TECHNICAL MEMORANDUM

570

Impact on initialization strategies and observations on seasonal forecast skill

Magdalena A. Balmaseda and David L. T. Anderson

Research Department

August 2008

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Abstract

Seasonal forecasting is a rapidly developing field with considerable effort devoted to developing comprehensive coupled general circulation models. Initialising these models is just as important, however, as the skill depends strongly on the way the coupled model is initialised. Three commonly used strategies are evaluated using the ECMWF System3, a state of the art coupled model. The most skillful scheme is that which makes most use of atmospheric and oceanic data even though this may generate initial imbalances in the coupled state. At seasonal time scales, increasing the balance at the cost of being further from the observations can degrade the forecasts. The relative importance of different components of the ocean observing system, equatorial moorings, altimetery and Argo floats, on forecast skill is assessed.

1 Introduction

The potential for climate predictability at seasonal time scales resides in information provided by the initial conditions, in particular the upper ocean density structure. We consider three initialisation strategies which differ in the way data are used to constrain the initial conditions. Method (i), the most common practice at operational centres such as ECMWF, is to initialise the atmosphere and ocean separately, assimilating all available meteorological and oceanic observations. The coupled system thus starts close to the observed state but whether this leads to the most skillful forecasts is less obvious as the method can have undesirable initialization shocks. In analysing the ocean, the model is forced by surface stress and surface fluxes of heat and fresh water obtained from atmospheric analyses and the quality of these fluxes has a strong influence on the quality of the ocean analyses.

An alternative initialisation method (ii) is to take the same atmospheric fluxes as in (i) and to force the ocean with them but not to assimilate ocean data. In this method the coupled model may be in closer balance since there is no ocean data to push the ocean away from being in balance with the winds. Luo *et al* 2005, used yet a different strategy (iii) in which the fluxes were obtained from the atmospheric model by forcing it with observed SST. The fluxes so obtained were then used to force the ocean during the ocean data assimilation. This strategy can reduce the initialization shock, since the atmosphere and ocean models will be in closer balance at the start of the integrations. Luo *et al* 2005 claim that their approach gives good forecast results but the results of Keenlyside *et al* 2005, using a similar method were not convincing. Comparison shows the skill of Luo *et al* 2005 to be lower than that of the ECMWF system S3 which uses (i), but the coupled models were very different and so no firm conclusions can be drawn as to the relative merits of the different initialisation strategies. To date no comparison of the three methods using a state of the art CGCM has been made. This paper evaluates the impact of the above initialization strategies on the coupled model drift, amplitude of interannual variability and skill of seasonal forecasts. An additional series of Observing System Experiments (OSES) is then conducted to assess the relative importance of different components of the ocean observing systems.

2 Evaluation of different initialization strategies

2.1 Methodology, Model and Experiments

Testing the impact of ocean initial conditions on the seasonal forecast skill requires generation of a comprehensive set of forecasts from the different ocean states. The baseline experiment for this study is the ECMWF S3 seasonal forecasting system (Anderson *et al* 2006, Molteni *et al* 2007), where the ocean initial conditions are created using the ECMWF ocean reanalysis system ORA-S3 (Balmaseda *et al* 2008). All available observations of temperature, salinity and altimetric sea level anomalies are used. The assimilation of altimeter data needs the

Table 1: Description of experiments

Experiment	Information in the Ocean Initial Conditions
i ALL	SST + Atmos obs + Ocean Obs
ii NO-OCOBS	SST + Atmos obs
iii SST-ONLY	SST
NOMOOR	ALL except moorings
NOALTI	ALL except altimeter
NOARGO	ALL except Argo
MDT0	As ALL, but using MDT0

prescription of the mean dynamic topography (MDT), which in ORA-S3 is derived from a previous assimilation run where subsurface temperature and salinity were used (MDT1). In methods (i) and (ii) the atmospheric fluxes are from the ERA40 reanalysis for the period January 1959 to June 2002 and NWP operational analyses thereafter (ERA/OPS). In all methods the ocean model SSTs are strongly relaxed to analyzed daily SST maps from the OIv2 SST product (Reynolds *et al*, 2002) from 1982 onwards. In methods ii and iii, no other ocean observations are used. The ocean model has a horizontal resolution of 1^o x 1^o with equatorial refinement. For further details see Balmaseda *et al.* (2008).

