How complex climate networks complement eigen techniques for the statistical analysis of climatological data

Jonathan F. Donges · Irina Petrova · Alexander Loew · Norbert Marwan · Jürgen Kurths

Received: / Accepted: date

Abstract Eigen techniques such as empirical orthogonal function (EOF) or coupled pattern (CP) / maximum covariance analysis have been frequently used for detecting patterns in multivariate climatological data sets. Recently, statistical methods originating from the theory of complex networks have been employed for the very same purpose of spatio-temporal analysis. This climate network (CN) analysis is usually based on the same set of similarity matrices as is used in classical EOF or CP analysis, *e.g.*, the correlation matrix of a single climatological field or the cross-correlation matrix between two distinct climatological fields. In this study, formal relationships as well as conceptual differences between both eigen and network approaches are derived and illustrated using global precipitation, evaporation and surface air temperature data sets. These results allow us to pinpoint that CN analysis can complement classical eigen techniques and provides additional information on the higher-order structure of statistical interrelationships in climatological data. Hence, CNs are a valuable supplement to the statistical toolbox of the climatologist, particularly for making sense out of very large data sets such as those generated by satellite observations and climate model intercomparison exercises.

Keywords climate networks · empirical orthogonal functions · coupled patterns · maximum covariance analysis · climate data analysis

Jonathan F. Donges · Norbert Marwan · Jürgen Kurths
Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany
E-mail: donges@pik-potsdam.de

Jonathan F. Donges
Stockholm Resilience Center, Stockholm University, Kräftriket 2B, 114 19 Stockholm, Sweden

Irina Petrova
Max-Planck-Institute for Meteorology, KlimaCampus, 20146 Hamburg, Germany

Alexander Loew
Department of Geography, University of Munich (LMU), Luisenstr. 37, 80333 Munich, Germany

Jürgen Kurths
Department of Physics, Humboldt University, Newtonstr. 15, 12489 Berlin, Germany
Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB243UE, United Kingdom
Department of Control Theory, Nizhny Novgorod State University, 603950 Nizhny Novgorod, Russia

1 Introduction

Climatologists have long been interested in studying correlations between climatological variables for gaining an understanding of the Earth's climate system's large-scale dynamics (Katz 2002). Pioneering work in this field was done by Sir Gilbert T. Walker in the beginning of the 20th century while attempting to find precursory patterns for Indian monsoon events using statistical methods (Walker 1910), which culminated in the discovery of the tropical Walker circulation and the Pacific Southern Oscillation (a part of the El Niño-Southern Oscillation known as ENSO). Later, new measurement devices as well as the rapid increase in available computing power allowed to investigate statistical interdependency structures of global or regional climatological fields $\mathbf{x}(t) = \{x_i(t)\}_{i=1}^{N}$ such as surface air temperature, pressure, or geopotential height (Fukuoka 1951; Lorenz 1956) (here, i is a spatial index, *e.g.*, labeling N meteorological measurement stations or grid points in an aggregated data set, and t denotes time).

Nowadays, techniques of eigenanalysis such as empirical orthogonal functions (EOFs) (Kutzbach 1967; Wallace and Gutzler 1981; Hannachi et al 2007) and coupled patterns (CPs) (Bretherton et al 1992) are standard tools for finding spatial as well as temporal patterns in climatological

data (von Storch and Zwiers 2003). Their applications range from statistical predictions (Lorenz 1956; Brunet and Vautard 1996; Repelli and Nobre 2004), over the definition of climate indices (Power et al 1999; Leroy and Wheeler 2008) to evaluating the performance of climate model simulation runs (Handorf and Dethloff 2009, 2012). While numerous linear and nonlinear extensions have been proposed (Ghil and Malanotte-Rizzoli 1991; Ghil et al 2002), e.g., rotated or simplified EOFs (Hannachi et al 2007) and other methods of dimensionality reduction such as neural network-based nonlinear principal component analysis (PCA) (Hsieh 2004) or isometric feature mapping (ISOMAP) (Tenenbaum et al 2000; Gámez et al 2004), classical EOF and CP analysis have remained among the most popular statistical techniques applied in climatology so far.

In the last decade, complex network theory has been introduced as a powerful framework for extracting information from large volumes of high-dimensional data (Newman 2003; Boccaletti et al 2006; Newman 2010; Cohen and Havlin 2010) such as those generated by neurophysiological or biochemical measurements, quantitative social science as well as climatological observations and modeling campaigns. While EOFs, CPs, and related methods effectively rely on a dimensionality reduction, network techniques allow to study the full complexity of the statistical interdependency structure within a multivariate data set. In these climate networks (CNs), which were first introduced by Tsonis and Roebber (2004); Tsonis et al (2006), nodes correspond to time series of climate variability at grid points or observational stations and links indicate a relevant statistical association between two such time series. For quantifying statistical associations, linear covariance or Pearson correlation can be used analogously to EOF and CP analysis (Tsonis and Roebber 2004; Tsonis and Swanson 2008; Yamasaki et al 2008), but nonlinear measures such as mutual information (Donges et al 2009a,b; Barreiro et al 2011) or transfer entropy (Runge et al 2012a) may be employed as well with care (Hlinka et al 2014). Among other applications, CNs have been used to uncover global impacts of El Niño events (Tsonis and Swanson 2008; Yamasaki et al 2008; Gozolchiani et al 2011; Martin et al 2013; Radebach et al 2013), trace the flow of energy and matter in the surface air temperature field (Donges et al 2009a), unravel the complex dynamics of the Indian summer monsoon (Malik et al 2012; Stolbova et al 2014), detect community structure enabling statistical prediction of climate indices (Tsonis et al 2011; Steinhaeuser et al 2011, 2012) as well as intercomparisons between climate models and observations (Steinhaeuser and Tsonis in press; Feldhoff et al 2014), and study large-scale circulation patterns and prominent modes of variability in the atmosphere (Tsonis et al 2008; Donges et al 2011c; Ebert-Uphoff and Deng 2012a,b). Furthermore, CN analysis has recently been employed to improve forecasting of El Niño episodes (Ludescher et al 2013, 2014), predict extreme precipitation events over South America (Boers et al 2014a) and to derive early warning indicators for the collapse of the Atlantic meridional overturning circulation (Mheen et al 2013). Extending upon the majority of studies focussing on recent climate variability, the CN approach has also been applied to study late Holocene Asian summer monsoon dynamics based on data from paleoclimate archives (Rehfeld et al 2013)

The main aim of this contribution is to put the recent CN approach into context with standard eigenanalysis, since both classes of methods are often based on the same set of statistical similarity matrices. We briefly review both classes of techniques to establish a common notation. Formal relationships are then derived between empirical orthogonal functions or coupled patterns and frequently used CN measures such as degree or cross-degree, respectively. These relationships are illustrated empirically using global satellite observations of precipitation and evaporation fields as well as surface air temperature reanalysis data. We furthermore illustrate and argue in which settings higher-order CN measures such as betweenness may contain information complementing classical eigenanalysis. For example, betweenness can be interpreted as approximating the flow of energy and matter within a climatological field and is particularly useful for identifying bottlenecks that may be particularly vulnerable to perturbations such as volcanic eruptions or anthropogenic influences (Donges et al 2009a, 2011c; Boers et al 2013; Molkenthin et al 2014a). Hence, by transferring insights and tools from complex network theory and complexity science to climate research, CNs meet the need for novel techniques of climate data analysis facing quickly increasing data volumes generated by growing observational networks and model intercomparison exercises like the coupled model intercomparison project (CMIP) (Meehl et al 2005; Taylor et al 2012).

This article is structured as follows: After describing the data to be analyzed (Section 2), we introduce eigen (Section 3) and network (Section 4) techniques for the statistical analysis of climatological data. Relationships between both approaches are formally derived and empirically demonstrated using observational climate data in Section 5. This leads us to pinpoint the added value of CN analysis (Section 6), before concluding in Section 7.

2 Data

Imperfect retrieval algorithms and data merging of atmospheric fields that are involved in the generation of reanalysis data sets may cause uncertainties and lower quality of the final product of data analysis. In order to obtain consistent and representative precipitation and evaporation fields, in this study, the fully satellite-based HOAPS-3 (Hamburg

Ocean Atmosphere Parameters and Fluxes from Satellite Data, http://www.hoaps.org, Andersson et al (2010b, 2011)) and combined HOAPS-3/ GPCC (Global Precipitation Climatology Center, http://www.gpcc.dwd.de, Andersson et al (2010a)) data sets are used. Regardless of the improved retrieval algorithms and high quality output product, the uniqueness of the HOAPS data set consists in utilization of only one satellite data set for retrieval of both, evaporation, and precipitation parameters. Originally available at the resolution of 0.5 degrees in latitude and longitude, monthly mean precipitation ($\mathbf{x}$(t)) and evaporation ($\mathbf{y}$(t)) anomaly fields (1992–2005) were resampled to T63 resolution (≈ 1.8 degrees) to reduce computational costs. Furthermore, areas with sea-ice coverage were excluded from the set of raw time series. This results in $N_P = 13,834$ and $N_E = 7,986$ grid points (or network nodes) and $M = 168$ samples for each time series for the global precipitation and evaporation data sets, respectively. The smaller number of nodes in the evaporation field arises because the data are only available over the oceans, but not over land. We use the full global data sets for comparing univariate techniques of climate data analysis, but for clarity restrict ourselves to the North Atlantic Ocean region for the multivariate methods.

Additionally, to put our work into context with earlier work on CN analysis (Tsonis and Swanson 2008; Yamasaki et al 2008; Donges et al 2009a; Steinhaeuser et al 2012), we study global monthly averaged surface air temperature (SAT) field data covering the years 01/1948–12/2007 taken from the reanalysis I project provided by the National Center for Environmental Prediction / National Center for Atmospheric Research (NCEP/NCAR, Kistler et al (2001)). This data set consists of $N_T = 10,224$ grid points (network nodes) and $M = 720$ samples for each time series.

3 Eigenanalysis

This section serves to introduce the mathematics of eigenanalysis necessary for the deductions made below. Specifically, standard EOF analysis of single climatological fields (e.g., the precipitation field) as well as coupled patterns based on a singular value decomposition of the cross-correlation matrix (also termed *maximum covariance analysis* (MCA) in von Storch and Zwiers (2003)) for studying statistical relationships between two climatological fields (e.g., the precipitation and evaporation fields) are discussed. Of all the variants of eigenanalysis (Hannachi et al 2007), these two approaches appear to be the most frequently used and are also most closely related to CN and coupled CN analysis, respectively, as will be elaborated on in Section 5. For further details, the reader is referred to Bretherton et al (1992), von Storch and Zwiers (2003) or Hannachi et al (2007).

Note, that for consistency with the CN literature (see Section 4), we define EOFs (CPs) based on the correlation (cross-correlation) instead of the covariance (cross-covariance) matrix. The results and conclusions presented in Sections 5 and 6 would not change qualitatively if the covariance (cross-covariance) matrix would be used for both eigenanalysis and CN construction.

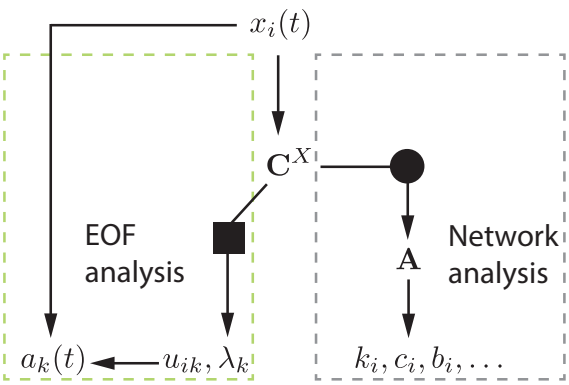

Fig. 1 A schematic outline of the relationship between univariate EOF and climate network analysis in the spirit of the diagrams in Bretherton et al (1992). The eigen decomposition (PCA) operation is represented by the square, the thresholding operation by the disc. All vectors are written in component form.

3.1 Empirical orthogonal function analysis

Given a set of normalized time series $\mathbf{x}(t) = \{x_i(t)\}_{i=1}^N$ with zero mean and unit standard deviation, the *correlation matrix* $\mathbf{C}^X = \{C_{ij}^X\}_{ij}$ is defined by

$$C_{ij}^X = \frac{1}{M} \sum_{t=1}^{M} x_i(t) x_j(t), \tag{1}$$

where M is the length (number of samples) of each time series.

