Interannual to Decadal Climate Predictability: A Multi-Perfect-Model-Ensemble Study

M. Collins (1), M. Botzet (2), A. Carril (3), H. Drange (4,5), A. Jouzeau (6), M. Latif (2,7), O.H. Otteraa (4), H. Pohlmann (2), A. Sorteberg (5), R. Sutton (8), L. Terray (6)

(1) Hadley Centre, Met Office, Exeter, UK.
(2) Max-Planck-Institut fuer Meteorologie, Hamburg, Germany.
(3) Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy.
(4) Nansen Environmental and Remote Sensing Center, Bergen, Norway.
(5) Bjerknes Centre for Climate Research, Bergen, Norway.
(6) CERFACS, Toulouse, France.
(7) Leibniz-Institute fur Meereswissenschaften, Kiel, Germany
(8) Centre for Global Atmospheric Modelling, Reading, UK

Abstract

Perfect model ensemble experiments are performed with five coupled atmosphere-ocean models to investigate the potential for initial-value climate forecasts on interannual to decadal time scales. Experiments are started from similar initial states and common diagnostics of predictability are used. We find that; variations in the ocean Meridional Overturning Circulation are potentially predictable on interannual to decadal time scales, a more consistent picture of the surface temperature impact of decadal variations in the MOC is now apparent, and variations of surface air temperatures in the N. Atlantic are also potentially predictable on interannual to decadal time scales, albeit with potential skill levels which are less than those seen for MOC variations. This inter-comparison represents a step forward in assessing the robustness of model estimates of potential skill and is a pre-requisite for the development of any operational forecasting system.

Introduction

Predictions of the future state of the climate system are of potential benefit to society. The ability to predict (here we consider the *potential* ability to predict) can also give insight into the physical aspects of the climate system which are not simply the averaged or integrated effects of chaotic, unpredictable weather “noise”. Restricting attention to variations in climate which are purely internally generated, predictability in the system hints at processes which have long time scales or which possibly have periodic behaviour. Quantifying the predictability associated with such processes can lead to a greater understanding of the climate system.

Operational predictions of climate on seasonal to interannual time scales associated with the El Nino Southern Oscillation (ENSO) are now commonplace (e.g. Goddard et al., 2001). Prediction systems for other seasonal-interannual “modes” of climate are also emerging (e.g. Rodwell and Folland, 2004). Here we consider the predictability of interannual to decadal variations in climate. On these time scales, both the initial conditions (principally the initial state of the ocean) and the boundary conditions (associated with both natural and anthropogenic forcing of the system) are important (Collins and Allen 2002; Collins 2002) but here we focus solely on the initial value problem of the predictability of internally generated interannual to decadal climate variability.

The Atlantic Meridional Overturning Circulation (MOC) is the main northward heat carrying component of the ocean part of the climate system (e.g. Trenberth and Caron 2001). Coupled atmosphere-ocean models (AOGCMs) exhibit internally generated variations the strength of the MOC and associated heat transport (e.g. Dong and Sutton 2001) and the surface climate impact of those variations have also been seen in historical (Latif et al. 2004) and palaeo-climate records (Delworth and Mann 2000). Shorter records of ocean observations (Dickson et al., 1996; Curry et al., 2003; Marsh, 2000), also exhibit variations which have been linked with the MOC. Variations in the MOC thus represent an ideal candidate for the study of interannual to decadal climate predictability.

Predictability studies with AOGCMs in which ensembles of simulations with small perturbations to the initial conditions are performed have revealed the potential predictability in these MOC variations and in related surface and atmosphere variables (Griffies and Bryan 1997; Grotzner et al., 1999; Boer, 2000; Collins and Sinha 2003; Pohlmann et al., 2004). While all studies show some level of potential predictability, it is difficult to form robust conclusions because of the range of complexity (and hence realism) of the different models used, because of the range of different initial states considered and because of subtle differences in the measures of predictability employed. For example, it is well known in weather forecasting that predictive skill can vary considerably with different initial conditions. Clearly it is important to quantify the potential skill-level of interannual-decadal climate forecasts prior to the expensive development of operational prediction schemes and the deployment of operational observing systems.

Here we present a step-forward in making a robust estimate of the potential predictive skill of interannual to decadal climate predictions associated with internally generated variations in the MOC. A coordinated set of potential predictability experiments have been performed with five recently developed complex AOGCMs. An attempt is made to initiate the experiments from similar ocean states and a common set of measures of potential skill are used. This "multi-model" approach has proved useful in other areas of weather and climate prediction. Here the emphasis is on a comparison of the levels of potential predictability seen in the different models. Other publications discuss the individual model results (e.g. Collins and Sinha, 2003; Pohlmann et al., 2004; Pohlmann et al. 2004) in more detail.

The Ensemble Experiments

Five coupled atmosphere-ocean models are used (see table 1):

The ARPEGE3-ORCALIM has an atmosphere component (Déqué et al., 1994) with a horizontal resolution of T63 with 31 levels in the vertical (20 in the troposphere). The ocean component, ORCA2, is the global configuration of the OPA8 Ocean model (Madec et al., 1998) with a horizontal resolution of 2° in longitude and 0.5° to 2° in latitude. It includes a dynamic-thermodynamic sea-ice model (Fichefet and Morales Maqueda 1997). The components are coupled through OASIS 2.5 (Valcke et al., 2000), which ensures the time synchronization and performs spatial interpolation from one grid to another.

