A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer

Imme Ebert-Uphoff[1] and Yi Deng[2]

Received 25 July 2012; revised 23 August 2012; accepted 28 August 2012; published 2 October 2012.

[1] In this paper we introduce a new type of climate network based on temporal probabilistic graphical models. This new method is able to distinguish between *direct* and *indirect* connections and thus can eliminate indirect connections in the network. Furthermore, while correlation-based climate networks focus on *similarity* between nodes, this new method provides an alternative viewpoint by focusing on *information flow* within the network over time. We build a prototype of this new network utilizing daily values of 500 mb geopotential height over the entire globe during the period 1948 to 2011. The basic network features are presented and compared between boreal winter and summer in terms of intra-location properties that measure *local memory* at a grid point and inter-location properties that quantify *remote impact* of a grid point. Results suggest that synoptic-scale, sub-weekly disturbances act as the main information carrier in this network and their intrinsic timescale limits the extent to which a grid point can influence its nearby locations. The frequent passage of these disturbances over storm track regions also uniquely determines the timescale of height fluctuations thus *local memory* at a grid point. The poleward retreat of synoptic-scale disturbances in boreal summer is largely responsible for a corresponding poleward shift of local maxima in *local memory* and *remote impact*, which is most evident in the North Pacific sector. For the NH as a whole, both *local memory* and *remote impact* strengthen from winter to summer leading to intensified information flow and more tightly-coupled network nodes during the latter period. **Citation:** Ebert-Uphoff, I., and Y. Deng (2012), A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer, *Geophys. Res. Lett.*, *39*, L19701, doi:10.1029/2012GL053269.

1. Introduction

[2] In their seminal papers *Tsonis and Roebber* [2004] and *Tsonis et al.* [2006] introduced the idea of climate networks, which brought tools from network analysis to the field of climate science. Their basic idea is to use atmospheric fields - or other physical quantities - to define a correlation network of nodes, where each node represents a point on a global grid. Any two nodes are connected if the cross-correlation of the data associated with those two nodes is beyond a threshold. Since these correlation networks were introduced to climate science in 2004, there has been a flurry of research activity in this area, discussing definition, calculation, evaluation and interpretation of climate networks [*Tsonis et al.*, 2006; *Donges et al.*, 2009]. Several research groups related global network changes over a longtime scale to El Niño activity [*Tsonis et al.*, 2007; *Tsonis and Swanson*, 2008; *Gozolchiani et al.*, 2008; *Yamasaki et al.*, 2008, 2009]. *Steinhaeuser et al.* [2012] identify multivariate and multiscale dependence structures in correlation networks. A summary of the progress, opportunities and challenges of networks in climate science was presented by *Steinhaeuser et al.* [2010].

[3] While most climate networks are defined as correlation networks, two other definitions have recently been proposed, MI networks [*Donges et al.*, 2009] and phase synchronization networks [*Yamasaki et al.*, 2009]. All three network definitions, however, decide whether an edge exists between two nodes in the network based only on a test involving those two nodes and the results are fairly similar for all three.

[4] We seek to bring an additional set of tools to climate networks, namely the framework of causal discovery and specifically structure learning for probabilistic graphical models. Structure learning for graphical models was developed within the framework of causal discovery and books on the topic abound, see *Pearl* [1988, 2000], *Spirtes et al.* [1993, 2000], *Neapolitan* [2003] and *Koller and Friedman* [2009]. Probabilistic graphical models have already been used in several applications related to climate research, for example to predict severe weather events [*Abramson et al.*, 1996], precipitation [*Cofino et al.*, 2002], seabreezes [*Kennett et al.*, 2001] and daily pollution levels [*Cossention et al.*, 2001]. *Bannerjee* [2011] uses them to detect abrupt climate changes and for prediction of land variables based on ocean variables. Graphical models are also used as an aid for climate related decision making [*Catenacci and Giuppomi*, 2009] and to assess climate change effects [*Peter et al.*, 2009]. *Chu et al.* [2005] use the related concept of Graphical Granger Models to study teleconnections.

[5] In *Ebert-Uphoff and Deng* [2012] we provide an introduction to structure learning for climate researchers and as an example apply them to derive hypotheses of causal relationships between four prominent modes of atmospheric low-frequency variability in boreal winter. In this paper we show how these methods can be applied to the new application of climate networks. Note that while the application of climate networks was already proposed in the future work section of *Ebert-Uphoff and Deng* [2012], no work was reported on that topic in that paper.

[6] The key of structure learning is to use *conditional independence tests* involving *sets* of variables (rather than just tests involving *pairs* of variables) to distinguish between

[1]Electrical and Computer Engineering, Colorado State University, Fort Collins, Colorado, USA.
[2]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.

Corresponding author: Y. Deng, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332-0340, USA. (yi.deng@eas.gatech.edu)

0094-8276/12/2012GL053269

direct and indirect connections. Thus we can determine networks that only include *direct* connections between geographic locations, and disregard indirect connections, thus reducing the number of links in the network and making it more tractable.