The aim of EOF analysis (also termed principal component analysis in the statistical literature (Preisendorfer and Mobley 1988)) is a dimensional reduction achieved by decomposing the data into linearly independent linear combinations of the different variables that explain maximum variance (Hannachi et al 2007). The EOFs $\mathbf{u}_k$ are obtained as solutions of the eigenvalue problem

$$\mathbf{C}^X \mathbf{u}_k = \lambda_k \mathbf{u}_k. \tag{2}$$

The k–th EOF $\mathbf{u}_k$ is the eigenvector corresponding to the k–th largest eigenvalue λ_k, where u_{ik} denotes the i–th component of the k–th EOF (Fig. 1). The EOFs are sorted according to the ordering of their associated non-negative eigenvalues λ_k such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_R$ (R is the rank of $\mathbf{C}^X$). Hence, $\mathbf{u}_1$ associated with the largest eigenvalue λ_1 is

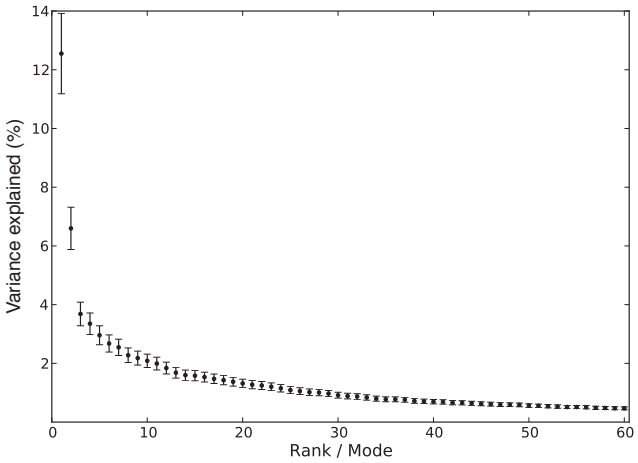

Fig. 2 Percentage of variance $\lambda_k / \sum_{l=1}^{R} \lambda_l$ explained by EOFs $\mathbf{u}_k$ for the HOAPS-3 / GPCC precipitation data set. Error bars were estimated using North's rule of thumb (North et al 1982).

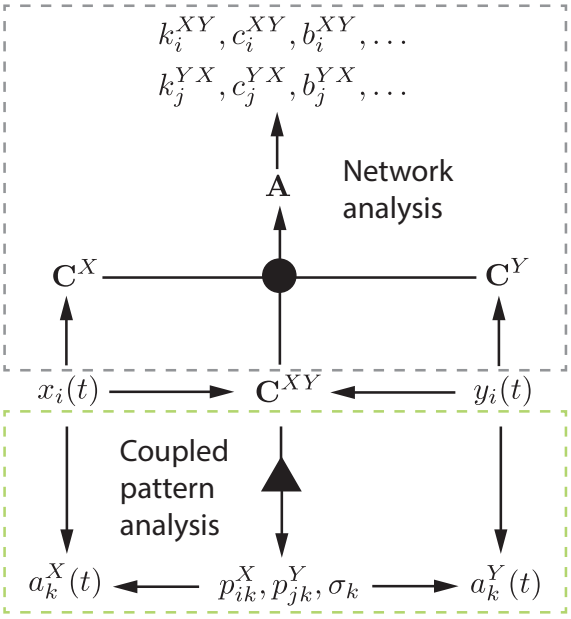

Fig. 3 A schematic outline of the relationship between coupled pattern (maximum covariance) and coupled climate network analysis in the spirit of the diagrams in Bretherton et al (1992). The singular value decomposition (SVD) operation is represented by the triangle, the thresholding operation by the disc. All vectors are written in component form.

called the *leading EOF* of the underlying data set and represents the one-dimensional projection of the data with the largest possible variance.

The normalized data $x_i(t)$ can be decomposed as (Fig. 1)

$$x_i(t) = \sum_{k=1}^{R} \lambda_k a_k(t) u_{ik}, \tag{3}$$

where $a_k(t)$ is the t–th component of the k–th principal component $\mathbf{a}_k$ (PC) (temporal pattern) associated with the k–th EOF $\mathbf{u}_k$ (spatial pattern) with

$$a_k(t) = \sum_{j=1}^{N} u_{kj} x_j(t). \tag{4}$$

For many climatological data sets such as the precipitation and evaporation fields studied here, most of the variance in the data $\mathbf{x}(t)$ can be explained by a small number of EOFs, i.e., the eigenvalues λ_k decay quickly with increasing rank k (Fig. 2). Equation (3) shows that in this situation, only a few EOFs and PCs are needed to closely approximate the data which allows the dimensionality reduction of high-dimensional data sets.

3.2 Coupled pattern (maximum covariance) analysis

Given two sets of normalized time series $\mathbf{x}(t) = \{x_i(t)\}_{i=1}^{N_X}$, and $\mathbf{y}(t) = \{y_j(t)\}_{j=1}^{N_Y}$ the *cross-correlation matrix* $\mathbf{C}^{XY} = \{C_{ij}^{XY}\}_{ij}$ is defined by

$$C_{ij}^{XY} = \frac{1}{M} \sum_{t=1}^{M} x_i(t) y_j(t), \tag{5}$$

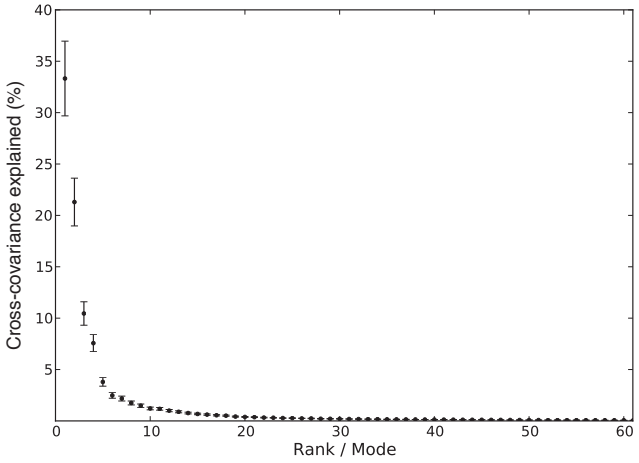

Fig. 4 Percentage of squared covariance $\sigma_k^2 / \sum_{l=1}^{R} \sigma_l^2$ between HOAPS-3 / GPCC precipitation (X) and HOAPS-3 evaporation (Y) data sets over the North Atlantic region (see Fig. 7) that is explained by pairs of coupled patterns $\mathbf{p}_k^X$, $\mathbf{p}_k^Y$. Most of the data sets' cross-covariance is captured by a small number of modes with the largest singular values σ_k. Error bars were estimated using North's rule of thumb (North et al 1982).

where M is the length (number of samples) of each time series. R in the following denotes the rank of $\mathbf{C}^{XY}$.

Maximum covariance analysis identifies spatially orthonormal pairs of coupled patterns $\mathbf{p}_k^X = \{p_{ik}^X\}_{i=1}^{N_X}$, $\mathbf{p}_k^Y = \{p_{jk}^Y\}_{j=1}^{N_Y}$ that explain as much as possible of the temporal

covariance between the two fields $\mathbf{x}(t)$ and $\mathbf{y}(t)$ (Bretherton et al 1992; von Storch and Zwiers 2003). The coupled patterns can be found by solving the system of equations

$$(\mathbf{C}^{XY})^T \mathbf{p}_k^X = \sigma_k \mathbf{p}_k^Y$$
$$\mathbf{C}^{XY} \mathbf{p}_k^Y = \sigma_k \mathbf{p}_k^X \quad (6)$$

by means of a singular value decomposition of $\mathbf{C}^{XY}$ (Fig. 3). Here, the $\mathbf{p}_k^X$ are an orthonormal set of R vectors called *left singular vectors*, the $\mathbf{p}_k^Y$ are an orthonormal set of R vectors called *right singular vectors*, and the σ_k are non-negative numbers called *singular values*, ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R$. Here, R denotes the rank of $\mathbf{C}^{XY}$. The total squared covariance explained by a certain pair of patterns $\mathbf{p}_k^X$, $\mathbf{p}_k^Y$ is σ_k^2. Therefore, the *leading coupled patterns* $\mathbf{p}_1^X$, $\mathbf{p}_1^Y$ explain the largest fraction of squared covariance between the two fields of interest. In our example, taking into account only a few pairs of coupled patterns with the largest σ_k already explains most of the covariance between the precipitation and evaporation fields (Fig. 4).

The fields $\mathbf{x}(t), \mathbf{y}(t)$ can be expanded in terms of the coupled patterns as

$$x_i(t) = \sum_{k=1}^{R} a_k^X(t) p_{ik}^X, \quad (7)$$

$$y_i(t) = \sum_{k=1}^{R} a_k^Y(t) p_{ik}^Y. \quad (8)$$

The expansion coefficients are obtained by projecting

$$a_k^X(t) = \sum_{i=1}^{R} p_{ik}^X x_i(t), \quad (9)$$

$$a_k^Y(t) = \sum_{i=1}^{R} p_{ik}^Y y_i(t). \quad (10)$$

4 Network techniques

Complex network analysis offers a general framework for studying the structure of associations (links) between objects (nodes) that are of interest in many disciplines. Typical examples include the internet or world wide web in computer science, road networks and power grids in engineering, food webs in biology or social networks in sociology (Newman 2003; Boccaletti et al 2006; Newman 2010; Cohen and Havlin 2010). It has become popular recently in several fields of science to apply the wealth of concepts and measures from complex network theory for the analysis of data that is even not given explicitly in network form. In network-based data analysis, a data set at hand, *e.g.*, consisting of time series such as electroencephalogram, climate records, or spatiotemporal point events such as earthquake aftershock swarms, first has to be transformed to a network representation by means of a suitable algorithm or mathematical mapping. The resulting networks are referred to as *functional networks* to distinguish them from *structural networks* that are derived from systems with a more obvious graph structure, *e.g.*, social networks or power grids. Examples of functional networks include gene regulatory networks in biology (Hempel et al 2011), functional brain networks in neuroscience (Bullmore and Sporns 2009), CNs in climatology (Donges et al 2009a,b, 2011c), or networks of earthquake aftershocks in seismology (Davidsen et al 2008). Forming a distinct class of methods, techniques for the network-based analysis of single or multiple time series such as recurrence networks (Xu et al 2008; Marwan et al 2009; Donner et al 2010) and visibility graphs (Lacasa et al 2008) have recently been studied intensively with a focus on (paleo-)climatological applications (Donges et al 2011a,b; Hirata et al 2011; Donner and Donges 2012; Feldhoff et al 2012).

The first functional network analysis of fields of climatological time series $\mathbf{x}(t)$ was presented by Tsonis and Roebber (2004), introducing the term *climate network*[1]. Climate network analysis offers novel insights by transferring the toolbox of measures and algorithms from complex network theory to the study of climate system dynamics. Climate networks are simple graphs (i.e., there are no self-loops and at most one link between each pair of nodes) consisting of N spatially embedded nodes i that correspond to time series $x_i(t)$ representing observations, reanalyses, or simulations of climatological variables at fixed measurement stations, grid cells, or certain predefined regions. Links $\{i, j\}$ represent particularly strong or significant statistical interdependencies between two climate time series $x_i(t)$, $x_j(t)$, where usually a filtering procedure is applied first to reduce the effects of the annual cycle (Donner et al 2008).

Put differently, for a pairwise measure of statistical association S_{ij} such as Pearson correlation (Tsonis and Roebber 2004; Tsonis et al 2006), mutual information (Donges et al 2009b,a; Paluš et al 2011), transfer entropy (Runge et al 2012a), or event synchronization (Malik et al 2012; Boers et al 2013; Stolbova et al 2014; Boers et al 2014b), a CN's adjacency matrix is given by

$$A_{ij} = \begin{cases} \Theta\left(S_{ij} - T_{ij}\right) & \text{if } i \neq j, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

[1] Note that the term *climate network* is also used in distinct contexts that are unrelated to graph theory or data analysis, *e.g.*, for describing collections of climatological/weather observation stations like the *Greenland climate network* (Steffen and Box 2001) or associations of political organizations dealing with anthropogenic climate change such as the *Climate Network Europe* (Raustiala 2001).

where $\Theta(\cdot)$ is the Heaviside function, T_{ij} denotes a threshold parameter, and $A_{ii} = 0$ is set for all nodes i to exclude self-loops. Usually, the threshold is fixed globally, *i.e.*, $T_{ij} = T$ for all node pairs (i, j). However, T_{ij} may also be set for each pair individually to only include links with values of S_{ij} exceeding a prescribed significance level, e.g., determined from a statistical test using surrogate time series (Paluš et al 2011). In most studies, symmetric measures of statistical interdependency $S_{ij} = S_{ji}$ have been considered, leading to undirected CNs. However, Gozolchiani et al (2011), Malik et al (2012) and Boers et al (2014b) exploited asymmetries in the cross-correlation function as well as in a measure of event synchronization to reconstruct directed CNs.

