COMMUNITY WHITE PAPER

Initialization for Seasonal and Decadal Forecasts

Lead author: Magdalena A. Balmaseda
ECMWF, Shinfield Park, Reading RG2 9AX, UK
Email: Magdalena.Balmaseda@ecmwf.int;
Ph +44 118 9499112, Fax +44 1189 96????

Contributing authors:

Oscar Alves	CAWCR (Australia)	o.alves@bom.gov.au
Toshiyuki Awaji	JAMSTEC/Kyoto University (Japan)	awaji@kugi.kyoto-u.ac.jp
David Behringer	NCEP (USA)	David.Behringer@noaa.gov
Nicolas Ferry	Mercator-Ocean (France)	nferry@mercator-ocean.fr
Yosuke Fujii	MRI (Japan)	yfujii@mri-jma.go.jp
Tong Lee	NASA JPL (USA)	Tong.Lee@jpl.nasa.gov
Michele Rienecker	GMAO, NASA (USA)	michele.rienecker@.nasa.gov
Tony Rosati	NOAA/GFDL (USA)	Tony.Rosati@noaa.gov
Detlef Stammer	University of Hamburg (Germany)	detlef.stammer@zmaw.de
Doug Smith	Met Office (UK)	doug.smith@metoffice.gov.uk
Franco Molteni	ECMWF (UK)	Franco.Molteni@ecmwf.int

The potential for predicting natural internal climate variability on seasonal and decadal time scales resides mainly in information provided by the ocean initial conditions. While seasonal forecasting is a consolidated activity, with several operational centers around the world issuing routine seasonal forecasts, decadal forecasting is still in its infancy. This paper discusses the role of the ocean observing system in the initialization of seasonal and decadal predictions. It is shown that the assimilation of ocean observations reduces the uncertainty in the estimation of the upper ocean thermal structure and improves the skill of seasonal forecasts. Results from observing system experiments conducted with different seasonal forecast systems indicate that no observing system is redundant, highlighting their complementary nature. Results from synthetic observing system experiments conducted assuming sustainability of the current observing system indicate that Argo data has the potential for constraining the deep ocean variables responsible for decadal variability.

1. Introduction

The prediction of natural internal climate variability on seasonal and decadal time scales relies mainly on the ocean initial conditions. Seasonal forecasting is currently a routine activity in several operational centres, with a growing number of economic and societal applications such as agriculture, health and energy. Decadal forecasting is still in its infancy, but it is a rapidly growing field, envisaged to develop fast in the next decade. This paper discusses the role of the ocean observing system in the initialization of seasonal and decadal forecasts.

The consolidation of seasonal forecasting as a routine operational activity during the last decades has been possible thanks to the improvement in coupled models and data assimilation methods, availability of forcing fluxes from atmospheric reanalysis, and the development of the ocean observing system. In particular, the implementation of the TAO/TRITON array in the Equatorial Pacific in the early 1990's has been instrumental for ENSO prediction, which is still considered the corner stone of seasonal forecasting. The skill of the seasonal forecasts has further improved with the advent of altimeter and Argo data. There is potential for improving the prediction of other modes of interannual variability, such as the Indian Ocean Dipole (IOD), which will benefit from the on-going development in the Indian Ocean observing system.

The sustainability of the current observing system is paramount to continued progress in seasonal forecasting.

There is growing demand for predictions at decadal time scales, especially on sensitive topics such as the possibility of abrupt climate change, hurricane activity, drought and sea level rise. Although decadal prediction is still in its infancy, recent development show promising results, highlighting the role of the ocean initial conditions. The initialization of the ocean for decadal predictions is a major challenge for the next decade.

This paper summarizes the advances in ocean initialization during the past decade and offers a perspective of the major challenges ahead. The role of ocean observations in seasonal forecasting is discussed in section 2, showing a decade of progress in seasonal forecasts and ending with an evaluation of the current ocean observing system. Section 3 presents current research on decadal forecasting and discusses the observational needs. The final section offers a summary and discussion of future perspectives in climate forecasting.

2. Initialization of seasonal forecasts

Seasonal forecasting is now a routine activity in several operational centers around the world, requiring near-real time knowledge of the state of the ocean. Seasonal forecasting systems are based on coupled ocean-atmosphere general circulation models that predict both the sea surface temperatures (SSTs) and their impact on the atmospheric circulation. The chaotic nature of the atmosphere is taken into account by issuing probabilistic forecasts, obtained by performing an ensemble of coupled integrations. Because of the deficiencies in coupled models, the forecasts need calibration before the forecast is issued. The calibration is done by conducting a series of past seasonal hindcasts, which in turn requires ocean initial conditions for a historical period (typically 15-25 years), equivalent to an ocean reanalysis. The hindcasts are also needed for skill assessment.