The same ocean model that is used for the analyses is used for the coupled forecasts, coupled daily to the atmospheric model, IFS cycle 31r1 at T159 horizontal resolution with 62 levels in the vertical, extending up to 5hPa. Table 1 gives a summary of the different experiments conducted. The hindcast, sometimes called reforecast experiments consist of 80 different start times, spanning the period 1987-2006 and sampled every three months (Jan, Apr, Jul and Oct). For each date, an ensemble of 5 coupled forecasts with perturbed initial conditions is integrated to 7-months ahead.

The three initialisation strategies can also be viewed as OSE type experiments in which atmospheric and oceanic data are withdrawn as can be seen in table 1. Differences in forecast skill between ALL (method i) and NO-OCOBS (method ii) are indicative of the impact of ocean observations, between NO-OCOBS and SST-ONLY (method iii) of the impact of atmospheric data that went into the ERA40/OPS. The combined impact of ocean and atmosphere information can be gauged by the differences between ALL and SST-ONLY. All skill scores have been cross-validated.

2.2 Assessment of skill of different initialisation strategies

The evolution of the SST bias in the coupled model is shown in the upper panels of fig 1 for regions NINO3 and NINO4 (defined in table 2) for the 3 experiments. The amplitude of the interannual variability of the coupled model as a function of lead time is shown in the lower panels. This latter is calculated as the ratio between the standard deviation of the interannual anomalies of the coupled model (computed separately for each ensemble member) and that of the observed SST.

Both the model bias and the amplitude of the interannual variability are sensitive to the initial conditions. In the Eastern Pacific (NINO3, left panels in fig 1), ALL shows the strongest warm bias for forecasts initialized in April, July and October. The warm bias is symptomatic of the existence of initialization shock: the coupled model is not able to maintain the slope of the thermocline in the initial conditions, and fast dynamic adjustment takes place through a downwelling Kelvin wave which depresses the thermocline in the Eastern Pacific, shutting down the upwelling and producing surface warming. The bias is close to zero in experiment NO-OCOBS, where the initial conditions have a flatter thermocline, and consequently the dynamic Kelvin wave adjustment

