Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project

Thematic Priority: Global Change and Ecosystems

Deliverable D 1.10

Hindcasts using the lagged average method and earlier start dates carried out with GloSea following the stream 1 setup

Due date of deliverable: month 36
Actual submission date: month 36

Start date of project: 1 September 2004 Duration: 60 Months

Organisation name of lead contractor for this deliverable: ECMWF

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

ENSEMBLES

Deliverable D 1.10

Authors: Matt Huddleston, Malcolm MacVean, Ian Culverwell, Bruce Ingleby, Alberto Arribas and Chris Harris

Due: month 36

Hindcasts using the lagged average method and earlier start dates carried out with GloSea following the stream 1 setup

Summary

The first 18 months of the ENSEMBLES project have produced a number of sets of hindcasts for the 1991-2001 period using a variety of coupled models and ensemble techniques. The upgraded Hadley Centre GloSea system has improved performance over the current operational system and a first intercomparison of the relative benefits of the decadal forecasting system (DePreSys) compared to GloSea is shown.

A range of ensemble methodologies have been explored within the EU FP6 ENSEMBLES integrated project. These include multi-model, perturbed physics and stochastic physics technologies. Currently, the multi-model method appears to slightly outperform the other single-model techniques but there are large uncertainties. In particular the performance of the various systems varies significantly from region to region, with start date and with forecast lead time. This suggests that other uncertainties are not being sampled and the length of the hindcasts is currently inadequate to get reliable forecast statistics.

A key recommendation is thus that GCM performance should be enhanced and ways sought to account for missing areas of forecast uncertainty, whilst also continuing to develop the various streams of complementary ensemble methodologies.

Ongoing research will guide both the design of the ENSEMBLES climate forecasting system and the next generation of operational seasonal forecasting systems at the Met Office.

Introduction

The chaotic nature of the climate system make climate forecasts sensitivities to small errors in the initial state of forecasts, model formulation and boundary forcing (e.g. solar irradiance). Thus, individual forecasts are of limited use and ensembles of forecasts are used to assess the range of possibilities in future climate on a range of timescales.

Four main techniques for the generation of ensembles for climate forecasting have been proposed and explored to varying degrees in several projects over the last decade:

1) Uncertainties in single model ***initial conditions*** are accounted for by generating an ensemble of analysis (of atmosphere and/or ocean conditions). The perturbations can of an optimal statistical nature or based on insight into the dynamics of the physical system.

2) The ***multi-model*** method samples errors that occur due to structural inadequacy in climate models (e.g. different model formulation, bugs) by using different coupled models together in an ensemble (e.g. Palmer et al., 2004). This method inherently also uses multi-analyses to initialise the multi-model which also samples analysis errors.

3) There are uncertainties in the specification of the parameters that are used in the parameterizations in climate models and by perturbing these structural parameters within a single model, errors in specific model formulations can be accounted for (e.g. Lin and Neelin, 2000, Collins et al., 2006). Currently, ***perturbed parameter*** formulations are implemented in one of two ways: (a) each ensemble member has a fixed set of parameters (as in the DEFRA QUMP project) or (b) parameters can vary randomly during a forecast (as developed in NWP).

4) There are errors that occur due to having unresolved scales in the models, and energy is dissipated at these truncation scales erroneously due to this sub-grid scale uncertainty. The impact of the unresolved processes on larger scales is also missing. The impact of the unresolved scales (e.g. on potential vorticity) can be approximated in the model by random (***stochastic***) perturbations of physical fields (e.g. Gray and Shutts, 2002).

Currently, seasonal forecasting systems at the Met Office use a combination of initial value perturbations and multi-model strategies. Clearly, a combination of these four alternatives may be desirable in any forecasting system, and even then there may be sources of error unaccounted for (e.g. physical processes no model represents). Research is ongoing to account for the remaining sources of uncertainty in ENSEMBLES and these techniques will be available for seasonal forecasting at a future stage.

The purpose of the EU FP6 ENSEMBLES integrated project, in which the Hadley Centre is taking a lead, is to build an improved probabilistic climate forecasting system for Europe and the globe, linking the developments through to specific applications. A key aim in phase 1 (first 18 months ending March 2006) is to reduce uncertainty in climate forecasts across the range of forecast timescales using the latest technology. This report details the contributions to those aims by the long-range forecasting group in the Hadley Centre.