The Bergen Climate Model (BCM) (Furevik et al., 2003, Bentsen et al., 2004) uses the Miami Isopycnic Coordinate Ocean Model (Bleck et al.1992) coupled to a dynamic-thermodynamic sea ice module. The ocean mesh is formulated on a Mercator projection with a nominal resolution of 2.4 degrees, and 24 vertical layers. The atmospheric component is version three of the ARPEGE model with a horizontal resolution of T63 and 31 layers in the vertical – essentially the same atmosphere that is used in ARPEGE3-ORCALIM. Fresh water and heat flux adjustments are applied.

ECHAM5/MPI-OM (Latif et al. 2004) uses version 5 of the European Centre-Hamburg atmosphere model (ECHAM5, Roeckner et al. 2004) at T42 resolution with 19 vertical layers. The oceanic component, the Max Planck Institute Ocean Model (MPI-OM, Marsland et al. 2003) is run on a curvilinear grid with equatorial refinement and 23 vertical levels. A dynamic-thermodynamic sea ice model and a river runoff scheme are included.

Version 3 of the Hadley Centre Climate Model (HadCM3 – Gordon et al., 2000; Collins et al., 2001) uses an ocean component with a horizontal resolution of 1.25º longitude by 1.25º latitude and 20 levels in the vertical. The atmospheric component uses a grid-point formulation with a horizontal resolution of 3.75ºx2.5º in longitude and latitude with 19 unequally spaced vertical levels (Pope et al., 2000). A simple thermodynamic sea-ice scheme is used.

The INGV model uses the ECHAM4 model (Roeckner, 1996) at T42 resolution with 19 vertical levels. The ocean component is essentially the same as that used in the ARPEGE3-ORCALIM model. More details can be found in Gualdi (2003).

Ensemble experiments are performed from initial states of anomalously high and anomalously low MOC taken from a control (i.e. unforced) run of each model (figure 1). In addition, some models were used to perform experiments with initial states near the time-mean value of overturning. Perturbations to the initial conditions were made using the common method of taking different atmospheric start conditions with the same ocean start condition (the "perfect model" approach e.g. Collins and Sinha 2003). While this perturbation methodology is in no way optimal in terms of, for example, sampling the likely range of atmosphere-ocean analysis error, it is sufficient to generate ensemble spread on the time scales of interest.

The availability of computer resources limited the number of ensemble members and experiments that could be performed: nevertheless all experiments were integrated out to at least 20 years. The experiments correspond to a total 1340 simulated years for the predictability experiments combined with a total of 3100 simulated years for the control experiments used to assess background variability. Annual mean diagnostics are examined because of the focus on interannual to decadal time scales.

Potential Predictability of MOC variations

The first point to note is the wide range of time scales and magnitudes of MOC variability in the different models (figure 2). The ECHAM5/MPI-OM model shows the largest variations in MOC strength with clear interdecadal variability present. HadCM3 and BCM also show interdecadal variations but at a reduced level in comparison. The ARPEGE3-ORCALIM model has the lowest level of variability but decadal-interdecadal time scales are still clearly present in the time series. The large trend seen in the INGV model is almost certainly due to a drift seen in this particular control experiment - the model has yet to reach equilibrium and we do not attempt to extract quantitative measures of predictability. Although not calculated, diagnostic measures of predictability/variability (e.g. Boer, 2000) would clearly show a range of different levels of MOC potential predictability in these models. However, the only reliable way to assess predictability is to perform ensemble experiments.

The perfect model ensemble experiments are also shown in figure 2. Potential predictability is evident when the ensemble spread is small in comparison with the total level of variability in the control time series, or even if the ensemble spread is relatively large but the centre of gravity of the ensemble is displaced significantly with respect mean of the control (e.g. Collins, 2001). We may imagine a background or climatological distribution which, in the absence of a forecast, would be all the information we would have to form an assessment of the future strength of the MOC. A forecasts may allow us to reduce the potential range (low ensemble spread) or shift the mean of the distribution (displaced ensemble), or both. Both types of (potential) predictability are seen on interannual to decadal time scales in the experiments shown in figure 2. For example, the first HadCM3 ensemble (anomalously strong MOC initial conditions) has relatively small ensemble spread in the first decade of the experiment and the ensemble is significantly shifted to stronger values with respect to the mean with no ensemble members indicating weaker than average overturning (see Collins and Sinha (2001) for more details). Other examples are clear.

There are a wide range of measures which may be used for forecast verification (here we measure the potential skill of a perfect model forecast – an upper limit). We examine two of the most-simple measures of forecast skill to quantify levels of potential predictability; the anomaly correlation (ACC) and normalised root mean squared error (RMSE). Formulas are given in Collins (2001) for the perfect model case.

Figure 3 shows both measures for the MOC in the ensemble experiments discussed above. For the strong MOC initial states, the ACC is "high" for approximately the first decade in all the model experiments, with "high" being above 0.6 – a commonly used cut-off value in weather forecasting. The RMSE is correspondingly low. After the first decade, the ARPEGE3 model predictability drops off rapidly whereas for the other models the ACC drops off slowly to low values by the end of the 20 year experiments. The RMSE similarly saturates in 20 years. For the weak MOC initial states, error growth and loss of predictability seems to happen sooner in the ensemble experiments, although there is some noise in these measures because of small ensemble sizes. ACC and RMSE are not shown for the normal initial states because of the small sample size.