In the following, univariate and coupled CNs are introduced for studying the statistical interdependency structure within single fields as well as between two fields, respectively, together with graph-theoretical measures that are typically used for their quantification. For consistency with eigenanalysis (see Section 3), we restrict ourselves to linear Pearson correlation at zero lag as the measure of statistical association, *i.e.*, $S_{ij} = |C_{ij}|$.

4.1 Univariate climate networks

Given a climatological field $\mathbf{x}(t)$, the *adjacency matrix* $\mathbf{A} = \{A_{ij}\}_{ij}$ of the associated *climate network* is given by

$$A_{ij} = \Theta(\left|C_{ij}^X\right| - T) - \delta_{ij} \tag{12}$$

with a prescribed global threshold $0 \leq T \leq 1$, where δ_{ij} denotes Kronecker's delta (see Eq. (1) for the definition of C_{ij}^X). The absolute value of Pearson correlation $\left|C_{ij}^X\right|$ is commonly used, typically because negative correlations are considered equally important as positive ones (Tsonis and Roebber 2004). Among others, univariate CNs have been studied by Tsonis et al (2006); Tsonis and Swanson (2008); Tsonis et al (2008); Yamasaki et al (2008); Gozolchiani et al (2008); Yamasaki et al (2009); Donges et al (2009a,b); Tsonis et al (2011); Berezin et al (2012); Gozolchiani et al (2011); Guez et al (2012); Paluš et al (2011); Donges et al (2011c); Tominski et al (2011); Zou et al (2011); Malik et al (2012); Rheinwalt et al (2012); Rehfeld et al (2013).

The *degree* k_i is the most frequently applied measure for studying CNs. It gives the number of network neighbors for each node i and is defined as

$$k_i = \sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} \Theta(\left|C_{ij}^X\right| - T) - 1. \tag{13}$$

Maxima in the spatial pattern $\mathbf{k}$ with values of the degree that are much larger than average are referred to as *supernodes* or *hubs* (Tsonis and Roebber 2004; Tsonis et al 2006). These super-nodes indicate regions in the underlying field that are particularly strongly correlated to many other parts of the globe which are typically related to teleconnection patterns (Tsonis et al 2008). For example, in the HOAPS-3 / GPCC precipitation data the most strongly connected region in the tropical Pacific (Fig. 5B) corresponds to the El Niño-Southern Oscillation that is known to display global teleconnections (Ropelewski and Halpert 1987; Halpert and Ropelewski 1992; Tsonis et al 2008).

Path-based centrality measures from network theory reveal higher-order patterns in the statistical interdependency structure of a climatological field (Donges et al 2009a,b; Paluš et al 2011). High-order, in this context, refers to structures such as paths or network motifs that consist of two or more links, in contrast to the degree that is restricted to counting pairwise relationships between nodes. In this study, shortest-path closeness and betweenness are considered. *Closeness centrality* $\mathbf{c} = \{c_i\}_{i=1}^N$ (CC) measures the inverse mean network distance of node i to all other nodes via shortest paths and is defined as

$$c_i = \frac{N-1}{\sum_{j=1}^{N} l_{ij}}, \tag{14}$$

where l_{ij} denotes the length of a shortest (or geodesic) path connecting nodes i and j, *i.e.*, the smallest number of links that are passed when traveling from i to j in the CN. In contrast, *betweenness* $\mathbf{b} = \{b_i\}_{i=1}^N$ (BC) counts the relative number of shortest paths connecting any pair of nodes j, k that include node i and is defined as

$$b_i = \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{n_{jk}(i)}{n_{jk}}. \tag{15}$$

Here, n_{jk} denotes the total number of shortest paths between j, k. $n_{jk}(i)$ gives the size of the subset of these paths that include i. CC and BC have been applied for comparing different types of CNs (Donges et al 2009b), revealing a backbone of energy flow in the surface air temperature field (Donges et al 2009a), unraveling the complex dynamics of the precipitation field during the Indian summer monsoon (Malik et al 2012), and studying the signatures of El Niño and La Niña events (Paluš et al 2011). See Section 6 for a more in depth discussion of the interpretation of these CN measures.

4.2 Coupled climate networks

One option for condensing information from more than one climatological observable in a CN is to define links based on statistical interdependencies between multivariate time series describing the dynamics of multiple observables recorded at the same locations/nodes. For example, Steinhaeuser

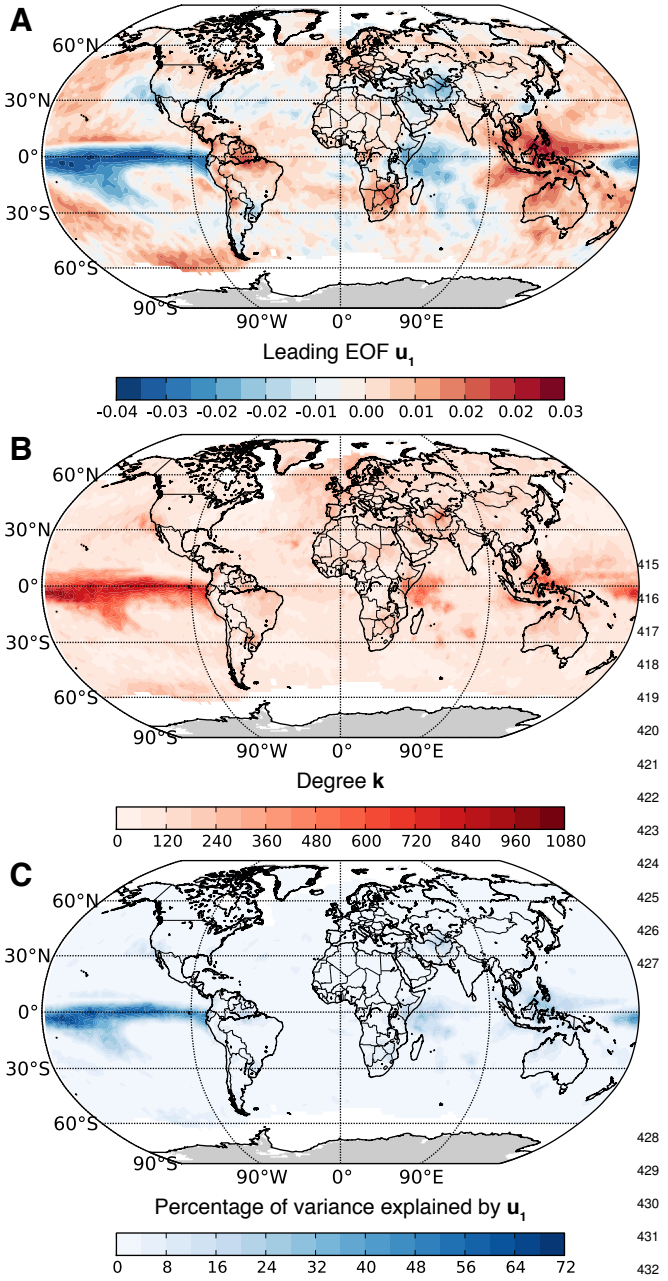

Fig. 5 Maps of (A) first EOF $\mathbf{u}_1$, (B) climate network degree field $\mathbf{k}$, and (C) local percentage of variance explained by first EOF $\mathbf{u}_1$, $100 \times \mathrm{Corr}(x_i(t), a_1(t))^2$ (homogeneous correlation map, see Björnsson and Venegas (1997)), for the global HOAPS-3 / GPCC precipitation data set. The climate network construction threshold $T = 0.27$ was chosen to yield a link density of $\rho = 0.01$ (Eq. (25)). Note the similarity in the patterns displayed in panels (A)–(C) that is explained in Section 5.

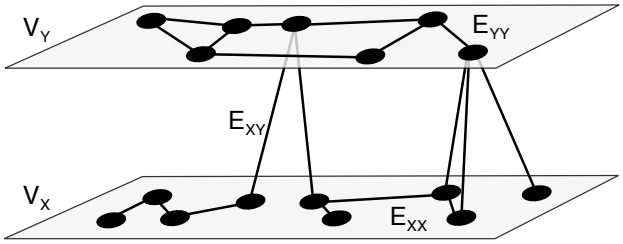

Fig. 6 A coupled climate network as it is constructed in this work, where V_X and V_Y denote the set of nodes in the subnetworks corresponding to grid points in data sets $\mathbf{x}(t)$ and $\mathbf{y}(t)$, respectively. E_{XX} and E_{YY} are sets of internal links within the subnetworks describing statistical relationships within each climatological field, while E_{XY} contains information on their mutual statistical interdependencies. Figure is adapted from (Donges et al 2011c).

et al (2010) analyzed a CN constructed from surface air temperature, pressure, relative humidity, and precipitable water to extract regions of related climate variability. In contrast to this multivariate approach, coupled CNs are designed to represent statistical dependencies within and between two climatological fields $\mathbf{x}(t) = \{x_i(t)\}_{i=1}^{N_X}, \mathbf{y}(t) = \{y_j(t)\}_{j=1}^{N_Y}$ or within and between different regions (Donges et al 2011c). For this purpose, all time series from each of the involved climatological fields are associated to $N_X + N_Y$ nodes in the resulting network (Fig. 6). A coupled CN is defined by its adjacency matrix $\mathbf{A}$ that is obtained by thresholding the correlation matrix $\mathbf{C}$ of the concatenated fields $\mathbf{x}(t), \mathbf{y}(t)$, analogously to Eq. (12). Decomposing $\mathbf{C}$ as

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^X & \mathbf{C}^{XY} \\ (\mathbf{C}^{XY})^T & \mathbf{C}^Y \end{pmatrix} \tag{16}$$

suggests to view coupled CNs as networks of networks or multilayer networks (Zhou et al 2006; Buldyrev et al 2010; Gao et al 2011; Boccaletti et al 2014), where subnetworks (network layers) $G_X = (V_X, E_{XX})$ and $G_Y = (V_Y, E_{YY})$ are the induced subgraphs of the sets of nodes V_X, V_Y belonging to data sets $\mathbf{x}(t)$, $\mathbf{y}(t)$, respectively (Fig. 6). While the edge sets E_{XX}, E_{YY} describe the fields' internal correlation structure based on the correlation matrices $\mathbf{C}^X, \mathbf{C}^Y$, the set of cross-edges E_{XY} captures dependencies between both fields and is based on the cross-correlation matrix $\mathbf{C}^{XY}$ (Fig. 3). Coupled CNs have been applied for studying the Earth's atmosphere's general circulation structure (Donges et al 2011c), processes linking climate variability in the North Atlantic and North Pacific regions via the Arctic (Wiedermann et al 2013, in prep.), global atmosphere-ocean interactions (Feng et al 2012). Also, the coupled CN approach underlies the method developed in Ludescher et al (2013, 2014) for forecasting El Niño events.