[7] We suggested the use of structure learning for climate networks earlier [*Ebert-Uphoff and Deng*, 2010], but it took almost two years to work out all the computational details and make it happen. In this paper we present the first results obtained using this method. To the best of our knowledge this work presents the first time that structure learning of graphical models has been applied to climate networks. While correlation-based climate networks focus on *similarity* between nodes, this new method provides a very different approach by focusing on *information flow* within the global network over time.

2. Method

[8] We employ structure learning for probabilistic graphical models as basis for our climate networks. The type of graphical model used here is a directed graph, also known as *Bayesian network.* Conditional independence tests involving sets of variables are used to distinguish between direct and indirect connections. They can also be used to determine the *direction* of connections (i.e., the direction in which information is flowing) in many cases.

[9] We develop a *temporal* graphical model that models the interactions among daily 500 mb geopotential height at individual geographical locations over the course of ~2 weeks. The geopotential height data is obtained from the NCEP/NCAR Reanalysis covering the period 1948 to 2011 [*Kalnay et al.*, 1996; *Kistler et al.*, 2001]. We first define a set of N *grid points* that describe geographical locations on the globe and interpolate/extrapolate geopotential height data onto this grid. Next we choose a time span over which we want to develop the model, in our case S = 15 slices, with one day between slices. Our climate network model then consists of (N × S) nodes, where each network node represents a geographical location combined with a delay ranging from 0 to (S-1) days.

[10] The network structure, i.e., which nodes are connected by an edge and in which direction, is calculated using the PC algorithm developed by *Spirtes and Glymour* [1991]. We use the implementation of the PC algorithm in the TETRAD package (Version 4.3.10-3, available at http://www.phil.cmu.edu/projects/tetrad/) and choose Fisher's Z-test for the conditional independence tests with a significance level of $\alpha = 0.1$. As a priori knowledge we use only that nodes at time t can only have an influence on nodes at the same or a later time. While the procedure above basically follows the methodology in *Ebert-Uphoff and Deng* [2012], its application to climate networks presents additional challenges, which are described in the next subsection.

2.1. Defining the Grid Points

[11] The primary challenge is to find a set of suitable grid points. First of all, the *number* of grid points, N, is limited due to computational complexity, since the PC algorithm must determine the structure of a network with (N × S) nodes. The computational complexity of the PC algorithm is exponential in the number of nodes. It can be reduced to polynomial complexity by restricting the maximal number of parents per node. After choosing S = 15 slices we limited N to 200 nodes, so that the temporal climate network has a total of 3000 nodes. TETRAD required about 4.5 days (on a MacBook Pro, 2.3GHz Core i7) to complete the calculations for 3000 nodes using the unrestricted version of the PC algorithm.

[12] Secondly, the *spacing* of the grid points is extremely important for the quality of the results. In fact we learned by experiment that unequal spacing between neighboring grid points tends to create residual patterns in the network that overshadow the actual climate signals we want to detect. The reason is that proximity of grid points affects the amount of information flow between them, so for our algorithm – which is based on information flow and seeks to only include *strong*, *direct* connections - to be objective we need all neighboring points to have about the same distance. Standard grids in climate science, such as equal area grids, do not satisfy this requirement. In fact there is no closed-form solution to determine a grid on a sphere where neighboring grid points have equal distance. Instead we calculated our grid points as *Fekete points* using an algorithm that starts out with random locations and then moves them iteratively to maximize the minimal distance between any two points [*Bendito et al.*, 2007]. Using this method we generated three equally spaced grids on the globe with N = 200 points. To check for grid variability we calculated all of our results for all three grids.

2.2. Graphical Representations of Climate Networks

[13] We define two types of plots, *network plots* that show all the edges of the network for specific delays, and *contour plots* that show specific network properties at the grid points.

[14] The result of the structure learning is a list of all the edges in the temporal climate network that spans N physical locations and S different time slices (days). Dropping the first time slice of the model proved to be sufficient to deal with the typical initialization problem of temporal Bayesian networks [see *Ebert-Uphoff and Deng*, 2012]. Thus we have 14 time slices left. To find out which edges are repeated across time slices, we group the edges as follows. Any edges that connect from the same start location to the same end location, that have identical arrow direction (→, ↔, or −), *and* identical time delay, are placed into the same group and the number of their occurrence is counted across time slices. For example, an edge that goes from grid point 10 to grid point 15 from time slice 2 to 3, is placed in the same group as an edge that goes from grid point 10 to grid point 15 from time slice 5 to 6. Its occurrence is then at least 2. Then we generate separate network plots for a delay of 0, 1, 2, … days, displaying all arrows with that delay. Occurrence of each edge group, which is a good indication for its strength, is shown as follows in the plots. Edges that occur once are not shown. Edges that occur twice (3–4 times) are shown with a dotted (dashed) line. Edges that occur 5 or more times are shown with a solid line, with thickness increasing for 7 or more occurrences.