While the number of ensemble experiments is small, we may attempt to draw some conclusions about the multi-model estimate of potential predictability of MOC variability in these experiments (figure 3 – thick solid line). The multi-model ensemble indicates potential predictability of interannual-decadal MOC variations for 1-2 decades into the future. It also indicates that initial states which have anomalously strong overturning are more predictable than those with anomalously weak overturning. This latter result is intriguing, but is subject to some uncertainty because of the relative small number of models and ensemble experiments included in the multi-model analysis. Nevertheless, some consensus is emerging in contrast to the previous situation in which a large range of predictability is seen in the literature. It would be safe to conclude that there is a robust signal of potential predictability of variations in the MOC on interannual to decadal time scales.

Potential Predictability of Surface Climate Variations

Predictions of MOC variability may be of interest to scientists, but they would be of little relevance to society unless they are accompanied by predictions of surface climate variables. A simple measure of the impact of MOC variations can be obtained be performing a regression between decadal-averaged MOC strength and decadal-averaged surface air temperature (SAT) in the different models (figure 4). The general impression in all the models is of a warmer Northern Hemisphere when the MOC is stronger and is transporting more heat polewards. Differing levels of statistical significance seen in figure 4 may be interpreted as resulting from different levels of signal to noise in the sense that in models with larger variations in MOC, the surface signal has a better chance of overwhelming the noise of unrelated random climate variations. What is interesting is that the magnitude of the surface response (in K/Sv) is similar across all models.

The North Atlantic ocean is a region in all the models in which there is a significant relationship between decadal variations in SAT (and underlying SST) and the MOC. Time series of annual mean SAT from the control and ensemble experiments averaged over a region of the North Atlantic (used in Collins and Sinha (2001) and Pohlmann et al. (2004)) are shown in figure 5. Strong similarities between these time series and those shown in figure 2 for the MOC are evident, although there is clearly more noise in this variable as a result of unrelated random variability.

ACC and RMSE measures of ensemble spread (figure 6) for N. Atlantic SAT are similar to those computed for MOC variations (figure 3) but the levels of potential predictability are clearly less and the differences between ensemble members greater. It may be possible to find greater levels of potential predictability for each individual model by adjusting the boundaries of the region chosen but here we compare the models on an equal footing. Also, the effects of interannual noise which are more prominent in this variable may be reduced by taking averages over a greater number of years. Nevertheless, the picture of potentially predictable surface climate variations associated with variations in the MOC appears consistent.

Discussion

Whereas previously it has been difficult to assess the potential for making interannual to decadal forecasts of climate due to different studies indicating different levels of predictability, a more complete picture of the predictability is emerging. This intercomparison study shows that;

1. variations in the ocean Meridional Overturning Circulation are potentially predictable on interannual to decadal time scales,
2. a more consistent picture of the surface temperature impact of decadal variations in the MOC is now apparent, and
3. variations of surface air temperatures in the N. Atlantic are also potentially predictable on interannual to decadal time scales, albeit with potential skill levels which are less than those seen for MOC variations.

Perhaps the biggest difference between the models is in the wide range of strengths of decadal variability evident in figure 2. In general, models with greater decadal MOC variability have greater levels of potential predictability – despite the fact that the ACC and RMSE are signal-to-noise measures and thus allow for a differences in background natural variability. Investigation into the mechanisms responsible for the different levels of variability would seem a priority.

The far more pertinent question is, of course, that of the (potential) prediction of surface climate variations over land. The simple measures used in this study do not reveal robustly predictable land signals. Collins and Sinha (2001) and Pohlmann et al. (2004) investigate probabilistic techniques more commonly used in medium-range and seasonal forecasting in the context of the interannual-decadal problem with some limited success. However, the application and verification of such measures (here the assessment of potential skill) requires much larger ensemble sizes and many more ensemble simulations than used here. Hopefully such ensembles will be performed in future. In addition, the number of modelling, initialisation and observational issues that need to be addressed before we routinely produce interannual-decadal climate forecasts are numerous.

References

Bentsen, M., H. Drange, T. Furevik and T. Zhou, 2004: Simulated variability of the Atlantic meridional overturning circulation Clim. Dyn., doi: 10.1007/s00382-004-0397-x

Bleck, R., C. Rooth, D. Hu, L. T. Smith, 1992: Salinity-driven thermohaline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic, J. Phys. Oceanogr., 22, 1486–1515.

Boer, G.J., 2000: A study of atmosphere-ocean predictability on long timescales. Clim. Dyn., 16, 469-477.

Collins, M., 2002: Climate Predictability on Interannual to Decadal Time Scales: The Initial Value Problem. Clim. Dyn. 19, 671-692.

Collins, M, and Allen, M. R., 2002: On assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability J. Climate, 21, 3104-3109.

Collins, M. and Sinha, B., 2003: Predictable Decadal Variations in the Thermohaline Circulation and Climate. Geophys. Res. Letts., vol. 30, no. 6, 10.1029/2002GLO16504.