GloSea ensemble methodology

Following from our contribution to the EU FP4 DEMETER project and our subsequent construction of an operational multi-model seasonal forecasting system (EURO-SIP) with ECMWF and Météo-France, the latest coupled models have been installed at ECMWF to form a benchmark multi-model baseline ensemble forecasting system for seasonal to decadal timescales (see Huddleston et al., 2006, GMR V.c).

As a contribution to ENSEMBLES, the upgraded GloSea ocean analysis and CGCM forecasting system has been used to produce a number of 9 members ensemble hindcast sets for May and November starts, 1991-2001:

a) DEMETER style wind-stress perturbed 3-member ocean analysis ensemble, with SST perturbations added at forecast time to generate a 9 member ensemble This is the benchmark method (subsequently referred to as z01a).
b) Ocean initial states are taken from the 3 days leading up to the forecast date to give a *lagged average* ocean ensemble. These are then perturbed with SST perturbations (as per -a- and subsequently referred to as z01c).
c) As per method -a-, but started from dumps valid 2 weeks earlier. This is to investigate sensitivity to initial conditions by making a small change to the seasonal cycle, as well as allowing information retention studies when used in comparison with the other ensemble sets (subsequently referred to as z01d).

Note that method (a) above necessarily perturbs the *best* estimate of the ocean state giving perturbed ensemble members that tend to perform worse than the control. This is a common result in ensemble forecasting systems that impose perturbations, be they pragmatic (as in this case where the perturbations used follow from insight into the dynamics of the equatorial oceanic system) or “optimal” (e.g. by perturbing the fastest growing modes as used at ECMWF for medium range weather forecasting).

Method (b) potentially overcomes this issue as each of the 3 ocean states is in balance, but has instead sampled time-related uncertainty. This method could also reduce the cost of generating an ocean ensemble as only one ocean analysis is needed.

A description of the research and production underpinning the generation of the ocean ensembles is in Annex 1.

Currently, perturbed physics ensembles have not been attempted with GloSea. In the QUMP framework established in the Hadley Centre, individual members are flux-corrected and need a model climatology (Collins et al., 2006). It is not clear how this would be implemented in the current GloSea system. The DePreSys decadal forecasting system is being used to perform a set of hindcasts using this methodology and preliminary results from this are presented below. It may be that constantly varying physics perturbations would overcome the problems with implementing such a system in GloSea. The RP (random-parameter) method for perturbing parameterisations developed in NWP, will be available for use in the next generation of seasonal forecasting model HadGEM1a.

A stochastic physics package is also available for the new Hadley Centre climate model, HadGEM1a (Stochastic Energy Backscatter Scheme), and this is being further developed to include stochastic perturbations to convective systems. This development is likely to impact positively on tropical performance.

Seasonal to decadal experiments within ENSEMBLES in phase 1.

Seasonal hindcast experiments for phase 1 have been performed in a common structure with the other ENSEMBLES participants. Hindcasts have been performed for May and November start dates for the eleven year period 1991 to 2001 in 9 member ensembles out to 7, 12 or 14 months. Additionally, two individual decadal hindcast experiments have been made from November 1965 and 1994. These experiments have necessitated the production of a 3 member ensemble of ocean analyses for 1987-2001 and 1961-1966 (see Annex 1).

Results from hindcast experiments

A preliminary assessment of the various ensemble methodologies was made for an ENSEMBLES milestone report for March 2006. The conclusions are that the multi-model technique performs slightly better than stochastic physics ensembles. Note that significant improvements are expected in the coming years in stochastic schemes. The perturbed physics ensembles had not been fully assessed at the time of this milestone report, but this methodology is likely to improve the single-model concerned and therefore the multi-model system.

In general, the performance of the GloSea system is improved over the current operational system (see, as an example, the SST forecast statistics for the central Pacific in figure A). This improvement may be due to the inclusion of greenhouse gases, a contribution that has been identified as improving the detection of trends in seasonal forecasts (Doblas-Reyes et al., 2006), but may also be due to the improved ocean analyses and ensemble generation strategy.

