GEOPHYSICAL RESEARCH LETTERS, VOL. 40, 2714–2719, doi:10.1002/grl.50515, 2013

# Interaction network based early warning indicators for the Atlantic MOC collapse

Mirjam van der Mheen,[1] Henk A. Dijkstra,[1] Avi Gozolchiani,[2] Matthijs den Toom,[1] Qingyi Feng,[1] Jürgen Kurths,[3] and Emilio Hernandez-Garcia[4]

Received 14 March 2013; revised 26 April 2013; accepted 26 April 2013; published 4 June 2013.

[1] Early warning indicators of the collapse of the Atlantic Meridional Overturning Circulation (MOC) have up to now mostly been based on temporal correlations in single time series. Here, we propose new indicators based on spatial correlations in the time series of the Atlantic temperature field. To demonstrate the performance of these indicators, we use a meridional-depth model of the MOC for which the critical conditions for collapse can be explicitly computed. An interaction network approach is used to monitor changes in spatial correlations in the model temperature time series as the critical transition is approached. The new early warning indicators are based on changes in topological properties of the network, in particular changes in the distribution functions of the degree and the clustering coefficient. **Citation:** van der Mheen, M., H. A. Dijkstra, A. Gozolchiani, M. den Toom, Q. Feng, J. Kurths, and E. Hernandez-Garcia (2013), Interaction network based early warning indicators for the Atlantic MOC collapse, *Geophys. Res. Lett.*, *40*, 2714–2719, doi:10.1002/grl.50515.

## 1. Introduction

[2] Since Stommel's seminal paper [*Stommel*, 1961], many model studies have shown that the Atlantic Meridional Overturning Circulation (MOC) may be sensitive to changes in the freshwater balance of the northern North Atlantic [*Bryan*, 1986]. When an anomalous freshwater flux is gradually applied over a broad swath in the subpolar North Atlantic, the MOC collapses in many ocean-climate models [*Rahmstorf et al.*, 2005]. Freshening of the surface waters in the Nordic and Labrador Seas inhibits the production of North Atlantic Deep Water (NADW), which feeds the deep southward branch of the MOC. The MOC collapse occurs due to the existence of a tipping point associated with the salt-advection feedback [*Dijkstra and Ghil*, 2005; *Lenton*, 2011].

[3] Over the last few years, there has been a growing interest in developing early warning indicators for proximity of tipping points in systems for which the external conditions change very slowly with time [*Scheffer et al.*, 2009; *Kuehn*, 2011; *Scheffer et al.*, 2012; *Barnosky et al.*, 2012]. For regime shifts in ecosystems, early warning indicators have been based on either temporal or spatial correlations [*Donangelo et al.*, 2010].

[4] For the MOC collapse problem, the slow change is caused by the gradually varying freshwater forcing in the northern North Atlantic.

[5] Early warning indicators of the approach to the tipping point for the MOC have so far been mostly based on temporal characteristics of the single time series [*Held and Kleinen*, 2004] or an ensemble of such time series [*Livina and Lenton*, 2007]. The techniques currently used (for an overview, see *Lenton* [2011]) are based on the concepts of critical slowdown (degenerate fingerprinting [*Held and Kleinen*, 2004], detrended fluctuation analysis [*Livina and Lenton*, 2007]), and the existence of multiple equilibria (potential analysis [*Livina et al.*, 2010, 2011]).

[6] In this paper, we present new early warning indicators for the collapse of the MOC which are based on the spatial correlations of the Atlantic temperature field. Such changes in spatial correlations can be effectively detected by an interaction network approach [*Tsonis and Roebber*, 2004; *Donges et al.*, 2009; *Bialonski et al.*, 2011; *Gozolchiani et al.*, 2011; *Tsonis and Swanson*, 2012]. A link between different locations in such a network is established when a correlation measure of their covarying time series exceeds a certain threshold. To demonstrate the application potential of the new indicators, we apply the interaction network methodology to temperature time series from a meridional-depth model of the Atlantic MOC.

## 2. Model, Simulations, and Network Construction

[7] To be able to test the quality of an early warning indicator for approaching a tipping point of the MOC, it is beneficial to have a model where such tipping points can be explicitly computed. In *den Toom et al*. [2011], this is accomplished for a two-dimensional (meridional-depth) model of the Atlantic MOC. The formulation of this model is for convenience presented in part 1 of the supporting information (SI).