The statistical interdependency structure between fields $\mathbf{x}(t)$, $\mathbf{y}(t)$ can be quantified with a set of graph-theoretical

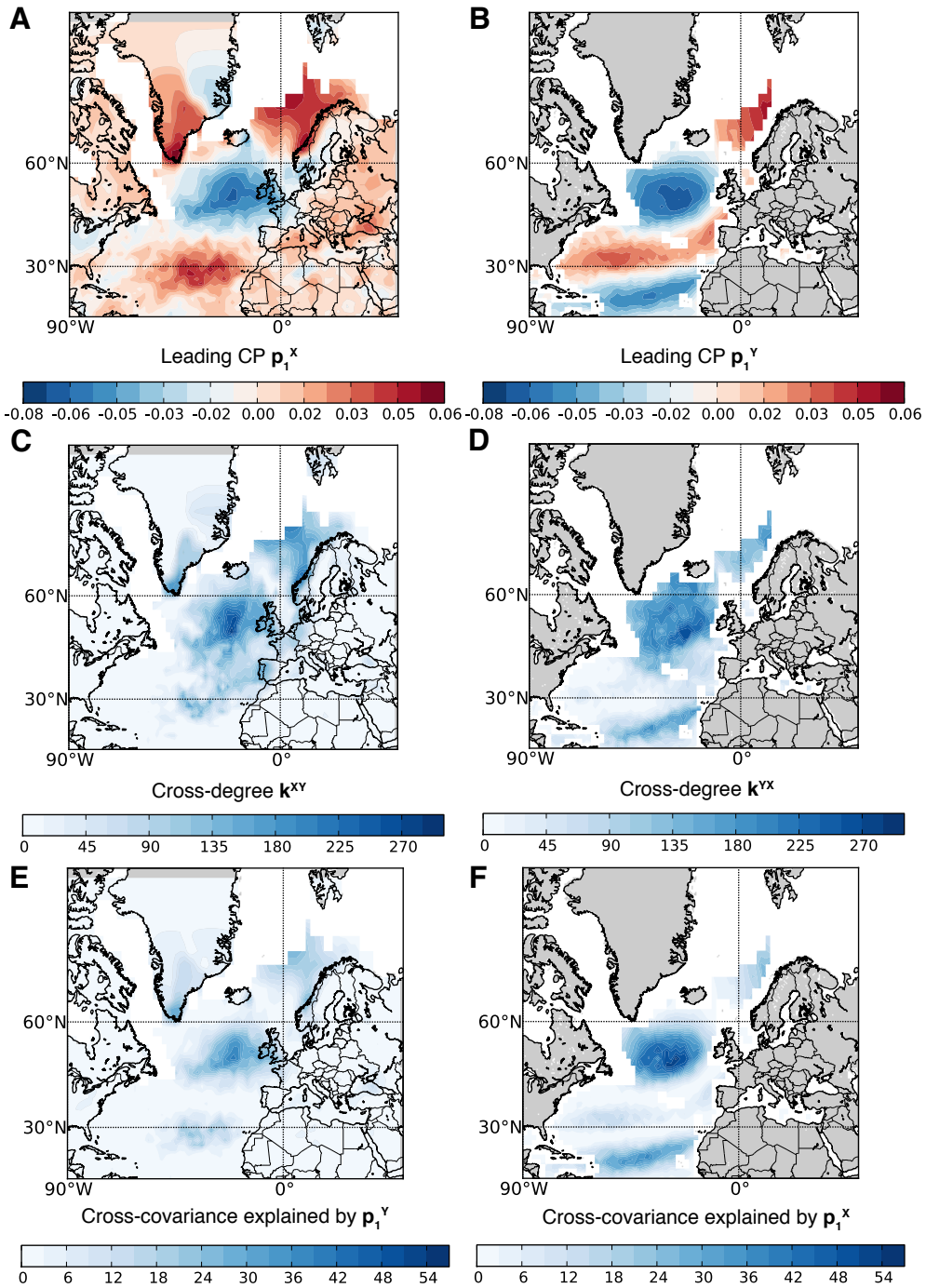

Fig. 7 Maps of leading pair of coupled patterns (A) $\mathbf{p}_1^X$ and (B) $\mathbf{p}_1^Y$, coupled climate network cross-degree fields (C) $\mathbf{k}^{XY}$ and (D) $\mathbf{k}^{YX}$, and percentage of cross-covariance explained by first pair of coupled patterns (E) $\mathbf{p}_1^Y$, $100 \times \mathrm{Corr}(x_i(t), a_1^Y(t))^2$, and (F) $\mathbf{p}_1^X$, $100 \times \mathrm{Corr}(y_i(t), a_1^X(t))^2$ (heterogeneous correlation maps, see Björnsson and Venegas (1997)), for the HOAPS-3 / GPCC precipitation (X) and HOAPS-3 evaporation (Y) data sets over the North Atlantic. For constructing the coupled climate network, a threshold $T = 0.47$ was chosen to yield a cross-link density of $\rho_{XY} = 0.01$ (Eq. (31)) resulting in internal link densities $\rho_X = 0.01$ and $\rho_Y = 0.06$ (Donges et al 2011c).

measures developed for investigating the topology of networks of interacting networks (Donges et al 2011c). The *cross-degree* $\mathbf{k}^{XY} = \{k_i^{XY}\}_{i=1}^{N_X}$ is the number of neighbors of node $i \in V_X$ in subnetwork G_Y:

$$k_i^{XY} = \sum_{j \in V_Y} A_{ij} = \sum_{j=1}^{N_Y} A_{ij}^{XY} = \sum_{j=1}^{N_Y} \Theta(|C_{ij}^{XY}| - T). \quad (17)$$

Analogously, the cross-degree $\mathbf{k}^{YX} = \{k_j^{YX}\}_{j=1}^{N_Y}$ is given by

$$k_j^{YX} = \sum_{i \in V_X} A_{ij} = \sum_{i=1}^{N_X} A_{ij}^{XY} = \sum_{i=1}^{N_X} \Theta(|C_{ij}^{XY}| - T). \quad (18)$$

Similarly to degree in univariate climate networks, regions i in field $\mathbf{x}(t)$ with a large cross-degree k_i^{XY} are considered to be strongly dynamically interrelated with many locations in field $\mathbf{y}(t)$ and vice versa. For the precipitation and evaporation data sets (Fig. 7C,D), such regions with high cross-connectivity correspond to major covariability areas of evaporation and precipitation fields driven by the North-Atlantic Oscillation (NAO) (Andersson et al 2010b; Petrova 2012).

Furthermore, analogously to univariate climate networks, generalizations of path-based measures for network of networks can be derived (Donges et al 2011c). Here, cross-closeness and cross-betweenness are considered. *Cross-closeness* $\mathbf{c}^{XY} = \{c_i^{XY}\}_{i=1}^{N_X}$ (cross-CC) measures the inverse mean network distance of node $i \in V_X$ to all nodes $j \in V_Y$ via shortest paths and is defined as

$$c_i^{XY} = \frac{N_X + N_Y - 1}{\sum_{j \in V_Y} l_{ij}}. \quad (19)$$

Cross-betweenness $\mathbf{b}^{XY} = \{b_i^{XY}\}_{i=1}^{N_X}$ (cross-BC) counts the relative number of shortest paths connecting any pair of nodes $j \in V_X, k \in V_Y$ that include node $i \in V_X$ and is defined as

$$b_i^{XY} = \sum_{j \in V_X} \sum_{k \in V_Y} \frac{n_{jk}(i)}{n_{jk}}. \quad (20)$$

For nodes j in field $\mathbf{y}(t)$, the measures $\mathbf{c}^{YX} = \{c_j^{YX}\}_{j=1}^{N_Y}$ and $\mathbf{b}^{YX} = \{b_j^{YX}\}_{j=1}^{N_Y}$ are obtained from analogous expressions following Donges et al (2011c). Interpretations of coupled CN measures will be discussed in Section 6.

5 Relationships between eigen and climate network analysis

Comparing the results of eigen and CN analysis, notable similarities become apparent, e.g., in the leading EOF $\mathbf{u}_1$ and CN degree $\mathbf{k}$ for the HOAPS-3 / GPCC precipitation data (Fig. 5). Analogous relations are observed when inspecting leading coupled patterns and coupled CN cross-degree for HOAPS-3 / GPCC precipitation and HOAPS-3 evaporation data (Fig. 7). To explain these similarities, in this section, formal relationships between patterns from eigen and CN analysis are derived and illustrated empirically for global precipitation and evaporation data sets. Relations between single field (EOFs and univariate CN measures, Section 5.1) as well as multiple field patterns (coupled patterns and coupled CN measures, Section 5.2), and temporal patterns are discussed. Note that similar relationships hold when both eigen and network analysis are based on a type of symmetric similarity matrix that is different from linear correlation at zero lag, *e.g.*, considering mutual information (Donges et al 2009a,b) or the ISOMAP algorithm (Tenenbaum et al 2000; Gámez et al 2004).

5.1 Single field patterns

As the correlation matrix $\mathbf{C}^X$ is symmetric and, hence, diagonalizable, it can be decomposed with respect to its eigensystem such that

$$C_{ij}^X = \sum_{k=1}^{R} u_{ik} \lambda_k u_{jk}. \quad (21)$$

If the leading EOF $\mathbf{u}_1$ explains a large fraction of the total variance, *i.e.*, if $\lambda_1 \gg \lambda_2$, then C_{ij}^X can be approximated as

$$C_{ij}^X \approx \lambda_1 u_{i1} u_{j1}. \quad (22)$$

Inserting this expression into the definition of CN degree (Eq. (13)) yields

$$k_i \approx \sum_{j=1}^{N} \Theta(\lambda_1 |u_{i1} u_{j1}| - T) - 1. \quad (23)$$

This approximation explains the empirically observed similarity between degree $\mathbf{k}$ and the leading EOF $\mathbf{u}_1$ (compare Fig. 5, panels A and B, for the precipitation data set) in the following way: All nodes j with $|u_{j1}| > \frac{T}{\lambda_1 |u_{i1}|}$ contribute to the degree k_i at node i, hence, a larger $|u_{i1}|$ typically leads to more positive contributions to the sum in Eq. (23) and, therefore, to a larger degree k_i. Consequently, CN degree $\mathbf{k}$

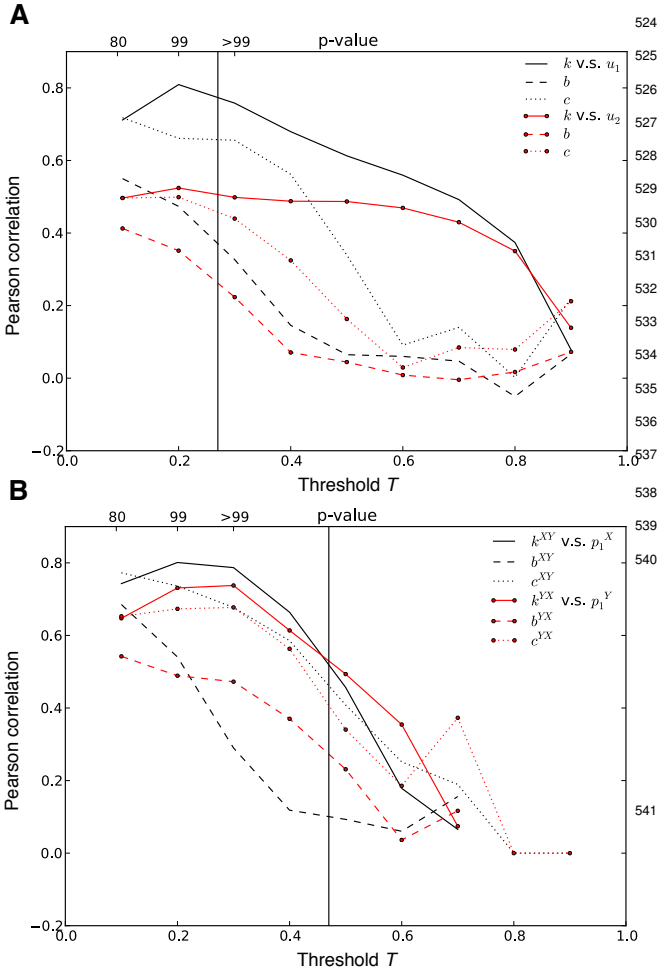

Fig. 8 Linear correlations between spatial patterns from eigen and network techniques for climate data analysis. Pearson correlation between (A) the absolute values of the first two EOFs $|\mathbf{u}_1|, |\mathbf{u}_2|$ and CN measures degree $\mathbf{k}$, closeness $\mathbf{c}$ and betweenness $\mathbf{b}$ for HOAPS-3 / GPCC precipitation data as well as (B) the first coupled patterns $\mathbf{p}_1^X, \mathbf{p}_1^Y$ and coupled CN measures cross-degree $\mathbf{k}^{XY}, \mathbf{k}^{YX}$, cross-closeness $\mathbf{c}^{XY}, \mathbf{c}^{YX}$, and cross-betweenness $\mathbf{b}^{XY}, \mathbf{b}^{YX}$ for HOAPS-3 / GPCC precipitation (X) and HOAPS-3 evaporation data. In both panels, correlations are displayed for varying network construction threshold T, where the corresponding p-value according to the Student's t-test is given on the upper horizontal axis. Vertical lines in panels (A) and (B) indicate the thresholds used in Figs. 5 and 7, respectively.

and the vector of absolute values of the leading EOF's elements $|\mathbf{u}_1|$ are expected to be positively correlated.

For the global precipitation data set, a large positive correlation between $\mathbf{k}$ and $|\mathbf{u}_1|$ is indeed detected for intermediate thresholds T of the order where CNs are typically constructed (Donges et al 2009b), while for smaller and larger thresholds, the correlation decreases (Fig. 8A). The latter is expected, since both for $T \to 0$ (fully connected network) and $T \to 1$ (network devoid of links), the CN contains no information about the climatological field anymore and the degree field is constant with $k_i \to N-1$ and $k_i \to 0$ for all nodes i, respectively. Hence, maximum pattern correspondence is expected for intermediate thresholds T (for these as well as computational reasons, results for $T = 0$ and $T = 1$ are not included in Fig. 8). Notably, selecting T as maximizing the correlation between degree $\mathbf{k}$ and the leading EOF $|\mathbf{u}_1|$ could provide a criterion for an informed choice of the threshold T. Such a choice would approximate a situation where the information that the CN contains on linear statistical interdependencies in the field of interest is maximized. Further work is needed to develop more suitable criteria for defining binary CNs with maximum information content. Furthermore and as expected, the correlation between degree $\mathbf{k}$ and the second EOF $|\mathbf{u}_2|$ is mostly smaller than that between degree and leading EOF (Fig. 8A).