[15] To obtain contour plots we calculate certain properties of the network. *Intra-location properties* are derived considering only edges between one grid point and earlier or later instances of the same grid point. Thus it is a measure of *local memory* at a geographical location. *Inter-location properties* consider only edges between two different grid points and serve as a measure of *remote impact* of a

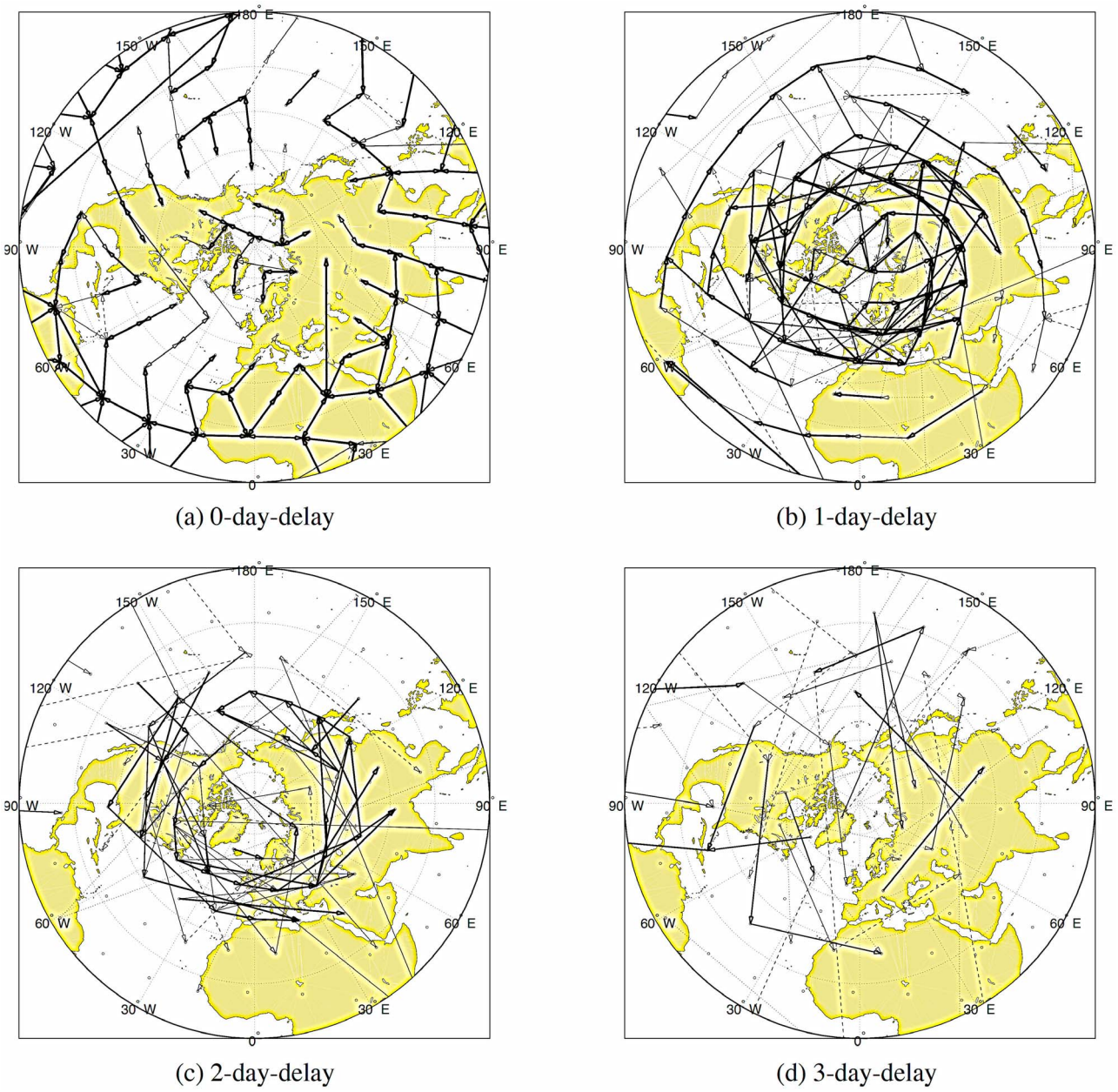

Figure 1. Typical network plots for delay of (a) 0, (b) 1, (c) 2, and (d) 3 days obtained for DJF based on Grid 1.

geographical location. We defined 4 properties, namely (1) intra-max-delay, (2) intra-ave-delay, (3) inter-max-delay, and (4) inter-ave-delay. Intra-max-delay records the largest time-delay (in terms of number of days) through which the information of the initial geopotential height value at a grid point can still be significantly "felt" at the same grid point. Intra-ave-delay, on the other hand, is an average of all these significant "memory-measuring" time-delays found at a single grid point. Inter-max (ave)-delay is the same as intra-max (ave)-delay except that it documents the largest (average) time-delay through which the impact of the height value at a geographical location can be felt at a *different* location. Contour plots are generated by interpolation from these 4 property values at individual grid points. Further averaging over the three grids is done to improve robustness of the results.

