Impacto de las observaciones oceánicas en predicciones climáticas decadales realizadas con el sistema de predicción acoplado del ECMWF

Impact of ocean observations on decadal climate prediction with the ECMWF coupled forecast system

Francisco Javier DOBLAS-REYES

European Centre for Medium-Range Weather Forecasts (ECMWF), United Kingdom
f.doblas-reyes@ecmwf.int

Received: 17 June 2009
Accepted: 18 September 2009

RESUMEN
La predicción decadal es un nuevo campo que intenta proporcionar información climática interanual útil maximizando la información de las observaciones. La inicialización de los modelos acoplados océano-atmósfera utilizados en la predicción decadal dinámica es un requisito primordial. En este trabajo se han empleado dos estrategias típicas de la inicialización usada en predicción estacional para ejecutar una serie de predicciones por conjuntos de diez años con el modelo acoplado del ECMWF. El sistema es semejante al usado en predicción estacional, aunque con un mejor tratamiento de la concentración de gases de efecto invernadero y de los aerosoles sulfatados. Las predicciones interanuales formuladas a partir del conjunto de simulaciones demuestran que el sistema tiene capacidad predictiva para la temperatura a escala global y regional con una antelación de varios años. Esto ocurre especialmente sobre los trópicos, donde la influencia de una buena predicción del contenido de calor influye directamente sobre la atmósfera. La interpretación de las estimaciones del impacto positivo de las observaciones oceánicas sobre las predicciones tropicales y de ciertos aspectos del océano se ven limitadas por el pequeño tamaño de la muestra (que es típico de los experimentos previstos para el quinto Coupled Model Intercomparison Project), la falta de observaciones fiables y el error sistemático.
Palabras clave: Predicción decadal; predicción por conjuntos; contenido de calor del océano; circulación termohalina.

ABSTRACT
Decadal forecasting is a new field that aims at providing useful interannual climate information making the most of the best available observations. Initialising the ocean-atmosphere coupled models employed in dynamical decadal forecasting is a requirement. Two commonly used strategies for initializing seasonal forecasts have been employed in a set of ten-year ensemble reforecasts carried out with the ECMWF coupled model. The forecast system is similar to the one used for seasonal forecasting, except for a refined treatment of greenhouse gas concentrations and sulphate aerosols. The interannual predictions formulated from the set of simulations show

that the system has skill in predicting global and regional mean temperature up to several years in the future, especially over the tropics, where the influence of the skilfully predicted ocean heat content directly influences the atmosphere aloft. The interpretation of the estimates of the beneficial impact of the ocean observations found over the tropics and for certain ocean variables suffers from the shortness of the sample (typical of the experiments devised for the fifth Coupled Model Intercomparison Project), the lack of reliable observations and the sizeable model systematic error.
Key words: Decadal prediction; ensemble forecasting; ocean heat content; thermohaline circulation.

SUMMARY: 1. Introduction. 2. Experimental setup. 3. Forecast quality assessment. 4. Mean state. 5. Interannual forecast quality. 6. Conclusions. 7. Acknowledgements. 8. References

1. INTRODUCTION

Climate-change projections and near-term climate prediction (also known as decadal prediction) try to satisfy a growing demand for climate information for this century. Decadal prediction explores the ability of the climate models employed in the Intergovernmental Panel for Climate Change (IPCC) assessments to predict regional climate changes in the near future by exploiting both initial-condition information and changes in atmospheric composition. It is well established that, based on knowledge of the initial conditions, important aspects of climate are predictable up to a year ahead. This predictability is primarily, though not solely, associated with the El Niño Southern Oscillation (ENSO). However, how far ahead is the mean climate predictable at regional spatial scales with some useful level of skill? And to what extent does a better knowledge of the initial conditions contribute to the quality of these forecasts?.

To explore the relative usefulness of initialized decadal predictions, the next Coupled Model Intercomparison Experiment, known as CMIP5[1] (a project sponsored by IPCC) includes a set of ensemble-based dynamical experiments where the forecast systems will be initialized once every five years and run for at least ten years in the future. This experimental setup is based on the decadal re-forecast experiment of the ENSEMBLES project[2].