For the May start dates, the performance of the GloSea system, for the first time, can not only be compared with those of ECMWF and Météo France, but also with the Hadley Centre DePreSys decadal forecasting system (Smith et al., 2006). This DePreSys system uses HadCM3 in an *anomaly coupled* mode to avoid forecast drift where the model is initialised with climate anomalies relative to it's own climatology, and the members of the ensemble use physics perturbations (necessitating the calibration of each forecast ensemble member by its own model climatology) following the QUMP research project (Smith et al., 2006, and Collins et al., 2006). The DePreSys performance appears to be better at longer lead time for this region and start month. It is not clear at this time whether these differences are due to the initialisation method, the model structure or the ensemble methodology, although the rate of growth of the spread of the ensemble for DePreSys and GloSea is similar (both models being based on a common atmosphere model and a similar ocean).

In general it can be seen that ensemble spread is too small in comparison to the RMS errors of the forecasts for all the forecasting systems, although this problem is particularly apparent for the ECMWF and Météo-France hindcasts.

Figure A shows that the lagged average (z01c) and wind-stress perturbation methods (z01a) for generating initial condition perturbations and the DePreSys hindcasts give complementary results, with the wind-stress perturbation method performing better in the NINO3.4 region of the central equatorial Pacific for lead times less than 6 months, and DePreSys for longer lead times. However these results vary markedly and are therefore not robust; one method produces best results at certain lead times in certain regions for certain start dates, the other methods perform better on the other occasions.

This highlights the non-stationary performance of the systems and suggest that other systematic errors not being sampled may be dominant (e.g. no coupled model is adequately representing all the key climate process including the quasi-biennial oscillation, the Madden-Julian oscillation and the El Nino-Southern Oscillation) and also that the hindcast sample is statistically quite small. A key recommendation is thus that GCM performance should be enhanced and ways sought to account for missing areas of forecast uncertainty, whilst also continuing to develop the various streams of complementary ensemble methodologies.

References:

Doblas-Reyes, F, R. Hagedorn, T. Palmer, and J.-J. Morcette, 2006: Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. Geophys. Res. Lett., 33, L07708, doi:10.1029/2005GL025061.

Collins, M., Booth, B. B., Harris, G. R., Murphy, J. M., Sexton, D. M. H.and Webb, M. J., 2006: "Towards quantifying uncertainty in transient climate change". Climate Dynamics, 27, 127-147.

Gray, M.E.B. and Shutts, G.J., 2002: A stochastic scheme for representing convectively generated vorticity sources in general circulation models. APR Turbulence and Diffusion Note No. 285. Met Office, UK.

Ingleby, B. and Huddleston, M., 2006: Quality control of ocean temperature and salinity profiles - historical and real-time data. Journal of Marine Systems, 65, 158-175.

Lin, J. W.-B., and J.D. Neelin, 2000: Influence of a stochastic moist convective parameterization on tropical climate variability. Geophys. Res. Lett., 27, 3691-3694.

Palmer, T.N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Diez, F.J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J.-M. Terres, M. C. Thomson, 2004. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). Bulletin of the American Meteorological Society, 85, 853-872

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K. and Murphy, J.M.: Improved surface temperature prediction for the coming decade from a coupled climate model. Science, 2007, in press.

Troccoli A., Haines, K. 1999 Use of temperature-salinity relation in a data assimilation context. J. Atmos. Ocean. Tech., 16, 2011-2025.