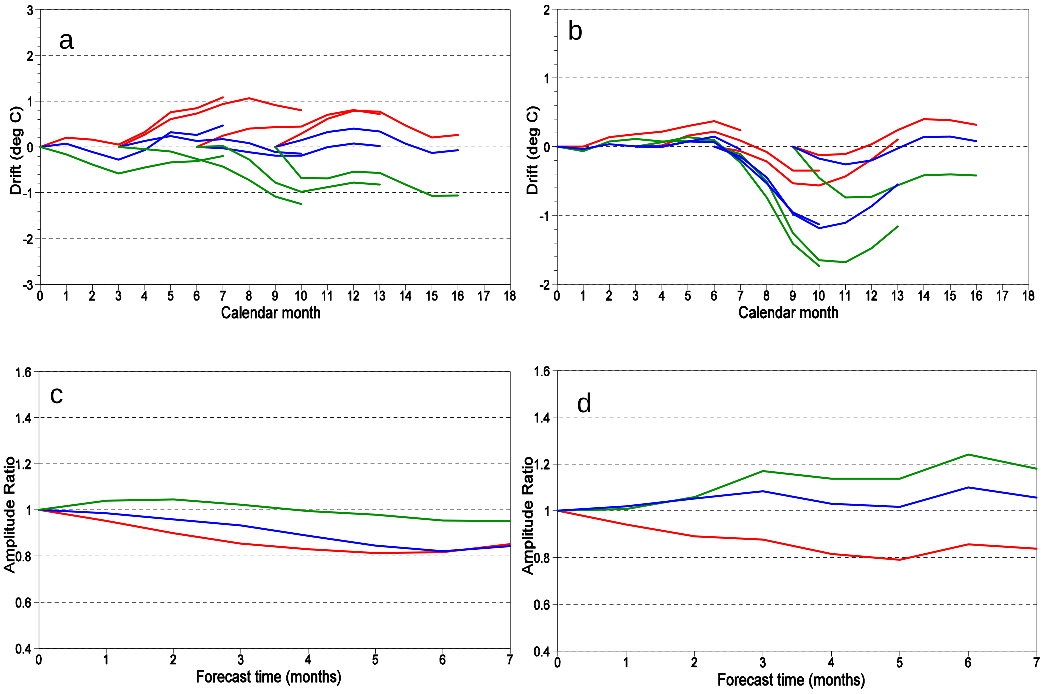


Figure 1: Top: forecast drift as a function of forecast lead time for 4 start months in regions NINO3 (a) and NINO4 (b) for experiments ALL (red), NO-OCOBS (blue) and SST-ONLY (green). Bottom: variance ratio as a function of lead time for the same experiments averaged over all start months in NINO3 (c) and NINO4 (d).

Table 2: Definition of area averaged indices

AREA NAME	latitudes	longitudes
NINO3	5N - 5S	150W-90W
NINO4	5N - 5S	160E-150W
EQ3	5N - 5S	150E-170W
EQPAC	5N - 5S	130E-80W
EQIND	5N - 5S	40E-120E
WTIO	10N - 10S	50E-70E
SETIO	0N - 10S	90E-110E
EQATL	5N - 5S	70W-30E
NSTRATL	28N-5N	80W-20E

is weaker. The cold bias in SST-ONLY, which develops especially fast for October starts is likely to be related to the thermocline being too shallow in the initial conditions, leading to an overestimation of the cooling by upwelling and the development of a cold bias as soon as the strong relaxation to SST used in the initialization process is turned off.

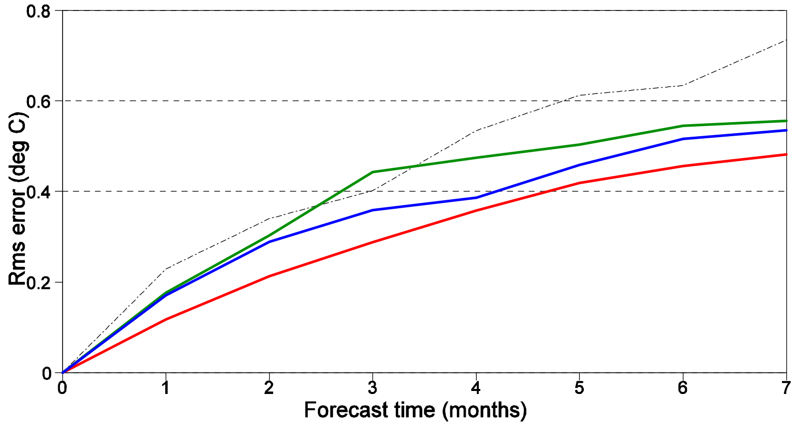
The amplitude of the interannual variability seems to be related to the magnitude of the bias, the least activity occurring in the presence of warm bias. This is to be expected if convective processes set an upper limit on how large values of SST can be and could explain why the amplitude of the interannual variability in ALL is low. However, it does not explain the low levels of activity in NO-OCOBS and SST-ONLY, suggesting that the underestimation of the interannual variability in NINO3 is not only related to the initial conditions, but stems from other sources of error in the coupled model.