[8] The surface temperature restoring forcing is taken similar to the one in *den Toom et al*. [2011]; see equation (3a) in the SI. The surface freshwater forcing $F_S$, which is applied as a virtual salt flux (equation (3b) in the SI), is chosen as

$$F_S(\phi) = \frac{(\gamma+\eta)}{\cos\phi}\cos\pi\frac{\phi}{\phi_N} + \beta F_p(\phi) - Q \qquad (1)$$

Additional supporting information may be found in the online version of this article.

[1]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands.
[2]Department of Physics, Bar-Ilan University, Ramat-Gan, Israel.
[3]Potsdam-Institut für Klimafolgenforschung, Potsdam, Germany.
[4]IFISC (CSIC-UIB) Campus Universitat de les Illes Balears, Palma de Mallorca, Spain.

Corresponding author: H. A. Dijkstra, Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands. (H.A.Dijkstra@uu.nl)


0094-8276/13/10.1002/grl.50515

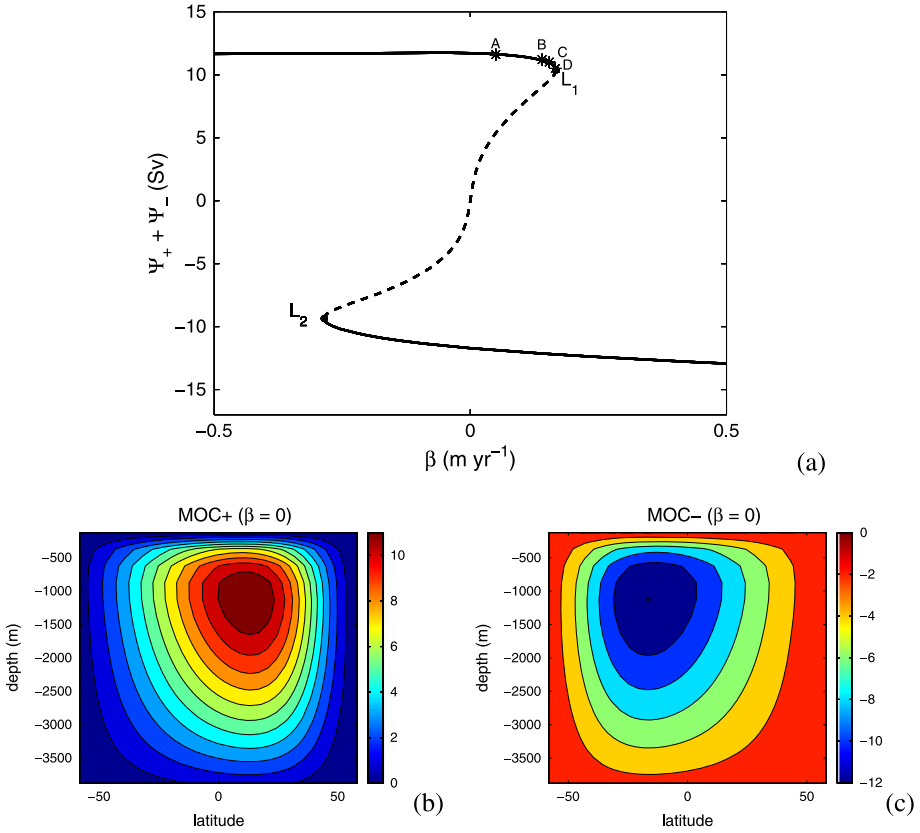


**Figure 1.** (a) Bifurcation diagram in which the sum of the maximum ($\Psi_+$) meridional overturning stream function (see part 1 of the SI) and minimum ($\Psi_-$) meridional overturning stream function values is plotted versus the anomalous freshwater flux amplitude $\beta$. Solid (dashed) branches indicate stable (unstable) steady states. (b) Pattern of the meridional overturning stream function ($\Psi$) at the upper stable branch in Figure 1a for $\beta = 0$. (c) Pattern of the meridional overturning stream function at the lower stable branch in Figure 1a for $\beta = 0$.