Assimilation of observations into an ocean model forced by prescribed atmospheric fluxes is the most common practice for initialization of the ocean component of a coupled model. The emphasis is on the initialization of the upper ocean thermal structure, particularly in the tropics, where SST anomalies have a strong influence on the atmospheric circulation. Observational information about SST is essential. Most of the initialization systems also use subsurface temperature (from XBT's, TAO/TRITON/PIRATA and Argo), most recently also salinity (mainly from Argo), and altimeter derived sea-level anomalies (SLA). The latter usually needs the prescription of an external Mean Dynamic Topography (MDT), which can be derived indirectly from gravity missions such as GRACE and, in the near future, GOCE. Some of the initialization systems use an on-line bias correction scheme or relaxation to climatology to control the mean state.

Table 1 provides a summary of the ocean analyses used for initialization of operational or quasi-operational seasonal forecast systems. In all these systems, the initialization of the ocean and atmosphere is done separately, aiming at generating the best analyses of the atmosphere and ocean through comprehensive data assimilation schemes in both media. The quality of the ocean analyses is strongly dependent on the atmospheric fluxes used to drive the ocean model in the production of the first guess. The forcing fluxes are usually provided by atmospheric re-analyses (ERA40, JRA-25, NCEP), and by Numerical Weather Prediction systems. The production of seasonal forecasts is resource-demanding: the need for ensembles and calibration implies the integration of the coupled model for several hundreds of years. This computational burden limits the practical resolution of the ocean model, which is typically of the order of 1 degree with some equatorial refinement in the horizontal, and about 10 meters in the vertical in the upper ocean.

2.1 Estimating the upper thermal structure

The simplest way to initialize the ocean is to run an ocean model forced with atmospheric fluxes and with a strong relaxation of the model SST to observations. This technique (CNTL in what follows) would be satisfactory if errors in the forcing fields and ocean model were small. However, surface flux products and ocean models are both known to have significant errors. The uncertainty induced in the upper

2

ocean by using different wind products can be as large as the interannual variability. Assimilation of ocean observations is then used to constrain the estimation of the ocean state.

To evaluate the impact of data assimilation on the estimation of the ocean state it is necessary to verify against independent data, such as velocity, which is usually not assimilated. This exercise has led to major improvements in data assimilation methods, eventually increasing the information content of the ocean observations. Figure 1 illustrates this idea by showing results from the current and a previous generation of ocean analysis. The previous (first) generation (green line) assimilated only temperature data with poor or no multivariate relationship. The current (second) generation of ocean analysis use physically based multivariate balance relationships and often assimilate salinity and altimeter data (black and red lines). All analyses systems in Table 1 belong to this second generation. In the first generation of analyses the temperature error was improved with respect to the CNTL (blue line), but at the expense of degrading the ocean currents and salinity. In the second generation, exemplified in the study by the PEDOAS analysis (Alves and Robert, 2005), the both temperature and salinity errors are greatly reduced with respect to the CNTL without significant degradation of the ocean currents.

The importance of salinity observations is discussed in Fujii et al (2009). Figure 2 shows results from experiments conducted with the MOVE/MRI.COM-G system (Usui *et al.* 2006), and indicate that only when salinity observations are assimilated is it possible to represent the strong meridional salinity gradient in the Western Equatorial Pacific, with low salinity waters north of the equator (upper row). Results also show that without the balance relationship between temperature and salinity it is not possible to represent the high salinity of the South Pacific Tropical Water (lower row), leading to the erosion of the vertical stratification and eventual degradation of the barrier layer. The barrier layer can influence the representation of the interannual variability of SST, since it acts as a reservoir of warm water (above 28ºC), which can influence the development of ENSO when propagated eastwards by westerly winds.

2.2 Impact of ocean observations on seasonal forecast skill

The skill of the seasonal forecasts is often used to gauge the quality of the ocean initial conditions. This may not always be an appropriate measure, since the quality of the coupled model will determine the precision of the assessment - if the major source of forecast error comes from the coupled model, improvements in the ocean initial conditions would have little impact on the forecast skill. This is something to bear in mind when interpreting results of the impact of the ocean data assimilation on seasonal forecasts.

Several studies have demonstrated the benefit of assimilating ocean data on the prediction of ENSO (Alves *et al.* 2003, Dommenget *et al* 2004, Cazes-Boezio *et al* 2008, among others). The benefits are less clear in other areas, such as the equatorial Atlantic, where model errors are large. Balmaseda and Anderson (2009), using the latest version of the ECMWF seasonal forecasting system (S3), evaluate three different initialization strategies, each of which uses different observational information. Strategy i) uses ocean, atmospheric and SST information, strategy ii) uses atmospheric information and SST, and strategy iii) uses only SST. Differences between i) and ii) are indicative of the impact of ocean observations, and comparison of ii) and iii) are indicative of the impact of the atmospheric observation. A series of 5-member ensemble coupled hindcasts spanning the period 1987-2006 initialized 3 months apart, was performed with initial conditions from each method. Results show that the ocean initialization has a significant impact on the mean state, variability and skill of coupled forecasts at the seasonal time scale. Figure 3 shows the relative reduction in the monthly mean absolute error (MAE) resulting from adding information from the ocean and/or atmosphere observations for the 1-3 month forecast range   in the regions defined in Table 2. The observational information has the largest impact in the Western Pacific (EQ3): the combined information of ocean and atmospheric observations can reduce the MAE more than 50% in the first 3 months. With the exception of the Equatorial Atlantic (EQATL), the best scores are achieved by strategy i). This means that for the ECMWF system, the benefits of the ocean data assimilation and the use of fluxes from atmospheric (re)analyses more than offset problems arising from initialization shock.