In the Central Western Pacific (NINO4, right panels in fig 1) the initial conditions also have a large impact on the model bias and interannual variability. Here, ALL shows the smallest bias, followed by NO-OCOBS. The cold bias for SST-ONLY is the largest. The cold biases in NO-OCOBS and SST-ONLY are especially large during the second half of the year, consistent with the cold tongue penetrating too far west. In this region the amplitude of the interannual variability is related to the mean state and to the initialization procedure. For instance, overactive upwelling, characteristic of a cold tongue regime in this area, will produce an overestimation of the interannual variability, as happens in experiment SST-ONLY. The amplitude is underestimated in experiment ALL, even when the bias is small. The underestimation of the interannual variability in NINO4 for experiment ALL, and in NINO3 for all the experiments, suggests the existence of errors not corrected with the initialization which could be of atmospheric origin.

Balmaseda *et al* 2008 show that the assimilation of ocean observations has two main effects on the ocean mean state: it increases the heat content of the Equatorial Pacific by deepening the thermocline and increases the slope of the thermocline. Results shown in fig 1 suggest that while the first correction is maintained during the forecast, thus avoiding the westward penetration of the cold tongue and the cold bias, the slope of the thermocline is difficult to maintain, and is lost by rapid dynamical adjustment leading to the warm bias in the Eastern Pacific.

The impact of initialization strategies in the forecast skill appears in fig. 2 as a function of lead time for region NINO4. In this region the most skillful forecast at all lead times is obtained by method (i), and the worst by method (iii). There is a clear advantage from assimilating ocean observations. The results hold for both RMS error (fig 2a) and anomaly correlation (fig 2b).

Figure 3 shows the impact on forecast skill for various regions in table 2. The relative reduction in the monthly mean absolute error (MAE) resulting from adding information from the ocean and/or atmosphere observations



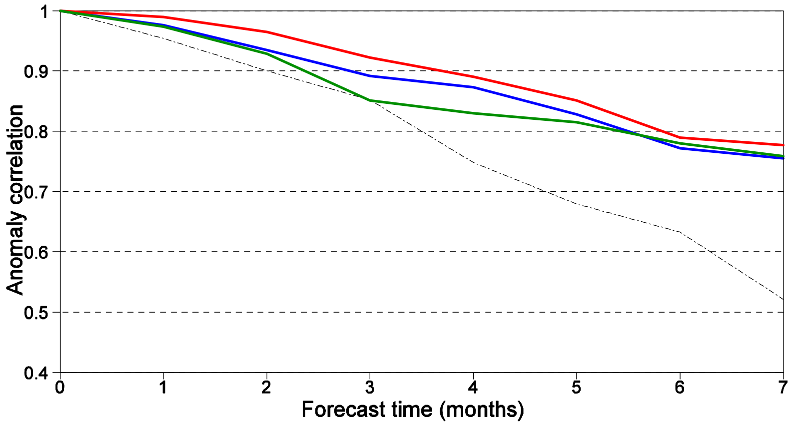


Figure 2: Impact of initialization strategies in forecast skill as a fuction of lead time in region NINO4, in terms of RMS error (top) and anomaly correlation (bottom). The best skill is obtained by experiment ALL (red) and the worst by SST-ONLY (green). Experiment NO-OCOBS is shown in blue.

for forecast range 1-7 months appears in fig 3a. For example, in the EQ3 region the impact of not using ocean obs is to increase the rms error by 12% (blue), of using neither ocean nor atmospheric observations is about 28% (grey). The impact of not using atmospheric observations is close to 15% (red). With the exception of EQATL, the best scores are achieved by experiment ALL. This means that for the ECMWF system, which uses i, the benefits of the ocean data assimilation and the use of fluxes from atmospheric (re)analyses more than offset problems arising from initialization shock. In the first 3 months of the forecast (not shown), the combined information of oceanic and atmospheric observations reduces the error by more than 40% in the different areas of the Equatorial Pacific (EQ3, NINO4, NINO3). Atmospheric observations are the main contributor to the reduction of forecast error. The contribution of the ocean observations is largest in the Central Western Pacific (13% in EQ3), but is negative in EQATL.

The contribution of oceanic and atmospheric observations seem to be cumulative in the reduction for MAE error at all lead times. The grey bars measure the difference in skill between (i) and (iii) confirming that the assimilation of atmospheric and oceanic data is markedly better than using just SST, suggesting that the Luo *et al* 2005 approach is not the best, at least not at the forecast ranges considered here.