Collins, M., Tett, S. F. B. and Cooper, C., 2002: The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. Clim. Dyn., 17, 61-81.

Curry, R., Dickson, R., Yashayaev, I., 2003: A change in the freshwater balance of the Atlantic Ocean over the past four decades. Nature 426, 826–829, doi:10.1038/nature02206.

Delworth, T. L. and Mann, M. E., 2002: Observed and simulated multidecadal variability in the Northern Hemisphere. Clim. Dyn. 16, 661-676.

Déqué, M., C. Dreveton, A. Braun and D. Cariolle, 1994: The climate version of ARPEGE/IFS: a contribution to the French community climate modelling. Climate Dyn., 10, 249-266.

Dickson, R., Lazier J., Meincke J., Rhines P., Swift J., 1996: Long term coordinated changes in the convective activity of the North Atlantic. Progress in Oceanography, 38, 241-295, 1996.

Dong, B. and Sutton, R. T., 2001: The dominant mechanisms of variability in Atlantic ocean heat transport in a coupled ocean-atmosphere GCM. Geophys. Res. Letts. 28, 2445-2448.

Fichefet T. and M. A. Morales Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. J. Geophys. Res. 102(C6),12609-12646.

Furevik, T., M. Bentsen, H. Drange, I. K. T. Kindem, N. G. Kvamstø and A. Sorteberg, 2003: Description and validation of the Bergen Climate Model: ARPEGE coupled with MICOM, Clim. Dyn., 21, 27-51, doi:10.1007/s00382-003-0317-5.

Goddard, L., S. Mason, S. Zebiak, C. Ropelewski, R. Basher, and M. Cane, 2001: Current approaches to seasonal to interannual climate predictions. Int. J. Climatol., 21, 1111–1152.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transport in a version of the Hadley Centre coupled model without flux adjustments. Clim. Dyn. 16, 147-168.

Grötzner, A., Latif, M., Timmermann, A. and Voss, R. Interannual to decadal predictability in a coupled ocean atmosphere general circulation model., 1999: J. Climate 12, 2607-2624.

Griffies, S.M. and Bryan, K., 1997: Predictability of North Atlantic multidecadal climate variability. Science 275, 181-184.

Gualdi S., E. Guilyardi, A. Navarra, S. Masina and P. Delecluse, 2003. The Interannual Variability in the Tropical Indian Ocean as Simulated by a CGCM. Clim. Dyn., 20, 567-582.

Latif, M., E. Roeckner, M. Botzet, M. Esch, H. Haak, S. Hagemann, J. Jungclaus, S. Legutke, S. Marsland, U. Mikolajewicz, and J. Mitchell, 2004: Reconstructing, Monitoring, and Predicting Decadal- Scale Changes in the North Atlantic Thermohaline Circulation with Sea Surface Temperature. J. Climate, 17, 1605-1614.

Madec, G., P. Delecluse, M. Imbard and C. Lévy, 1998: OPA 8.1 Ocean General Circulation Model reference manual. Note du Pôle de modélisation, Institut Pierre-Simon Laplace, N°11, pp 91.

Marsh, R., 2000: Recent variability of the N. Atlantic Thermohaline Circulation inferred from surface heat and freshwater fluxes, J. Climate, 13, 3239-3260.

Marsland, S., H. Haak, J. Jungclaus, M. Latif and F. Röske, 2003: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. Ocean Modelling, 5, 91-127.

Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M. And Tschuck, P., 2004: Estimating the decadal predictability of a coupled AOGCM. J. Climate, in press.

Pohlmann, H., Sienz, F., and Latif, M., 2004: Influence of the multidecadal Atlantic meridional circulation variability on European climate. J. Climate, submitted.

Pope, V. D., Gallani, M. L., Rowntree, P. R. and Stratton, R. A., 2000: The impact pf new physical parametrizations in the Hadley Centre climate model – HadAM3. Clim. Dyn. 16, 123-146.

Roeckner, E., G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, A. Rhodin, U. Schlese, U. Schulzweida, A. Tompkins, 2003: The atmospheric general circulation model ECHAM 5. PART I: Model description. MPI Rep. 349, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 127 pp.

Rodwell, M.-J., and C.-K. Folland, 2002: Atlantic air-sea interaction and seasonal predictability. Quart. J. R. Met. Soc., 128, 1413-1444.

Roullet, G. and G. Madec, 2000: Salt conservation, free surface and varying volume: a new formulation for Ocean GCMs. J. Geophys. Res., 105, 23927-23942.

Trenberth, K. E. and Caron, J. M., 2001: Estimates of meridional atmosphere and ocean heat transports. J. Climate, 14, 3433-3443.

Valcke, S., L. Terray, and A. Piacentini, 2000: The OASIS coupled user guide version 2.4. Tech. Rep. TR/CMGC/00-10, CERFACS, 85 pp.

Figures and Tables

Model	Number of ensemble experiments	Number of ensemble members in each experiment	Length of ensemble experiments (years)	Length of control run (years)
ARPEGE3-ORCALIM	2	6(+1)	25	200
BCM	2	3(+1)	20	300
ECHAM5/MPI-OM	3	6(+1)	20	500
HadCM3	3	8(+1)	20	2000
INGV	2	2(+1)	20	100

Table 1: A summary of the AOGCMs used in the perfect model potential predictability experiments. The numbers in column 3 of the form 6(+1) indicate that 6 ensemble members were performed from a state taken from the control run but that the section of the control run may also be viewed as an additional ensemble member.