The seasonal forecast skill can also be used to evaluate the ocean observing system. Fujii *et al* (2008), evaluate the impact of the TAO/TRITON array and Argo float data on the JMA seasonal forecasting

system by conducting data retention experiments: all available observations are assimilated in ALL; data from the TAO/TRITON array is excluded in NTT, and data from Argo is excluded in NAF. Sets of 11-member ensemble forecasts were conducted, starting 4 times per year for the period 2004-2007. Figure 4 shows Root Mean Square Errors (RMSEs) of SST forecasts (1-7 month range) in the regions defined in Table 2. The RMSEs are normalized by the equivalent RMSEs for persistence forecasts. Results show that TAO/TRITON data improves the forecast of SST in the eastern equatorial Pacific (NINO3, NINO4), and that Argo floats are essential observations for the prediction of the SST in tropical Pacific and Indian Oceans. Similar results have been obtained with the ECMWF seasonal forecasting system (Balmaseda *et al.* 2009).

3. Initialization of decadal forecasts

There is increasing interest in the possibility that decadal climate predictions may be able to provide information enabling better adaptation to climate change. On decadal timescales, the magnitude of natural internal variability of the climate system can be similar to, or greater than, anthropogenically-forced changes. Decadal climate predictions therefore attempt to capitalise on predictability of natural internal variability arising from relatively slow processes in the oceans by assimilating ocean observations (Smith *et al.* 2007, Keenlyside *et al.* 2008, Pohlmann *et al.* 2009).

It is likely that decadal climate predictions will be substantially improved by Argo data. However, quantifying any improvement is complicated by the fact that Argo data have only been available for a few years. One possible way to evaluate the observing system is to conduct synthetic observing system experiments using a perfect model scenario. Zhang et al 2007 use a coupled model simulation of the 20th century as "truth", and assimilating "observations" from that "truth" experiment into a second, independent experiment using a Coupled Data Assimilation system (CDA) based on an Ensemble Kalman filter. "Observations" are taken from the model at places and times where real observations were taken. These observations are then assimilated into a coupled model ensemble that is subjected only to preindustrial forcings. The assimilation is able to successfully reconstruct twentieth-century ocean heat content variability and trends in most locations (not shown). However, the Atlantic Meridional Overturning Circulation (AMOC) was difficult to constrain using the twentieth-century observational network. The AMOC is only successfully reproduced when the Argo observing system is simulated (figure 5). These results indicate that Argo profiles are critical for developing a nowcasting capability for the AMOC.

A series of multi-model decadal hindcasts have been conducted as part of the ENSEMBLES project, spanning the period 1960 to 2005. Figure 6 shows results from hindcasts using the MetOffice decadal prediction system (DePreSys, Smith *et al.* 2007), with improved correlations between observed and hindcast near-surface temperatures in recent years. Whether this improvement is caused by the increased observational coverage due to Argo floats is still open, due to the limited sample.

4. Summary

Assimilation of ocean observations for the ocean initialization in seasonal forecasts has become a common practice, with several institutions around the world producing routine ocean re-analyses to initialise their operational seasonal forecasts. Most initialization systems assimilate SST, sea level anomalies from altimeter data, subsurface temperature and salinity and sea level via multivariate schemes, imposing physical and dynamical constraints among different variables. The resulting ocean reanalyses are a valuable data resource for climate variability studies and have the advantage of being continuously brought up to real time. They are being used experimentally for the initialization of decadal forecasts.

Results from several initialization systems show that the assimilation of ocean data reduces the uncertainty in the upper ocean thermal structure and improves seasonal forecast skill, although ultimately, the impact of initialization in a seasonal forecasting system will depend on the quality of the coupled model.

Assessing the ocean observing system from the perspective of decadal forecasts is more challenging: the observing system has not been stationary enough to even monitor accurately decadal variability. However, results from synthetic observing system experiments conducted assuming sustainability of the current observing system indicate that Argo data has the potential for constraining the

mid-deep ocean variability responsible for decadal variability. Decadal hindcasts initialized during the Argo period are slightly better than those initialized prior to Argo, but the sampling size is obviously too small for any claim of statistical significance. Sustained observations over the full depth of the ocean are also expected to be important in constraining the estimate of decadal and longer variability that may affect decadal prediction.