Using the full eigen-decomposition of $\mathbf{C}^X$, an exact relationship between the degree $\mathbf{k}$ and all EOFs $\mathbf{u}_k$ together with their associated eigenvalues λ_k can be derived as

$$k_i = \sum_{j=1}^{N} \Theta\left(\left| \sum_{k=1}^{R} u_{ik} \lambda_k u_{jk} \right| - T \right) - 1. \tag{24}$$

Using this expression, the scalar link density

$$\rho = \frac{\langle k_i \rangle_{i=1}^{N}}{N-1} \tag{25}$$

can likewise be expanded or approximated, where $\langle \cdot \rangle$ denotes the arithmetic mean. Similarly, a relationship between area-weighted EOFs (Hannachi et al 2007), the area-weighted degree (Heitzig et al 2012) (also called area weighted connectivity (Tsonis et al 2006)) and all other network measures directly expressible in terms of the adjacency matrix A_{ij} can be derived.

5.2 Coupled patterns

The cross-correlation matrix $\mathbf{C}^{XY}$ can be decomposed in terms of singular values and coupled patterns as (Fig. 3)

$$C_{ij}^{XY} = \sum_{k=1}^{R} \sigma_k p_{ik}^X p_{jk}^Y. \tag{26}$$

The relationship between cross-degree $\mathbf{k}^{XY}, \mathbf{k}^{YX}$ and coupled patterns $\mathbf{p}_k^X, \mathbf{p}_k^Y$ can then be derived as above:

$$k_i^{XY} = \sum_{j=1}^{N_Y} \Theta\left(\left|\sum_{k=1}^{R} \sigma_k p_{ik}^X p_{jk}^Y\right| - T\right) \quad (27)$$

$$\approx \sum_{j=1}^{N_Y} \Theta\left(\sigma_1 \left|p_{i1}^X p_{j1}^Y\right| - T\right), \quad (28)$$

$$k_j^{YX} = \sum_{i=1}^{N_X} \Theta\left(\left|\sum_{k=1}^{R} \sigma_k p_{ik}^X p_{jk}^Y\right| - T\right) \quad (29)$$

$$\approx \sum_{i=1}^{N_X} \Theta\left(\sigma_1 \left|p_{i1}^X p_{j1}^Y\right| - T\right). \quad (30)$$

The approximations hold for the maximum singular value fulfilling $\sigma_1 \gg \sigma_2 \geq \cdots \geq \sigma_R$. R is the rank of the cross-correlation matrix $\mathbf{C}^{XY}$. By a similar argument as given above this shows that $\mathbf{k}^{XY}$ and $|\mathbf{p}_1^X|$ ($\mathbf{k}^{YX}$ and $|\mathbf{p}_1^Y|$) are expected to be positively correlated which is consistent with our results regarding the interdependency structure between precipitation and evaporation fields. While in our example, the correspondence between the resulting patterns is somewhat less pronounced than in the single-field setting (Fig. 8B), still regions with a strongly negative loading in the leading coupled patterns $\mathbf{p}_1^X$ and $\mathbf{p}_1^Y$ appear as super nodal structures in the cross-degree fields (Fig. 7). When studying varying network construction thresholds T, as in the case of single-field patterns, the correlation between the absolute values of the leading pair of coupled patterns and cross-degree fields is maximum for intermediate T and decreases for $T \to 0$ and $T \to 1$ (Fig. 8B). Also, consistently with Eqs. (27) and (29), the correlation between the second pair of coupled patterns and cross-degree fields is always smaller than that observed for the leading pair of coupled patterns (results not shown).

The scalar cross-link densities (Donges et al 2011c)

$$\rho_{XY} = \frac{\left\langle k_i^{XY} \right\rangle_{i=1}^{N_X}}{N_Y}$$

$$\rho_{YX} = \frac{\left\langle k_j^{YX} \right\rangle_{j=1}^{N_Y}}{N_X} \quad (31)$$

can also be expanded and approximated in terms of CPs and singular values using the above expressions. Analogously, area-weighted coupled patterns (von Storch and Zwiers 2003) are related to the area-weighted cross-degree introduced by Feng et al (2012) and Wiedermann et al (2013).

5.3 Temporal patterns

In EOF analysis, temporal patterns (principal components) $a_k(t)$ describing the evolution of their associated spatial patterns $\mathbf{u}_k$ are easily obtained by projecting the data $\mathbf{x}(t)$ onto the latter patterns $\mathbf{u}_k$ (Eq. (4)). Analogously, the same holds for multivariate extensions such as coupled pattern analysis (Bretherton et al 1992; von Storch and Zwiers 2003), see Section 3. In CN analysis, however, the temporal evolution of spatial network measure patterns such as the degree $\mathbf{k}$ or betweenness $\mathbf{b}$ cannot be directly obtained from the adjacency matrix $\mathbf{A}$ and $\mathbf{x}(t)$. To allow the study of non-stationarities in the statistical interdependence structure of climatological fields, several authors have investigated the evolving local (*e.g.*, $\mathbf{k}(t)$ or $\mathbf{b}$(t)) and global properties of CNs $\mathbf{A}(t)$ constructed from temporal windows sliding over the time series data (Gozolchiani et al 2008; Yamasaki et al 2008, 2009; Gozolchiani et al 2011; Guez et al 2012; Berezin et al 2012; Carpi et al 2012; Martin et al 2013; Radebach et al 2013; Ludescher et al 2013, 2014). A similar strategy could be applied to coupled CN analysis.

It should be noted that unlike in the above sections, no direct relationship can be derived linking temporal patterns from eigen and network analysis. The reason for this is twofold. First, temporal patterns $a_k(t)$ of standard EOF analysis are based on the full data set $\mathbf{x}(t)$, while the evolving spatial network patterns are computed from subsets (defined by temporal windows) of $\mathbf{x}(t)$. Second, since temporal patterns $a_k(t)$ of eigenanalysis are merely scalar prefactors in the expansion Eq. (3) (see Figs. 1 and 3), the spatial EOF patterns $\mathbf{u}_k$ are time-independent, whereas evolving CN measures such as $\mathbf{k}(t)$ can vary independently at every location i. Hence, in contrast to standard EOF patterns, the spatial patterns in the network properties derived from evolving CNs are explicitly time-dependent. The latter case is analogous to extended EOF analysis, where standard EOF analysis is applied in a sliding-window mode as well (Fraedrich et al 1997).

6 Discussion

The relationships derived in the previous section provide guidance on deciding how and in which applications CN analysis can be expected to yield information that is complementary to the results of eigenanalysis. Particularly, we will focus on a discussion and climatological interpretation of single field and coupled patterns derived from precipitation and evaporation data (Section 6.1) and relate this to a study of single field patterns for global surface air temperature data (Section 6.2). Based on these insights, we point out some methodological as well as practical potentials of CN analysis of climatological fields (Section 6.3).

6.1 Precipitation and evaporation data

For the HOAPS-3 / GPCC precipitation and HOAPS-3 evaporation data sets, pronounced similarities between the fea-

tures observed in the degree or cross-degree fields and those in the leading EOF or coupled patterns that are derived from the same data have been described and explained mathematically (Section 5). More specifically, active regions displaying strong correlations with many other locations, and, hence, a large degree or cross-degree (termed super-nodes in the context of CN analysis (Tsonis and Roebber 2004; Tsonis et al 2006; Barreiro et al 2011)) correspond to regions with large positive or negative loading in the leading EOF or coupled patterns. For example, this can be observed for the equatorial Pacific in the precipitation data (Fig. 5A,B). The spatial similarity between the amplitude of the leading EOF and CN degree field reveals the well-known ENSO variability pattern (Ropelewski and Halpert 1987). Particularly, the patterns in the explained variance fraction (Fig. 5C) closely resemble high connectivity areas of the CN resembling most prominent ENSO teleconnections (Andersson et al 2010b; Halpert and Ropelewski 1992; Ropelewski and Halpert 1987). Additional dipole information described by the EOF is typically preserved by neighbors of the network's major super-nodes (not shown here, see Petrova (2012) and Kawale et al (2013)).

Considering the bivariate analysis of precipitation and evaporation data over North Atlantic (Fig. 7), regions with a strongly negative loading in the leading pair of coupled patterns appear as super nodal structures in the cross-degree fields obtained from coupled CN analysis. Areas with a high fraction of explained cross-covariance (Fig. 7E,F) well correspond to the coupled network topology as indicated by the cross-degree fields (Fig. 7C,D) and all together depict major covariability areas of evaporation and precipitation driven by the NAO. The cross-degree field $\mathbf{k}^{XY}$ (Fig. 7C), displaying the number of strong correlations between precipitation variability at a certain location with evaporation dynamics at all other grid points, reveals teleconnections associated to the NAO over the southern tip of Greenland as well as a positive NAO signal over Portugal and a negative NAO signal over Norway (Andersson et al 2010b). In turn, the cross-degree field $\mathbf{k}^{YX}$ (Fig. 7D), showing the number of strong correlations between evaporation dynamics at one point and precipitation variability at all other locations, is only available over the ocean and follows the covariance structure of the main evaporation determinant parameters with NAO (Cayan 1992; Marshall et al 2001).

Beyond the frequently studied degree $\mathbf{k}$, complex network theory provides a wealth of additional measures that can be used to study higher-order properties of the statistical interdependency structure within and between climatological fields. For example, the mentioned measures based on the properties of shortest paths in (coupled) CNs such as (cross-) closeness $\mathbf{c}$ ($\mathbf{c}^{XY}, \mathbf{c}^{YX}$) and (cross-) betweenness $\mathbf{b}$ ($\mathbf{b}^{XY}, \mathbf{b}^{YX}$) (Fig. 9) have been argued to give insights on the local speed of propagation as well as the preferred pathways for the spread of perturbations within or between the studied fields, respectively (Donges et al 2009a,b, 2011c; Malik et al 2012; Molkenthin et al 2014a). In this way, CN analysis has the potential to unveil information on climate dynamics from climatological field data that conceptually supplements the results of eigenanalysis.

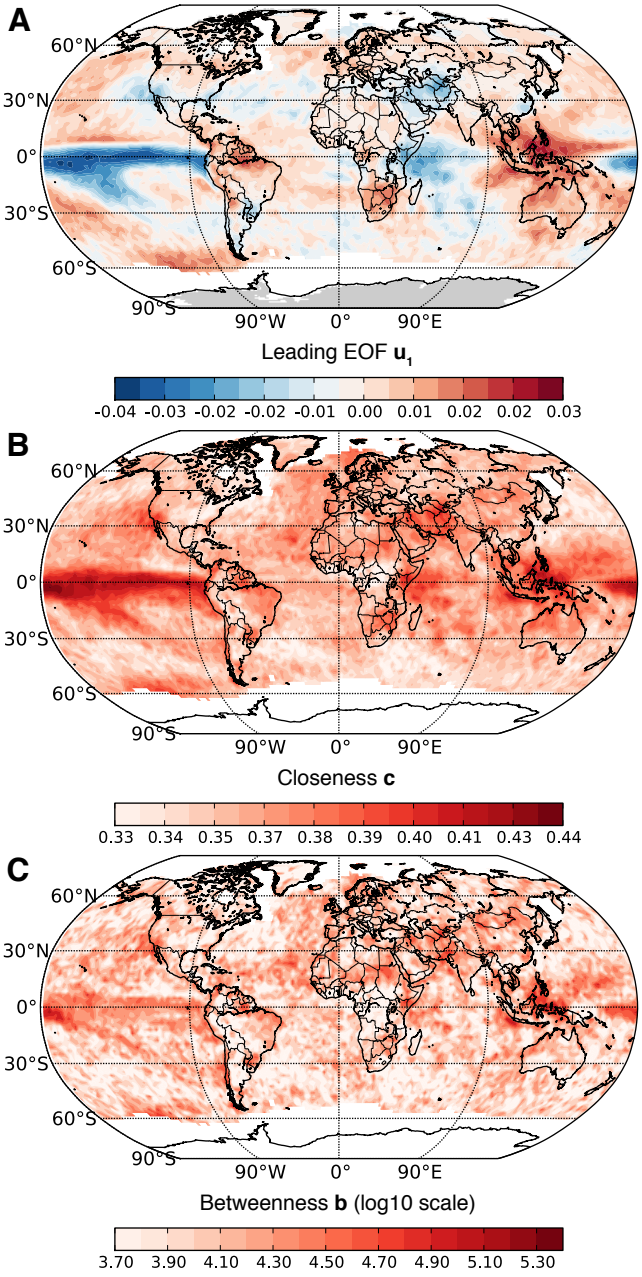

Fig. 9 Maps of (A) leading EOF $\mathbf{u}_1$, (B) closeness field $\mathbf{c}$, and (C) betweenness field $\mathbf{b}$ for the global HOAPS-3 / GPCC precipitation climate network. The network construction threshold $T = 0.27$ was chosen to yield a link density of $\rho = 0.01$.

