

## Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations

A. Schmittner,<sup>1</sup> M. Latif,<sup>2</sup> and B. Schneider<sup>2</sup>

Received 11 August 2005; revised 21 October 2005; accepted 27 October 2005; published 14 December 2005.

[1] Most climate models predict a weakening of the North Atlantic thermohaline circulation for the 21st century when forced by increasing levels of greenhouse gas concentrations. The model spread, however, is rather large, even when the forcing scenario is identical, indicating a large uncertainty in the response to forcing. In order to reduce the model uncertainties a weighting procedure is applied considering the skill of each model in simulating hydrographic properties and observation-based circulation estimates. This procedure yields a “best estimate” for the evolution of the North Atlantic THC during the 21st century by taking into account a measure of model quality. Using 28 projections from 9 different coupled global climate models of a scenario of future CO<sub>2</sub> increase (SRESA1B) performed for the upcoming fourth assessment report of the Intergovernmental Panel on Climate Change, the analysis predicts a gradual weakening of the North Atlantic THC by 25(±25)% until 2100. **Citation:** Schmittner, A., M. Latif, and B. Schneider (2005), Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations, *Geophys. Res. Lett.*, 32, L23710, doi:10.1029/2005GL024368.

### 1. Introduction

[2] The thermohaline circulation (THC) is a global 3-dimensional belt of ocean currents that transports large amounts of heat and freshwater around the world [Manabe and Stouffer, 1999]. In the North Atlantic, it is manifested in a meridional overturning circulation (AMOC) which, through its northward transport of warm tropical waters by the Gulf Stream and North Atlantic Current, effectively contributes to the warming of Northern Europe [Trenberth and Caron, 2001; Rahmstorf, 2003]. Previous model projections [e.g., Manabe and Stouffer, 1993; Stocker and Schmittner, 1997; Intergovernmental Panel on Climate Change (IPCC), 2001] suggested that global warming may lead to a strong weakening or even to a complete disappearance of the AMOC, which would have serious impacts on the climate, the ecology and the economy of many countries surrounding the North Atlantic.

[3] The value of ensemble prediction is well established in numerical weather forecasting and seasonal climate prediction. An important outcome from the field of seasonal climate prediction is the demonstration of the superiority of

the multi-model ensemble over any single model. This feature is quite universal and not restricted to any particular region or variable [Palmer *et al.*, 2004]. Thus a multi-model ensemble is an effective method for sampling model uncertainties and for making more reliable forecasts. When applied to global change prediction, multi-model ensembles, however, yield a large uncertainty for the climate of the 21st century both globally and regionally [IPCC, 2001]. In particular, the future evolution of the AMOC is characterized by a large model spread: While some models simulate a rather strong weakening of the AMOC, other models are relatively stable and simulate either only a moderate or no change [IPCC, 2001; Gregory *et al.*, 2005]. Here we investigate the behavior of the AMOC in the most recent greenhouse simulations conducted for the upcoming Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC). Our aim is to improve the model projections of the AMOC by reducing the uncertainty. We do this by taking into account the models' skills in simulating observation based circulation estimates and observed climatological hydrographic conditions in the assessment of the multi-model ensemble.

[4] Such an assessment methodology based on model skill to obtain more reliable forecasts has a long history in weather and seasonal forecasting [e.g., Fraedrich and Leslie, 1987; Fraedrich and Smith, 1989; Metzger *et al.*, 2004]. A similar approach was followed by Murphy *et al.* [2004], who analyzed surface air temperature in an ensemble of greenhouse simulations. They constrained the different model versions by a multi-variate climate prediction index derived from observations. A major result of this study is that the weighted probability density function of climate sensitivity based on model performance is narrower than the unweighted one, thus decreasing the uncertainty. Knutti *et al.* [2003] investigated an ensemble of reduced complexity models using observed surface warming and ocean heat uptake as constraints and found a very broad range of AMOC responses.