Independent observations, not directly assimilated in the analysis process, have provided invaluable contributions to the improvement of assimilation methods, thus increasing the information content of the assimilated observations. There is still evidence that the current generation of initialization systems do not make optimal use of the existing observations. For instance, in most of the existing operational systems, the initialization is done in uncoupled mode: by using forcing fluxes from atmospheric re-analysis, the uncoupled initialization has the advantage of incorporating relevant atmospheric variability, such as westerly wind bursts. However, this strategy can lead to initialization shock, which is likely to be larger in those regions where model errors are large. This is the case of the Equatorial Atlantic, where the existence of initialization shock seems to be detrimental to the forecast skill. A third generation of initialization systems is being developed, where the oceanic and atmospheric initial conditions are generated simultaneously using a coupled model and so have the potential of retaining the balances relevant for the coupled system (Sugiura *et al* 2009, Zhang *et al* 2007, among others).

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.

Alves. O. and C. Robert (2005). Tropical Pacific ocean model error covariances from Monte Carlo simulations. *Q. J. R. Meteorol. Soc.*, **131**, 3643-3658.

Balmaseda, M.A., Arthur Vidard and David Anderson (2008). The ECMWF ORA-S3 ocean analysis system. *Mon. Wea. Rev*, 136, 3018-3034.

Balmaseda, M.A., and D. Anderson. 2009. Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters 36(L01701):doi:10.1029/2008GL035561.*

Behringer, D. W. (2007). The Global Ocean Data Assimilation System at NCEP. *11th Symposium on Integrated Observing and Assimilation Systems for Atmosphere,Oceans,and Land Surface, AMS 87th Annual Meeting , San Antonio, Texas, 12pp*

Cazes-Boezio, G., D. Menemenlis D, C.R. Mechoso, 2008: Impact of ECCO ocean-state estimates on the initialization of seasonal climate forecasts. *J. Clim.*, **21**,1929-1947.

Dommenget, D. and D. Stammer, 2004: Assessing ENSO Simulations and Predictions Using Adjoint Ocean State Estimation. *J. Climate*, 17, No. 22, 4301-4315.

Fujii, Y., S. Matsumoto, M. Kamachi, and S. Ishizaki (2009) Estimation of the equatorial Pacific salinity field using ocean data assimilation system. Submitted to *Adv. in Geosciences*.

Fujii, Y., T. Yasuda, S. Matsumoto, M. Kamachi, and K. Ando (2008) Observing System Evaluation (OSE) using the El Niño forecasting system in Japan Meteorological Agency. Proceedings of the oceanographic society of Japan fall meeting (in Japanese).

Fukumori, I., 2002: A partitioned Kalman filter and smoother. *Mon. Wea. Rev.*, **130**, 1370-1383.

Keppenne, C.L., M.M. Rienecker, J.P. Jacob, and R. Kovach (2008) Error Covariance Modeling in the GMAO Ocean Ensemble Kalman Filter. *Mon. Wea. Rev.*, **136**, 2964-2982, doi:10.1175/2007MWR2243.1.

Keenlyside N, M Latif, J Jungclaus, L Kornblueh, E Roeckner, 2008, Advancing decadal-scale climate prediction in the North Atlantic Sector, *Nature*, 453, 84-88.

Martin, M.J., A. Hines, M.J. Bell. 2007. Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Quarterly Journal of the Royal Meteorological Society* **133**:981-995.

Murphy, J.M., D.M.H. Sexton, D.N. Barnett, G.S. Jones, M.J. Webb, M. Collins and D.A. Stainforth, 2004, Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768-772.

Pham D. T., J. Verron and M. C. Roubaud, (1998) A singular evolutive extended Kalman filter for data assimilation in oceanography, *J. Mar.Syst.*, **16**, 323-340.

Pohlmann, H., J. Jungclaus, J. Marotzke, A. Köhl, and D. Stammer, 2009, Improving Predictability through the Initialization of a Coupled Climate Model with Global Oceanic Reanalysis, *J. Climate*, accepted.

Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa (2008) Development of a 4-dimensional variational coupled data assimilation system for enhanced analysisn and prediction of seasonal to interannual climate variations. *J. Geophys. Res. .* **113**, C10017, doi:10.1029/2008JC004741.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris and 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

Usui, N., S. Ishizaki, Y. Fujii, H. Tsujino, T. Yasuda, and M. Kamachi (2006) Meteorological Research Institute Multivariate Ocean Variational Estimation (MOVE) System: Some early results. *Adv. Spa. Res.*, **37**, 806-822.

Zhang, S., M.J. Harrison, A. Rosati, and A. Wittenberg (2007) System Design and Evaluation of Coupled Ensemble Data Assimilation for Global Oceanic Studies. *Mon. Wea. Rev.*, **135**, 3541-3564.

Table 1: Summary of different ocean assimilation systems used in the initialization of operational and quasi-operational seasonal forecasts.