Focussing on the precipitation data to further investigate this aspect, we find that the correlation of CC and BC to the first two EOFs obtained from the data are systematically and

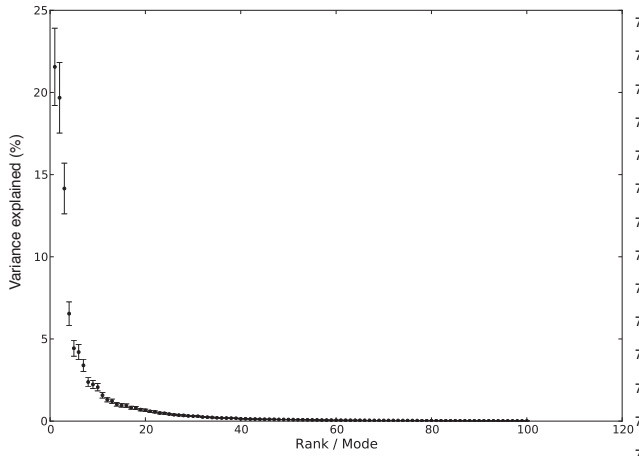

Fig. 10 Percentage of variance $\lambda_k / \sum_{l=1}^{R} \lambda_l$ explained by EOFs $\mathbf{u}_k$ for the NCEP/NCAR surface air temperature data set. Error bars were estimated using North's rule of thumb (North et al 1982).

significantly smaller than that between the degree field and the same EOFs (Fig. 8A). Similarly, in the bivariate case, the correlations of cross-CC and cross-BC with the leading coupled pattern are considerably smaller than those between the latter and cross-degree for most thresholds T (Fig. 8B). However, for the HOAPS-3 / GPCC precipitation data, the patterns observed in the leading EOF resemble those found in the CC and BC fields (Fig. 9) as well as those in the degree field (Fig. 5). These results can be explained from a network point of view by considering that precipitation fields are typically only correlated on short spatial scales and display a smaller degree of spatial coherency when compared to other atmospheric variables such as pressure or temperature (Feldhoff et al 2014). In turn, this leads to a larger degree of randomness in the structure of CNs constructed from this data. In random networks, correlations between centrality measures such as degree, closeness and betweenness arise (Boccaletti et al 2006). In other words, spatially incoherent climatological fields can give rise to CNs with a notable degree of disorder in the placement of links between different nodes which induces correlations between network centrality measures. For the precipitation data set at hand, the first eigenvalue separates from the remaining spectrum (Fig. 2) leading to a pronounced correlation between the leading EOF $\mathbf{u}_1$ and the degree field (see Eq. 23), and, hence, to correlations between $\mathbf{u}_1$ and CC, BC.

6.2 Surface air temperature data

Next, we investigate the NCEP/NCAR reanalysis I surface air temperature (SAT) field as another frequently studied data set. The properties of this data are complementary to those of the precipitation field discussed above in two aspects: (i) for the SAT data, the leading two EOFs explain approximately the same amount of variance (Fig. 10), while the leading eigenvalue separates more markedly from the remainder of the spectrum in the case of the precipitation data (Fig. 2), and, (ii) the SAT field is known to display a stronger degree of spatial coherency than the precipitation field. In the light of the discussion in Section 6.1, these two properties are reflected when comparing the leading three EOFs and network properties for the SAT data set (Fig. 11). Firstly, the degree field resembles the leading EOF less than in case of precipitation data (Fig. 11A,D), which is expected due to the weaker separation of the leading eigenvalues (Section 5.1 and Eq. 23). Consistently, the degree field displays an even less pronounced similarity to the second and third EOFs (Fig. 11B,C,D). While the patterns found in the CC field (Fig. 11E) still partly resembles those in the degree field (Fig. 11D) as well as those in the leading two EOFs (Fig. 11A,B), the BC field displays markedly distinct features (Fig. 11F). Only in a few regions, these structures of high betweenness appear to coincide with patterns of large EOF loadings, e.g., high betweenness structures found along the West coasts of North and South America correspond to large positive loadings in the second and third EOFs, respectively.

The observed linear wave-like structures of large BC in the SAT field have been interpreted as signatures of the transport of temperature anomalies in strong surface ocean currents (Donges et al 2009a,b). For example, the large betweenness structures resemble strong western boundary currents such as the Kuroshio of the east coast of Japan or Eastern boundary currents such the Canary current off the African west coast. It should be noted that while some of the structures in the BC field such as the one resembling the North Atlantic's subtropical gyre appear blurred, the logarithmic color scale in Fig. 11F implies that even small changes in color correspond to exponentially large changes in BC. This interpretation of high betweenness structures in CNs constructed based on Pearson correlation as advective structures such as strong currents is supported by recent analytical studies that are based on well-known fluid dynamical model systems (Molkenthin et al 2014a,b). Further evidence that is also consistent with this interpretation of betweenness was found in a study of vertical interactions in the atmospheric geopotential height field, where regions of large cross-BC in the Arctic suggest that vertical air induced by the Arctic vortex is important for mediating the propagation of wind field anomalies between different isobaric surfaces (Donges et al 2011c). Also, Boers et al (2013) employ BC and further network measures for precipitation data over South America to highlight the importance of atmospheric structures such as the South American low level jet for the propagation of extreme rainfall events, specifically over long distances.

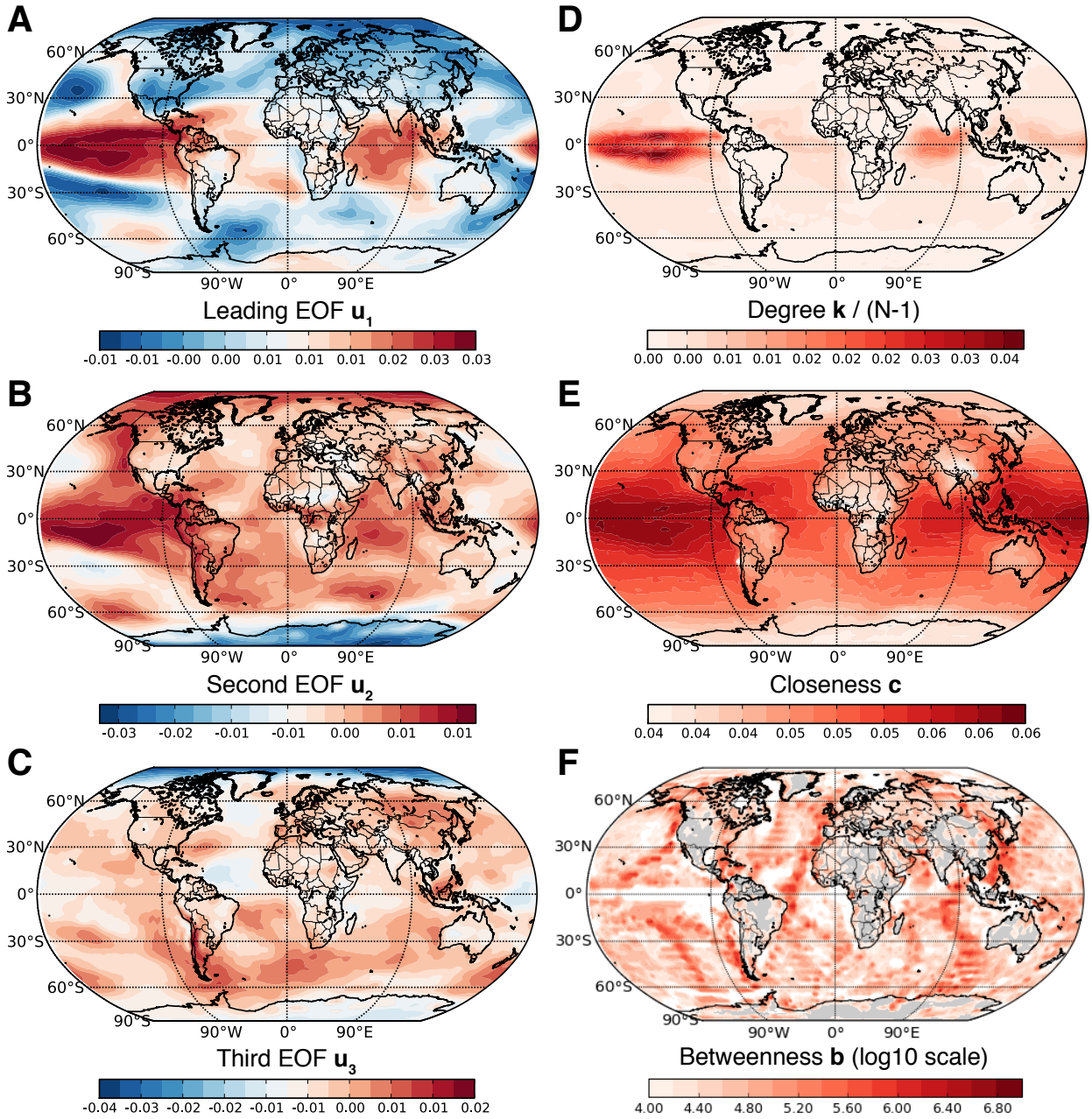

Fig. 11 Maps of (A,B,C) the leading three EOFs $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, (D) normalized degree field $\mathbf{k}/(N-1)$, (E) closeness field $\mathbf{c}$, and (F) betweenness field $\mathbf{b}$ for the global NCEP/NCAR surface air temperature climate network. The network construction threshold $T = 0.67$ was chosen to yield a link density of $\rho = 0.01$. In panel (F), gray shading indicates regions with betweenness values smaller than 10^4.

6.3 Potentials of climate network analysis

The examples discussed above suggest that CN analysis may be particularly useful in situations where (i) a dominant EOF (pair of coupled patterns) explaining significantly more variance (cross-covariance) in the data than further modes does not exist and (ii) the climatological field of interest displays a certain degree of spatial coherence reflecting, e.g., winds, ocean currents or long-range teleconnections. Such rules could be useful in practice when deciding on which methodology should be applied to a data set of interest. While future research beyond the scope of this work is needed to address these suggestions, we move on to discuss the potentials of CN analysis from a methodological point of view.

Considering higher-order network properties, approximate and exact relationships akin to Eqs. (23) and (24) can be derived for other (coupled) CN measures of interest like

the local clustering coefficient (Donges et al 2009b; Malik et al 2012)

$$\mathcal{C}_i = \frac{\sum_{j,k=1}^{N} A_{ij} A_{jk} A_{ki}}{\sum_{j,k=1}^{N} A_{ij} A_{ik}} \tag{32}$$

by plugging in the approximation $A_{ij} \approx \Theta(|\lambda_1 u_{i1} u_{j1}| - T) - \delta_{ij}$ or the full expansion of A_{ij} in terms of EOFs (Section 5.1). However, the resulting lengthy expressions, particularly for path-based network measures such as CC and BC (Heitzig et al 2012), hardly help to gain further understanding other than that both eigen and network approaches are based on the same underlying similarity matrix (Figs. 1 and 3). In contrast, taking the local clustering coefficient as an example illustrates the added value of the complex network point of view: Eq. (32) can be easily understood as a local measure for transitivity in the correlation structure of a climatological field (Donges et al 2009b, 2011c), while the same measure viewed as some function of all EOFs $\mathbf{u}_k$ would be considered hard to interpret or meaningless in terms of eigenanalysis alone. In that sense, the network approach allows insights into the correlation structure of climatological fields that go beyond and complement those obtainable by EOF analysis.

It has been shown in earlier studies that the statistical information provided by CN analysis is valuable for complementing standard techniques of eigenanalysis for tasks like model tuning, model validation (Feldhoff et al 2014), model and model-data intercomparison (Petrova 2012; Steinhaeuser and Tsonis in press; Fountalis et al 2013; Feldhoff et al 2014), statistical forecasting (Steinhaeuser et al 2011), and explorative data analysis (Steinhaeuser et al 2010, 2012). Furthermore, the network approach allows to employ advanced algorithms for pattern recognition (Kawale et al 2013), spatial coarse-graining (Fountalis et al 2013) or community detection (Tsonis et al 2011; Steinhaeuser et al 2011; Steinhaeuser and Tsonis 2014). Recently, a series of studies based on well-defined fluid-dynamical model systems has provided deeper insights into the structure of CNs, particularly into how the latter is related to the dynamics of the underlying physical system, as well as fostered the interpretation of CN measures (Molkenthin et al 2014a,b; Tupikina et al 2014).