where $\phi$ indicates latitude, $\phi_N$ is the northern boundary of the equatorially symmetric domain, $\gamma$ (in m/yr) is the strength of the background freshwater forcing, and $\eta$ is a white noise term, with $< \eta(t)\eta(t') >= \eta_0\delta(t - t')$. Furthermore, $\beta$ is the strength of an anomalous freshwater flux which is only applied over the area [40°N, 60°N] (where $F_p = 1$; it is 0 elsewhere). With the area given over which the freshwater is applied, the total freshwater flux in $m^3$/yr is easily converted to Sv. The constant $Q$ appears due to the requirement that the surface-integrated salt flux is 0 for all parameter values.

[9] The model equations were discretized on a 32 × 16 meridional-depth spatial grid, and standard parameters used are provided in Table S1 of the SI. For $\eta_0 = 0$ and $\gamma = 0.3$ m/yr, the so-called bifurcation diagram [*Dijkstra and Ghil*, 2005] in $\beta$ is shown in Figure 1a. In this diagram, a measure of the steady solutions of the model is plotted versus the forcing parameter $\beta$. The diagram contains two so-called saddle-node bifurcations (indicated by $L_1$ and $L_2$) connected by a branch of unstable steady states. The MOC patterns of the stable steady states for $\beta = 0$ are shown in Figures 1b and 1c. We will focus on the upper tipping point indicated in Figure 1a as the saddle-node bifurcation $L_1$.

[10] The next four values of $\beta$ are chosen close to $L_1$, as indicated by the labels (A, B, C, and D) in Figure 1a. Starting from the steady state solution, transient computations over 500 years were performed for these values of $\beta$ with a noise amplitude $\eta_0 = 0.1$. In these simulations, the MOC fluctuates with a typical peak-to-peak amplitude of about 5 Sv. In the following, we analyze the spatial correlations of the annual mean temperature field of these transient simulations.

[11] An (undirected and unweighted) interaction network is reconstructed from each temperature field time series using a threshold value $\tau$ of the Pearson correlation coefficient between each pair of the 32 × 16 grid points. If this correlation exceeds $\tau$, two locations (nodes) are assumed "connected" [*Tsonis and Roebber*, 2004; *Donges et al.*, 2009]. In the results below, we will analyze the topological properties of these networks using $\tau = 0.7$ (details are provided in part 2 of the SI).

## 3. Results

[12] An overview of local and global measures to analyze the topology of a network is for example given in *Donges et al.* [2009]. We will here mainly focus on the degree distribution of the network and provide also results for the distribution of the clustering coefficient (see part 3 of the SI). The degree $d$ of a node is the number of connections from that node to other nodes, and the local clustering coefficient of a node provides information on the connections between neighbors of that node.

