



POTSDAM-INSTITUT FÜR
KLIMAFOLGENFORSCHUNG

Originally published as:

Feldhoff, J. H., Lange, S., Volkholz, J., Donges, J. F., Kurths, J., Gerstengarbe, F.-W. (2015): Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate. - *Climate Dynamics*, 44, 5-6, 1567-1581

DOI: [10.1007/s00382-014-2182-9](https://doi.org/10.1007/s00382-014-2182-9)

Available at <http://link.springer.com>

© Springer

Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate

Jan H. Feldhoff · Stefan Lange · Jan Volkholz ·
Jonathan F. Donges · Jürgen Kurths ·
Friedrich-Wilhelm Gerstengarbe

Received: 28 October 2013 / Accepted: 14 May 2014 / Published online: 31 May 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract In this study we introduce two new node-weighted difference measures on complex networks as a tool for climate model evaluation. The approach facilitates the quantification of a model's ability to reproduce the spatial covariability structure of climatological time series. We apply our methodology to compare the performance of a statistical and a dynamical regional climate model simulating the South American climate, as represented by the variables 2 m temperature, precipitation, sea level pressure, and geopotential height field at 500 hPa. For each variable, networks are constructed from the model outputs and evaluated against a reference network, derived from the ERA-Interim reanalysis, which also drives the models. We compare two network characteristics, the (linear) adjacency

structure and the (nonlinear) clustering structure, and relate our findings to conventional methods of model evaluation. To set a benchmark, we construct different types of random networks and compare them alongside the climate model networks. Our main findings are: (1) The linear network structure is better reproduced by the statistical model statistical analogue resampling scheme (STARS) in summer and winter for all variables except the geopotential height field, where the dynamical model CCLM prevails. (2) For the nonlinear comparison, the seasonal differences are more pronounced and CCLM performs almost as well as STARS in summer (except for sea level pressure), while STARS performs better in winter for all variables.

Keywords Climate model evaluation · Complex networks · South American climate · Network comparison

We dedicate this manuscript to our friend and colleague Jan H. Feldhoff who sadly passed away during the review process.

J. H. Feldhoff · S. Lange (✉) · J. Volkholz ·
J. F. Donges · J. Kurths · F.-W. Gerstengarbe
Potsdam Institute for Climate Impact Research,
P.O. Box 60 12 03, 14412 Potsdam, Germany
e-mail: slange@pik-potsdam.de

J. H. Feldhoff · S. Lange · J. Kurths
Department of Physics, Humboldt University, Newtonstr. 15,
12489 Berlin, Germany

J. F. Donges
Stockholm Resilience Center, Stockholm University,
Kräftriket 2B, 11419 Stockholm, Sweden

J. Kurths
Institute for Complex Systems and Mathematical Biology,
University of Aberdeen, Aberdeen AB243UE, United Kingdom

F.-W. Gerstengarbe
Department of Geography, Humboldt University,
Rudower Chaussee 16, 12489 Berlin, Germany

1 Introduction

Almost a decade after the introduction of complex network methods into climate science (Tsonis and Roebber 2004), network-based model validation techniques are still few and far between. Climate networks associate geographic locations with nodes (also called vertices) of a mathematical object called network or graph (Newman 2009). The connections between nodes (called links or edges) represent similarities between climatological time series at those locations, derived mostly from reanalyses or remote sensing data. The mathematical field of complex network theory has thrived over the past decades (Strogatz 2001) and now offers a variety of methods to uncover different aspects of the topological structure of networks (Newman 2003; Cohen and Havlin 2008).

Applied to the climate system, these methods have already lead to several substantial new insights: From the identification of dynamical transitions (Tsonis et al. 2007) and teleconnections (Donges et al. 2009b), via the study of El Niño (Yamasaki et al. 2008; Gozolchiani et al. 2011) and monsoon systems (Malik et al. 2011; Boers et al. 2013), to actual predictive power (Ludescher et al. 2013). It has been demonstrated that data-based climate networks show remarkable versatility. The model-based branch of the framework is far less developed, although recent attempts by Steinhäuser and Tsonis (2013) and Fountalis et al. (2013) show the growing interest in the use of network methods for climate model intercomparison and climate model analysis (van der Mheen et al. 2013).