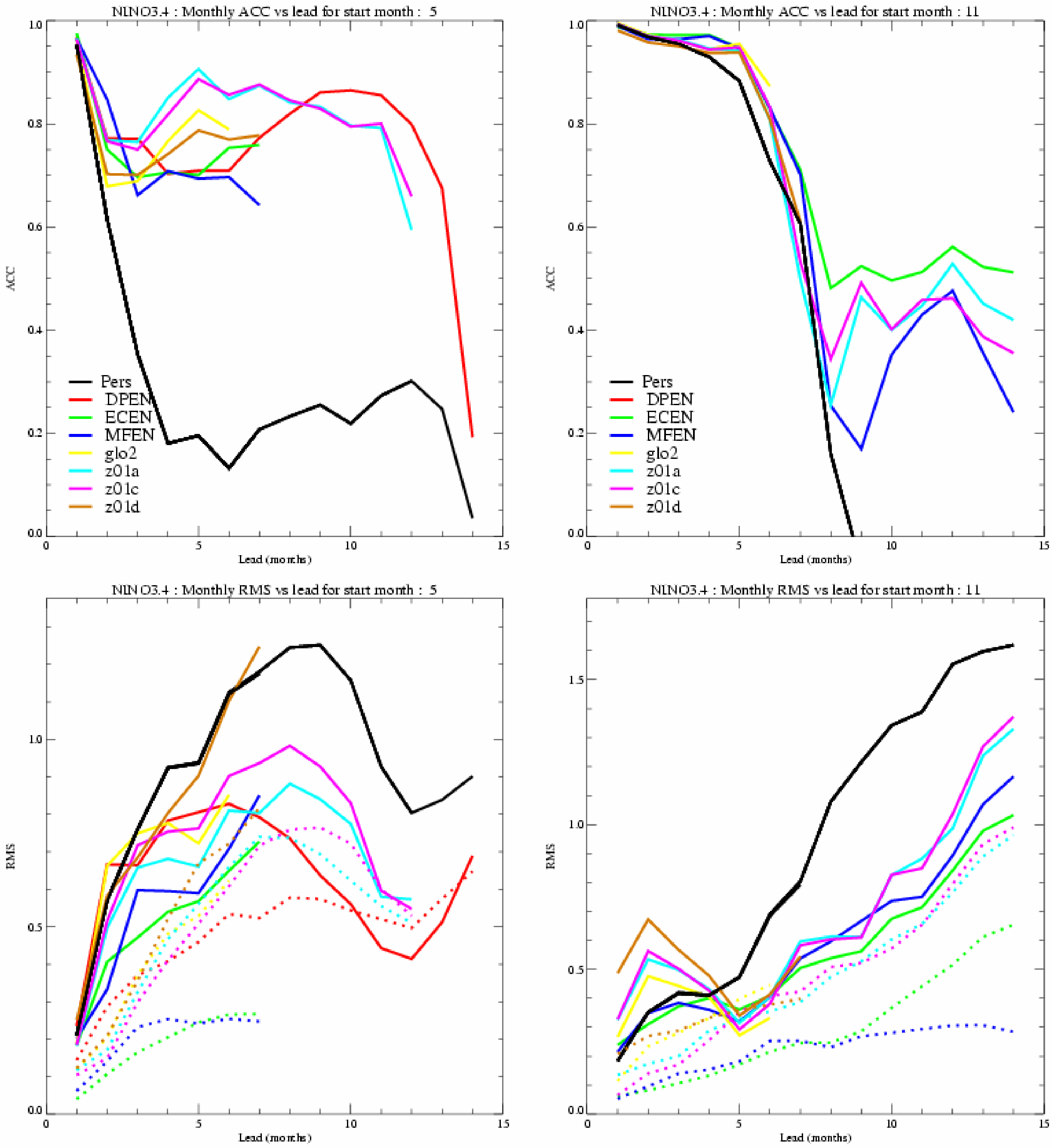

Figure A: Anomaly correlation (top) and RMS statistics (bottom) for sea surface temperature for the ensemble mean a number of hindcasts started in May (left) & Nov (right), 1991-2001, for the NINO3.4 central equatorial Pacific region compared with HadISST observations at different lead times. The solid black line gives the result for a persistence (Pers) forecast where anomalies at the start of the forecast continue. Red is for forecasts from the Hadley Centre decadal forecasting system DePreSys (DPEN – May only), green is for the ECMWF system (ECEN), dark blue for the CNRM/Météo France (MFEN) system, yellow is for the current Met Office operational GloSea2 system, and cyan, pink and brown are for variants of the upgraded GloSea system (z01a, z01c and z01d). Dotted lines in the lower plot show the ensemble spread which should ideally match RMS errors for the system concerned. Note that DePreSys hindcasts are calibrated using the individual member model climatology, whilst all others use the forecast climatology of all forecasts for the whole period (following experiment design).