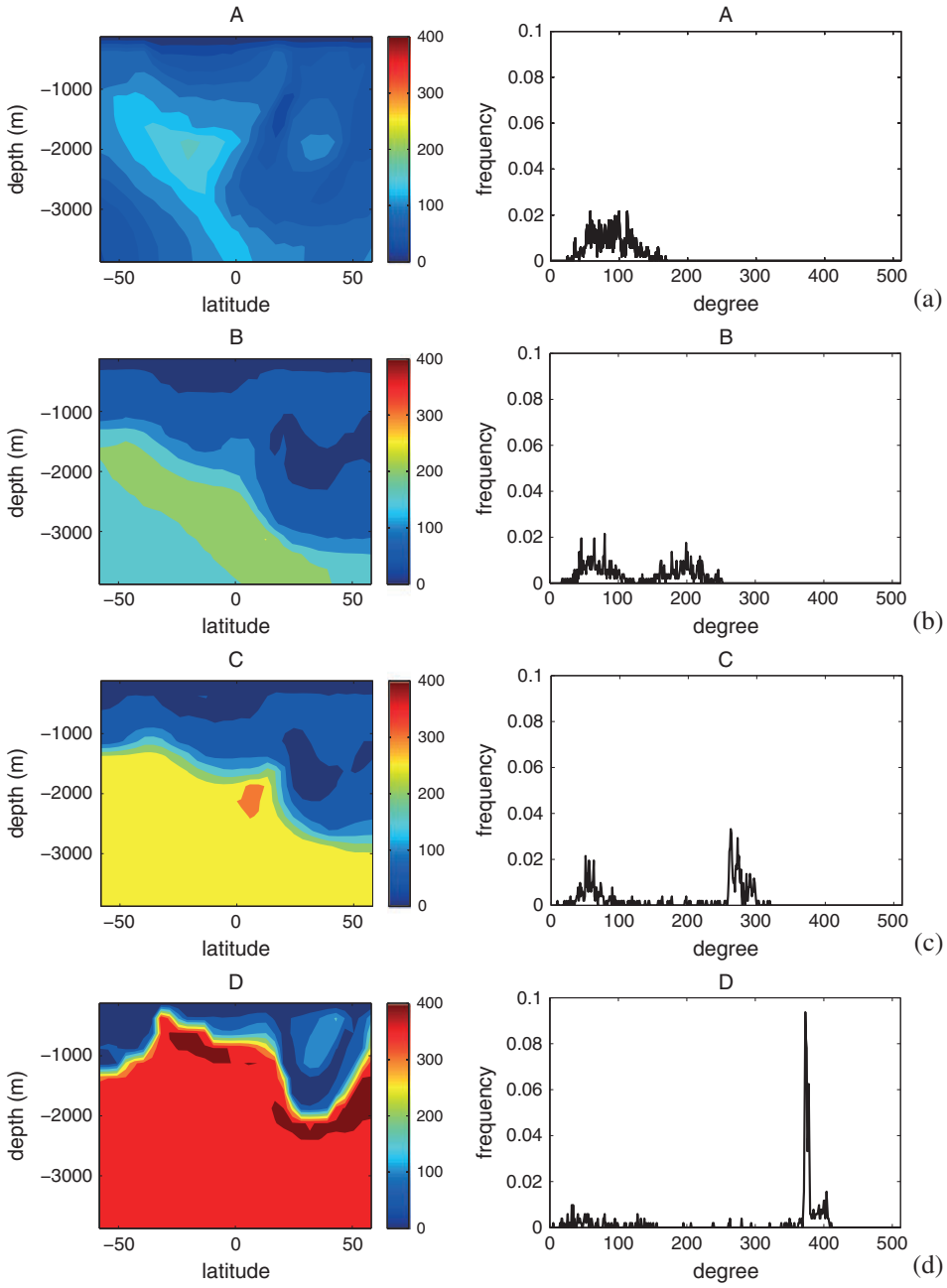


**Figure 2.** Degree fields (left panel) and degree distributions (right panel) of the undirected, unweighted network (with a threshold $\tau = 0.7$) for the temperature field time series obtained at different locations A, B, C, and D as indicated in Figure 1a under transient noise ($\eta = 0.1$) forcing and $\gamma = 0.3$. (a) Point A: $\beta = 0.05$ m/yr; (b) Point B: $\beta = 0.139$ m/yr; (c) Point C: $\beta = 0.154$ m/yr; (d) Point D: $\beta = 0.166$ m/yr.

### 3.1. Equilibrium Flows

[13] The degree fields of the networks constructed at the different values of $\beta$ (locations A, B, C, and D in Figure 1a) are shown in Figure 2 (left panels). Note that the maximum degree of a node (connections to other nodes) is $d_{\max} = 32 \times 16 - 1 = 511$. Relatively far from $L_1$, all temperature grid points have a relatively small degree. When $\beta$ is increased, high degree first appears at nodes in the deep ocean. Close to $L_1$, there is a large area (below ~1000 m) where the degree values are very large and a relatively low degree only occurs for nodes located in the upper ocean, north of 30°N.

[14] The histograms of the degree fields (the degree distributions) for different $\beta$ are also plotted in Figure 2 (right panels). For relatively small $\beta$, the degree distribution is unimodal, and the mean is located near $d = 100$. A second peak for $d = 200$ appears in Figure 2b, where the degree distribution becomes bimodal, and the maximum degree of the first peak shifts to slightly smaller values of $d$. This tendency strengthens in Figure 2c, where the maximum of the second peak shifts to an even higher degree. For location D, near $L_1$, about 10% of the grid points have a high degree of $d = 370$ (Figure 2d). Also, the clustering coefficient fields

and the clustering coefficient distribution show the same pattern changes and peak shift with $\beta$ (see part 2 of the SI and Figure S1).