From a theoretical point of view, evaluating the network structure of modeled climate data constitutes a promising extension to conventional evaluation techniques like the comparison of annual cycles or seasonal means. While the latter approaches investigate properties of time series at each geographical location individually, climate networks describe their covariability and thus represent an essentially different aspect of spatio-temporal climate variability.

The intention of this paper is to propose a new method to evaluate climate models by means of complex networks. For this purpose, we compare the structural similarity of climate networks obtained from models to those obtained from reanalysis data (the *reference networks*). The differences between them are considered proxies for the quality of the underlying modeling, here the simulation of the South American climate as performed by a dynamical (CCLM) and a statistical (STARS) regional climate model (RCM). This work is to be seen as both an introduction of the methodology, and as a case study on the feasibility of the application of CCLM and STARS to South American climate.

With our methodology, we aim at a direct comparison of the network structure, which has the advantage of including all available information about the complex networks under study. Any kind of preprocessing of the networks, e.g. a clustering of nodes, bears the inherent danger of adding spurious information or diminishing the complexity of the network structure, possibly stripping it of relevant features.

One of the applied network difference measures is a modification of the Hamming distance, which is rooted in information theory (Hamming 1950) and has found plenty of applications, often in combination with complex network analysis (Donges et al. 2009a; Zhou et al. 2006; Ciliberti et al. 2007). We also compare the clustering structures (Watts and Strogatz 1998) of the observed and modeled networks by computing the root-mean-square distance of their respective fields of local clustering

coefficients in order to evaluate the recreation of nonlinear dependencies by the models.

It should be noted that, comparing the spatial statistical interdependency structure within climatological fields, our method is related to approaches based on empirical orthogonal functions and teleconnection patterns (Handorf and Dethloff 2012; Stoner et al. 2009), yet distinct due to the inclusion of information about nonlinear interrelations (Donges et al. 2013b).

In the next section, we outline the key features of the two particular RCMs under study and describe the simulation setup (Sect. 2). The methodology of network-based model evaluation is presented in Sect. 3 and its results are given in Sect. 4, where we compare the output of simulations of the regional climate of South America, followed by a discussion on the robustness of the method. Finally, we draw conclusions and give an outlook on possible further applications in Sect. 5.