ANNEX 1

Ocean analysis ensemble calibration for EU FP6 ENSEMBLES

At the start of the ENSEMBLES project, significant efforts were made to examine the hindcast sets that were generated for the EU FP5 project ENACT to find the best ocean analysis scheme for CGCM hindcast production. The wind stress and SST perturbations provided by ECMWF for FP6 ENSEMBLES were then combined with the data assimilation development done within the ENACT project and an updated set of ocean observations (now including over 7 million observations – see Ingleby & Huddleston, 2006) used to produce a new set of ocean analyses. A new 3 member ocean analysis ensemble was then produced for 1987-2001 and 1961-1966 to initialize ENSEMBLES seasonal to decadal hindcasts.

The model analysis scheme was tuned to improve the mean state and variability. Figure B shows an example of the improvements made by including the water mass preservation scheme (Troccoli and Haines, 1999) for salinity concentrated in the Atlantic, and while figure C shows the impact the tuning has had on a Taylor diagram. The largest impact has been on reducing subsurface salinity variance by including the assimilation of salinity data and the water-mass preservation scheme. The latest analyses have significantly better statistics than the current operational ocean analysis.

Finally, careful calibration of the ocean ensemble perturbations was made to ensure that subsurface intra-ensemble anomalies were representative of known uncertainties. Figures D and E show the observed (Ingleby and Huddleston 2006), and intra-ensemble standard deviation of temperature and salinity. For temperature, the spatial pattern of variability is well represented (at least for the equatorial regions) and the intra-ensemble spread is of the order of 20% of the observed variability. Note the areas of highest variability will have the highest errors associated with them and these areas will be highly correlated. Given that it is the errors in the variability we are trying to represent in the ensemble and not the variability itself, this figure seems reasonable (excepting the too-large intra-ensemble spread in the Atlantic) . This suggests that wind-stress perturbation methodology for perturbing initial conditions is valid for the tropical oceans.

For salinity, the subsurface tropical intra-ensemble variability has a different distribution to that of temperature. The wind stress methodology, surprisingly, does generate intra-ensemble variability at more than 20% of the observed variability; a result similar to the spread in the temperature ensemble. The location of the maximum variability in the west Pacific is too far west however, and there is anomalous spread in the ensemble in the equatorial Indian Ocean. Intra-annual sub-surface variability in salinity is much less well observed however and so it is not easy to ascertain an appropriate level at this time.

