



Investigation of rotated PCA from the perspective of network communities applied to climate data

David Hartman, Jaroslav Hlinka, Martin Vejmelka, and Milan Palus

Academy of Sciences of the Czech Republic, Institute of Computer Science, Prague 8, Czech Republic (mp@cs.cas.cz)

Applications of the rotated principal component analysis (RPCA) have a long history in climatology usually due to efforts of finding specific circulation patterns (Barnston and Livezey 1987). Using this approach several well known patterns like the North Atlantic Oscillation (NAO) or the Pacific/North American Pattern (PNA) can be identified (Barnston and Livezey 1987; Feldstein 2000). Applied to the whole globe this method gives several weakly related components that can be suspected of being important modes of climate variability. On the other hand, a relatively new topic in climate research is that of community detection and analysis (Tsonis et al. 2011), although the detection of communities in complex networks is a well established scientific field itself (Fortunato 2010; Girvan and Newman 2002). To analyze community structure one has to consider the climate system as a complex network (Tsonis and Swanson 2012), i.e. as a set of nodes represented by a climate-related variable on specific globe positions and a set of edges mutually connecting these nodes according to chosen measure of coherence (Hlinka et al. preprint). Determination of optimal community structure is well known to be a hard problem and there are several methods excelling in specific situations (Fortunato 2010) and several ways of measuring quality of resulting community structure such as modularity (Newman and Girvan 2004). Following the fact that RPCA gives us a set of components that can be represented as a community structure we investigate the potential of RPCA in community-detection context. For this purpose we use data from global National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kistler et al. 2001), more specifically surface air temperature (SAT) and surface pressure level (SPL).

Acknowledgement: This study is supported by the Czech Science Foundation, Project No. P103/11/J068.

Barnston, AG; Livezey RE (1987) Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review* 115:1083–1126.

Feldstein, SB (2000) The timescale, power spectra, and climate noise properties of teleconnection patterns. *Journal of Climate* 13(24), 4430–4440.

Fortunato, S (2010) Community detection in graphs. *Physics Report-Review Section of Physics Letters* 486(3-5), 75–174.

Kistler, R and Coauthors (2001) The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bulletin of American Meteorological Society* 82, 247–267.

Hlinka J, Hartman D, Vejmelka M, Novotna D and Palus M. Non-linear dependence and teleconnections in climate data: sources, relevance, nonstationarity. Submitted preprint, available at arXiv:1211.6688.

Girvan, M ; Newman, MEJ (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Science of the United State of America* 99(12), 7821–7826.

Newman, MEJ and Girvan N (2004) Finding and evaluating community structure in networks. *Physical Review E* 69(2), 026113.

Tsonis, AA; Swanson, KL (2012) On the origins of decadal climate variability: a network perspective. *Nonlinear Processes in Geophysics* 19(5), 559–568.

Tsonis, AA; Wang, G; Swanson, KL ; Rodrigues, FA; Costa, LD (2011) Community structure and dynamics in climate networks. *Climate Dynamics* 37(5-6), 933–940.