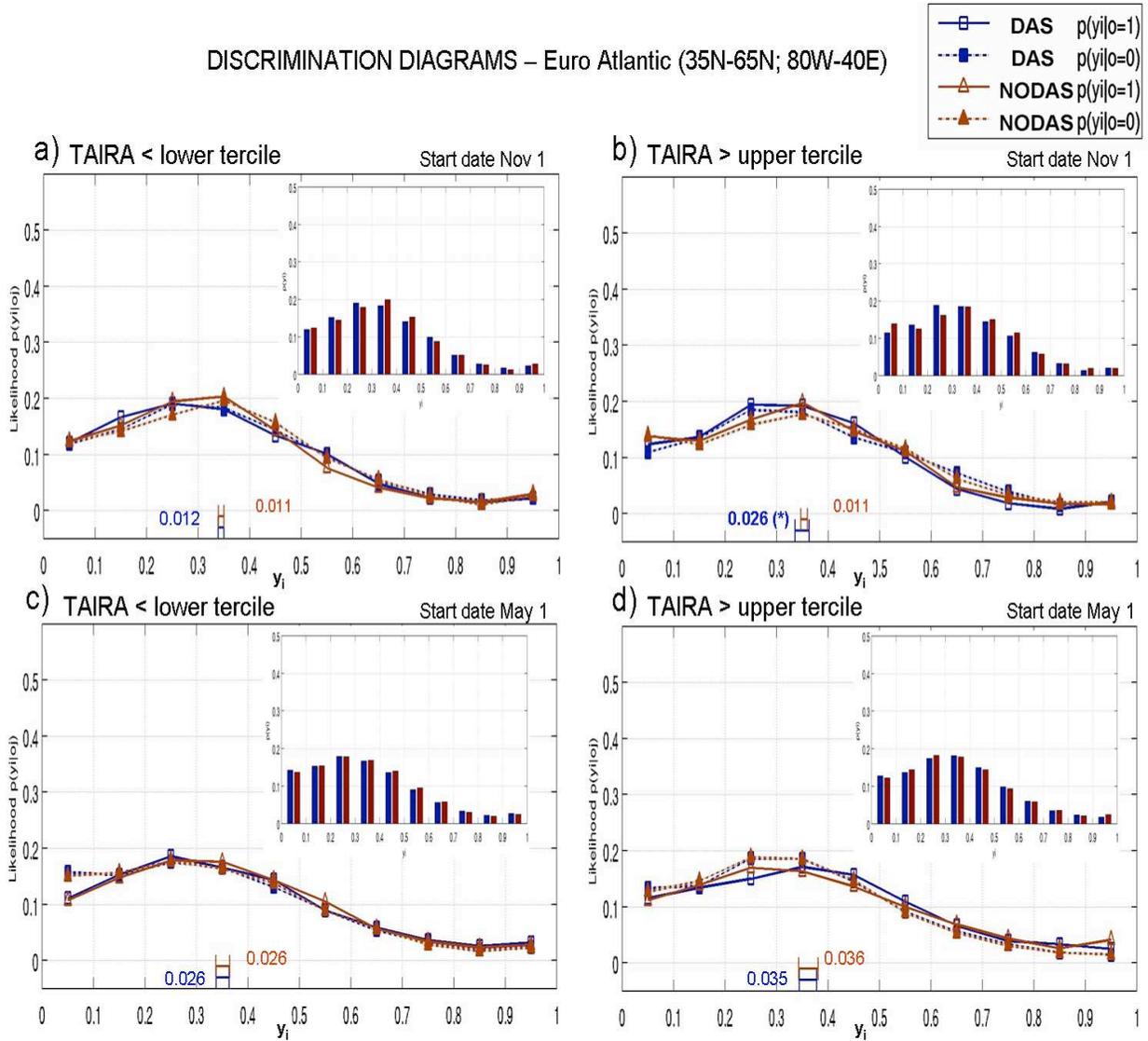


FIG. 14. As in Fig. 13 but for the Euro-Atlantic region (35-65N; 80W-40E).