### 2. Model Data and Observations

[5] We obtained results from 9 global climate models which were integrated as part of the AR4 of IPCC. All models were integrated using observed concentrations of greenhouse gases/aerosols from 1850 to present (scenario 20C3M). Future concentrations are prescribed according to IPCC-scenario SRESA1B until 2100. Carbon dioxide concentrations rise up to about 700 ppm until 2100 in this scenario leading to a radiative forcing of about 6 W/m<sup>2</sup>. Globally averaged surface air temperature increases by about 3°C until 2100 in a simplified model with an intermediate climate sensitivity [IPCC, 2001]. If multiple

<sup>1</sup>College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.

<sup>2</sup>Ocean Circulation and Climate Dynamics, Leibniz-Institut für Meereswissenschaften, Kiel, Germany.

**Table 1.** RMS Errors for the Individual Models and the Resulting Weight  $W^a$ 

	$T_{\text{global}}$	$S_{\text{global}}$	$D_{\text{global}}$	$SST_{\text{NAatl}}$	$SSS_{\text{NAatl}}$	$D_{\text{NAatl}}$	$AMOC_{\text{MAX}} (17.7)$	$AMOC_{48\text{N}} (16.0)$	$AMOC_{24\text{N}} (15.8)$	Weight
$g_i$	1	1	1	2	2	2	2	2	3	
1 (CCCMA) 5,5 <sup>b</sup>	0.24	0.68	0.66	<b>0.31</b>	1.00	0.98	0.59 (7.2)	0.59 (6.5)	0.59 (6.4)	0.20
2 (GFDL-2.0) 2,1 <sup>b</sup>	<b>0.20</b>	<b>0.43</b>	<b>0.57</b>	0.34	<b>0.53</b>	<b>0.75</b>	0.31 (23.2)	0.34 (21.5)	0.16 (18.3)	<b>2.75</b>
3 (GISS-AOM) 2,2 <sup>b</sup>	<i>0.66</i>	0.75	2.29	0.43	0.79	3.48	<i>0.79</i> (31.7)	0.53 (24.5)	0.22 (19.2)	<i>0.03</i>
4 (GISS-EH) 5,3 <sup>b</sup>	0.31	<i>0.76</i>	1.57	0.61	1.12	1.85	0.56 (27.7)	<i>0.65</i> (26.5)	0.34 (21.1)	0.46
5 (GISS-ER) 4,5 <sup>b</sup>	<i>0.69</i>	<i>0.82</i>	<i>2.06</i>	<i>0.65</i>	1.11	2.40	0.75 (30.9)	0.47 (23.5)	0.13 (17.9)	0.16
6 (IAP) 3,3 <sup>b</sup>	0.24	0.57	0.98	<i>1.74</i>	2.69	2.10	<i>0.79</i> (3.8)	<i>0.98</i> (0.3)	<i>0.81</i> (2.9)	<i>0.04</i>
7 (MIROC-HI) 1,1 <sup>b</sup>	<b>0.22</b>	<b>0.43</b>	<b>0.63</b>	0.42	<i>1.46</i>	0.80	0.22 (13.7)	0.24 (12.2)	0.35 (10.3)	1.62
8 (MIROC-MED) 3,3 <sup>b</sup>	0.25	0.56	0.79	0.50	0.91	<b>0.79</b>	<b>0.12</b> (19.7)	<b>0.09</b> (14.4)	<b>0.02</b> (15.5)	<b>2.29</b>
9 (MRI) 5,5 <sup>b</sup>	0.26	<i>0.76</i>	0.86	<b>0.35</b>	<b>0.46</b>	0.86	<b>0.07</b> (16.5)	<b>0.09</b> (14.5)	<b>0.03</b> (16.2)	1.46

<sup>a</sup>RMS errors are normalized by the standard deviation of the observations. The weights  $g_i$  for the individual variables in the calculation of the total weight are given. Bold numbers give the two best; italic numbers give the two worst models. Numbers in parentheses in columns 8–10 denote the absolute value of the circulation in Sv.

<sup>b</sup>Number of ensemble runs for the 20C3M and SRESA1B scenarios, respectively.

(ensemble) runs were available for an individual model we used the ensemble mean in the assessment.