3. Results

[16] Here we use the Northern Hemisphere (NH) as a starting point to outline the most basic features of this new climate network and contrast those features between boreal winter and summer, defined respectively as the three-month period of December-January-February (DJF) and June-July-August (JJA). Figure 1 shows the *network plots* defined previously for NH DJF based on the first set of grid (Grid 1) and for a delay of 0, 1, 2, and 3 days. Note that for 0-day-delay the PC algorithm is unable to determine the direction for most edges (Figure 1a), while for a delay of 1 or more

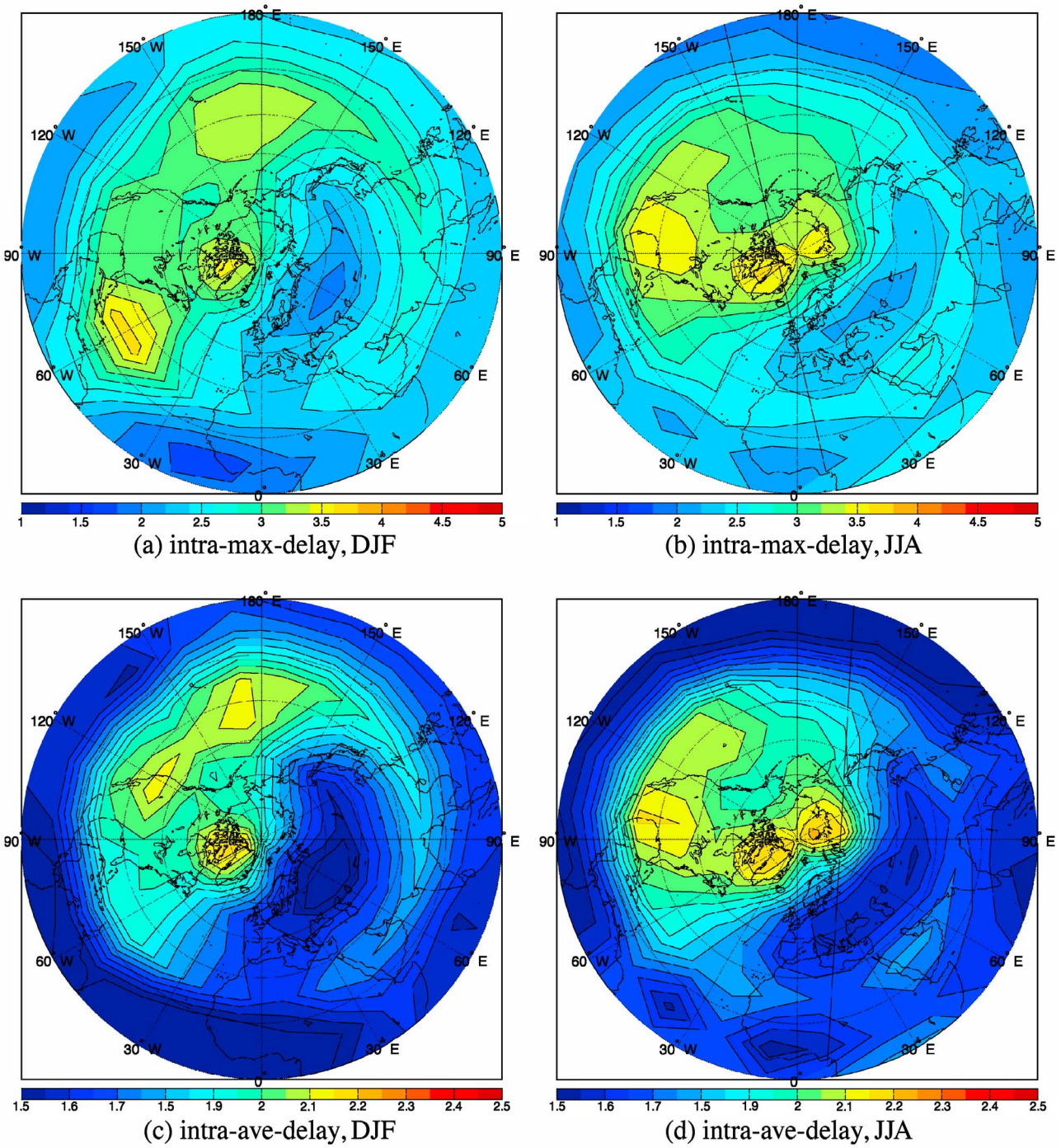

Figure 2. NH distribution of intra-max-delay in boreal (a) winter and (b) summer. (c and d) Same as Figures 2a and 2b except for intra-ave-delay. Unit: day.

