Geosci. Model Dev. Discuss., 8, C36–C39, 2015
www.geosci-model-dev-discuss.net/8/C36/2015/

Geoscientific Model Development Discussions Open Access

Interactive comment on "Par@Graph – a parallel toolbox for the construction and analysis of large complex climate networks" *by* H. Ihshaish et al.

J.F. Donges (Referee)

donges@pik-potsdam.de

Received and published: 14 February 2015

The work under consideration is an excellent paper presenting considerable and much needed advances in the network theoretical analysis of climatological data, e.g., from atmospheric or oceanographic data sources. The Par©Graph toolbox allows to efficiently construct and analyze climate networks from very large data sets including up to 10ˆ6 time series. This feature now allows to address new research questions related to the structure of climatic dynamics on small spatial scales or in three dimensions (e.g., three-dimensional atmospheric or oceanic climate networks). Also, the study of climate networks composed of a number of relevant variables such as temperature, pressure or salinity (coupled climate networks) promises novel insights into the processes governing climate dynamics and their complex interactions.

Beyond the information already given in the paper, I as a reader would be interested in learning about the following points:

- In what programming language is the toolbox implemented? Is there an easy to use interface in a scripting language such as Python?

- It would be important to know where the software can be obtained for interested researchers and whether it will be available as an open source code.

- Also, how flexible is the toolbox with regards to including new measures of time series similarity (such as event synchronization [1]) to the Network Constructor as well as additional measures for network quantification (e.g., node-weighted measures for taking into account inhomogeneous grids [2] or motif densities in directed networks [3,4])?

- What are the steps required towards including recently proposed methodologies for reconstructing statistically causal climate networks [5,6] in a parallel computing way using the toolbox? The challenge here is that links do not simply arise from a pairwise computation involving two time series, but involve a larger set of time series each because essentially, partial correlations and conditional probabilities are taken into account to exclude the potentially misleading effects of indirect couplings or common drivers.

- How is the toolbox equipped to deal with coupled climate networks [7] including links between time series in different climatological fields? In this case, it is important to allow for multiple thresholds for network construction, applying different thresholds within each field as well as between fields.

More detailed comments:

- p. 322, footnote 3: Please use the following URL for referring to pyunicorn: http://tocsy.pik-potsdam.de/pyunicorn.php. The suggested descriptive reference is: Donges, J.F., J. Heitzig, J. Runge, H.C. Schultz, M. Wiedermann, A. Zech, J.H. Feldhoff, A. Rheinwalt, H. Kutza, A. Radebach, N. Marwan and J. Kurths (2013, April).

GMDD
8, C36–C39, 2015

Advanced functional network analysis in the geosciences: The pyunicorn package. Geophysical Research Abstracts (Vol. 15, p. 3558).

- p. 322, line 24: change to -> "... developed at the Potsdam Institute ...".

- p. 323, line 4: in earlier works, also climate networks involving O(10ˆ4) have been studied. For example, the frequently analyzed NCEP/NCAR reanalysis surface air temperature data set involves more than 10,000 grid points or nodes.

- Section 2 and later Sections: there are some minor inconsistencies in the notation: a) Please use C_{ij} consistently (sometimes C_{i,j} is also used). b) N(i) is introduced on p. 326 to denote the neighborhood of node i, but later, sometimes k(i) is used for the same purpose, even though k(i) also denotes the degree of node i. c) Use subscript or brackets consistently for denoting the node i for a local network measure. In the paper, e.g., C_i is used for the clustering coefficient, but H(i) is used for the entropy measure.

- p. 326, Equation 5: a factor of 2 is missing in the nominator. This is because in the definition of the clustering coefficient, the number of links \rho_i between neighbors of i is divided by the maximimum possible number of such links which is (k_i over 2) = k_i(k_i-1) / 2.

- p. 327, line 18: for consistency, use BC_i instead of BC(v).

- p. 329, line 5: change to -> "..., then the pair (i,j) is copied to a ..."

- p. 329, line 7: explain what p_0 is here (the master compute node).

- There appear to be some capitalization issues on the front page p. 319 (e.g., look at the names of the institutions authors are affiliated to).

- p. 334, line 5/6: change to -> "... repeated annual cycle ..."

- p.334, line 24: capitalization: change to -> "Intel".

References:

GMDD

8, C36–C39, 2015

[1] Boers, N., Bookhagen, B., Barbosa, H. M. J., Marwan, N., Kurths, J., & Marengo, J. A. (2014). Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. Nature communications, 5.

[2] Heitzig, J., Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2012). Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes. The European Physical Journal B-Condensed Matter and Complex Systems, 85(1), 1-22.

[3] Zemp, D. C., Wiedermann, M., Kurths, J., Rammig, A., & Donges, J. F. (2014). Node-weighted measures for complex networks with directed and weighted edges for studying continental moisture recycling. EPL (Europhysics Letters), 107(5), 58005.

[4] Zemp, D. C., Schleussner, C. F., Barbosa, H. M. J., Van der Ent, R. J., Donges, J. F., Heinke, J., ... & Rammig, A. (2014). On the importance of cascading moisture recycling in South America. Atmospheric Chemistry and Physics, 14(23), 13337-13359.

[5] Runge, J., Heitzig, J., Petoukhov, V., & Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy. Physical review letters, 108(25), 258701.

[6] Ebert-Uphoff, I., & Deng, Y. (2012). Causal discovery for climate research using graphical models. Journal of Climate, 25(17), 5648-5665.

[7] Donges, J. F., Schultz, H. C. H., Marwan, N., Zou, Y., & Kurths, J. (2011). Investigating the topology of interacting networks-Theory and application to coupled climate subnetworks. The European Physical Journal B, 84(4), 635-651.

Interactive comment on Geosci. Model Dev. Discuss., 8, 319, 2015.