[6] Observation-based estimates of the AMOC at 24°N from *Ganachaud and Wunsch* [2000] and *Lumpkin and Speer* [2003], at 48°N from *Ganachaud* [2003], and its maximum value in the North Atlantic from *Smethie and Fine* [2001] and *Talley et al.* [2003], as well as temperature, salinity, and pycnocline depth observations from the World Ocean Atlas 2001 [*Conkright et al.*, 2002] are used to evaluate the climate models.

### 3. Weighting Methodology

[7] We evaluate the ocean component of the climate models in terms of the simulated global temperature ( $T$ ), salinity ( $S$ ) and pycnocline depth ( $D = \int(\rho_{\text{max}} - \rho)dz / \int(\rho_{\text{max}} - \rho)dz$ ) distributions during the period 1981–2000. The pycnocline depth is a dynamically important variable controlling the upper ocean flow [*Gnanadesikan et al.*, 2002]. Since our goal is an assessment of the simulation of the Atlantic overturning circulation we use additional measures that characterize the Atlantic circulation. Sea surface temperature (SST) and salinity (SSS) in the North Atlantic depend strongly on the AMOC for its effect on the northward advection of warm and salty subtropical surface waters which is most pronounced between 40–70°N [e.g., *Schmittner et al.*, 2002]. Therefore we use SST and SSS in this region as well as the pycnocline depth restricted to the Atlantic basin north of 35°S, since theory [*Marotzke and Klinger*, 2000] and model results [*Hughes and Weaver*, 1994] suggest that the density gradients within the Atlantic drive the overturning. Additionally, observation-based estimates of the mass flux at 24°N, 48°N and its maximum value are used as controls in the model assessment (see Table 1).

[8] The skill score  $S$  is a combination of the normalized (by the standard deviation of the observations) root mean square (rms) errors of the above described variables weighted by  $g_i$  (Table 1) in order to emphasize circulation estimates and the tracer distributions in the North Atlantic:

$$S^2 = \Sigma_i g_i \text{rms}_i^2 / \Sigma_i g_i.$$

We have tested different versions of the skill score, e.g. including different choices for the  $g_i$ , considering correlation coefficients and pattern rms errors. The main results

were similar and therefore we restrict our discussion to the above formulation.

[9] The weights  $W$  used in the assessment of the models are calculated based on a probabilistic approach assuming Gaussian statistics as from *Murphy et al.* [2004]:

$$W = \exp(-2S^2).$$

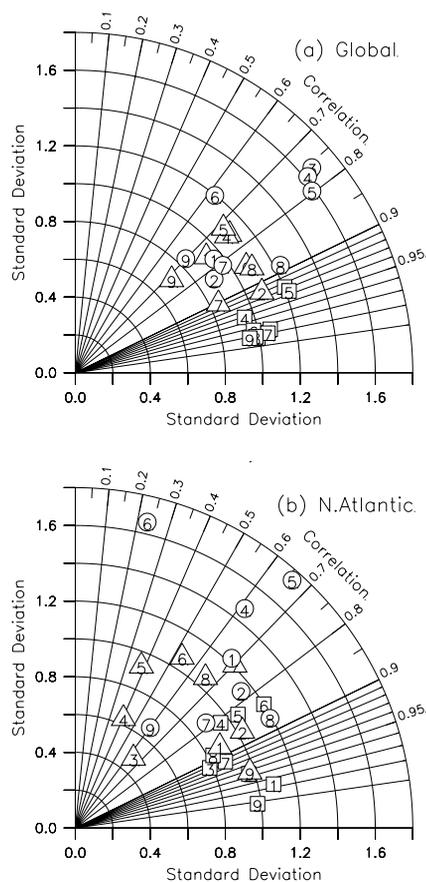
In order to account for the fact that flux corrections may influence the transient model response and artificially increase the correspondence with the observations, we penalized model 1 (global flux correction) by multiplying the rms errors of  $T, S$  and  $D$  by 2 and those of model 9 (tropical flux correction) by 1.3.