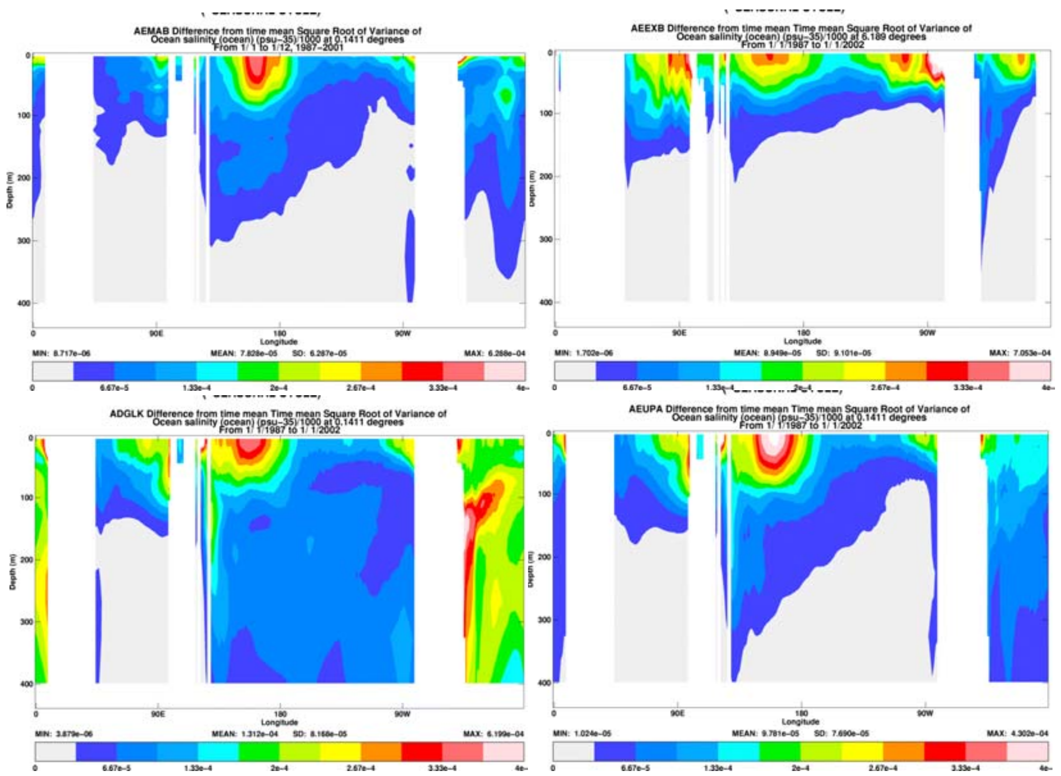

Figure B: A comparison of sub-surface salinity variance in equatorial-depth sections for the ENSEMBLES ocean analyses generated for 1987-2001. Top left is monthly salinity variance in an objective model-free analysis (i.e. close to observations), top right is for a forced ocean model, bottom left is a forced model with data assimilation, and bottom right is a forced model with data assimilation and a module to preserve water mass properties.

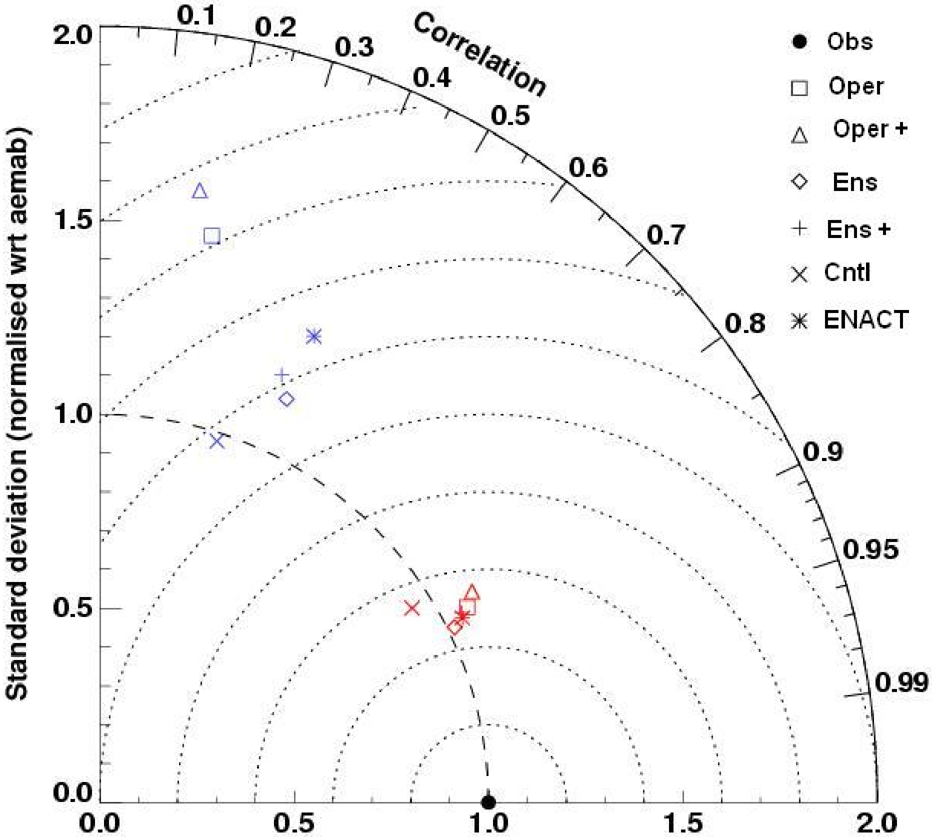

Figure C: Comparison of upper ocean (average of top 200m) monthly mean temperature and salinity anomalies along the equator for 1987-2001. Red symbols are for temperature, blue ones are for salinity. *Obs* are gridded observations from an objective analyses (Ingleby and Huddleston, 2005). The *oper* is the current operational ocean analysis. The *Ens* is from the new ENSEMBLES ocean analysis. *Cntl* is from a forced ocean model (no ocean data assimilation) and *ENACT* is the equivalent ENACT ocean analysis. Names with a **+** following them indicate that the analyses were perturbed using time varying wind perturbations.