days the a priori knowledge helps to uniquely determine all directions in the plot (Figures 1b–1d). It is interesting to see that the distribution density of the 0-day-delay inter-location edges decreases rapidly with latitude (Figure 1a). This clustering of 0-day-delay edges in the tropics and subtropics points to the existence of hidden "common" causes (e.g., sea surface temperature (SST)) that potentially regulate 500 mb height values at lower latitudes and have yet to be incorporated into the model. With non-zero time-delay, inter-location edges become increasingly clustered in the northern midlatitudes (Figures 1b–1d). A majority of the midlatitude edges are nearly zonally-oriented and directed from west to east, reflecting that synoptic-scale disturbances (with surface manifestations as extratropical cyclones/anticyclones) are the *information carrier* in the mid-troposphere during winter time (Figures 1b–1d). The role of synoptic-scale disturbances in establishing the climate network is further confirmed by the relative maxima of edge density found over eastern North Pacific, North America and eastern North Atlantic, where storm track activity tends to peak in DJF [e.g., *Chang et al.*, 2002; *Deng and Mak*, 2006; *Mak and Deng*, 2007].

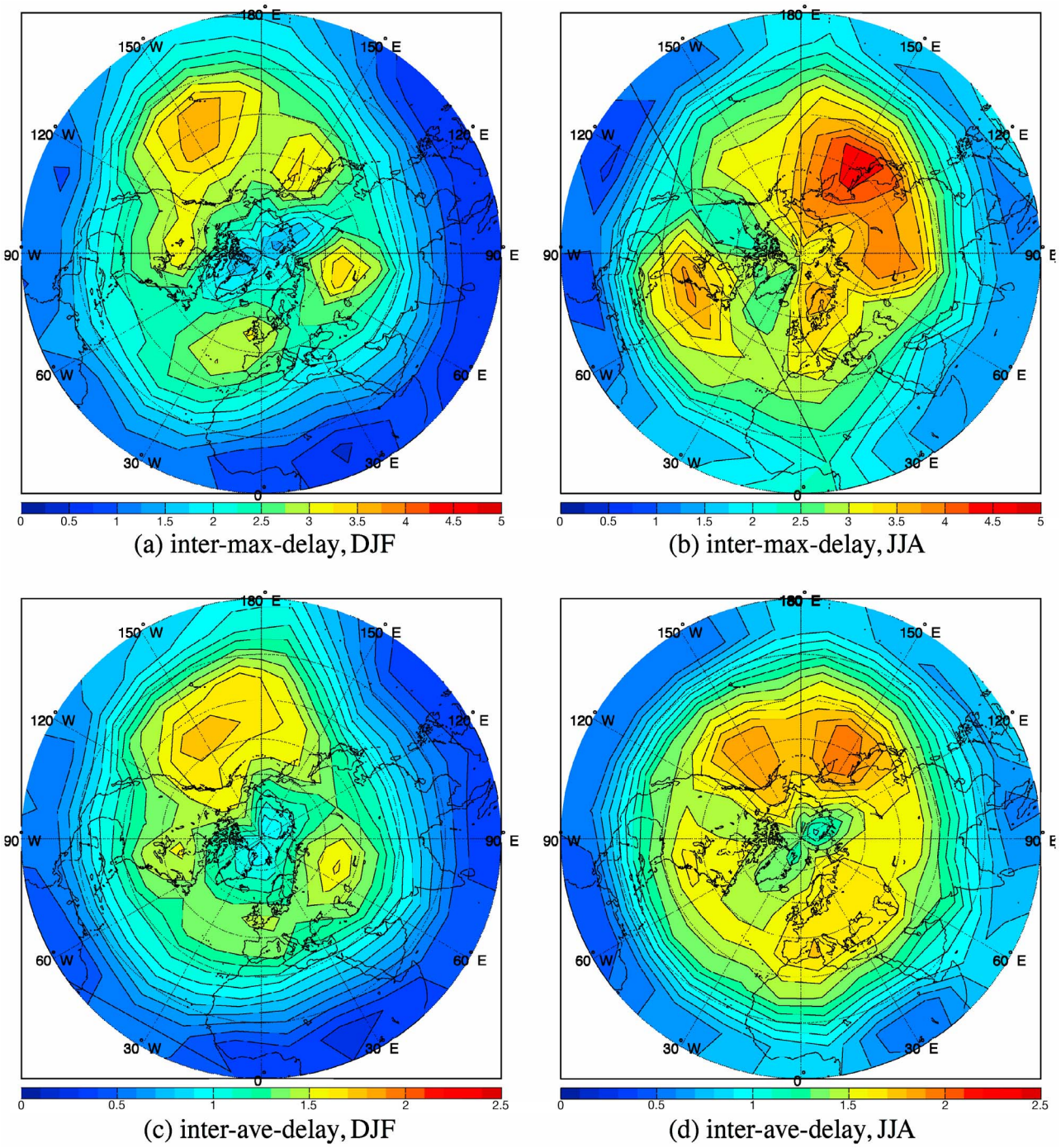

Figure 3. NH distribution of inter-max-delay in boreal (a) winter and (b) summer. (c) and (d) are the same except for inter-ave-delay. Unit: day.