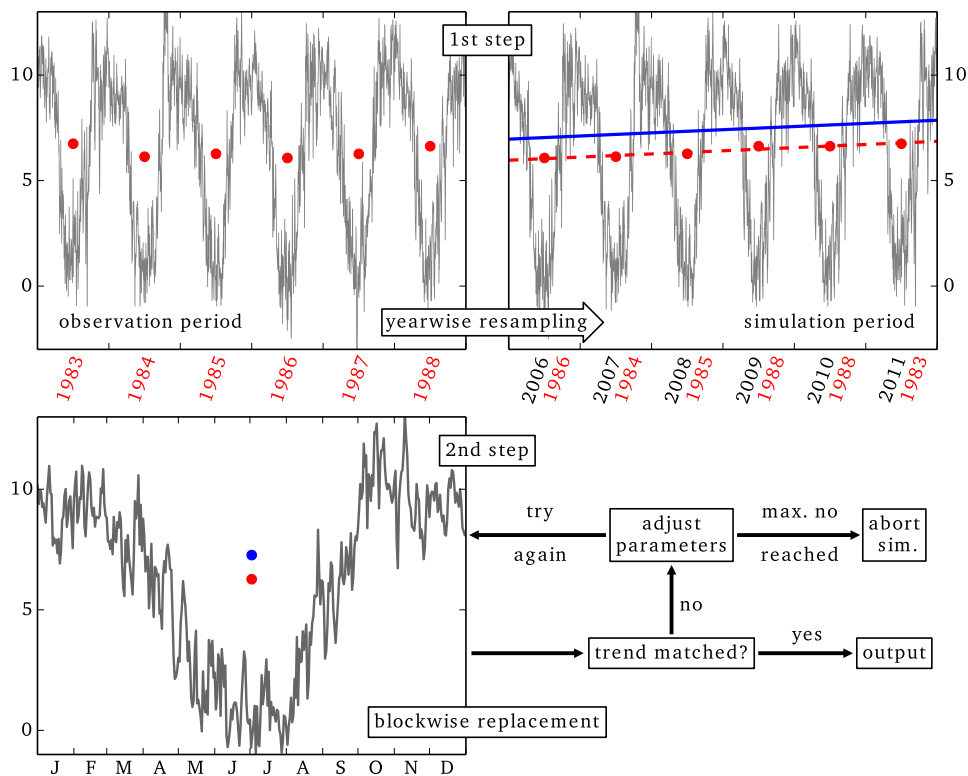
2 Regional climate modeling

Simulating meteorological processes on the mesoscale and below, regional climate models bridge the gap between general circulation models (GCMs), which operate on a global scale at rather coarse horizontal resolution (~ 100 km), and climate impact models, which focus on specific processes or features in a confined region such as hydrology, agricultural production, forestry, etc. (Gutsch et al. 2011; Reyer et al. 2013). For climate projections, impact models are typically driven by RCMs, which in turn downscale GCM data (Stocker et al. 2013). Although, with ever growing computational power, the border between global and regional modeling might become blurry and eventually disappear, for the time being, RCMs are still frequently found in the impact modeling chain. To illustrate our evaluation method we examine RCM simulations. Since this work is about the evaluation procedure rather than climate projections, we only produce hindcasts driven by the ERA-Interim reanalysis data (Dee et al. 2011).

2.1 The statistical approach: STARS

The statistical analogue resampling scheme (STARS) was originally developed in order to provide climate realizations for impact models (Werner and Gerstengarbe 1997), but has since been successfully applied for regional climate projections (Orlowsky and Fraedrich 2008; Orlowsky et al. 2008, 2010; Lutz et al. 2013). The general idea is to stochastically resample meteorological data according to a given trend of some meteorological variable. Typically, temperature is chosen as the trend variable since this is the natural choice in the context of global warming and since

Fig. 1 Basic principle of STARS: at first (*top panel*), entire years from the observation period are resampled for their yearly means (*red dots*) to approximate a prescribed trend line (*blue*). Then, by iteratively replacing 12-day blocks (*bottom panel*), the resulting time series is further tuned to improve the matching of the actual (*red dot*) and prescribed (*blue dot*) yearly mean values



trends of other variables like precipitation are often less robust.

A sketch of the model’s workflow is shown in Fig. 1. At first, as a form of biased bootstrapping, observations (or, in this case, reanalysis data) are resampled as entire years in order to approximately match a prescribed trend line. The match is further improved by replacing blocks of 12 days in an iterative process. The resulting date-to-date mapping is then applied to all variables and all locations. Thereby, the simulation output is guaranteed to be physically consistent, within the limits of the input data’s consistency.

STARS can only produce output which has been observed rather than create entirely new situations like a dynamical model. For example, no new extreme values can be simulated. It should also be noted that uncertainties in the input data will propagate into the simulation. Apart from the quality, the availability of data is also a key constraint on the applicability of STARS. As a rule of thumb, the length of the simulation period should not be longer than the observation period in order to prevent unnaturally low variability in the model output. Thus, since the ERA-Interim dataset starts in 1979, we used the input data from 1979 to 1995 to simulate the period from 1996 to 2011, prescribing the temperature trend of the reanalysis during the simulation period.