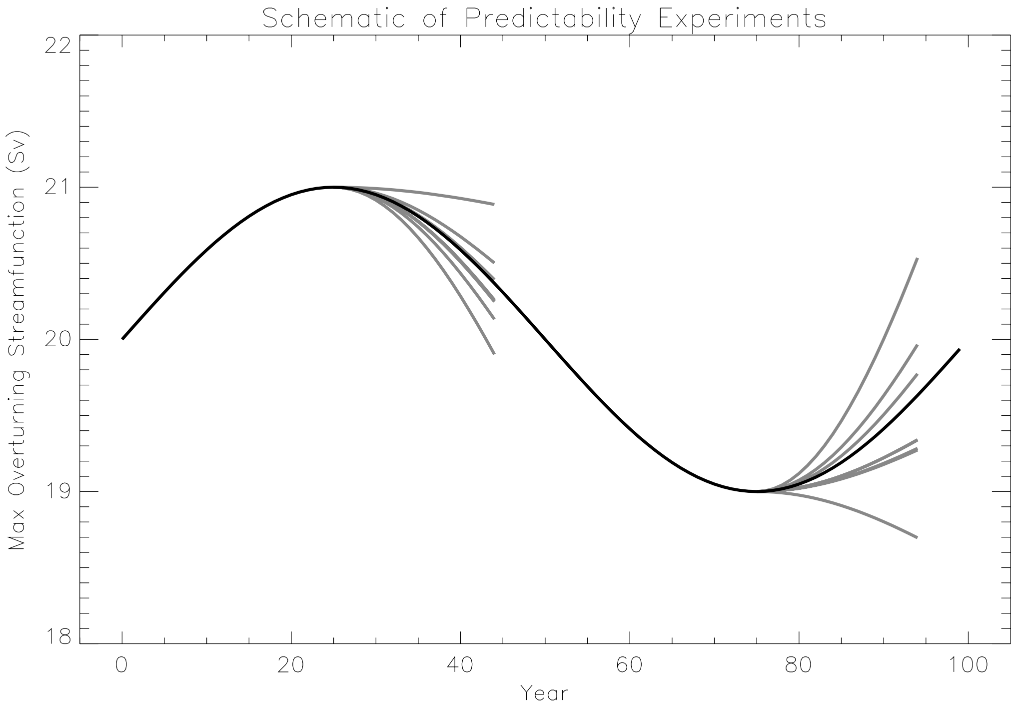

Figure 1: A schematic figure of the experimental design used in this study. The thick black line represents decadal-time scale internally generated variations in the strength of the Meridional Overturning Circulation (MOC) from a control run of a coupled atmosphere-ocean model. The grey lines represent "perfect model" ensemble experiments in which small perturbations to the initial conditions are made. For each of the models used in the study, we endeavoured to initiate the ensemble experiments from a state of relatively strong and relatively weak overturning. In addition, some models we use to initiate experiments from a state of relatively normal overturning.