MRI-JMA http://ds.data.jma.go.jp/tcc/tcc/products/elnino/index.html MOVE/MRI.COM-G. Multi-variate 3DVAR. Usui *et al.* 2006
ORA-S3 (ECMWF) www.ecmwf.int/ products/forecasts/d/charts/ocean **ECMWF System 3** http://www.ecmwf.int/products/forecasts/d/charts/seasonal/ HOPE/OI: Multivariate OI. Balmaseda et al 2008.
POAMA – PEODAS (CAWCR, Melbourne) http://poama.bom.gov.au/research/assim/index.htm Multivariate Ensemble OI. Robert and Alves 2005.
GODAS (NCEP) http://www.cpc.ncep.noaa.gov/products/GODAS/ MOMv3/3Dvar. Behringer 2007.
MERCATOR (Meteo France) http://bulletin.mercator-ocean.fr/html/welcome_en.jsp PSYS2G. Multivariate reduced order Kalman filter. Pham et al 1998.
MO (MetOffice) http://www.metoffice.gov.uk/research/seasonal/ Glosea3.OI. Martin *et al* 2007
GMAO ODAS-1 http://gmao.gsfc.nasa.gov/research/oceanassim/ODA_vis.php **GMAO Seasonal Forecasts:** http://gmao.gsfc.nasa.gov/cgi-bin/products/climateforecasts/index.cgi ODAS-1 : OI and EnKF Keppenne *et al*. 2008

Table 2: Definition of area average indices

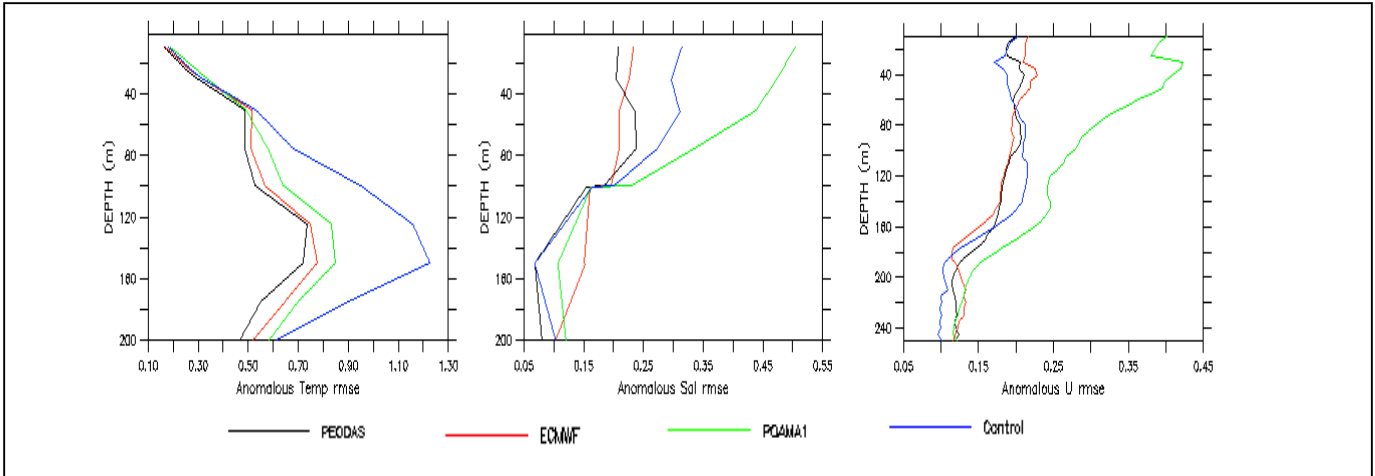|			
NINO12:	10-0ºS, 90-80ºW	EQPAC	5ºS-5ºN, 130ºE-80ºW
NINO3	5ºS-5ºN, 90-150ºW	EQIND	5ºS-5ºN, 40º-120ºE
NINO34	5ºS-5ºN, 170-120ºW	WTIO	10ºS-10ºN, 50º-70ºW
NINO4	5ºS-5ºN, 160ºE-150ºW	STIO	10ºS-0ºN, 90º-110ºE
EQ3	5ºS-5ºN, 150ºE-170ºW	EQATL	5ºS-5ºN, 70ºW-30ºE
NINO-W	0º-15ºN, 130º-150ºE	NSTRATL	5ºN-28ºN, 80ºW-20ºE