The degraded skill in the Equatorial Atlantic may be indicative of poor balance between the ocean initial conditions and the coupled model fluxes, perhaps symptomatic of coupled model errors (Davey *et al* 2001). It can also be indicative of errors in the ocean initial conditions, although comparison with independent data shows that assimilating ocean data does improve the quality of the ocean analysis (Balmaseda *et al* 2008). It could also reflect spurious variability produced by the non-stationarity of the ocean observing system. A closer look at the impact of the different ocean observing systems for more specific time periods is given in the next section.

3 Relative importance of individual ocean observing systems

In the previous section we showed that the most skillful forecasts come from analyses which make most use of ocean observations. A further set of experiments was conducted to assess the contributions to forecast skill from the three main components of the observing system. In experiment NOMOOR the information from the moored array TAO/TRITON/PIRATA was removed, and in experiment NOALTI, the altimeter-derived sea level anomalies were not assimilated. In experiment MDT0, the external MDT used in the assimilation of altimeter comes from experiment NO-OCOBS, i.e. it does not contain information about ocean observations. The impact of altimeter, moorings and MDT is measured by comparing experiments ALL, NOMOOR, NOALTI and MDT0, for the period 1993-2006 (56 initial dates), 1993 being the start of high quality altimetry. The positive impact of Argo on forecast skill has already been reported by Balmaseda *et al* 2007. To measure the relative impact of Argo compared to the other ocean observing system, here we use exactly the same experiment (NOARGO, where the information from the Argo floats was withdrawn), and compare it with NOMOOR and NOALTI. The comparison is only done for the period 2001-2006 (28 dates in total), as ARGO really only started in 2001.

Figure 3b shows the relative reduction in the forecast error by including information from the altimeter and moored arrays and external MDT. The statistics have been computed for the period 1993-2006 and for the forecast range 1-7 months. In experiment ALL, and NOMOOR, the external MDT used to assimilate altimeter data is estimated as the 1993-2001 mean dynamic topography of an experiment where subsurface data was assimilated. MOOR can be taken as indicative of the impact of assimilating the anomalies from the mooring array. MDT0 offers an indication of the impact of using two different estimates of the altimetric mean state on forecast skill.

The information from the mooring array is the dominant factor in improving the skill in the different regions of