It is not clear how current climate forecast systems loose information from the initial conditions in the first few years of the forecast, especially because the information reduction depends on what aspect of climate is considered: ENSO, global mean temperature, monsoon characteristics, North Atlantic temperature, etc. The relative importance of the initial conditions in climate prediction changes with the time scale, but is supposed to be a continuous function that decreases with forecast time, becoming negligible after several decades (Cox and Stephenson, 2007). Ocean initial conditions are more relevant than variations in atmospheric composition in seasonal forecasting, except after an explosive volcanic eruption, while re-

[1] http://www.clivar.org/organization/wgcm/references/Taylor_CMIP5.pdf
[2] http://www.ecmwf.int/research/EU_projects/ENSEMBLES/exp_setup/stream2.html

cent results of decadal forecast exercises show that atmospheric composition changes prime after several years over the initial conditions (Keenlyside et al., 2008). Previous work (Smith et al., 2007; Keenlyside et al., 2008) has shown evidence that the initial state of the ocean can influence forecasts a decade or more ahead. However, decadal skill is almost exclusively found for ocean rather than for atmospheric variables, a limitation mainly due to the sizeable systematic error of current forecast systems.

In this study, the degree of predictability of atmospheric and ocean variables in the forecast range one-to-ten years, i.e. beyond the generally accepted limit of ENSO-related predictability, is studied with the ECMWF system.

2. EXPERIMENTAL SETUP

Ensemble re-forecasts have been carried out with the IFS/HOPE coupled system. The IFS/HOPE forecast system used the atmospheric IFS cycle 33r1 (Bechtold et al., 2008) with a horizontal truncation of T_L159 and 60 vertical levels extending up to 5 hPa. The ocean model has a horizontal resolution of 1°, with an equatorial refinement of 0.3°, and 29 levels in the vertical. The coupler OASIS2 is used to interpolate the fields exchanged once per day between the oceanic and atmospheric grids. IFS uses a climatological annual cycle of four types of aerosol (sea-salt, desert dust, organic matter, black carbon). The system includes the prescribed interannual evolution of global mean annual values of greenhouse trace gases (CO_2, CH_4, N_2O and CFCs) and anthropogenic aerosols, as well as interannual variations of total solar irradiance.

The re-forecasts were started once every five years over the period 1960 to 2005, i.e. in 1960, 1965, and so on. The atmosphere and land surface was initialized from the ERA40 reanalysis (Uppala et al., 2005) before 2005 and the operational ECMWF analyses for the 2005 re-forecast. Each simulation started at 00 GMT on the 1st of November of each year and run for 120 months. No relaxation or flux correction was active during the forecast.

The baseline experiment uses ocean initial conditions from the ORA-S3 ocean re-analysis (Balmaseda et al., 2007). All available observations of temperature, salinity and altimetric sea level anomalies are used in this re-analysis. The atmospheric fluxes are from the ERA40 reanalysis for the period January 1959 to June 2002 and ECMWF operational analysis thereafter. The ocean model sea surface temperature (SST) is strongly relaxed to analyzed daily SST maps from the OIv2 SST (Reynolds et al., 2002) product from 1982 onwards. Re-forecasts initialized with this data will be referred to henceforth as Assim. An alternative initialization consists in using data from an ocean simulation where the ocean has been forced with the atmospheric fluxes, but no ocean data have been assimilated. With this method the coupled model may be in closer balance since there is no ocean data to push the ocean away from the ocean model attractor. Re-forecasts initialized with data from this second set of ocean initial conditions will be named NoObs.

3. FORECAST QUALITY ASSESSMENT

Various measures of forecast quality have been used to assess the relative merits of the three forecast systems. The scores include the anomaly correlation coefficient (ACC) and root mean square error (RMSE) of the ensemble mean. All forecast quality measures have used ERA40 and ERA-Interim as the atmospheric reference dataset. For ocean variables, the ORA-S3 ocean re-analysis (Balmaseda et al., 2007) has been used.

Every forecast quality measure has been computed taking into account the systematic error of the forecast systems. Forecast anomalies are estimated by removing the mean over a fixed period, using the re-forecasts for which there are reference data available in cross-validation mode. The anomalies for the reference dataset are estimated for the same calendar period. This method is different from what would be done in an operational context, where the anomalies would be computed using only past information. However, the shortness of the sample, which is made of just ten re-forecasts, prevents the computation of a more meaningful set of anomalies.