[15] The results in Figure 2 can be explained by looking at the changes in the first few Empirical Orthogonal Functions (EOFs) of the temperature field versus $\beta$. The patterns of the first three EOFs at locations $A$ to $D$ are shown in Figure S2 (with variances provided in Table S2). Temperature anomalies are most pronounced in the northern part of the domain and at the surface. When the variance would be completely dominated by only one EOF, it is easy to show that the network is fully connected (see part 3 of the SI), and hence the degree of each node is maximal. The deviation maximal degree at each node is hence related to the relative variance explained by each of the EOFs at this node. When $L_1$ is approached, the amount of variance explained becomes dominated by EOF1 (in particular in the deeper ocean), and hence many nodes in the network will get a high degree and also a high clustering coefficient.

[16] The results in Figure 2 strongly motivate to use properties of the degree distribution as early warning indicators of the approach to the tipping point. Although there are values of $\beta$ where the degree distribution is clearly bimodal, near the tipping point this distribution is changing substantially and again becomes more unimodal. Hence, the expectation value $E_d$ of the normalized degree distribution $d/d_{\max}$ is expected to show a strong shift toward larger values as $L_1$ is approached. A similar measure $E_c$ is computed from the distribution of the clustering coefficients, and both will be used as new indicators of early warning.

[17] In Figure 3, the indicator $E_d$ shows a strong increase with $\beta$ when the tipping point is approached (solid black curve). The same holds for the indicator $E_c$ plotted as the dashed black curve in Figure 3. As soon as $\beta > 0.14$ myr$^{-1}$, a new regime in spatial correlations appears where the degree and clustering distributions become dominated by high values in the deeper ocean. We also computed both indicators when only a part of the grid points are used to compute the degree distribution. The blue (red) curves in Figure 3 arise when only the northern and southern (top and bottom) grid points are taken into account to reconstruct the network. The indicator $E_d$ constructed from the north-south grid points also smoothly increases with $\beta$ (and has much better behavior than that for the top-bottom grid) as it captures the high spatial correlations in the deep ocean.

### 3.2. Slowly Forced Flows

[18] We next investigated the performance of both indicators $E_c$ and $E_d$ in a transient simulation. Over a period of 35,000 years, $\beta$ was linearly increased from 0.05 to 0.17 myr$^{-1}$ (Figure 4a), and again noise ($\eta_0 = 0.1$) was applied to the freshwater flux. When the zonal dimension of the basin is assumed to be 64°, this amounts to about $10^{-3}$ Sv/1000 yr, which is comparable to other model studies [*Rahmstorf et al.*, 2005]. A time series of the maximum value of the MOC (Figure 4a) indicates that the MOC starts collapsing after about 27,000 years (indicated by the vertical dashed line in Figure 4) when $\beta$ approaches $L_1$. At this time, the actual value of $\beta = 0.142$ yr$^{-1}$, which is slightly larger than the value of point $B$ in Figure 1a. Note that because point $B$ is in the multiple equilibrium regime, the MOC can already decrease substantially under the presence of noise even before reaching $L_1$.

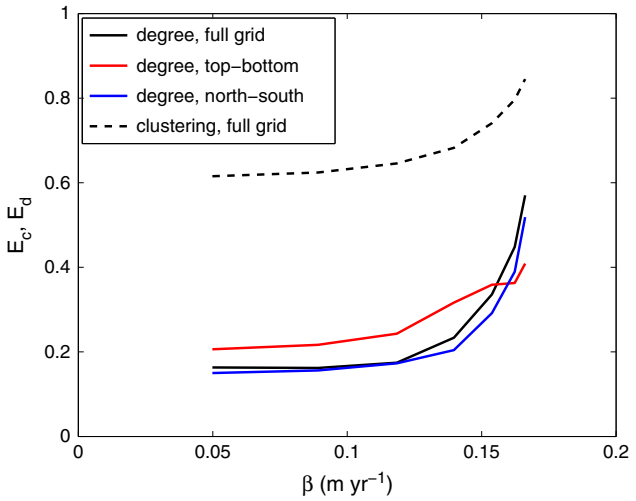


**Figure 3.** Results from the equilibrium flow simulations. The expectation value $E_d$ (solid) of the normalized degree distribution ($d/d_{\max}$) and the expectation value $E_c$ (dashed) of the distribution of the clustering coefficient versus $\beta$. The black curves are for all grid points. The red (blue) curve for the degree shows the indicators for a temperature network where only the bottom and top (northern and southern) boundary grid points of the domain are considered.