[17] Figure 2 presents *contour plots* that illustrate the spatial distribution of intra-max-delay and intra-ave-delay, both measures of *local memory*, in DJF (Figures 2a and 2c) and JJA (Figures 2b and 2d). In either season, the two network properties share a similar spatial structure with the value of intra-max-delay typically 1 day greater than that of intra-ave-delay. In terms of intra-ave-delay, Eastern North Pacific and a region west of Greenland are marked with "longest" local memory (>2 days) in winter (Figure 2c), largely a result of frequent passages of synoptic-scale disturbances that help determine the sub-weekly-scale fluctuations at affected grid locations. In summer (Figure 2d), due to the poleward retreat of storm tracks, the eastern North Pacific is occupied by sporadic and shorter-scale disturbances. Correspondingly, long local memory is now found over regions closer to the North Pole where storm tracks are located in JJA and over southern North America where relatively stagnant atmospheric conditions contribute to slow variations in the 500 mb geopotential height.

[18] The *remote impact* of an individual location, quantified through inter-max-delay and inter-ave-delay, are shown in Figure 3 for DJF (Figures 3a and 3c) and JJA (Figures 3b

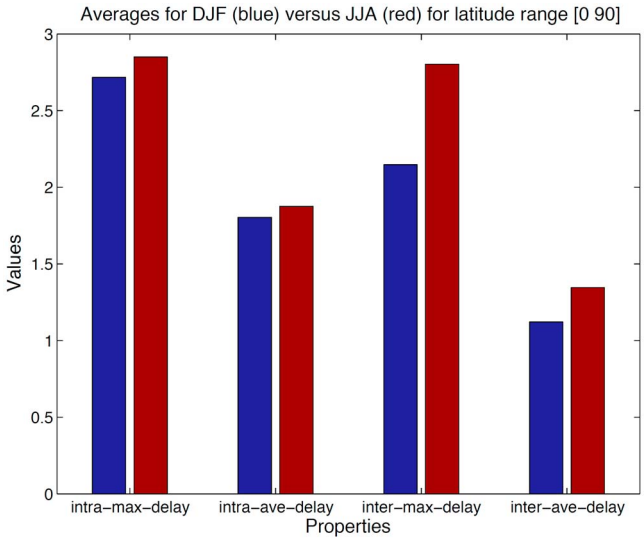

Figure 4. Changes of the NH-averaged local memory (intra-max (ave)-delay) and remote impact (inter-max (ave)-delay) from boreal winter (DJF) to summer (JJA). Unit: day.

and 3d). Like the local memory measure, the spatial distributions of these two inter-location properties share a certain degree of similarity, particularly in winter (Figures 3a and 3c). The values of inter-max-delay on average are approximately 1.5 days greater than those of inter-ave-delay. Consistent with the previous argument that synoptic-scale disturbances act as the main information carrier in winter (Figures 3a and 3c), the locations exhibiting largest *remote impact* ($\sim$2 days in inter-ave-delay and $\sim$4 days in inter-max-delay) are over the eastern North Pacific and appear as a zonal belt extending from the western North Atlantic to northern Eurasian continent, corresponding well to the North Pacific and North Atlantic storm track. During summer the poleward retreat of storm tracks drives regions of large *remote impact* further north, most obviously in the North Pacific sector (Figure 3d). There is also a moderate increase of $\sim$0.5 days from DJF (Figure 3c) to JJA (Figure 3d) in the value of the local maxima of *remote impact*, probably related to a reduced occurrence frequency and/or increased life-spans of synoptic-scale disturbances in the Arctic during boreal summer. Over the North-Atlantic-Europe sector, the meridional shift of zones of large *remote impact* is not pronounced, although moderate increases in the values of inter-ave-delay from DJF to JJA are clearly present (Figures 3c and 3d).

[19] Figure 4 compares the 4 network properties averaged over the entire NH in winter and summer. Measured through intra-ave-delay and inter-ave-delay, a NH grid point's memory of its 500 mb height value goes back to approximately 1.8 days on average while the mean time-span through which nearby locations can feel the impact of a particular grid point is approximately 1.3 days. For the NH as a whole, *local memory* slightly increases in summer compared to in winter – the JJA value of intra-max (ave)-delay is $\sim$0.2 (0.1) days greater than the corresponding DJF value. Similar but more substantial increases are found for hemispherically-averaged *remote impact*, where inter-max (ave)-delay in JJA is $\sim$0.75 (0.25) days greater than that in DJF. Further calculations indicate that the observed seasonal cycle in the values of the intra-location and inter-location properties occurs consistently over the tropics, midlatitudes and polar region. In summary, this new climate network derived through conditional independence tests is more robust in boreal summer with individual geographical nodes possessing longer local memory and information flow among different geographical nodes more steady and intense.