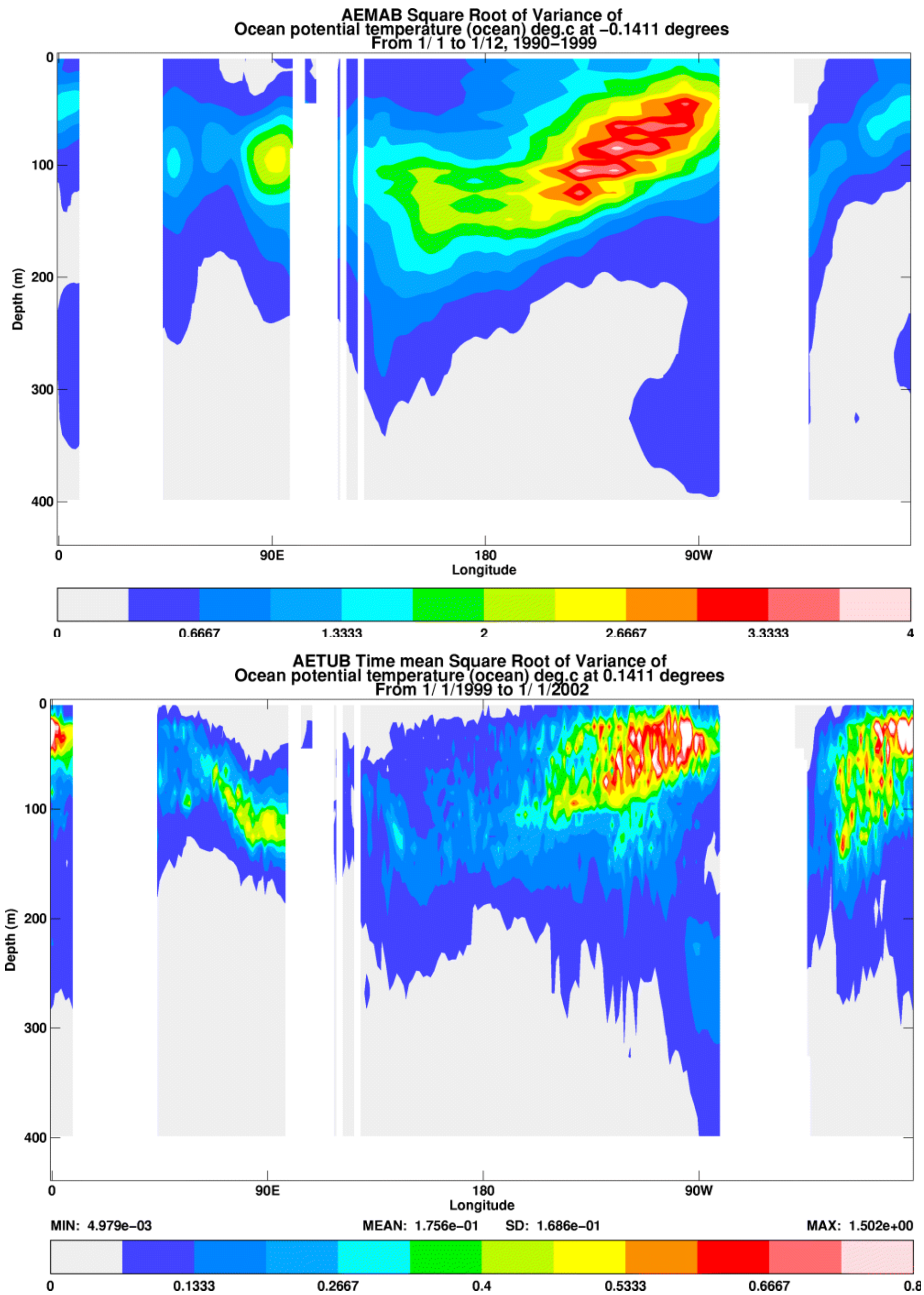

Figure D: Upper panel shows the standard deviation of monthly mean anomalies from an objective analysis for 1990-1999 for potential temperature on an equatorial-depth section. Note the striping is due to the fixed location of the TAO array moorings. The lower panel shows the intra-ensemble standard deviation between a perturbed ensemble member and the control for 1999-2001. Note different scales.