[19] Previously suggested early warning indicators of the transition in this MOC time series are plotted in Figures 4c–4f. Below we present only the results for a sliding window of 5000 years, but the robustness of the results with respect to the sliding window (5000–15,000) years was tested by using the Mann-Kendall trend test [*Hamed and Rao*, 1998; *Lenton et al.*, 2012].

[20] In Figure 4c, the (rescaled ) lag – 1 autocorrelation of the projection of the time series onto the first EOF (degenerate fingerprinting) is plotted. Here, we used an aggregation window of 50 years. The first EOF is generated from results of the first 5000 years of the simulation, and the lagged autocorrelation is calculated in a time window of 5000 years. The coefficient reaches zero indeed near the transition showing a critical slowdown of the MOC. However, the indicator is not monotonic and also increases when the MOC is still far from the transition.

[21] Figure 4d shows the power of the fluctuation function, as determined by the DFA procedure, using linear and quadratic detrending. The largest window taken in fitting the fluctuation function to a power law is 100 years. In most of the cases, the fitting quality coefficient is close to 1 (perfect fit), and the typical error in the coefficient is 0.5%. The quality of the sampling is assessed with respect to fluctuations of the measure during time. We have checked different orders of polynomial detrending in the DFA procedure, between 1 (linear) and 4 (quartic); the qualitative behavior is similar. The average power coefficient is 1.82, 1.85, 1.71, and 1.62, with respect to the four detrending orders 1–4, and there is no warning of a transition.

[22] Figure 4e gives the standard deviation of the MOC record. Each point is calculated using a sliding window of 5000 years with a shift of 200 years; it indeed shows a steady increase toward the transition, but it is difficult to set a threshold for an alarm. In Figure 4f, the fraction of