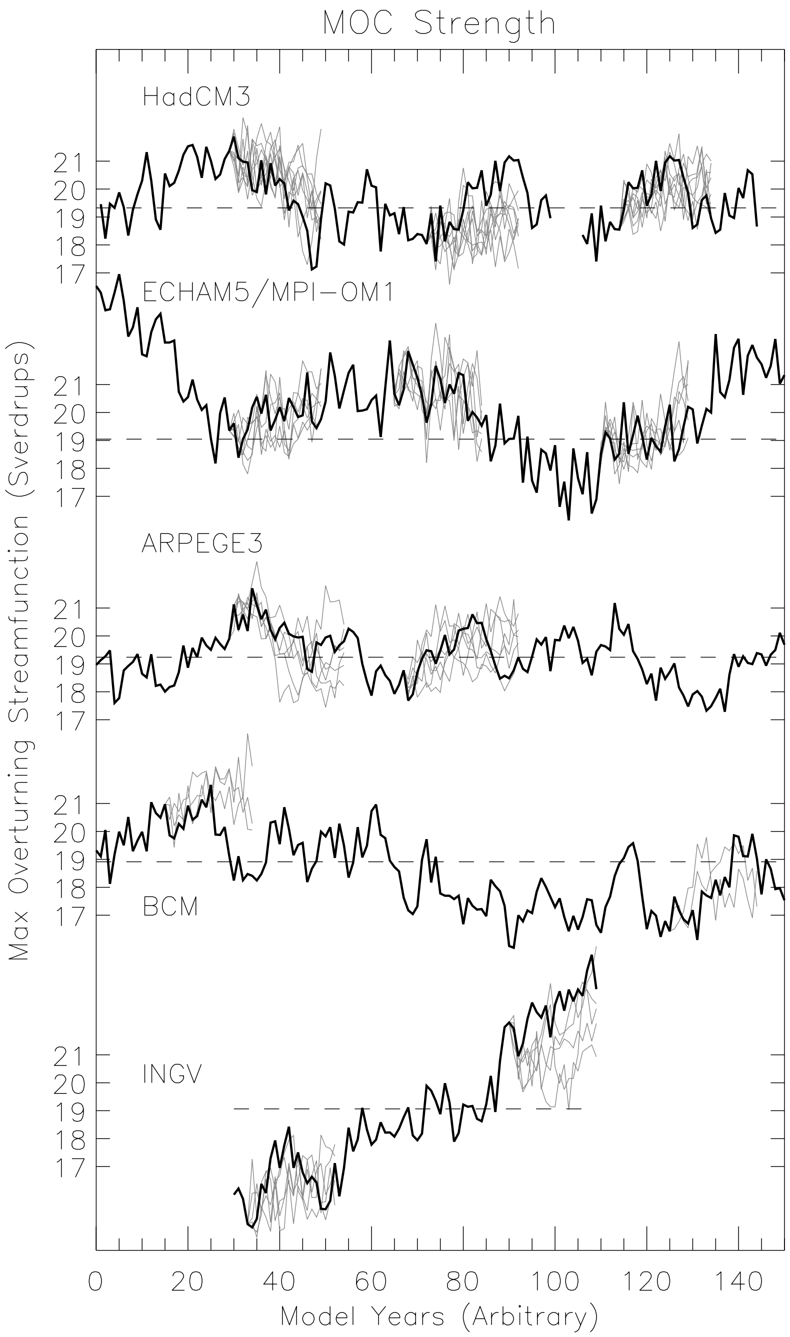

Figure 2: Time series of the strength of the MOC taken from the unforced control runs of five coupled atmosphere-ocean models (black lines - names indicate on the figure) and from the perfect model ensemble experiments (grey lines). MOC variations arise purely because of the internal dynamics of the coupled system and model years are arbitrary. The drift seen in the INGV model is a spin-up effect and the experiments are excluded from any quantitative analysis.

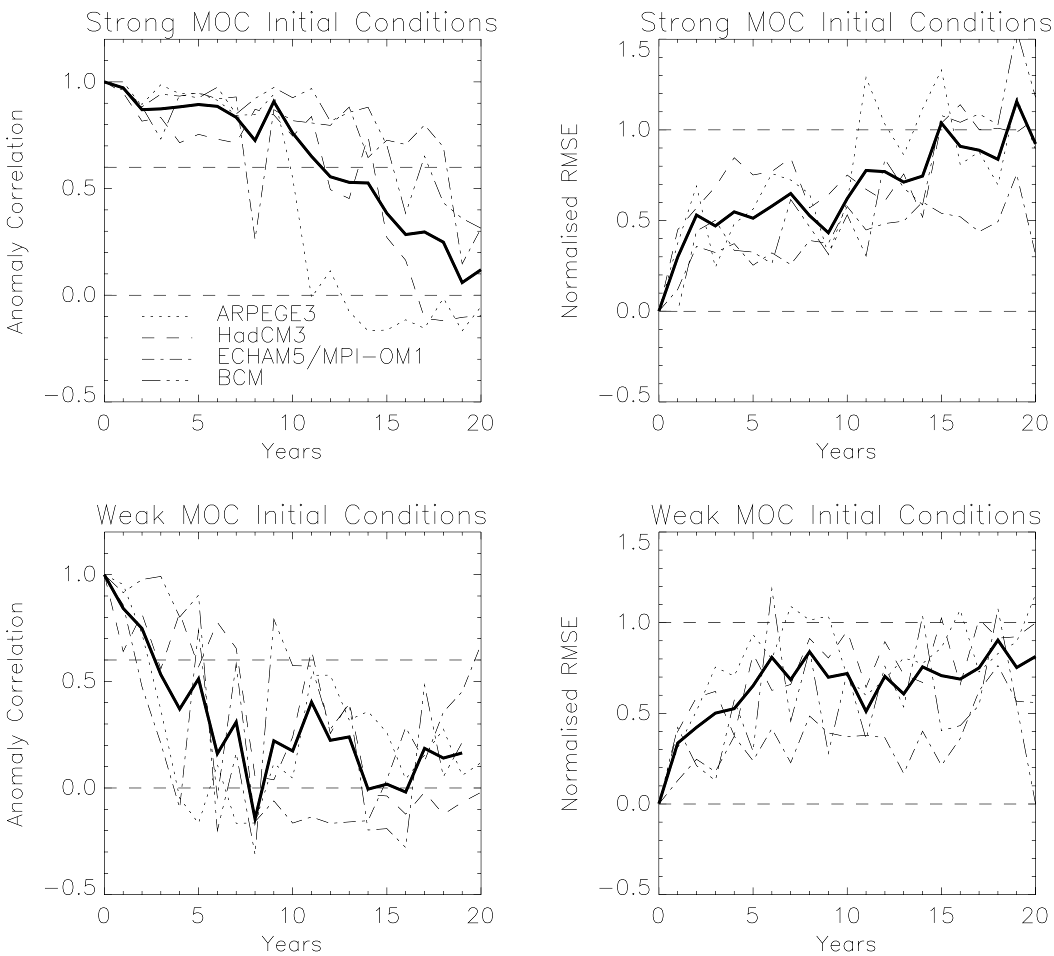

Figure 3: Measures of the potential predictability of variations in the strength of the MOC from four of the five coupled models (see legend). The left panels show the anomaly correlation coefficient (ACC - unity for perfect potential predictability, zero for no potential predictability) for strong MOC initial conditions (top panel), weak MOC initial conditions (middle panel) and normal MOC initial conditions (bottom panel). The right panels show the normalised root mean squared error (RMSE - zero for perfect potential predictability, unity for no potential predictability) in the same order. Also shown in the figures are the multi-model average ACC and RMSE (thick black line).