4. MEAN STATE

Model inadequacy cause forecasts to drift away from the observed climate towards an imperfect model climate. The non-linear interaction between the drift of the model and the representation of interannual variability makes important to discuss forecast quality on the basis of the knowledge of the unconditional model error.

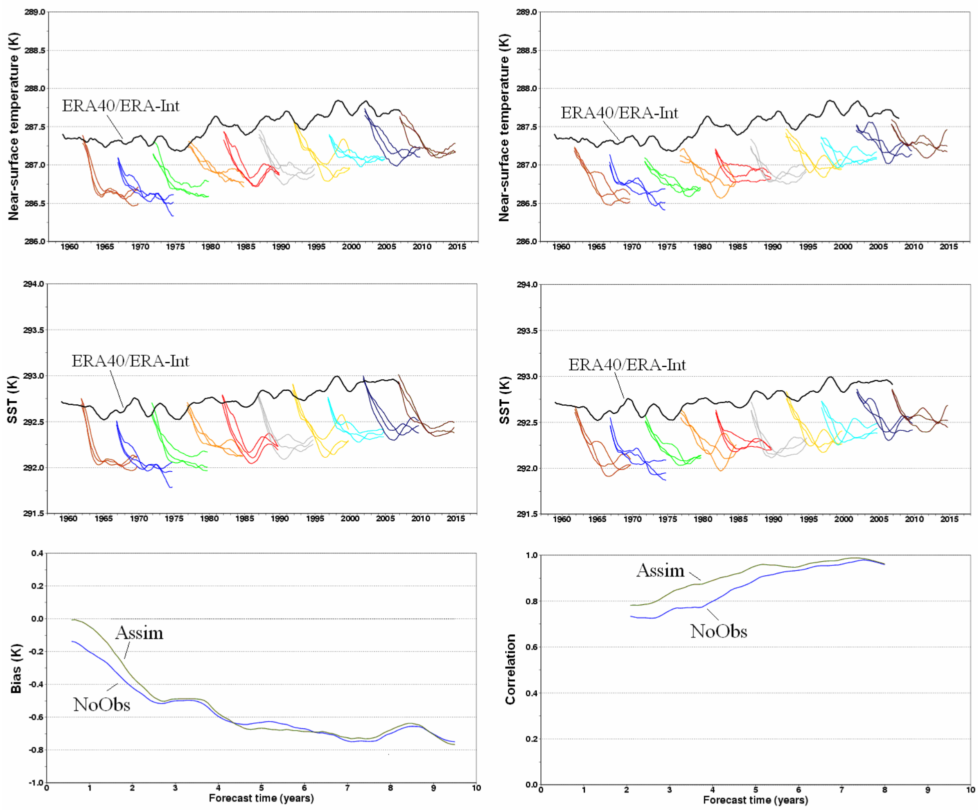

Figure 1. Global-mean near-surface air temperature (K, top row) and near-global (70ºN-60ºS) SST (K, middle row) for the ten three-member ensemble re-forecasts of the Assim (left) and NoObs (right) experiments. Each re-forecast is illustrated with lines of a different colour. Data from ERA40/ERA-Interim are shown in black. All time series have been smoothed out with a 24-month centred moving average, which removes data for the first and last years of each time series. The bottom row shows the bias (K, left) and the ensemble-mean correlation (right) of the global-mean near-surface air temperature of the Assim and NoObs experiments and have been computed using ERA40/ERA-Interim and three-member ensemble re-forecasts for the period 1960-1995. A 12-month moving average has been applied to the bias estimates, while the correlation has been computed with a moving window on four-year averaged anomalies.

Figure 1 shows the global-mean near surface temperature and the near-global (70°N-60°S) sea surface temperature (SST) for the Assim and NoObs experiments. The deviation of the model temperatures from the observed ones is fast, reaching -0.6 K for global-mean temperature and 0.4 K for the SST (not shown) after three to four years. The drift of both experiments is different in the first few years, although both reach a similar equilibrium level. The faster drift of the Assim experiment might be due to the stronger lack of balance with respect to the model attractor of its state after the first few months. The similarity of both experiments during the second half of the re-forecasts is a feature due to the strong drift; time series of model anomalies show larger differences.