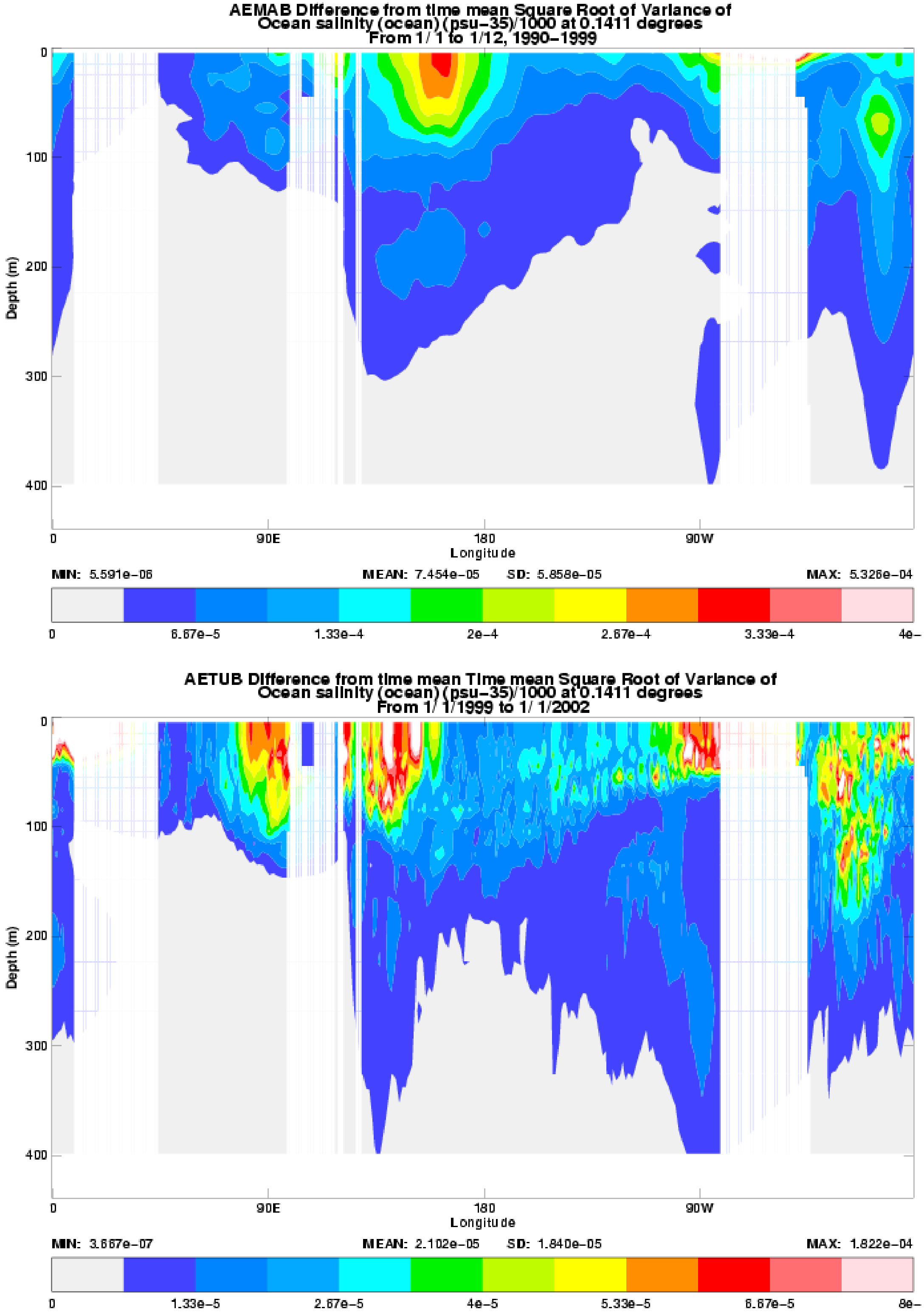

Figure E: Upper panel shows the standard deviation of monthly mean anomalies from an objective analysis for 1990-1999 for salinity on an equatorial-depth section. The lower panel shows the intra-ensemble standard deviation between a perturbed ensemble member and the control for 1999-2001. Note different scales.