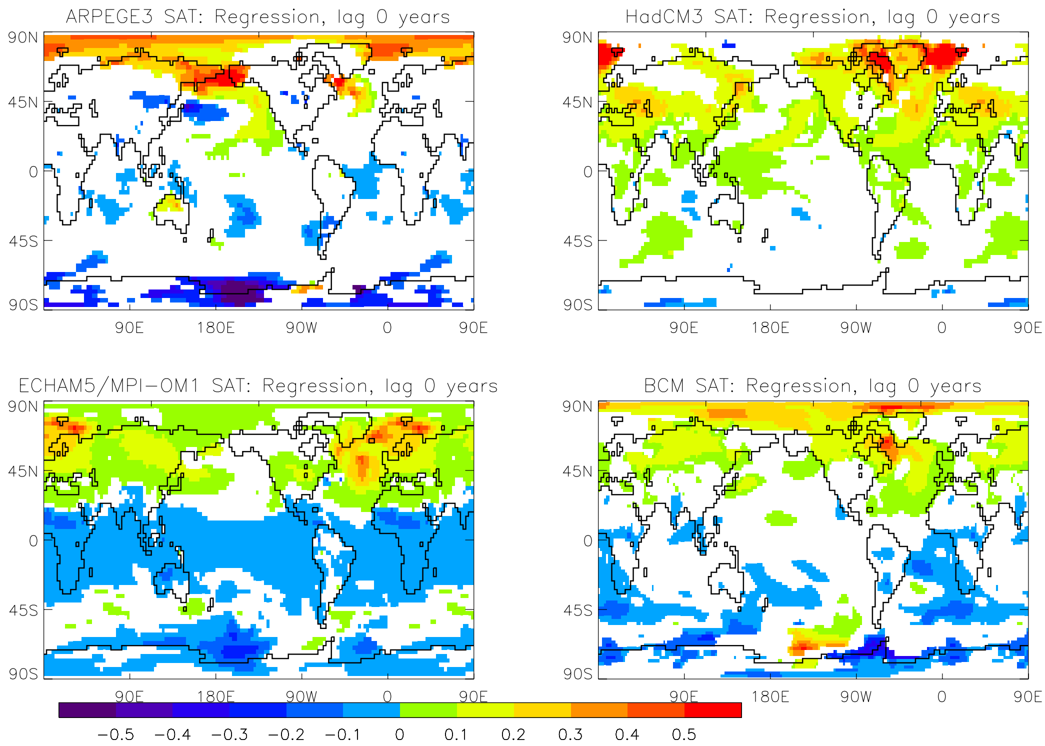

Figure 4: The coefficient of regression (degrees K per Sverdrup) of decadal mean surface air temperature against decadal mean MOC strength from four of the five coupled atmosphere-ocean models. Regions are shaded only where the coefficient is significantly different from zero at the 5% confidence level (based on an F-test).