The spatial distribution of the model drift is very similar in both experiments. Using re-forecasts with start dates within the period 1960-1995, which is the time interval with reference data available for all the forecast period, one finds that the model is particularly cooler than the reference. This is especially obvious over the tropical regions, although the cooling is slightly alleviated in Assim. Concerning other variables, no difference between the experiments can be found for mean sea level pressure, except for a reduction of the systematic error on the North Pacific. A broad similarity has also been observed for the interannual standard deviation of both experiments.

The model drift has been characterized in the ocean too. For instance, Figure 2 shows the global-mean ocean temperature averaged over the top 300 metres for the ten ensemble re-forecasts of both experiments and for the ORA-S3 ocean re-analysis. The re-analysis has a long-term upward trend starting around 1965. As with global-mean temperature, both experiments depict a similar degree of cooling, although in this case the final state differs between them, the Assim experiment showing a larger degree of interannual variability at the end of each re-forecast.

Another important example of how much the ocean state of both experiments can differ is illustrated by the meridional ocean circulation. The zonally integrated meridional water velocity across the Atlantic basin experiences in Assim a change after the first couple of years of forecast that shallows by ~500 m the Atlantic meridional overturning circulation (AMOC) cell, weakening the northward branch and strengthening the southward one. The drift is weaker in NoObs, for which the AMOC cell is already shallower than in the ORA-S3 analysis at the beginning of the simulations.

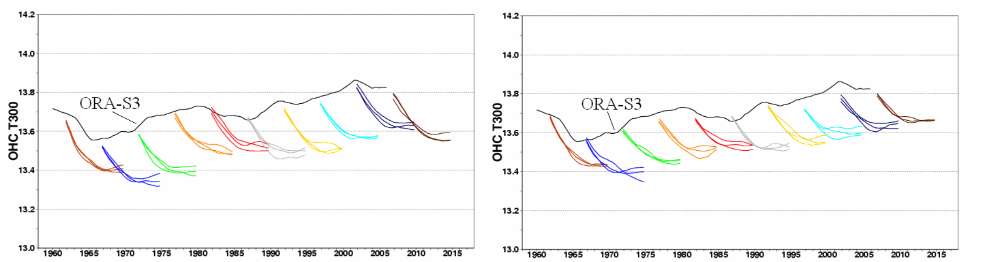

Figure 2. Global-mean ocean temperature (K) averaged over the top 300 metres for the ten three-member ensemble re-forecasts of the Assim (left) and NoObs (right) experiments. Each re-forecast is illustrated with lines of a different colour. Data from the ORA-S3 ocean re-analysis are shown in black. All time series have been smoothed out with a 24-month centred moving average, which removes data for the first and last years of each time series.

5. INTERANNUAL FORECAST QUALITY

The accuracy of the IFS/HOPE re-forecasts has been assessed using the eight re-forecasts with start dates within the period 1960-1995. The reason behind this choice is that there is not a complete reference dataset available to compare to the 2000 and 2005 start date simulations for the full forecast period. The ensemble-mean correlation of the global-mean temperature computed with these eight re-forecasts for four-year averages on a centred moving window is shown in Fig. 1. A four-year average allows predictions for forecast periods from two to eight years, of which Assim shows higher correlation than NoObs for all forecast periods, although the difference reduces with forecast time.

Large areas with positive skill appear in both experiments, a promising prospect for interannual and decadal forecasting. Assim seems to have better skill over the tropical band and large regions of the Southern Hemisphere. The tropical improvement is consistent with the behaviour found at the seasonal time scale (Balmaseda and Anderson, 2009). In several regions, such as the Northern Hemisphere continental areas and the North Pacific, the absence of ocean observations in the initial conditions increases the skill with respect to the Assim experiment. This effect might be partly due to the lack of robustness of the estimates, as correlations have been computed with a small sample of eight.