A particular advantage of CN analysis is that statistical methods originating from information and dynamical systems theory such as transfer entropy (Runge et al 2012a,b), probabilistic graphical models (Ebert-Uphoff and Deng 2012a,b), or event synchronization (Malik et al 2012) can be naturally used for network construction, and, hence, for identifying processes and patterns which are not accessible when studying linear correlation matrices alone. Applying these modern methods of time series analysis for network construction allows, among other applications, to study the synchronization of climatic extreme events (Malik et al 2012; Boers et al 2013, 2014b) or to suppress the misleading effects of auto-dependencies in time series, common drivers and indirect couplings by reconstructing causal interactions (in the statistical sense of information theory) between climatic sub-processes (Ebert-Uphoff and Deng 2012a; Runge et al 2012a,b, 2014). This in turn enables a more direct interpretation of the reconstructed network structures and resulting patterns in network structures in terms of climatic sub-processes and their interactions, avoiding the conceptual problems that arise in the interpretation of results from purely correlation-based techniques such as classical EOF or CP analysis / MCA (Dommenget and Latif 2002; Jolliffe 2003; Monahan et al 2009).

7 Conclusions

In summary, the main aim of this article has been to put the recently developed CN approach into context with standard eigenanalysis of climatological data, since both classes of methods are usually based on the same set of statistical similarity matrices, *i.e.*, the linear correlation and cross-correlation matrices at zero lag. We have derived formal relationships between empirical orthogonal functions or coupled patterns and frequently used first-order CN measures such as degree or cross-degree, respectively. These relations have been illustrated empirically using global satellite observations of precipitation and evaporation fields as well as reanalysis data for the global surface air temperature field. However, it has been shown that, and in which specific practical settings, higher-order CN measures such as closeness and betweenness may contain complementary statistical information with respect to classical eigenanalysis. We have argued that this information could be valuable for tasks such as model tuning, validation, and intercomparison as well as for improving statistical predictions of climate variability and explorative data analysis. Hence, by transferring insights and tools from complex network theory and complexity science to climate research, CNs meet the need for novel techniques of climate data analysis facing quickly increasing data volumes generated by growing observational networks and model intercomparison exercises like the coupled model intercomparison project (CMIP) (Taylor et al 2012). Furthermore, the application of advanced network-theoretical concepts and methods from fields like complexity science, information theory and machine learning promises novel and deep insights into Earth system dynamics, particularly considering the complex interactions of human societies with global climatic and biogeochemical processes.

Acknowledgements This work has been financially supported by the Leibniz association (project ECONS), the German National Academic Foundation, the Potsdam Institute for Climate Impact Research, the Stordalen Foundation, BMBF (project GLUES), the Max Planck Soci-

ety, and DFG grants KU34-1 and MA 4759/4-1. For climate network analysis, the software package `pyunicorn` was used that is available at `http://tocsy.pik-potsdam.de/pyunicorn.php` (Donges et al 2013). We thank Reik V. Donner and Doerthe Handorf for discussions and comments on an earlier version of the manuscript.

References

Andersson A, Bakan S, Graßl H (2010a) Satellite derived North Atlantic precipitation variability and its dependence on the NAO index. Tellus A 62(4):453–468, doi:10.1111/j.1600-0870.2010.00458.x

Andersson A, Fennig K, Klepp C, Bakan S, Graßl H, Schulz J (2010b) The Hamburg ocean atmosphere parameters and fluxes from satellite data – HOAPS-3. Earth Syst Sci Data 2:215–234, doi:10.5194/essd-2-215-2010

Andersson A, Klepp C, Fennig K, Bakan S, Grassl H, Schulz J (2011) Evaluation of HOAPS-3 ocean surface freshwater flux components. J Appl Meteor Climatol 50(2):379–398, doi:10.1175/2010JAMC2341.1

Barreiro M, Marti AC, Masoller C (2011) Inferring long memory processes in the climate network via ordinal pattern analysis. Chaos 21(1):13,101, doi:10.1063/1.3545273

Berezin Y, Gozolchiani A, Guez O, Havlin S (2012) Stability of climate networks with time. Sci Rep 2:666, doi:10.1038/srep00666

Björnsson H, Venegas SA (1997) A manual for EOF and SVD analysis of climatic data. Tech. Rep. C2GCR report No. 97-1, Department of Atmospheric and Oceanic Sciences, Centre for Climate and Global Change Research, McGill University

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: Structure and dynamics. Phys Rep 424(4-5):175–308, doi:10.1016/j.physrep.2005.10.009

Boccaletti S, Bianconi G, Criado R, Del Genio C, Gómez-Gardeñes J, Romance M, Sendina-Nadal I, Wang Z, Zanin M (2014) The structure and dynamics of multilayer networks. Physics Reports doi:10.1016/j.physrep.2014.07.001

Boers N, Bookhagen B, Marwan N, Kurths J, Marengo J (2013) Complex networks identify spatial patterns of extreme rainfall events of the South American monsoon system. Geophys Res Lett 40(16):4386–4392

Boers N, Bookhagen B, Barbosa H, Marwan N, Kurths J, Marengo J (2014a) Prediction of extreme floods in the Eastern Central Andes based on a complex networks approach. Nat Comm 5: 5199, doi:10.1038/ncomms6199

Boers N, Donner RV, Bookhagen B, Kurths J (2014b) Complex network analysis helps to identify impacts of the El Niño Southern Oscillation on moisture divergence in South America. Clim Dynam (online first) pp 1–14, doi:10.1007/s00382-014-2265-7

Bretherton CS, Smith C, Wallace JM (1992) An intercomparison of methods for finding coupled patterns in climate data. J Climate 5(6):541–560, doi:10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2

Brunet G, Vautard R (1996) Empirical normal modes versus empirical orthogonal functions for statistical prediction. J Atmos Sci 53(23):3468–3489, doi:10.1175/1520-0469(1996)053<3468:ENMVEO>2.0.CO;2

Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. Nature 464(7291):1025–1028, doi:10.1038/nature08932

Bullmore E, Sporns O (2009) Complex brain networks: Graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186–198, doi:10.1038/nrn2575

Carpi LC, Saco PM, Rosso OA, Ravetti MG (2012) Structural evolution of the tropical Pacific climate network. Eur Phys J B 85(11):1–7, doi:10.1140/epjb/e2012-30413-7

Cayan DR (1992) Latent and sensible heat flux anomalies over the northern oceans: The connection to monthly atmospheric circulation. J Climate 5(4):354–369, doi:10.1175/1520-0442(1992)005<0354:LASHFA>2.0.CO;2

Cohen R, Havlin S (2010) Complex networks: Structure, robustness and function. Cambridge University Press, Cambridge

Davidsen J, Grassberger P, Paczuski M (2008) Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity. Phys Rev E 77(6):066,104, doi:10.1103/PhysRevE.77.066104

Dommenget D, Latif M (2002) A cautionary note on the interpretation of EOFs. J Climate 15(2):216–225

Donges JF, Zou Y, Marwan N, Kurths J (2009a) The backbone of the climate network. Europhys Lett 87(4):48,007, doi:10.1209/0295-5075/87/48007

Donges JF, Zou Y, Marwan N, Kurths J (2009b) Complex networks in climate dynamics. Eur Phys J Spec Top 174(1):157–179, doi:10.1140/epjst/e2009-01098-2

Donges JF, Donner RV, Rehfeld K, Marwan N, Trauth M, Kurths J (2011a) Identification of dynamical transitions in marine palaeoclimate records by recurrence network analysis. Nonlinear Proc Geophys 18(5):545–562, doi:10.5194/npg-18-545-2011

Donges JF, Donner RV, Trauth MH, Marwan N, Schellnhuber HJ, Kurths J (2011b) Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. Proc Natl Acad Sci USA 108(51):20,422–20,427, doi:10.1073/pnas.1117052108

Donges JF, Schultz HCH, Marwan N, Zou Y, Kurths J (2011c) Investigating the topology of interacting networks – Theory and application to coupled climate subnetworks. Eur Phys J B 84(4):635–652, doi:10.1140/epjb/e2011-10795-8

Donges JF, Heitzig J, Runge J, Schultz HC, Wiedermann M, Zech A, Feldhoff J, Rheinwalt A, Kutza H, Radebach A, et al (2013) Advanced functional network analysis in the geosciences: The pyunicorn package. Geophysical Research Abstracts 15:3558

Donner RV, Donges JF (2012) Visibility graph analysis of geophysical time series: Potentials and possible pitfalls. Acta Geophys 60(3):589–623, doi:10.2478/s11600-012-0032-x

Donner RV, Sakamoto T, Tanizuka N (2008) Complexity of spatio-temporal correlations in Japanese air temperature records. In: Donner R, Barbosa S (eds) Nonlinear time series analysis in the geosciences: Applications in climatology, geodynamics and solar-terrestrial physics, Lecture Notes in Earth Science, vol 112, Springer, Berlin, pp 125–154, doi:10.1007/978-3-540-78938-3_7

Donner RV, Zou Y, Donges JF, Marwan N, Kurths J (2010) Recurrence networks – A novel paradigm for nonlinear time series analysis. New J Phys 12(3):033,205, doi:10.1088/1367-2630/12/3/033025

Ebert-Uphoff I, Deng Y (2012a) Causal discovery for climate research using graphical models. J Climate 25:5648–5665, doi:10.1175/JCLI-D-11-00387.1

Ebert-Uphoff I, Deng Y (2012b) A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer. Geophys Res Lett 39:L19,701, doi:10.1029/2012GL053269

Feldhoff JH, Donner RV, Donges JF, Marwan N, Kurths J (2012) Geometric detection of coupling directions by means of inter-system recurrence networks. Phys Lett A 376:3504–3513, doi:10.1016/j.physleta.2012.10.008

Feldhoff JH, Lange S, Volkholz J, Donges JF, Kurths J, Gerstengarbe FW (2014) Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate. Clim Dynam (online first) doi:10.1007/s00382-014-2182-9

Feng A, Gong Z, Wang Q, Feng G (2012) Three-dimensional air–sea interactions investigated with bilayer networks. Theor Appl Climatol 109(3-4):635–643, doi:10.1007/s00704-012-0600-7

Fountalis I, Bracco A, Dovrolis C (2013) Spatio-temporal network analysis for studying climate patterns. Clim Dynam (online first), doi:10.1007/s00382-013-1729-5

Fraedrich K, McBride JL, Frank WM, Wang R (1997) Extended EOF analysis of tropical disturbances: TOGA COARE. J Atmos Sci 54(19):2363–2372, doi:10.1175/1520-0469(1997)054<2363:EEAOTD>2.0.CO;2

Fukuoka A (1951) A study of 10-day forecast (a synthetic report). Geophys Mag: Tokyo 22:177–208

Gámez AJ, Zhou CS, Timmermann A, Kurths J (2004) Nonlinear dimensionality reduction in climate data. Nonlinear Proc Geophys 11(3):393–398, doi:10.5194/npg-11-393-2004

Gao J, Buldyrev SV, Stanley HE, Havlin S (2011) Networks formed from interdependent networks. Nat Phys 8(1):40–48, doi:10.1038/NPHYS2180

Ghil M, Malanotte-Rizzoli P (1991) Data assimilation in meteorology and oceanography. Adv Geophys 33:141–266, doi:10.1016/S0065-2687(08)60442-2

Ghil M, Allen M, Dettinger M, Ide K, Kondrashov D, Mann M, Robertson AW, Saunders A, Tian Y, Varadi F, et al (2002) Advanced spectral methods for climatic time series. Rev Geophys 40(1):1–1 – 1–41, doi:10.1029/2000RG000092

Gozolchiani A, Yamasaki K, Gazit O, Havlin S (2008) Pattern of climate network blinking links follows El Niño events. Europhys Lett 83(2):28,005, doi:10.1209/0295-5075/83/28005

Gozolchiani A, Havlin S, Yamasaki K (2011) Emergence of El Niño as an autonomous component in the climate network. Phys Rev Lett 107(14):148,501, doi:10.1103/PhysRevLett.107.148501

Guez O, Gozolchiani A, Berezin Y, Brenner S, Havlin S (2012) Climate network structure evolves with North Atlantic Oscillation phases. Europhys Lett 98:38,006, doi:10.1209/0295-5075/98/38006