[10] Finally, the question arises whether the model responses should be scaled by the climate sensitivity. We did not find, however, any systematic relationship between climate sensitivity and AMOC response.

### 4. Model Assessment

[11] *Taylor* [2001] diagrams display the correspondence of each model with the observations for the global fields of temperature, salinity and pycnocline depth (Figure 1a) and for North Atlantic SST, SSS and pycnocline depth (Figure 1b). The model data were normalized by the observed standard deviation. A “perfect” model would reside in the point (1,1) in the  $\sigma, R$ -plane of the Taylor diagram.

[12] In general, temperature is simulated more successfully by the models than salinity or pycnocline depth. Furthermore, global statistics exhibit less spread than those for the North Atlantic. In particular, all models simulate the global ocean temperature distribution quite realistically, with a normalized standard deviation close to unity and correlations above 0.9. Models 3 and 5 display a systematically too deep thermocline (not shown) resulting in larger rms errors (Table 1). The global salinity distribution is simulated less successfully than that for temperature: The correlations are much smaller and the standard deviations are off by at least 10% in most models. This suggests that the atmospheric hydrological cycle and/or sea ice are still not very well simulated in most models. Models 3 and 5 display a systematically too salty upper ocean and model 9 has no gradients below a few hundred meters depth (not shown). The correlations for the global field of pycnocline depth are similar to those for salinity but the rms errors are



**Figure 1.** Taylor diagrams for temperature (squares), salinity (triangles) and pycnocline depth (circles) for the global ocean (a) and the North Atlantic (b). The performance of the individual models is shown as a function of the normalized spatial standard deviation and the pattern correlation.

larger (Table 1), likely because errors for temperature and salinity may add. The North Atlantic statistics, which may be more relevant for the AMOC, exhibit a similar tendency: Temperature is simulated with more success than salinity, and the spread is larger for salinity. There appears to be no clear systematic relationship between global and North Atlantic statistics.

[13] The mass flux is inconsistent with the observations for models 6, 3 and 1. Model 6 has almost no deep water formation in the North Atlantic. The final weight  $W$  is almost zero for models 3 and 6 and very small for models 5 and 1. Models 2 and 8 are superior to the others.

## 5. Projection

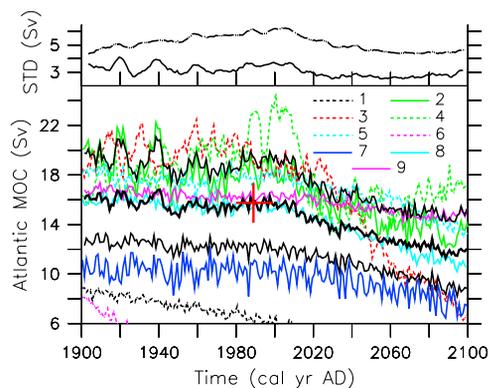
[14] How do these model statistics affect the projection? We show in Figure 2 the index of the AMOC at  $24^{\circ}\text{N}$ . As in the report by IPCC [2001], there is still a large spread in the model behavior. The initial states, the level of decadal variability, and the response to greenhouse warming, all three are rather different. The initial conditions, for instance, can differ by as much as 10 Sv ( $1\text{Sv} = 10^6\text{m}^3/\text{s}$ ). Likewise, the level of decadal variability varies from virtual no variability to a decadal standard deviation of several

Sverdrups (Sv). The weighted model mean from 1980–1999 is consistent with the observations.

[15] The projection based on the weighted mean shows a linear weakening of the circulation from  $15.7(\pm 3.5)$  Sv during the last decade of the last century (1990–1999) towards  $11.8(\pm 2.9)$  Sv during the decade 2090–2099. This presents a decrease by about  $25(\pm 25)\%$ . This result is remarkably robust with respect to alternative methods to calculate the model weights. The unweighted mean exhibits a similar reduction from  $14.2(\pm 6.0)$  Sv to  $10.3(\pm 4.6)$  Sv of 27%, suggesting that the response of AMOC does not depend much on the model performance. We can thus conclude that a considerable weakening of the AMOC can be expected until 2100. No individual model shows an abrupt collapse of the circulation during this century.