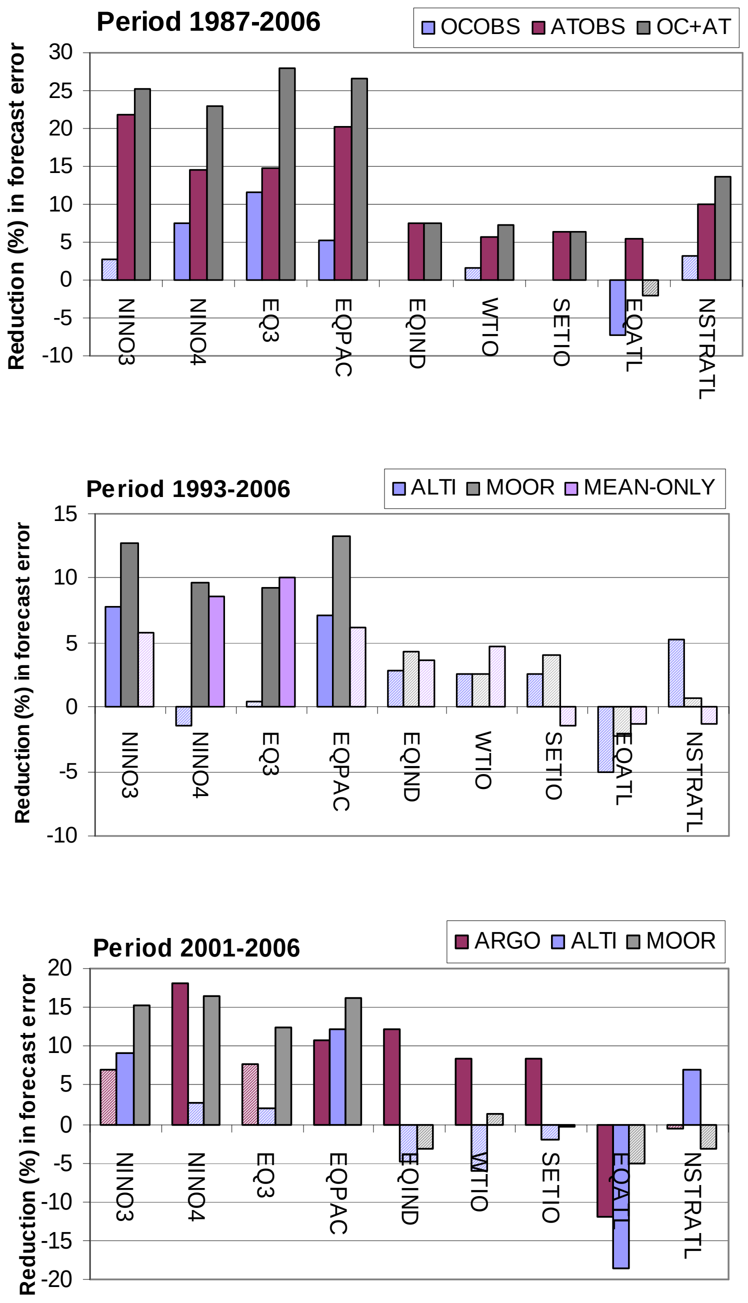


Figure 3: Impact of initialization in forecast skill for different regions, as measured by the reduction in mean absolute error for the forecast range 1-7 months. Solid bars indicate differences are above the 80% significant level. The upper panel compares initialization strategies for the period 1987-2006. Blue indicates the differences between strategy i and ii which differ in the use of ocean observations. Red (ATOBS) indicates differences between ii and iii, which differ in the use of atmospheric data, while grey (OC+AT) gives differences between i and iii and represents the combined impact of atmospheric and oceanic data. The middle panel compares altimeter, moorings and MDT for the period 1993-2006. ALTI indicates the difference in skill between NO-ALTI and ALL, and MOOR the difference between NO-MOOR and ALL. MEAN indicates the differences from using the different MDTs. The bottom panel shows the comparison between Argo, altimeter and moorings for the period 2001-2006. ARGO represents the difference between NO-ARGO and ALL.

the Equatorial Pacific and improves the skill in the Equatorial Indian Ocean. This is likely a remote effect, since there are few Indian Ocean moorings. The remote influence of the moorings in the Indian Ocean can occur via the ocean initial conditions (i.e. the remote effect of assimilating data in the Pacific propagating into the Indian Ocean via the Indonesian Through-flow), or via an atmospheric bridge, whereby an improvement in forecast skill in the Pacific leads to an improved atmospheric circulation which impacts forecasts of the Indian Ocean. The impact of the external MDT is also quite substantial in the Pacific, and to a lesser degree the Equatorial Indian Ocean. The effect of altimeter data is more noticeable in the NINO3 and NSTRATL, although here the impact does not reach the 80% significance level. Moorings, MDT and Altimeter have a positive impact on the WTIO, although the values are not significant at the 80% confidence level. The Equatorial Atlantic again stands out as the only region where the different observational information consistently has a detrimental effect. Even though the differences are not statistically significant, this result suggests that model or initialization errors are important.