Similar conclusions can be reached for other variables. Figure 3 depicts the anomaly correlation of 850 hPa temperature from both experiments over the Northern Hemisphere and the tropical band for different forecast periods. To obtain each anomaly correlation coefficient, the spatial variance/covariance between the forecasts and the corresponding reference is computed for each one of the eight re-forecasts. The eight variances and covariances are averaged before the final correlation is computed. In agreement with the results for near-surface temperature, most cases display positive skill, with much higher values for the tropical band than for the extratropics. While the Assim experiment is more skilful than NoObs in certain instances over the Northern Hemisphere (e.g. for winter over the 6-10 year forecast period), there are other cases in which NoObs is better (e.g., predictions for the first forecast year). Instead, Assim is overall better than NoObs for tropical temperatures. It is important to bear in mind that the relative improvement of Assim with respect to NoObs for predictions of the tropical free atmosphere can only be explained by the impact of the assimilation of ocean observations in the initial conditions.

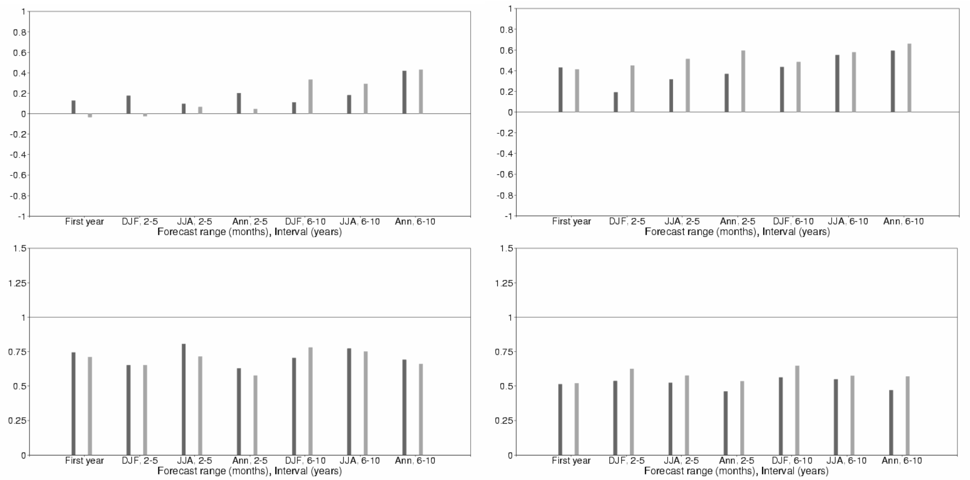

Figure 3. Ensemble-mean anomaly correlation (top row) and ratio between the spread and the root mean square error (bottom row) for 850 hPa temperature over the Northern Hemisphere (north of 30°N, left column) and the tropical band (20°N-20°S, right column) for different forecast periods of the Assim (dark grey bars) and NoObs (light grey bars) experiments. The pairs of bars correspond, from left to right, to the first calendar year of forecast (months 3-14), winter of the 2-5 year forecast period, summer of 2-5, annual mean of 2-5, winter of 6-10, summer of 6-10 and annual mean of 6-10. The estimates have been computed using ERA40/ERA-Interim and three-member ensemble re-forecasts for the period 1960-1995.

The ratio between the spread, computed as the standard deviation of the ensemble members around the ensemble mean, and the RMSE is a measure of the degree of reliability of the ensemble: for Gaussian-distributed variables, a reasonable approach for most annual averages, the ratio is expected to be close to one in a system where the ensemble spread is large enough to allow a full description of the variations in forecast error. In most cases, values lower than one have been found, which suggest that the ensemble spread is underestimated, although it should be reminded that estimates of the spread computed with just three members are far from robust. As in the case of the correlation, the Assim experiment performs better than NoObs over the tropics and worse over the Northern Hemisphere.