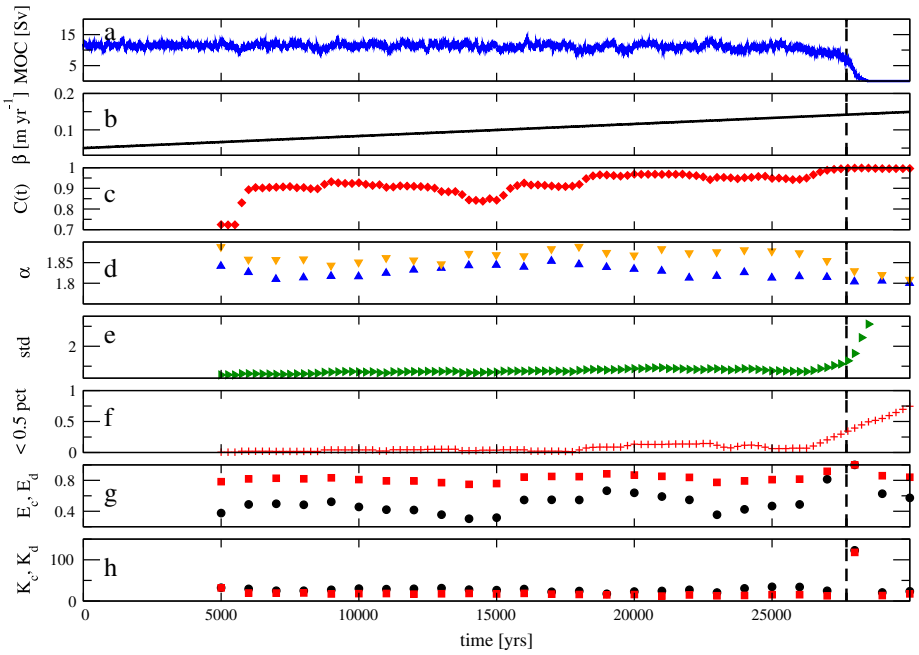


**Figure 4.** Results from the slowly forced flow simulations. (a) Time series of the maximum value of the MOC. (b) Transient change of $\beta$ versus time from 0.05 to 0.17 myr$^{-1}$ in 35,000 years. (c) The (rescaled) lag – 1 autocorrelation of the projection of the time series onto the first EOF. (d) Power of the fluctuation function, as determined by the DFA procedure, using linear (up marker) and quadratic (down marker) detrending. (e) Standard deviation of the MOC time series shown in Figure 4a. (f) Fraction of the time spent below the 0.5 percentile of the MOC amplitude. (g) Results from the network-based indicators $E_d$ (circles) and $E_c$ (squares) for the full network as in Figure 3. (h) Same as Figure 4g but now the kurtosis ($K_d$ (circles) and $K_c$ (squares)) is plotted. The vertical dashed line indicates the value of $\beta$ at the tipping point.

the time spent below the 0.5 percentile (estimated based on the first 5000 years) of the MOC strength is plotted using the same sliding window as in Figure 4e. This indicator displays a much sharper increase, but it is a rather ad hoc measure which also depends on a calibration in the far past.

[23] Finally, the network indicators $E_c$ and $E_d$ are plotted in Figure 4g and the kurtosis of the degree and clustering coefficient distributions in Figure 4h. Both are determined by networks constructed from the full spatial temperature field using a sliding window of 5000 years with a shift of 1000 years. Both indicators $E_c$ and $E_d$ exhibit similar behavior with an increase before the collapse. The kurtosis distributions (Figure 4h) are even more clearly showing a strong increase before collapse. By considering reconstructed networks for different sliding window lengths, it was found that the time at which the peak occurs in the indicators is not sensitive to a sliding window length in the range 5000–15,000 years. Similar peaks are found in the standard deviation and skewness of both distributions. Of all the indicators considered, the interaction network based kurtosis indicators are least susceptible to cause false alarms of the MOC collapse.

## 4. Summary and Discussion

[24] The strength of the Atlantic MOC is currently actively monitored [*Cunningham et al.*, 2007] with the motivation that the MOC may rapidly (within a few decades) change in strength. However, the time series of the MOC strength such as derived from the RAPID-MOCHA (Meridional Overturning and Heat Flux Array) program are certainly too short to be able to apply early warning indicators based on the slowdown and increased variance of the MOC [*Thompson and Sieber*, 2011; *Lenton*, 2011].

[25] Using an interaction network approach, we have developed here new indicators which are based on spatial correlations of the temperature variations in the Atlantic Ocean. We studied the performance of these indicators by looking at solutions of a meridional-depth model of the MOC for which the critical conditions for collapse could be computed in detail. We showed that the expectation value of the normalized degree distribution $E_d$ increases steeply and smoothly when the tipping point $L_1$ is approached. The explanation for this increase is based on the increased dominance of the first EOF in the variance of the solution in a large part of the domain. The disadvantage is that the indicator needs data fields as input (instead of single time series), although a partial set of measurements (for example only at the northern and southern boundaries) may also result in adequate early warning indicators.

[26] The results of the paper are applicable to output from general circulation models, where time series of fields are available. They furthermore show that to apply them to observations, a good spatial resolution in measurements is desirable. In a potential application to determine whether transitions in the MOC have been involved in the Dansgaard-Oeschger oscillations, well-synchronized records of ocean bottom temperatures are needed. Also, for the present-day MOC, the results in this paper strongly motivate to monitor the strength of the MOC at a more southerly location (such as the South Atlantic Meridional Overturning Circulation (SAMOC) project planned at 30°S).

[27] **Acknowledgments.** The authors thank Jonathan Donges, Norbert Marwan, and Reik Donner (PIK, Potsdam) for providing the software package (pyUnicorn) used in the network calculations here. The authors would like to acknowledge the support of the LINC project (no. 289447) funded by ECs Marie-Curie ITN program (FP7-PEOPLE-2011-ITN). H.D. also thanks the Burgers program of the University of Maryland, College Park, MD (USA), and in particular Kayo Ide for the support and hospitality during a visit where this paper was finished. We thank one of the anonymous referees for pointing out some inconsistencies in the first version of this paper.

[28] The Editor thanks Anastasios Tsonis and an anonymous reviewer for their assistance in evaluating this paper.