Figure1. RMS error of interannual anomalies of temperature (left), salinity (middle) and zonal current (right). Shown are the PEODAS reanalysis (black), the old POAMA re-analysis (green), the ECMWF ORA-S3 (red) and the Control simulation (blue). The verifying observations are from the TAO mooring at location 165°E.
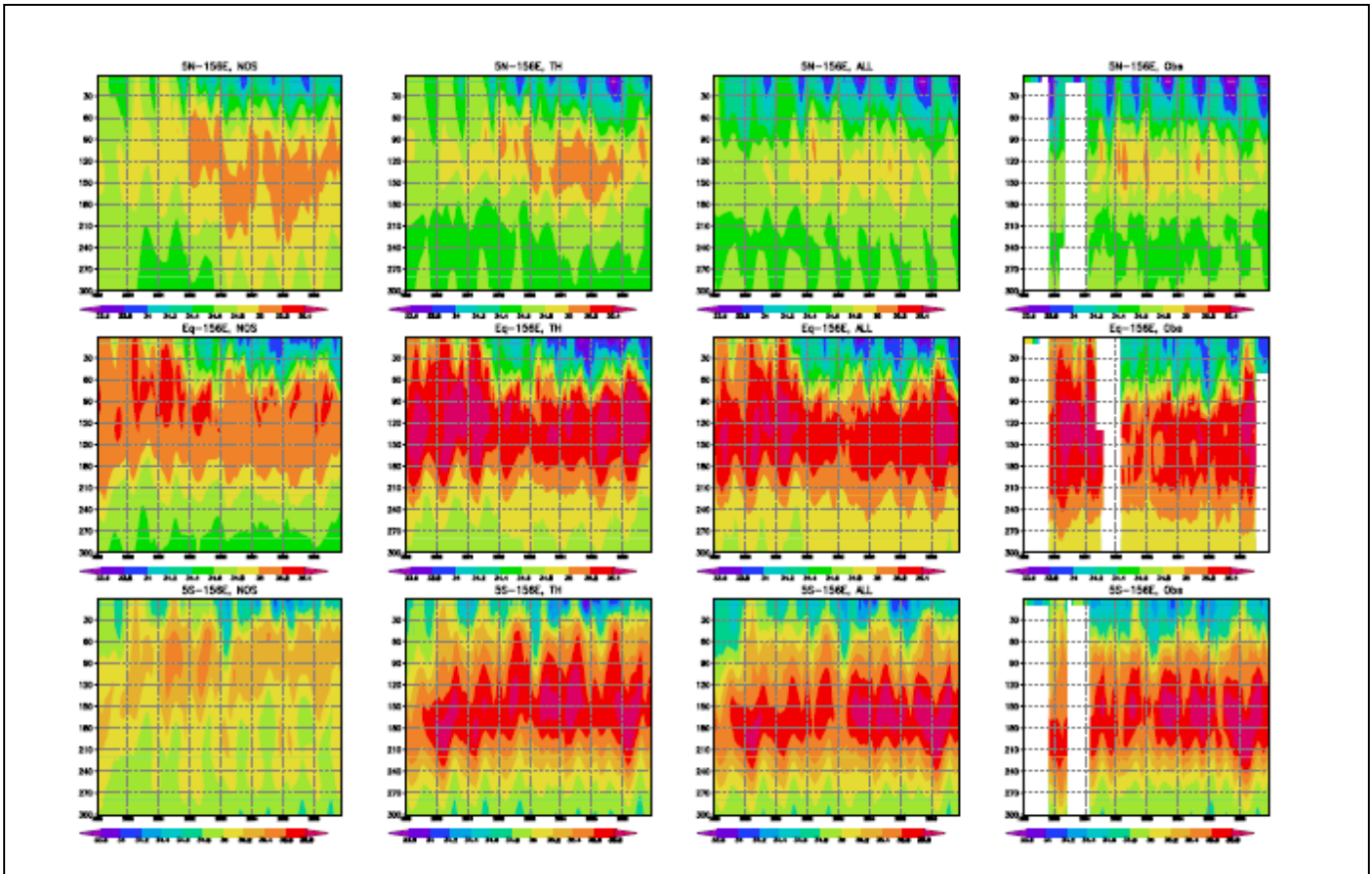

Figure 2: Depth-Time sections of salinity for the period 1999-2006 at 156E and 5N (upper row), Equator (central row) and 5S (lower row). The left column is for experiment NOS, where only temperature is assimilated and not balance T-S relationship is imposed. The second column, for experiment TH, is for the experiment when only temperature data is assimilated including the T-S relationship. The third column is for experiment ALL, where salinity and temperature are assimilated. The right column shows the observational value from the TRITON array. Vertical grid lines mark the beginning of each year. Horizonal grid lines interval is 30 m.

**Reduction (%) in SST forecast error.
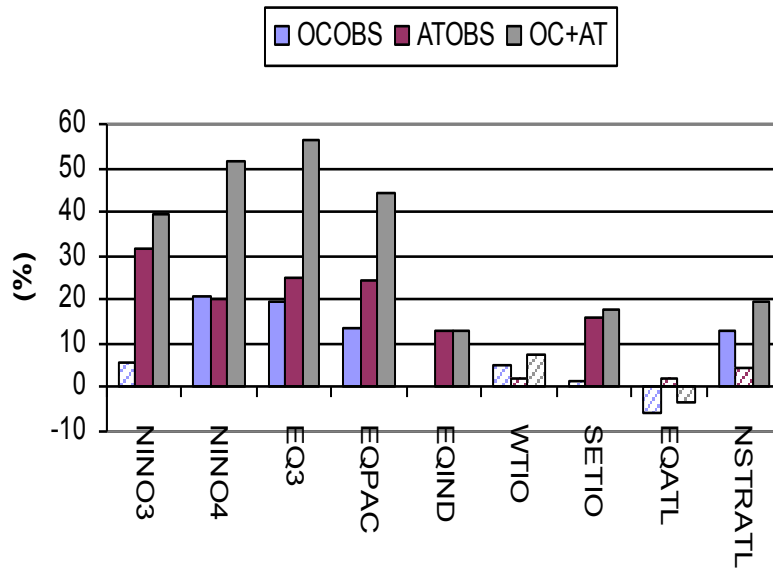Forecast Range: 1-3 months. Period 1987-
2000**