Halpert MS, Ropelewski CF (1992) Surface temperature patterns associated with the Southern Oscillation. J Climate 5(6):577–593, doi:10.1175/1520-0442(1992)005<0577:STPAWT>2.0.CO;2

Handorf D, Dethloff K (2009) Atmospheric teleconnections and flow regimes under future climate projections. Eur Phys J Spec Top 174:237–255, doi:10.1140/epjst/e2009-01104-9

Handorf D, Dethloff K (2012) How well do state-of-the-art atmosphere-ocean general circulation models reproduce atmospheric teleconnection patterns? Tellus A 64:19,777, doi:10.3402/tellusa.v64i0.19777

Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: a review. Int J Climatol 27:1119–1152, doi:10.1002/joc.1499

Heitzig J, Donges JF, Zou Y, Marwan N, Kurths J (2012) Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes. Eur Phys J B 85(1):38, doi:10.1140/epjb/e2011-20678-7

Hempel S, Koseska A, Kurths J, Nikoloski Z (2011) Inner composition alignment for inferring directed networks from short time series. Phys Rev Lett 107(5):54,101, doi:10.1103/PhysRevLett.107.054101

Hirata Y, Shimo Y, Tanaka HL, Aihara K (2011) Chaotic properties of the Arctic Oscillation Index. SOLA 7:33–36, doi:10.2151/sola.2011-009

Hlinka J, Hartman D, Vejmelka M, Novotná D, Paluš M (2014) Non-linear dependence and teleconnections in climate data: sources, relevance, nonstationarity. Climate Dynamics 42(7-8):1873–1886, doi:10.1007/s00382-013-1780-2

Hsieh WW (2004) Nonlinear multivariate and time series analysis by neural network methods. Rev Geophys 42(1):RG1003, doi:10.1029/2002RG000112

Jolliffe IT (2003) A cautionary note on artificial examples of EOFs. J Climate 16(7):1084–1086, doi:10.1175/1520-0442(2003)016<1084:ACNOAE>2.0.CO;2

Katz RW (2002) Sir Gilbert Walker and a connection between El Niño and statistics. Stat Sci 17(1):97–112, doi:10.1214/ss/1023799000

Kawale J, Liess S, Kumar A, Steinbach M, Snyder P, Kumar V, Ganguly AR, Samatova NF, Semazzi F (2013) A graph-based approach to find teleconnections in climate data. Statistical Analysis and Data Mining 6(3):158–179, doi:10.1002/sam.11181

Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, Dool HVD, Jenne R, Fiorino M (2001) The NCEP–NCAR 50–year reanalysis: Monthly means CD–ROM and documentation. Bull Amer Meteor Soc 82(2):247–268, doi:10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2

Kutzbach JE (1967) Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. J Appl Meteorol 6(5):791–802, doi:10.1175/1520-0450(1967)006<0791:EEOSLP>2.0.CO;2

Lacasa L, Luque B, Ballesteros F, Luque J, Nuno JC (2008) From time series to complex networks: The visibility graph. Proc Natl Acad Sci USA 105(13):4972–4975, doi:10.1073/pnas.0709247105

Leroy A, Wheeler MC (2008) Statistical prediction of weekly tropical cyclone activity in the southern hemisphere. Mo Wea Rev 136(10):3637–3654, doi:10.1175/2008MWR2426.1

Lorenz EN (1956) Empirical orthogonal functions and statistical weather predictions. Scientific report 1, Dep. of Met., MIT, Cambridge, Massachusetts

Ludescher J, Gozolchiani A, Bogachev MI, Bunde A, Havlin S, Schellnhuber HJ (2013) Improved El Niño forecasting by cooperativity detection. Proc Natl Acad Sci USA 110(29):11,742–11,745, doi:10.1073/pnas.1309353110

Ludescher J, Gozolchiani A, Bogachev MI, Bunde A, Havlin S, Schellnhuber HJ (2014) Very early warning of next El Niño. Proc Natl Acad Sci USA 111(6):2064–2066, doi:10.1073/pnas.1323058111

Malik N, Bookhagen B, Marwan N, Kurths J (2012) Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. Clim Dynam 39(3-4):971–987, doi:10.1007/s00382-011-1156-4

Marshall J, Kushnir Y, Battisti D, Chang P, Czaja A, Dickson R, Hurrell J, McCartney M, Saravanan R, Visbeck M (2001) North Atlantic climate variability: phenomena, impacts and mechanisms. Int J Climatol 21(15):1863–1898, doi:10.1002/joc.693

Martin E, Paczuski M, Davidsen J (2013) Interpretation of link fluctuations in climate networks during El Niño periods. Europhys Lett 102(4):48,003, doi:10.1209/0295-5075/102/48003

Marwan N, Donges JF, Zou Y, Donner RV, Kurths J (2009) Complex network approach for recurrence analysis of time series. Phys Lett A 373(46):4246–4254, doi:10.1016/j.physleta.2009.09.042

Meehl G, Covey C, McAvaney B, Latif M, Stouffer R (2005) Overview of the coupled model intercomparison project (CMIP). Bull Amer Meteor Soc 86(1):89–93, doi:10.1175/BAMS-86-1-89

Mheen M, Dijkstra HA, Gozolchiani A, Toom M, Feng Q, Kurths J, Hernandez-Garcia E (2013) Interaction network based early warning indicators for the atlantic MOC collapse. Geophys Res Lett 40(11):2714–2719, doi:10.1002/grl.50515

Molkenthin N, Rehfeld K, Marwan N, Kurths J (2014a) Networks from flows – from dynamics to topology. Scientific Reports 4:4119, doi:10.1038/srep04119

Molkenthin N, Rehfeld K, Stolbova V, Tupikina L, Kurths J (2014b) On the influence of spatial sampling on climate networks. Nonlinear Proc Geophys 21(3):651–657, doi:10.5194/npg-21-651-2014

Monahan AH, Fyfe JC, Ambaum MH, Stephenson DB, North GR (2009) Empirical orthogonal functions: The medium is the message. J Climate 22(24):6501–6514, doi:10.1175/2009JCLI3062.1

Newman M (2010) Networks: An Introduction. Oxford University Press, Oxford

Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256, doi:10.1137/S003614450342480

North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. Mo Wea Rev 110:699–706, doi:10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2

Paluš M, Hartman D, Hlinka J, Vejmelka M (2011) Discerning connectivity from dynamics in climate networks. Nonlinear Proc Geophys 18(5):751–763, doi:10.5194/npg-18-751-2011

Petrova I (2012) Structural interrelationships between evaporation and precipitation: Application of complex networks to satellite based fields. Master's thesis, University of Hamburg

Power S, Casey T, Folland C, Colman A, Mehta V (1999) Inter-decadal modulation of the impact of ENSO on Australia. Clim Dynam 15(5):319–324, doi:10.1007/s003820050284

Preisendorfer RW, Mobley CD (1988) Principal component analysis in meteorology and oceanography. Elsevier, Amsterdam

Radebach A, Donner RV, Runge J, Donges JF, Kurths J (2013) Disentangling different types of El Niño episodes by evolving climate network analysis. Phys Rev E 88(5):052,807, doi:10.1103/PhysRevE.88.052807

Raustiala K (2001) Nonstate actors in the global climate regime. In: Luterbacher U, Sprinz DF (eds) International relations and global climate change, MIT Press, Cambridge, Massachusetts, pp 95–117

Rehfeld K, Marwan N, Breitenbach SFM, Kurths J (2013) Late Holocene Asian summer monsoon dynamics from small but complex networks of paleoclimate data. Clim Dynam 41(1):3–19, doi:10.1007/s00382-012-1448-3

Repelli CA, Nobre P (2004) Statistical prediction of sea-surface temperature over the Tropical Atlantic. Int J Climatol 24(1):45–55, doi:10.1002/joc.982

Rheinwalt A, Marwan N, Kurths J, Werner P, Gerstengarbe FW (2012) Boundary effects in network measures of spatially embedded networks. Europhys Lett 100(2):28,002, doi:10.1209/0295-5075/100/28002

Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. Mo Wea Rev 115(8):1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2

Runge J, Heitzig J, Kurths J (2012a) Escaping the curse of dimensionality in estimating multivariate transfer entropy. Phys Rev Lett 108:258,701, doi:10.1103/PhysRevLett.108.258701

Runge J, Heitzig J, Marwan N, Kurths J (2012b) Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy. Phys Rev E 86(6):061,121

Runge J, Petoukhov V, Kurths J (2014) Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models. J Climate 27(2):720–739

Steffen K, Box J (2001) Surface climatology of the Greenland ice sheet: Greenland climate network 1995-1999. J Geophys Res 106(D24):33,951–33,964, doi:10.1029/2001JD900161

Steinhaeuser K, Tsonis AA (2014) A climate model intercomparison at the dynamics level. Clim Dynam 42(5-6):1665–1670, doi:10.1007/s00382-013-1761-5

Steinhaeuser K, Tsonis AA (in press) A climate model intercomparison at the dynamics level. Climate Dynamics

Steinhaeuser K, Chawla NV, Ganguly AR (2010) An exploration of climate data using complex networks. ACM SIGKDD Explorations 12(1):25–32, doi:10.1145/1882471.1882476

Steinhaeuser K, Chawla NV, Ganguly AR (2011) Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science. Statistical Analysis and Data Mining 4(5):497–511, doi:10.1002/sam.10100

Steinhaeuser K, Ganguly AR, Chawla NV (2012) Multivariate and multiscale dependence in the global climate system revealed through complex networks. Clim Dynam 39(3-4):889–895, doi:10.1007/s00382-011-1135-9

Stolbova V, Martin P, Bookhagen B, Marwan N, Kurths J (2014) Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka. Nonlinear Proc Geophys 21(4):901–917, doi:10.5194/npg-21-901-2014

von Storch H, Zwiers FW (2003) Statistical analysis in climate research. Cambridge University Press, Cambridge

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Amer Meteor Soc 93(4):485–498, doi:10.1175/BAMS-D-11-00094.1

Tenenbaum JB, De Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290(5500):2319–2323, doi:10.1126/science.290.5500.2319

Tominski C, Donges JF, Nocke T (2011) Information visualization in climate research. In: Proceedings of the International Conference Information Visualisation (IV), London, IEEE Computer Society, pp 298–305

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

Tsonis AA, Swanson KL (2008) Topology and predictability of El Niño and La Niña networks. Phys Rev Lett 100(22):228,502, doi:10.1103/PhysRevLett.100.228502

Tsonis AA, Swanson KL, Roebber PJ (2006) What do networks have to do with climate? Bull Amer Meteor Soc 87(5):585–595, doi:10.1175/BAMS-87-5-585

Tsonis AA, Swanson KL, Wang G (2008) On the role of atmospheric teleconnections in climate. J Climate 21(12):2990, doi:10.1175/2007JCLI1907.1

Tsonis AA, Wang G, Swanson KL, Rodrigues FA, Costa L (2011) Community structure and dynamics in climate networks. Clim Dynam 37(5-6):933–940, doi:10.1007/s00382-010-0874-3

Tupikina L, Rehfeld K, Molkenthin N, Stolbova V, Marwan N, Kurths J (2014) Characterizing the evolution of climate networks. Nonlinear Processes in Geophysics 21(3):705–711, doi:10.5194/npg-21-705-2014

Walker GT (1910) Correlation in seasonal variations of weather. II. Memoirs of the Indian Meteorological Department 21(2):22–45

Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the Northern Hemisphere winter. Mo Wea Rev 109(4):784–812, doi:10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2

Wiedermann M, Donges JF, Heitzig J, Kurths J (2013) Node-weighted interacting network measures improve the representation of real-world complex systems. Europhys Lett 102:28,007, doi:10.1209/0295-5075/102/28007

Wiedermann M, Donges JF, Donner RV, Handorf D, Kurths J (in prep.) Northern hemisphere ocean-atmosphere coupling from an interacting climate network perspective

Xu X, Zhang J, Small M (2008) Superfamily phenomena and motifs of networks induced from time series. Proc Natl Acad Sci USA 105(50):19,601–19,605, doi:10.1073/pnas.0806082105

Yamasaki K, Gozolchiani A, Havlin S (2008) Climate networks around the globe are significantly affected by El Niño. Phys Rev Lett 100(22):228,501, doi:10.1103/PhysRevLett.100.228501

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

Zhou CS, Zemanová L, Zamora-Lopéz G, Hilgetag CC, Kurths J (2006) Hierarchical organization unveiled by functional connectivity in complex brain networks. Phys Rev Lett 97(23):238,103, doi:10.1103/PhysRevLett.97.238103

Zou Y, Donges JF, Kurths J (2011) Recent advances in complex climate network analysis. Complex Systems and Complexity Science 8(1):27–38