## References

Barnosky, A. D., et al. (2012), Approaching a state shift in Earth's biosphere, *Nature*, *486*, 52–58.

Bialonski, S., M. Wendler, and K. Lehnertz (2011), Unraveling spurious properties of interaction networks with tailored random networks, *PLoS ONE*, *6*(8), e22826.

Bryan, F. O. (1986), High-latitude salinity effects and interhemispheric thermohaline circulations, *Nature*, *323*, 301–304.

Cunningham, S., et al. (2007), Temporal variability of the Atlantic meridional overturning circulation at 26.5°, *N. Science*, *317*, 935.

den Toom, M., H. A. Dijkstra, and F. W. Wubs (2011), Spurious multiple equilibria introduced by convective adjustment, *Ocean Model.*, *38*, 126–137.

Dijkstra, H. A., and M. Ghil (2005), Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach, *43*(3), RG3002.

Donangelo, R., H. Fort, V. Dakos, M. Scheffer, and E. Van Nes (2010), Early warnings for catastrophic shifts in ecosystems: Comparison between spatial and temporal indicators, *Int. J. Bifurcat. Chaos*, *20*, 315–321.

Donges, J. F., Y. Zou, N. Marwan, and J. Kurths (2009), Complex networks in climate dynamics, *The Eur. Phys. J. Special Topics*, *174*, 157–179.

Gozolchiani, A., S. Havlin, and K. Yamasaki (2011), Emergence of El Niño as an autonomous component in the climate network, *Phys. Rev. Lett.*, *107*, 148501.

Hamed, K. H., and A. R. Rao (1998), A modified Mann-Kendall trend test for autocorrelated data, *J. Hydrol.*, *204*, 182–196.

Held, H., and T. Kleinen (2004), Detection of climate system bifurcations by degenerate fingerprinting, *Geophys. Res. Lett.*, *31*, L23207, doi:10.1029/2004GL020972.

Kuehn, C. (2011), A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics, *Physica D.*, *240*, 1020–1035.

Lenton, T. M. (2011), Early warning of climate tipping points, *Nature Climate Change*, *1*, 201–209.

Lenton, T. M., V. N. Livina, V. Dakos, and M. Scheffer (2012), Climate bifurcation during the last deglaciation, *Clim. Past*, *8*, 1127–1139.

Livina, V. N., and T. M. Lenton (2007), A modified method for detecting incipient bifurcations in a dynamical system, *Geophys. Res. Lett.*, *34*, L03712, doi:10.1029/2006GL028672.

Livina, V. N., F. Kwasniok, and T. M. Lenton (2010), Potential analysis reveals changing number of climate states during the last 60 kyr, *Clim. Past*, *6*, 77–82.

Livina, V. N., F. Kwasniok, G. Lohmann, J. W. Kantelhardt, and T. M. Lenton (2011), Changing climate states and stability: From Pliocene to present, *Clim. Dynamics*, *37*, 2437–2453.

Rahmstorf, S., M. Crucifix, A. Ganopolski, H. Goosse, I. Kamenkovich, R. Knutti, G. Lohmann, R. Marsh, L. A. Mysak, Z. Wang, and A. J. Weaver (2005), Thermohaline circulation hysteresis: A model intercomparison, *Geophys. Res. Letters*, *32*, L23605, doi:10.1029/2005GL023655.

Scheffer, M., J. Bascompte, W. A. Brock, and V. Brovkin (2009), Early-warning signals for critical transitions, *Nature*, *461*, 53–59.

Scheffer, M., et al. (2012), Anticipating critical transitions, *Science*, *338*, 344–348.

Stommel, H. (1961), Thermohaline convection with two stable regimes of flow, *Tellus*, *2*, 244–230.

Thompson, J. M. T., and J. Sieber (2011), Predicting climate tipping as a noisy bifurcation: A review, *Int. J. Bifurcat. Chaos*, *21*, 399–423.

Tsonis, A. A., and P. J. Roebber (2004), The architecture of the climate network, *Physica A*, *333*, 497–504.

Tsonis, A. A., and K. L. Swanson (2012), On the origins of decadal climate variability: A network perspective, *Nonlinear Proc. Geoph.*, *19*(5), 559–568.