Due to the statistical nature and low computational demands of STARS, it is possible to obtain ensembles of climate realizations in very short time, which makes this

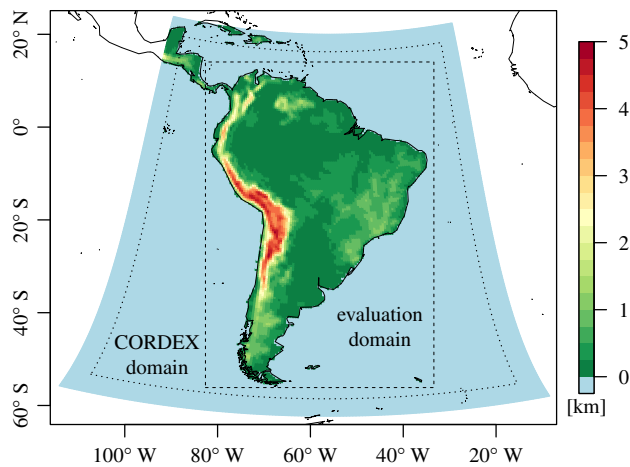


Fig. 2 CCLM’s domain of computation including the sponge frame (*colored*), the CORDEX-South-America domain (*dotted*), and the common domain of evaluation (*dashed*). Colors indicate surface height

model ideal for studying a whole range of possible scenarios. For this study, we generated an ensemble of 200 realizations using STARS version 2.4.

2.2 The dynamical approach: CCLM

The COSMO-CLM (CCLM, Rockel et al. 2008) is the climate version of the COSMO-Model (Baldauf et al.

2011), which is the operational numerical weather prediction model of the German Weather Service and other members of the COSMO consortium. The development of CCLM is steered by the CLM Community which has more than 50 member institutions from Europe, Asia, Africa and America. The model has been extensively applied to European domains (e.g. Jaeger et al. 2008; Zahn and Storch 2008; Hohenegger et al. 2009; Davin and Seneviratne 2011) but also to the Indian subcontinent (Dobler and Ahrens 2010), to CORDEX-East-Asia (Fischer et al. 2013), and to CORDEX-Africa (Nikulin et al. 2012). One of the very first applications was to South America (Böhm et al. 2003) but it has been run there rarely afterwards (Rockel and Geyer 2008; Wagner et al. 2011). CCLM is dynamical in the sense that it solves thermohydrodynamical equations describing the atmospheric circulation. The equations are discretized on a three-dimensional grid based on a rotated geographical coordinate system.

In this study the CCLM version 4.25.3 was used. Deviating from its default configuration, the model was run with 40 vertical levels, reaching up to 30 km above sea level and a Rayleigh damping height of 18 km, as has been suggested for tropical regions (Panitz et al. 2013). We set the bottom of the deepest hydrologically active soil layer to 8 m, since rain forest roots go down to such depths (Baker et al. 2008). The numerical integration was performed with a total variation diminishing Runge–Kutta scheme (Liu et al. 1994) and a Bott advection scheme (Bott 1989), since both are supposedly more accurate than their default alternatives. We employ an implementation of the ECMWF IFS Cy33r1 convection scheme (Bechtold et al. 2008) and diagnose subgrid-scale clouds by a normalized saturation deficit criterion (Sommeria and Deardorff 1977). Additionally, a few tuning parameters were adjusted during preceding sensitivity experiments. Particularly, changing the convective parametrization and the subgrid-scale cloud scheme led to major improvements of the model performance over South America. These findings are presented in detail in a separate paper (Lange et al. 2014).

We run the model on the CORDEX-South-America domain (Giorgi et al. 2009) as displayed in Fig. 2. This implies a horizontal resolution 0.44° and 166×187 grid points including a 10 grid points wide sponge frame. The simulation covers the years 1979–2011 where the first 17 years serve as spinup time, since the STARS output is only available from 1996.