## 6. Discussion

[16] There are some problems with our methodology. First, it remains to be shown that simulating correct climatological mass fluxes, and temperature, salinity or pycnocline depth patterns can really improve AMOC prediction. As small-scale processes such as convection in the Labrador Sea or the overflows across the Greenland-Iceland-Scotland ridge system may influence AMOC behaviour, an approach based more on the important physical mechanisms to evaluate the climate models would be desirable. As such our methodology can be regarded only as a first step in the direction of a more refined model evaluation. The lack of an advanced ocean observing system, however, makes this a challenge. Strategies developed for ocean model evaluation, e.g. natural or artificial tracer distributions as in the OCMIP exercises [Doney *et al.*, 2004], should be extended to coupled models. Evaluation of the time dependent AMOC



**Figure 2.** Evolution of the Atlantic MOC as defined by the maximum overturning at  $24^{\circ}\text{N}$  for the period 1900–2100. The MOC evolutions of runs with a skill score larger than one are shown as solid lines, those from models with a smaller skill score as dashed lines. The weighted ensemble mean is shown by the thick black curve together with the weighted standard deviations (thin black lines). Observational estimates of the circulation at  $24^{\circ}\text{N}$  ( $15.75 \pm 1.6$  Sv [Ganachaud and Wunsch, 2000; Lumpkin and Speer, 2003]) at the end of the last century are shown as the red cross centered at year 1989. The top narrow panel shows the weighted (solid) and unweighted (dashed) standard deviations.

response should be attempted in the future. This will be possible using direct observations for the last century, and for the more distant past by including the increasing data base of paleo ocean circulation changes (e.g. during the last glacial maximum).

[17] An interesting outcome of our study is the result that the weighted model means are very similar to the unweighted ensemble mean. Does this mean the effort of weighting the models is useless? Whereas the evolution of the weighted and unweighted means is similar, the weighted standard deviation is generally smaller than the unweighted (Figure 2, top panel). This suggests that our method of assessing the models with observations can indeed reduce the uncertainty in the projection.

[18] Finally, what can we conclude for the stability of the North Atlantic THC under increased levels of greenhouse gas concentrations? First, a significant weakening of the AMOC is to be expected until 2100. Second, this change will evolve gradually, no model simulates an abrupt change. These two findings are consistent with the study of Gregory *et al.* [2005], who analyzed another type of greenhouse warming simulations. Third, the anthropogenically induced change in the North Atlantic THC is unlikely to leave the range of natural variability during the next several decades. This was also concluded by Curry *et al.* [1998] by analyzing ocean observations of the last 50 years and M. Latif *et al.* (Is the thermohaline circulation changing, submitted to *Journal of Climate*, 2005) by investigating the SSTs of the last century. We note, however, that increased melting from the Greenland ice sheet, a process not included in present climate models, may induce an additional freshwater forcing for the North Atlantic and accelerate the weakening of AMOC during the 21st century.

[19] **Acknowledgments.** This work was supported by the German CLIVAR and European ENSEMBLES projects and the SFB 460. We acknowledge the international modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. The help of Frank Kösters in the early stages of this study is greatly acknowledged.