Figure 3: Impact of initialization in forecast skill for the different regions in Table 2, as measured by the reduction in mean absolute error for the forecast range 1-3 months. Solid bars indicate differences are above the 80% significance level. Blue (OCOBS) indicates the impact of ocean observations. Red (ATOBS) indicates the impact of atmospheric data, while grey (OC+AT) represents the combined impact of atmospheric and oceanic data.

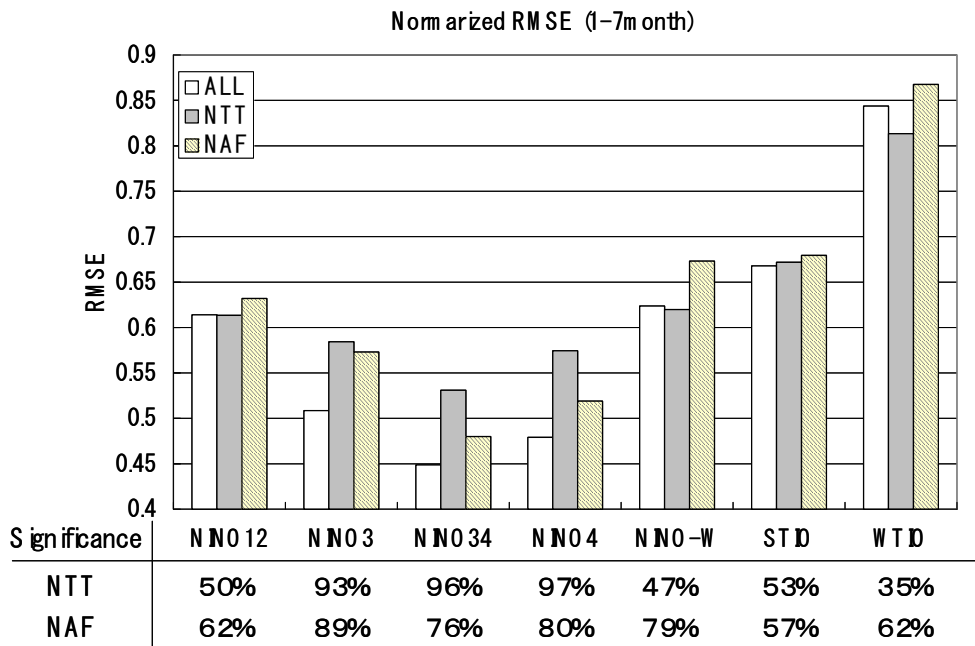| Significance	NINO12	NINO3	NINO34	NINO4	NINO-W	STIO	WTIO
NTT	50%	93%	96%	97%	47%	53%	35%
NAF	62%	89%	76%	80%	79%	57%	62%