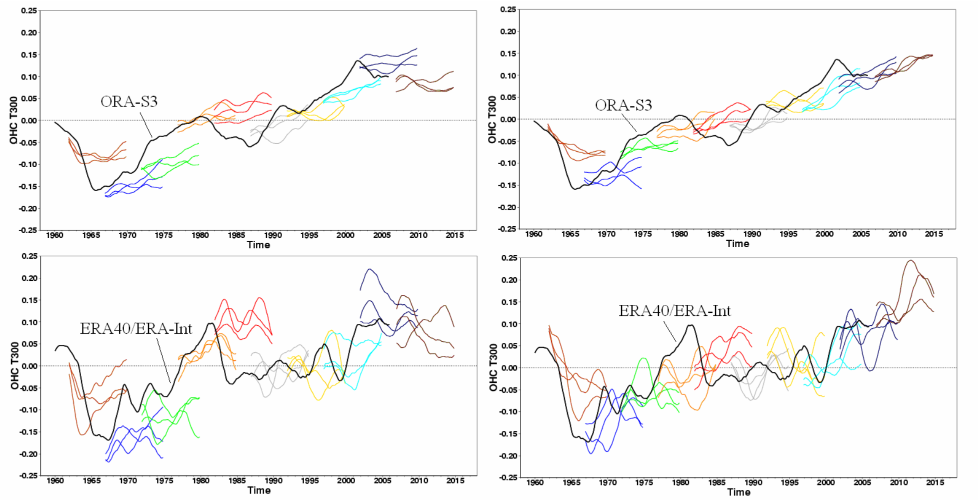

Figure 4. Anomalies of global-mean (top row) and tropical (20°N-20°S, bottom row) ocean temperature (K) averaged over the top 300 metres for the ten three-member ensemble re-forecasts of the Assim (left) and NoObs (right) experiments. Anomalies are computed with respect to the corresponding climate over the period 1960-2005 (eight re-forecasts). Each re-forecast is illustrated with lines of a different colour. Data from ORA-S3 (top) and ERA40/ERA-Interim (bottom) are shown in black. All time series have been smoothed out with a 24-month centred moving average, which removes data for the first and last years of each time series.

Part of the skill in predicting near-surface temperature comes from a skilful representation of the upper ocean heat content. Figure 4 displays anomalies of the global-mean ocean temperature averaged over the top 300 metres, a proxy for the top ocean heat content. Both sets of forecasts have a similar spread and upward trend in most predictions, except for the ones starting in 1960. The lack of interannual variability for a given member of a re-forecast suggests that in most cases the ocean heat content anomaly is persisted from the initial trend. The main difference between the experiments is that the Assim predictions have a larger variability from one re-forecast to the next one, to the point that they barely overlap as the NoObs ensembles do. The ocean initial conditions used for Assim and NoObs have different ocean heat content anomalies, something that is also manifested in the re-forecasts. This difference in behaviour is even clearer with tropical upper ocean mean temperature. Figure 4 also illustrates that Assim has much larger variations from the beginning of each re-forecast than NoObs. The larger variability from one forecast to the next one can only be explained by the differences in the initial conditions. While there are differences in the interannual variations of the anomalies, no large differences have been found between the experiments in the drift or its rate of change. Interestingly, the ensemble-mean correlation for the tropical upper ocean mean temperature is, though low, higher for Assim than for NoObs in the range of forecast periods between 2.5 and 5.5 years.

Previous studies (e.g. Collins et al., 2006) suggest that an accurate initialization of the AMOC could allow the Atlantic multidecadal variability to be predicted a few years in advance. Past AMOC fluctuations have been poorly observed and a large uncertainty exists among ocean analyses. The Assim experiment has been initialized with ocean states that reproduce the AMOC variability to the best of our knowledge (Balmaseda et al., 2007), while the ocean initial conditions used in the NoObs experiment underestimate this variability. However, little agreement has been found between the AMOC forecast anomalies and those from the analyses. The Assim anomalies have larger variability than the NoObs anomalies for the AMOC. Not surprisingly, the ensemble-mean correlation for the AMOC is negative and only reaches positive values for the second part of the re-forecasts of the Assim experiment. The systematic errors described in the previous section might limit the model ability to predict the ocean state.

6. CONCLUSIONS

The decadal forecast quality of the IFS/HOPE system described here is promising in light of the current model biases and taking into account that several climate processes, such as those related to sea-ice formation, export and melting, are not represented in the model. It is shown that the use of ocean sub-surface observations in the creation of the initial conditions increases the skill of atmospheric temperature over the tropics and the Southern Hemisphere, although it comes along with a decrease in skill over certain regions of the Northern Hemisphere. As a consequence, the experiment that uses assimilated ocean observations is more skilful in predicting interannual variations of global-mean temperature. The improvement in skill for atmospheric temperature is linked to a better prediction of the upper ocean mean temperature, especially over the tropics. However, the ability to predict interannual variations of the Atlantic meridional overturning circulation is very low.