## References

- Conkright, M. E., *et al.* (2002), World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures, CD-ROM documentation, 17 pp., NOAA, Silver Spring, Md.
- Curry, R. G., M. S. McCartney, and T. M. Joyce (1998), Oceanic transport of subpolar climate signals to mid-depth subtropical waters, *Nature*, 391, 575–577.
- Doney, S. C., *et al.* (2004), Evaluating global ocean carbon models: the importance of realistic physics, *Global Biogeochem. Cycles*, 18, GB3017, doi:10.1029/2003GB002150.
- Fraedrich, K., and L. M. Leslie (1987), Combining predictive schemes in short-term forecasting, *Mon. Weather Rev.*, 115, 1640–1644.
- Fraedrich, K., and N. R. Smith (1989), Combining predictive schemes in long-range forecasting, *J. Clim.*, 2, 291–294.
- Ganachaud, A. (2003), Large-scale mass transports, water mass formation, and diffusivities estimated from World Ocean Circulation Experiment (WOCE) hydrographic data, *J. Geophys. Res.*, 108(C7), 3213, doi:10.1029/2002JC001565.
- Ganachaud, A., and C. Wunsch (2000), Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data, *Nature*, 408, 453–457.
- Gnanadesikan, A., R. D. Slater, N. Gruber, and J. L. Sarmiento (2002), Oceanic vertical exchange and new production: A comparison between models and observations, *Deep Sea Res., Part II*, 49, 363–401.
- Gregory, J. M., *et al.* (2005), A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO<sub>2</sub> concentration, *Geophys. Res. Lett.*, 32, L12703, doi:10.1029/2005GL023209.
- Hughes, T. M. C., and A. J. Weaver (1994), Multiple equilibria of an asymmetric two-basin ocean model, *J. Phys. Oceanogr.*, 24, 619–637.
- Intergovernmental Panel on Climate Change (IPCC) (2001), *Climate Change 2001: The Scientific Basis. Technical Summary of the Working Group I Report*, Cambridge Univ. Press, New York.
- Knutti, R., T. F. Stocker, F. Joos, and G. K. Plattner (2003), Probabilistic climate change projections using neural networks, *Clim. Dyn.*, 21, 257–272.
- Lumpkin, R., and K. Speer (2003), Large-scale vertical and horizontal circulation in the North Atlantic Ocean, *J. Phys. Oceanogr.*, 33, 1902–1920.
- Manabe, S., and R. J. Stouffer (1993), Century-scale effects of increased atmospheric CO<sub>2</sub> on the ocean-atmosphere system, *Nature*, 364, 215–218.
- Manabe, S., and R. J. Stouffer (1999), The role of thermohaline circulation in climate, *Tellus, Ser. A*, 51, 91–109.
- Marotzke, J., and B. A. Klinger (2000), The dynamics of equatorially asymmetric thermohaline circulations, *J. Phys. Oceanogr.*, 30, 955–970.
- Metzger, S., M. Latif, and K. Fraedrich (2004), Combining ENSO forecasts: A feasibility study, *Mon. Weather Rev.*, 132, 456–472.
- Murphy, J. M., *et al.* (2004), Quantification of modelling uncertainties in a large ensemble of climate changes simulations, *Nature*, 430, 768–772.
- Palmer, T. N., *et al.* (2004), Development of a European Multi-Model Ensemble System for Seasonal to Interannual Prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, 85, 853–872.
- Rahmstorf, S. (2003), Thermohaline circulation: The current climate, *Nature*, 421, 699.
- Schmittner, A., K. J. Meissner, M. Eby, and A. J. Weaver (2002), Forcing of the deep ocean circulation in simulations of the Last Glacial Maximum, *Paleoceanography*, 17(2), 1015, doi:10.1029/2001PA000633.
- Smethie, W. M., Jr., and R. A. Fine (2001), Rates of North Atlantic Deep Water formation calculated from chlorofluorocarbon inventories, *Deep Sea Res., Part I*, 48, 189–215.
- Stocker, T. F., and A. Schmittner (1997), Influence of CO<sub>2</sub> emission rates on the stability of the thermohaline circulation, *Nature*, 388, 862–865.
- Talley, L. D., J. L. Raid, and P. E. Robbins (2003), Data-based meridional overturning streamfunctions for the global ocean, *J. Clim.*, 16, 3213–3226.
- Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106, 7183–7192.
- Trenberth, K. E., and J. M. Caron (2001), Estimates of meridional atmosphere and ocean heat transports, *J. Clim.*, 14, 3433–3443.
- M. Latif and B. Schneider, Ocean Circulation and Climate Dynamics, Leibniz-Institut für Meereswissenschaften, Duesternbrooker Weg 20, D-24105 Kiel, Germany. (mlatif@ifm-geomar.de)
- A. Schmittner, College of Oceanic and Atmospheric Sciences, Oregon State University, 104 Ocean Admin Bldg., Corvallis, OR 97331-5503, USA.