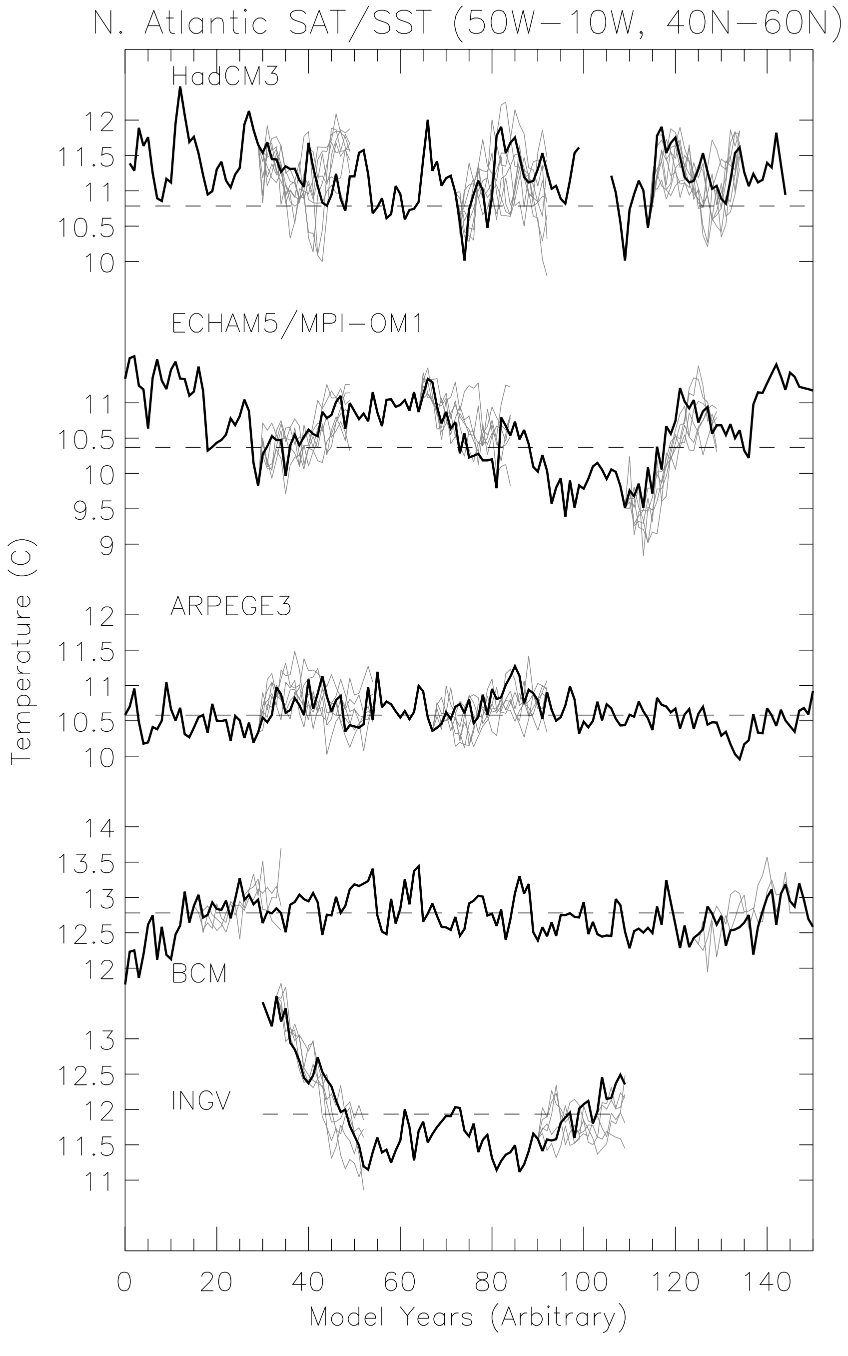

Figure 5: As in figure 2 but for surface air temperature averaged in the region 50ºW-10ºW, 40ºN-60ºN.

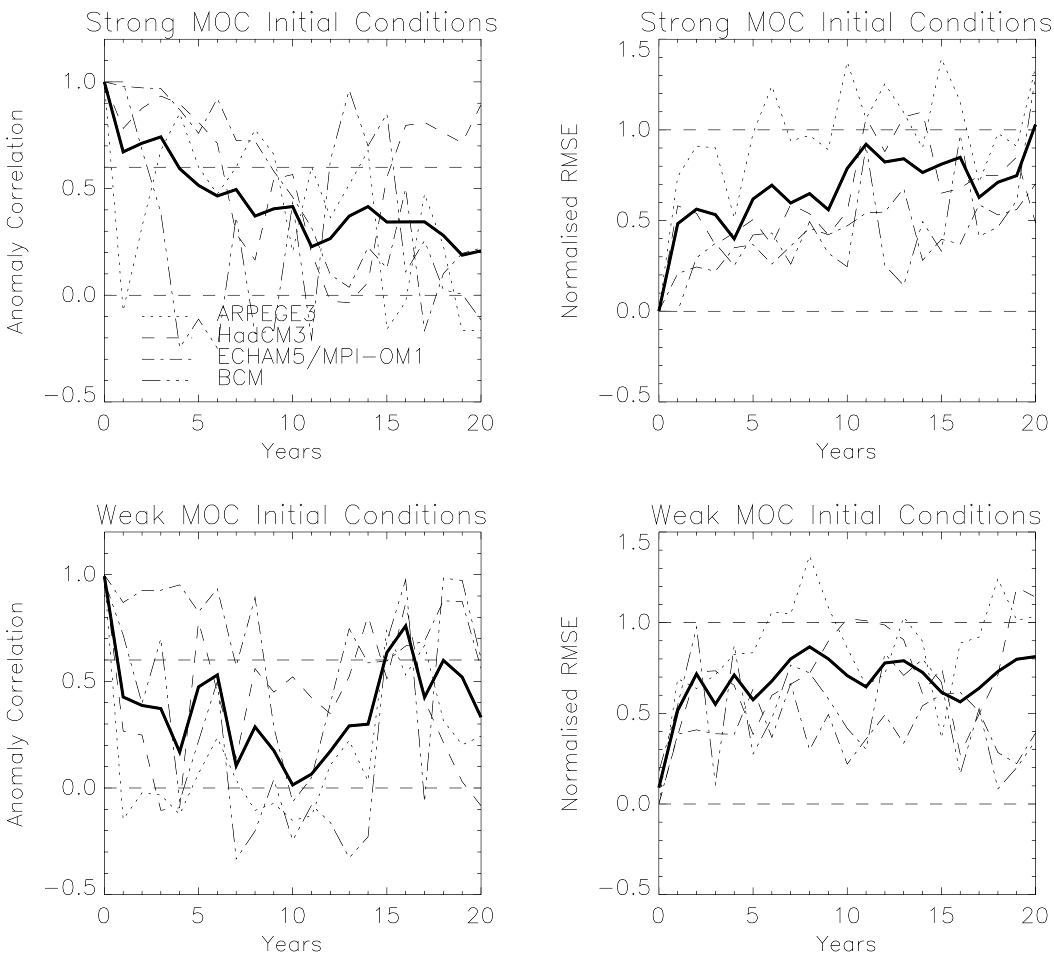

Figure 6: As in figure 3 but for surface air temperature averaged in the region 50ºW-10ºW, 40ºN-60ºN.