2.3 Common domain of evaluation

In order to construct climate networks of the same spatial embedding, both model outputs were to match in resolution and geographical boundaries. We chose a section of the native ERA-Interim grid, encompassing the South

American mainland (Fig. 2): 82.3°W – 33.8°W and 13.7°N – 55.8°S . The resolution is approximately 0.7° in both latitude ϕ and longitude λ . This makes for a bounding box of $N = N_\phi \times N_\lambda = 100 \times 70 = 7,000$ grid cells, which will be represented by nodes in the subsequently constructed climate networks. Since this is a regular Gaussian grid of considerable latitudinal extent, grid cells at different latitudes represent differently sized areas (about $78 \text{ km} \times 78 \text{ km} = 6,084 \text{ km}^2$ at the equator and $44 \text{ km} \times 78 \text{ km} = 3,432 \text{ km}^2$ at the southern boundary)—an effect we take into account by introducing area-proportional node weights (cf. Sect. 3.3).

Since STARS only resamples the input data, its output is already on the native ERA-Interim grid. CCLM output was remapped, conservatively (Jones 1999) in case of precipitation and bilinearly otherwise.

3 Methodology

After some introductory definitions from complex network theory (Newman 2003; Boccaletti et al. 2006), we move on to present the two essential parts of our methodology: The construction of spatially embedded networks and their comparison.

3.1 Basic definitions

A *network* (Newman 2009) or *graph* $G = (\mathcal{N}, \mathcal{E})$ consists of a *node set* $\mathcal{N} = \{1, \dots, N\}$ of *nodes* or *vertices*, potentially pairwise connected by *links* or *edges*, constituting an *edge set* $\mathcal{E} \subseteq \{\{i, j\} : i \neq j \in \mathcal{N}\}$. Connected nodes are called *neighbors*. The full information about G is contained in its binary *adjacency matrix* $(a_{ij})_{i, j \in \mathcal{N}}$, with $a_{ij} = 1$, if the nodes i and j are connected and $a_{ij} = 0$, otherwise.

These definitions imply that, within the scope of this article, we only work with *undirected, unweighted, simple* graphs (i.e. no directed links, no link weights, and no self-loops). We do, however, apply *node weights* w_i to enable the use of area-weighted measures (Sect. 3.3).

The number of connections of any node i is called the *degree* k_i of node i , with

$$k_i = \sum_{j=1}^N a_{ij}. \quad (1)$$

It is, in other words, the number of *neighbors* of that particular node, or the *cardinality* of its *neighborhood* $\mathcal{N}_i = \{j \in \mathcal{N} : a_{ij} = 1\}$, the set of nodes, connected to i . A measure of higher order is the (*local*) *clustering coefficient* c_i , which estimates the likelihood of any two