Figure 4. Impact of withholding TAO/TRITON (NTT) and Argo (NAF) from the initialization of seasonal forecasts in the skill of the JMA forecasting system. The bars show the RMSEs of the 1-7 month forecasts of monthly SST anomalies normalized by the RMSEs of persistence forecasts. Significance level of the positive impact is given below the bar graph. The regions are defined in Table 2. From Fujii *et al* . (2008).
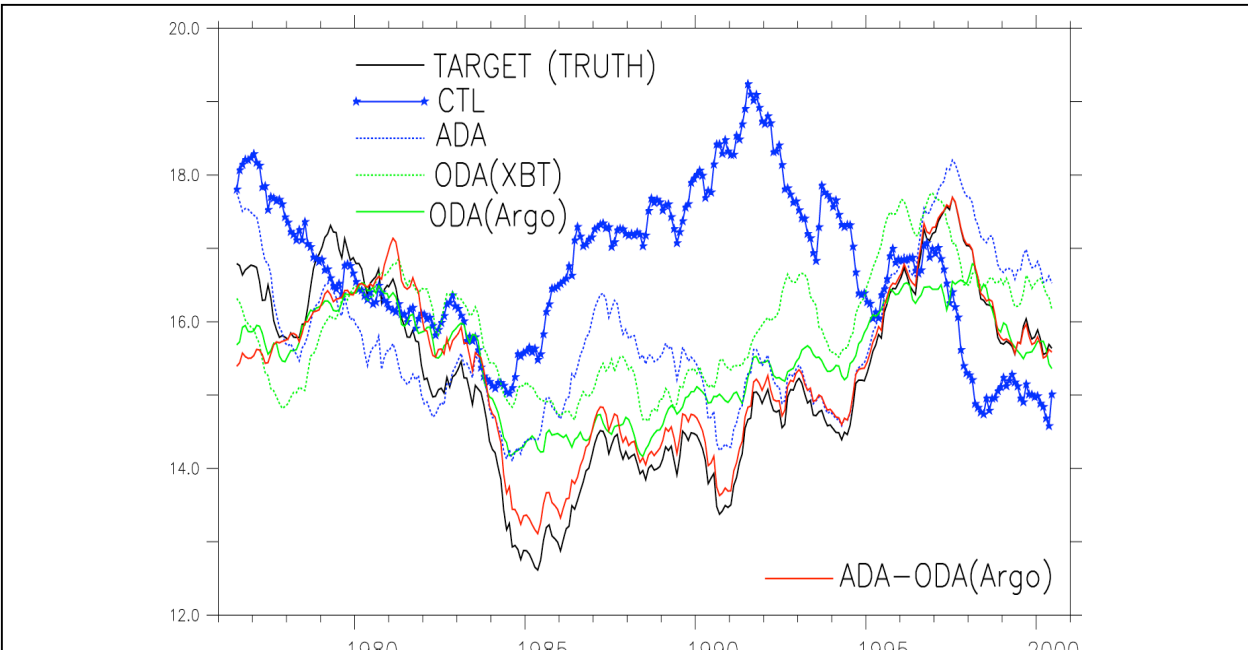
Figure 5 Perfect model studies of the capability of the CDA and observing systems to analyze the AMOC are being carried out at GFDL. The black curve represents a solution from a model simulation. The blue curve represents an independent simulation from which the CDA was initialized. The ADA curve represents only atmospheric data assimilation; the ODA curves represent only ocean data assimilation using the XBT network and the Argo network. The red curve is the coupled ADA and ODA(Argo) and shows the CDA fully recovers the targeted AMOC variability. The y-axis represents AMOC in Sv and x-axis is time in years.
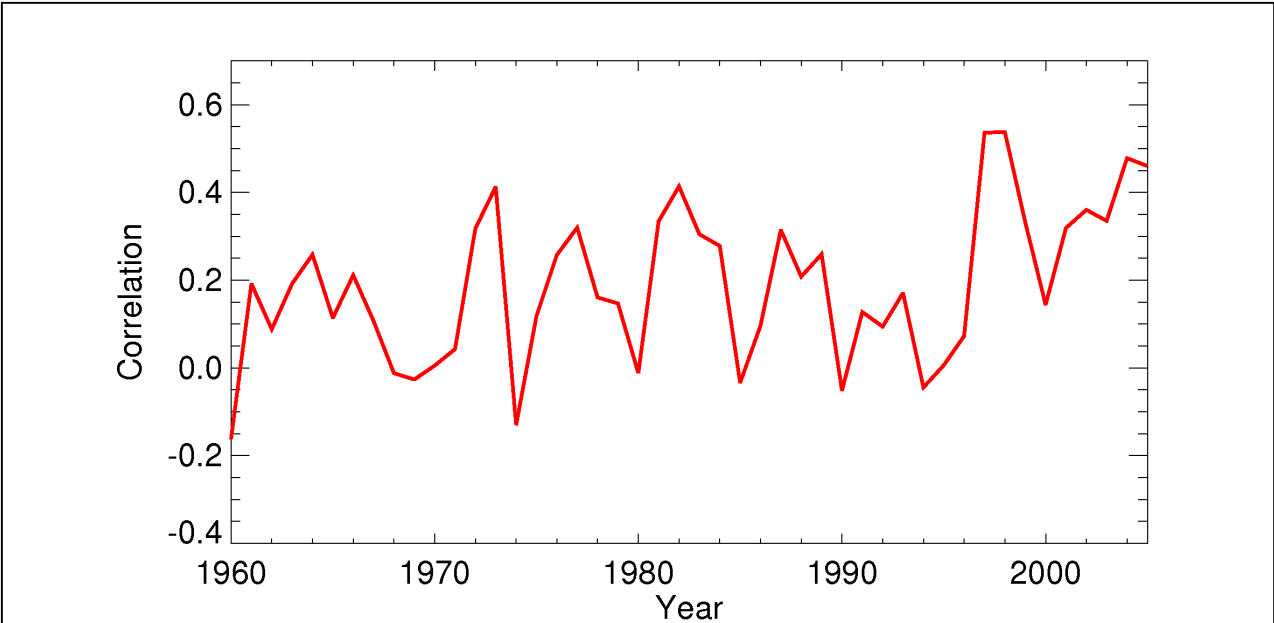

Figure 6: Spatial correlation between observed (HadCRUT3) and predicted (DePreSys) three-year mean surface temperature anomalies (relative to 1958-2001).

9