4. Concluding Remarks

[20] In this paper, we report basic properties of a new type of climate network obtained through structure learning of temporal probabilistic graphical models. By focusing on direct connections only and by analyzing information flow in the network over time, this new method provides a new way to identify and visualize network properties of the atmosphere – such as length of local memory and identification of pathways of remote impact - that were not available with previous methods.

[21] The contrast of network properties in the 500 mb geopotential height field between boreal winter and summer is discussed in terms of both intra-location properties that measure *local memory* at a grid point and inter-location properties that quantify *remote impact* of a grid point. Synoptic-scale disturbances are found to be the main information carrier in this network and their intrinsic timescale limits the impact of a geographical location on its nearby locations. They also uniquely determine the timescale of height fluctuations (thus *local memory*) at individual locations given frequent passages, particularly over storm track regions. The poleward retreat of synoptic-scale disturbances (thus storm tracks) in boreal summer is largely responsible for a corresponding poleward shift of local maxima in *local memory* and *remote impact*, which is most evident in the North Pacific sector. For the NH as a whole, both *local memory* and *remote impact* strengthen from winter to summer suggesting a more robust network and more tightly-coupled network nodes during boreal summer. One possible explanation for a less organized and more chaotic atmosphere in winter as revealed by this analysis is the predominance of baroclinic and barotropic instability of sheared flows in winter that spontaneously produce atmospheric disturbances (e.g., synoptic-scale disturbances) and thus add random noises to the atmosphere.

[22] The climate network introduced here has its root in causal discovery and emphasizes information flow within a system. We are currently exploring its potential of serving as a basic tool to detect impacts of long-term climate variability and change on the statistical properties of atmosphere at individual locations. Additional ongoing efforts include further examinations of the sensitivity of the network results to grid-spacing on the globe and investigations of the physical processes governing the variability of network properties across multiple timescales.

[23] **Acknowledgments.** We would like to thank the two anonymous reviewers for their insightful suggestions that led to significant improvement of the paper. The NCEP-NCAR reanalysis data used in this study was provided through the NOAA Climate Diagnostics Center. This research was in part supported by the DOE Office of Science Regional and Global Climate Modeling (RGCM) program under grant DE-SC0005596 and NASA Energy and Water Cycle Study (NEWS) program under grant NNX09AJ36G.

[24] The Editor thanks the two anonymous reviewers for assisting in the evaluation of this paper.

References

Abramson, B., J. Brown, W. Edwards, M. Murphy, and R. Winkler (1996), Hailfinder: A Bayesian system for forecasting severe weather, *Int. J. Forecast.*, *12*, 57–71, doi:10.1016/0169-2070(95)00664-8.

Bannerjee, A. (2011), Probabilistic graphical models for climate data analysis, paper presented at Supercomputing Workshop on Climate Knowledge Discovery (CKD-SC11), IEEE Comput. Soc., Seattle, Wash., 13 Nov.

Bendito, E., A. Carmona, A. M. Encinas, and J. M. Gesto (2007), Estimation of Fekete points, *J. Comput. Phys.*, *225*, 2354–2376, doi:10.1016/j.jcp.2007.03.017.

Catenacci, M., and C. Giuppomi (2009), Potentials of Bayesian networks to deal with uncertainty in climate change adaptation policies, *Tech. Rep. RP0070*, 29 pp., Cent. Euro-Mediterr. per i Cambiamenti Clim., Lecce, Italy.

Chang, E. K. M., S. Lee, and K. L. Swanson (2002), Storm track dynamics, *J. Clim.*, *15*, 2163–2183, doi:10.1175/1520-0442(2002)015<02163:STD>2.0.CO;2.

Chu, T., D. Danks, and C. Glymour (2005), Data driven methods for nonlinear granger causality: Climate teleconnection mechanisms, *Tech. Rep. CMU-PHIL*, 171 pp., Dep. of Philos., Carnegie Mellon Univ., Pittsburgh, Pa.

Cofino, A., R. Cano, C. Sordo, and J. Gutierrez (2002), Bayesian networks for probabilistic weather prediction, in *Proceedings of ECAI 2002: 15th European Conference on Artifical Intelligence*, edited by F. van Harmelen, pp. 695–700, IOS Press, Amsterdam.

Cossention, M., F. Raimondi, and M. Vitale (2001), Bayesian models of the pm 10 atmospheric urban pollution, in *Air Pollution IX*, *Adv. Air Pollut.*, vol. 10, edited by C. A. Brebbia and G. Latini, pp. 143–152, WIT Press, Southampton, U. K.

Deng, Y., and M. Mak (2006), Nature of the differences in the intraseasonal variability of the pacific and Atlantic storm tracks: A diagnostic study, *J. Atmos. Sci.*, *63*, 2602–2615, doi:10.1175/JAS3749.1.