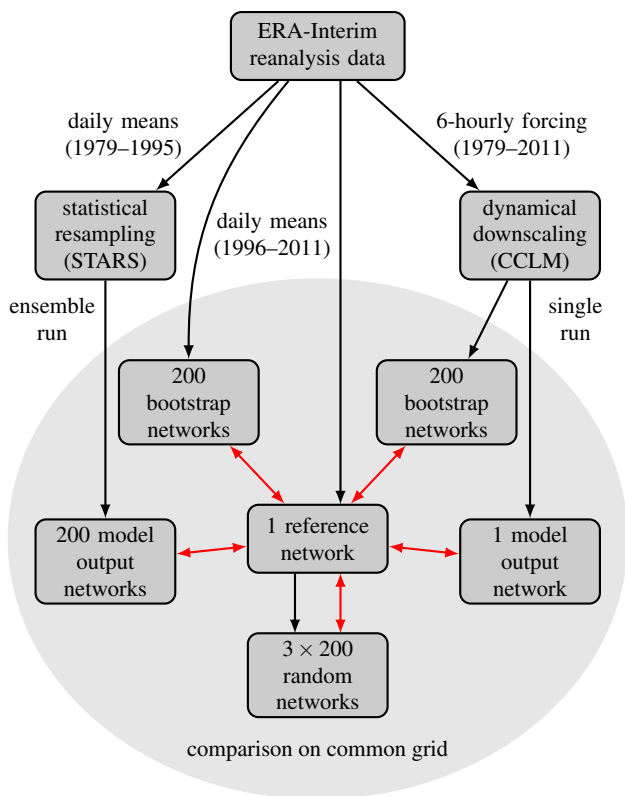


Fig. 3 Experiment design. For each climatological variable considered, complex networks are generated through different pathways (modeling, bootstrapping, random network generation) and compared to the respective *reference network*, generated directly from the ERA-Interim data. The *red double-headed arrows* indicate the application of the difference measures C^* (Eq. 7) and H^* (Eq. 8)

neighbors of node i also being connected (Watts and Strogatz 1998):

$$c_i = \frac{1}{k_i(k_i - 1)} \sum_{j,k \in \mathcal{N}_i} a_{jk}. \tag{2}$$

Apart from k_i and c_i , there are many more measures describing different aspects of network topology, but in this study we are going to restrict our analyses to the aforementioned ones.

3.2 Network construction

We demonstrate two methods of constructing networks: data-based climate networks from the model outputs and reanalysis data, and three types of random-based surrogate networks which inherit different features from the reference network in order to function as null models. The latter shall provide a useful benchmark for the results by imitating climate models which only reproduce certain features of the reference network structure.

3.2.1 Climate networks: model output and reanalysis data

The data-based climate networks are constructed from three sources: CCLM model output, STARS model output, and ERA-Interim reanalysis data, in each case for the daily means of four different variables: 2 m temperature (T2M), total precipitation (PREC), geopotential height at 500 hPa (GEO500), and sea level pressure (SLP). The temperature and precipitation variables represent major features of the climate system with high impact on biosphere and human society, as well as great importance for impact modeling. The sea level pressure and geopotential variables are representatives of the circulation system on the surface and in higher altitudes, respectively, and their faithful reproduction is equally essential for accurate climate modeling.

In all cases the general procedure of network construction is the same, following (Donges et al. 2009a, b):

1. Choose a time frame, such as austral summer (DJF), austral winter (JJA), or any other interesting time span within the evaluation period, to get N time series, one for each grid cell.
2. For each time series apply a moving average filter with a sliding window of length l . The default value used here was $l = 7$ days.
3. Produce climatological anomalies from each time series, i.e. remove the seasonality such that the resulting time series are approximately stationary in mean and variance.
4. Calculate similarities for all pairs of time series using, e.g. Pearson correlation or rank correlation, to obtain an $(N \times N)$ -similarity matrix.
5. In the similarity matrix, set those values to 1 which are greater than a chosen threshold and set those below to 0.
6. Use the thresholded similarity matrix as adjacency matrix of the climate network. Set the main diagonal to zero to exclude self-loops.
7. Finally, assign to each node i a *node weight* w_i proportional to the geographic area it represents, i.e. $w_i \propto \cos(\phi_i)$ (Heitzig et al. 2012; Wiedermann et al. 2013).

Steps 2–4 in the construction procedure need some special attention. To obtain the similarity matrix, we calculate all correlations at lag zero. Usually, weather phenomena distribute over a larger area in a matter of several days, so a time lag might be considered. We instead chose to apply a moving average to the daily values to account for this fact, avoiding the application of a fixed time lag. Unless stated otherwise, we average across of $l = 7$ days and discuss the sensitivity of our results with respect to l in Sect. 4.4. Other similarity measures, such as event synchronization (Malik

et al. 2011; Boers et al. 2013) also allow for dynamical time lags but are not subject of this work.

In order to conduct step 3 we need an approximation of the seasonal cycle in daily resolution, which we construct by calculating the long-term mean for each day and smoothing the resulting time series by a Gaussian filter to account for the rather short evaluation time of 16 years.