Figure 3c shows the impact on forecast skill of Argo, moorings and altimeter. The statistics have been calculated only for the (rather short) post-Argo period 2001-2006 and so the impacts are best considered as indicative rather than definitive. The figure shows the dominant impact of Argo in the NINO4 (18%), and EQIND (12%) regions. Argo is the only observing system with a significant impact on the WTIO region (8%) and while it has a positive impact (8%) in the SETIO region, the differences are not significant at the 80% confidence level. The information from the moorings is still dominant in the whole Equatorial Pacific (16%), NINO3 (15%) and EQ3(13%), and quite large in NINO4 (16%), although here they are less important than Argo. Meanwhile the altimeter has a significant positive impact in the Equatorial Pacific, and it is the only observing system with positive impact in the North Subtropical Atlantic (8%), although this latter does not reach the 80% significance level. Again, for this period, all the observing systems have a negative impact on the EQATL region. Whether this is related to the faulty sensors in the SOLO/FSI Argo floats needs to be investigated further.

4 Summary and conclusions

The most complex assimilation used operationally is uncoupled initialisation in which atmospheric observations and are assimilated into an atmospheric GCM, and ocean ocean observations are assimilated into an ocean model forced by fluxes from the atmospheric analysis. Complex assimilation methods are used in both media. Simpler methods without complex assimilation have also been advocated eg Luo et al 2005. These are appealing as they do not require large investment in assimilation strategies and may have reduced initialisation shock, since the coupled model should be in better balance at the start of the forecast. However, the results here, using a controlled methodology and a common coupled GCM, indicate that this is not so. The best forecasts come from the most comprehensive system. It is not even true that the model biases are smaller in the simpler systems as has been argued.

Observing system experiments have been used to gauge the contribution of different observation types to the forecast skill. All the different observing systems contribute to the improvement of seasonal forecast skill almost everywhere. The moorings having a dominant effect in the equatorial Pacific, altimeter data have a significant effect in the North Subtropical Atlantic and in the Eastern Pacific, while Argo has a larger effect in the Central-Western Pacific, and in the Indian Ocean. The equatorial Atlantic is a region where the forecasts are not improved through the use of ocean observations, probably indicative of model error.

More sophisticated assimilation methods that reduce the initialization shock but still make full use of the observations are desirable, though not yet developed. It has been shown that in the ECMWF seasonal forecasting System 3 the quality of the wind product used in the generation of the ocean initial conditions is instrumental for increasing the skill of the seasonal forecasts, with the winds from atmospheric reanalysis being far superior to those obtained by an atmospheric model forced by observed SST. This result should be taken into account in the ongoing efforts to design coupled model initialization strategies.

Acknowledgments The authors thank Franco Molteni for his useful comments in the preparation of this manuscript.

REFERENCES

Alves O., M. Balmaseda, D Anderson, T Stockdale, 2003, Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Q. J. R. Meteorol. Soc.*,130, Jan 2004, 647-668.

Anderson, D. L. T.,*et al*, 2006. Development of the ECMWF seasonal forecast System 3. ECMWF Technical Memorandum 503. Available from www.ecmwf.int.

Balmaseda, M.A., Arthur Vidard and David Anderson 2008. The ECMWF ORA-S3S ocean analysis system. *Mon. Wea. Rev.*, **136**, 3018-3034.

Balmaseda, M., D. Anderson, and A. Vidard, 2007. Impact of Argo on analyses of the global ocean. *Geophys. Res. Lett.*, **34**, L16605, doi:10.1029/2007GL030452.

Davey M., et al, 2001. STOIC: a study of coupled model climatology and variability in tropical ocean regions. *Clim. Dyn.*, **18**, 403-420.

Keenlyside, N.S., M. Latif, M. Botzet, J. Jungclaus, and U. Schulzweida, 2005. A coupled method for initializing El Nio Southern Oscillation forecasts using sea surface temperature. *Tellus*, **57A(3)**, 340-356.

Luo, J.J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005. Seasonal Climate Predictability in a Coupled OAGCM Using a Different Approach for Ensemble Forecasts. *J. Clim.*, **18**, 4474-4497.

Molteni F., et al 2007: ECMWF Seasonal Forecast System 3. CLIVAR EXCHANGES No 43.

Reynolds R., N Rayner, T Smith, D Stokes, W Wang, 2002. An improved in situ and satellite SST analysis for climate.*J. Clim.*, **15**, 1609-1625.