Donges, J., Y. Zou, N. Marwan, and J. Kurths (2009), The backbone of the climate network, *Europhys. Lett.*, *87*, 48007, doi:10.1209/0295-5075/87/48007.

Ebert-Uphoff, I., and Y. Deng (2010), Causal discovery for climate networks, *Res. Rep. GT-ME-2010-001*, 15 pp., Ga. Inst. of Technol., Atlanta. [Available at http://smartech.gatech.edu/handle/1853/36564.]

Ebert-Uphoff, I., and Y. Deng (2012), Causal discovery for climate research using graphical models, *J. Clim.*, doi:10.1175/JCLI-D-11-00387.1, in press.

Gozolchiani, A., K. Yamasako, O. Gazit, and S. Havlin (2008), Pattern of climate network blinking links follows El Niño events, *Europhys. Lett.*, *83*, 28005, doi:10.1209/0295-5075/83/28005.

Kalnay, E., et al. (1996), The NCEP / NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kennett, R. J., K. B. Korb, and A. E. Nicholson (2001), Seabreeze prediction using Bayesian networks, in *Proceedings of the Fifth Pacific-Asia Conference on Knowledge Discovery and Data Minung (PAKDD 2001)*, edited by J. Fong and M. K. Ng, pp. 148–153, City Univ. of Hong Kong, Hong Kong.

Kistler, R., et al. (2001), The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*, 247–267, doi:10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Koller, D., and N. Friedman (2009), *Probabilistic Graphical Models–Principles and Techniques*, 1st ed., 1280 pp., MIT Press, Cambridge, Mass.

Mak, M., and Y. Deng (2007), Diagnostic and dynamical analyses of two outstanding aspects of storm tracks, *Dyn. Atmos. Oceans*, *43*, 80–99, doi:10.1016/j.dynatmoce.2006.06.004.

Neapolitan, R. E. (2003), *Learning Bayesian Networks*, 647 pp., Prentice Hall, Upper Saddle River, N. J.

Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd ed., 552 pp., Morgan Kaufman, San Mateo, Calif.

Pearl, J. (2000), *Causality: Models, Reasoning, and Inference*, 400 pp., Cambridge Univ. Press, Cambridge, U. K.

Peter, C., W. de Lange, J. Musango, K. April, and A. Potgleter (2009), Applying Bayesian modelling to assess climate change effects on biofuel production, *Clim. Res.*, *40*, 249–260, doi:10.3354/cr00833.

Spirtes, P., and C. Glymour (1991), An algorithm for fast recovery of sparse causal graphs, *Soc. Sci. Comput. Rev.*, *9*, 62–72, doi:10.1177/089443939100900106.

Spirtes, P., C. Glymour, and R. Scheines (1993), *Causation, Prediction, and Search*, *Springer Lect. Notes Stat.*, vol. 81, 1st ed., 526 pp., Springer, New York.

Spirtes, P., C. Glymour, and R. Scheines (2000), *Causation, Prediction, and Search*, 2nd ed., 546 pp., MIT Press, Cambridge, Mass.

Steinhaeuser, K., N. V. Chawla, and A. R. Ganguly (2010), Complex networks in climate science: Progress, opportunities and challenges, in *Proceedings of the 2010 Conference on Intelligent Data Understanding*, pp. 16–26, NASA Ames Res. Cent., Moffett Field, Calif.

Steinhaeuser, K., A. R. Ganguly, and N. V. Chawla (2012), Multivariate and multiscale dependence in the global climate system revealed through complex networks, *Clim. Dyn.*, *39*, 889–895, doi:10.1007/s00382-011-1135-9.

Tsonis, A., and P. Roebber (2004), The architecture of the climate network, *Physica A*, *333*, 497–504, doi:10.1016/j.physa.2003.10.045.

Tsonis, A., and K. Swanson (2008), Topology and predictability of El Niño and La Niña networks, *Phys. Rev. Lett.*, *100*, 228502, doi:10.1103/PhysRevLett.100.228502.

Tsonis, A. A., K. L. Swanson, and P. J. Roebber (2006), What do networks have to do with climate?, *Bull. Am. Meteorol. Soc.*, *87*, 585–595, doi:10.1175/BAMS-87-5-585.

Tsonis, A. A., K. Swanson, and S. Kravtsov (2007), A new dynamical mechanism for major climate shifts, *Geophys. Res. Lett.*, *34*, L13705, doi:10.1029/2007GL030288.

Yamasaki, K., A. Gozolchiani, and S. Havlin (2008), Climate networks around the globe are significantly affected by El Niño, *Phys. Rev. Lett.*, *100*, 228501, doi:10.1103/PhysRevLett.100.228501.

Yamasaki, K., A. Gozolchiani, and S. Havlin (2009), Climate networks based on phase synchronization track El Niño, *Prog. Theor. Phys.*, *179*, Suppl., 178–188, doi:10.1143/PTPS.179.178.