For the unbound variables T2M, SLP, and GEO500 we obtained approximately Gaussian-distributed anomaly time series by subtracting the seasonal cycle from the smoothed daily values. Approximate homoscedasticity is assumed for these variables given that our analysis will be constrained to seasonal time series. The similarities are estimated by Pearson correlation.

Other variables, in our case PREC, have a natural lower bound of zero, and thus their probability distribution is more complicated, clearly non-Gaussian (frequently modeled as a Weibull-, Gamma-, or mixed-exponential distribution, (Li et al. 2012). In this case, we apply Spearman's rank correlation, which also works for non-Gaussian variables, to estimate the similarities. Also, to compute the anomaly values and approach homoscedasticity, we divide the smoothed daily values by the seasonal cycle instead of subtracting it. This yields an expectation value of 1 mm/day at all times and leaves zeros at zero. Simply subtracting the daily means would transform zeros into values which rise as the climatology falls. To avoid dividing by zero we define a minimal value of 0.1 mm/day which we divide by whenever it is underrun. We chose this value since it is usually referred to as the minimally measurable daily rain amount. In our data, this case actually occurs only in northern Chile (Atacama desert) and the adjacent part of the Pacific ocean.

The threshold for the similarity matrix (step 5) was applied adaptively such that a desired link density ρ in the resulting networks was achieved. Unless stated otherwise, we used $\rho = 0.01$, meaning that we included only the 1 % strongest correlations in our analysis, which is considered an effectual trade-off between structural richness and statistical significance (Donges et al. 2009a).

It should be noted that climate networks are often constructed by thresholding the matrix of the absolute values of correlation coefficients (Tsonis and Roebber 2004; Donges et al. 2009a, b). In the context of network comparison however, this could lead to the problematic situation, in which, for two networks with adjacency matrices (a_{ij}^A) and (a_{ij}^B) , $a_{ij}^A = 1$ is due to a positive correlation while $a_{ij}^B = 1$ is due to a negative correlation. Hence, although the relation of i to j is of wholly different nature in the two networks, a comparison of them would yield agreement. In order to prevent this case we focus on positive correlations here.

Using the above recipe, we construct climate networks from each of the 200 STARS realizations, as well as one from the CCLM simulation, and one from the ERA-Interim data. The latter network is assumed to be a close approximation of the real-world network structure and will thus be used as the *reference network* for all others to be compared to (Fig. 3).

3.2.2 Surrogate networks: bootstraps and random models

The large ensemble of networks from STARS output allows for assumptions being made on variations in the quality of the statistical modeling. Due to the considerable computational demand of the dynamical modeling, we have only one CCLM simulation available. In order to still be able to estimate the uncertainty of the dynamical modeling, a technique from the *bootstrapping* family of methods (Efron 1979), also known as *case resampling*, is applied: We bootstrap the CCLM output by randomly drawing entire seasons with replacement, such that the length of the bootstrapped time series equals the original, and apply this reordering synchronously to the whole output time series field to preserve spatial patterns and correlations. Repeating this procedure 200 times, we get a set of realizations, each of which we create a climate network from.

The same technique is applied to the reanalysis dataset, yielding 200 surrogate networks, which are supposed to closely resemble the reference network and thus form an upper bound for the performance of the climate models.

To also create lower bounds and thus add a sense of scale to our comparison, we further extend our comparison to include three different kinds of random models as surrogates to the reference network. Each type of random model demonstrates what performance we could expect if our climate models would only reproduce a specific feature of the reference network:

- Erdős–Rényi model (ER)
This most basic type of random network (Erdős and Rényi 1959) only conserves the total number of links (or equivalently, the link density ρ) of the reference network. Its edges are rewired completely at random. This can be seen as a worst-case model.
- Configuration model (CM)
Here the degree k_i of each node i is the same as in the reference network, while its neighborhood \mathcal{N}_i is randomized (Newman 2003). While more sophisticated than the ER approach, this model should still perform worse than the RCMs.
- Spatially embedded random network model (SERN)
This model was introduced to estimate the effects of spatial embedding on connectivity in random networks

(Barnett et al. 2007). It was also recently used to study the effect of boundaries on measures in regional climate networks (Rheinwalt et al. 2012). The algorithm constructs random networks with approximately the same distribution of geographic link lengths as the reference network and the same link density.