The gain in forecast quality obtained with the use of the ocean observations in the initial conditions is small. The small skill differences found between both experiments might be due to the important model drift induced by model error, which prevents the model from making the most of the additional information available in the Assim experiment. This problem has been faced in the past by the seasonal forecast community and is being addressed by significant investment in model improvement. Additional experimentation with versions of the model with reduced systematic error might help shedding light into the role of the systematic error in limiting the benefits of an improved observation system.

The experiments described in this paper use the experimental setup defined by the ENSEMBLES project. This experimental setup has been inherited by CMIP5 for its decadal prediction exercise. One of the main problems found is that the differences between experiments are so subtle and the uncertainty in the forecast quality estimates so large that it is difficult to come up with significant conclusions with the small samples considered. Another problem is related to the way the anomalies are computed. While long simulations of the Twentieth Century climate could be used to estimate the model climate, these simulations do not take into account the drift typical of all coupled models that do not use flux correction terms. As a consequence, the anomalies with respect to the long climate simulations could be biased and their information misleading.

7. ACKNOWLEDGEMENTS

This work was supported by the ENSEMBLES project (GOCE-CT-2003-505539). The authors acknowledge the significant contributions by David Anderson, Magdalena Balmaseda, Thomas Jung, Kristian Mogensen, Jean-Jacques Morcrette, Franco Molteni, James Murphy, Tim Palmer, Doug Smith, Tim Stockdale and Antje Weisheimer.

En particular, querría aprovechar la ocasión para agradacer la impresionante tarea que, al igual que muchos otros profesores universitarios, ha realizado Elvira Zurita, quien inevitablemente tuvo una influencia en mi formación que espero no perder nunca.

APPENDIX

ACC: Anomaly Correlation Coefficient
AMOC: Atlantic Meridional Overturning Circulation
CFCs: Chlorofluorocarbons
CMIP5: Phase five of the Coupled Model Intercomparison Project
ENSO: El Niño / Southern Oscillation
IFS/HOPE: Integrated Forecast System / Hamburg Ocean Primitive Equation model
IPCC: Intergovernmental Panel on Climate Change
OASIS2: Version 2 of the On-line Applicant Status and Information System
OIv2: Version 2 of the National Oceanic and Atmospheric Administration Optimum Interpolation sea surface temperature analysis
ORA-S3: Ocean Re-Analysis for System 3
RMSE: Root Mean Squared Error

8. REFERENCES

BALMASEDA M.A. & D. ANDERSON (2009). Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.*, 36, L01701, doi:10.1029/2008GL035561.

BALMASEDA M.A., G.C. SMITH, K. HAINES, D. ANDERSON, T.N. PALMER & A. VIDARD (2007). Historical reconstruction of the Atlantic meridional overturning circulation from the ECMWF operational ocean reanalysis. *Geophys. Res. Lett.*, 34, L23615, doi:10.1029/2007GL031645.

BECHTOLD, P., M. KOEHLER, T. JUNG, F.J. DOBLAS-REYES, M. LEUTBECHER, M.J. RODWELL, F. VITART & G. BALSAMO (2008). Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.,* 134, 1337-1351.

COLLINS, M. & 11 co-authors (2006). Interannual to decadal climate predictability in the North Atlantic: A multimodel-ensemble study. *J. Climate*, 19, 1195-1203.

COX, P. & D.B. STEPHENSON (2007). Climate change: A changing climate for prediction. *Science*, 317, 207-208, doi: 10.1126/science.1145956.

KEENLYSIDE N.S, M. LATIF, J. JUNGCLAUS, L. KORNBLUH & E. ROECKNER (2008). Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, 453, 84–88, doi:10.1038/nature06921.

REYNOLDS, R.W., N.A. RAYNER, T.M. SMITH, D.C. STOKES & W. WANG (2002). An improved in situ and satellite SST analysis for climate. *J. Climate*, 15, 1609–1625.

SMITH D.M, S. CUSACK, A.W. COLMAN, C.K. FOLLAND, G.R. HARRIS & J.M. MURPHY (2007). Improved surface temperature prediction for the coming decade from a global climate model. *Science*, 317, 796–799, doi:10.1126/science.1139540.

UPPALA, S.M. & 45 co-authors (2005). The ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.,* 131, 2961-3012.