For each random network type we generated ensembles of 200 realizations. We now have six sets of networks: Two from climate models (including 200 bootstraps from the CCLM simulation), three from random network models, and one from reanalysis bootstraps, all sharing the same spatial embedding, link density and node weights. By assessing their similarity to the reference network we can estimate the quality of the underlying modeling processes from the climate network perspective (Fig. 3).

3.3 Comparison of spatially embedded networks

One way to quantify the dissimilarity of graphs is the *Hamming distance* (Hamming 1950). For two unweighted simple graphs A and B with adjacency matrices (a_{ij}^A) and (a_{ij}^B) and a common set of N nodes, its normalized form is the fraction of edges that have to be changed in one graph in order to convert it into the other (Donges et al. 2009a):

$$H(A, B) = \frac{1}{N(N-1)} \sum_{i,j=1}^N a_{ij}^A \text{ XOR } a_{ij}^B, \tag{3}$$

with

$$a_{ij}^A \text{ XOR } a_{ij}^B = \begin{cases} 0 & a_{ij}^A = a_{ij}^B, \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

More generally defined as a distance measure on the set of binary strings of length L (here $L = N^2$), the Hamming distance is a metric (Hamming 1950).

To be able to compare spatially embedded networks of nodes representing differently sized grid cells, we apply the framework proposed by Heitzig et al. (2012), called node splitting invariance (n. s. i.). For example, while the degree k_i only counts the number of nodes connected to i , by summing up the area-proportional weights of the connected nodes, we can construct a measure which accounts for the *area* connected to i :

$$k_i^* = \sum_{j \in \mathcal{N}_i^+} w_j, \tag{5}$$

with $\mathcal{N}_i^+ = \mathcal{N}_i \cup \{i\}$, the *extended neighborhood* of i . Likewise, there is an n. s. i. version of the clustering coefficient:

$$c_i^* = \frac{1}{(k_i^*)^2} \sum_{j,k \in \mathcal{N}_i^+} w_j a_{jk}^+ w_k. \tag{6}$$

Here we have to sum over a_{jk}^+ , the entries of the *extended adjacency matrix* with $a_{i \neq j}^+ = a_{ij}$, $a_{ii}^+ = 1$. For details on why to use \mathcal{N}^+ and (a_{ij}^+) to abide the concept of node-splitting invariance, please refer to Heitzig et al. (2012). The n. s. i. clustering coefficient takes into account the area, represented by each neighbor of node i , and weights their contribution to the clustering coefficient accordingly.

To estimate how well the clustering fields of two given networks A and B match (in our case reference and model network), we determine their differences via the node-weighted root-mean-square error (RMSE) of clustering coefficients

$$C^*(A, B) = \sqrt{\frac{1}{W} \sum_{i=1}^N w_i (c_i^{A*} - c_i^{B*})^2}, \tag{7}$$

where $W = \sum_{i=1}^N w_i$ is the total node weight. This measure is just one example of how to compare the nonlinear properties of two networks. Analogously, one could compare other measures of higher order, such as *betweenness* or *closeness*. As a linear measure of mutual differences in the neighborhood structure, we introduce the node-weighted Hamming distance

$$H^*(A, B) = \frac{1}{W^2} \sum_{i,j=1}^N w_i (a_{ij}^A \text{ XOR } a_{ij}^B) w_j. \tag{8}$$

With C^* and H^* we now have two difference measures for spatially embedded networks, which account for differently weighted nodes, as in the case of nodes representing geographic areas of varying size.

3.4 Conventional area-weighted difference measures

To complete our methodology and relate to previous works of climate model evaluation, we apply area-weighted versions of two standard difference measures, the root-mean-square error E of the climatological mean field μ of a variable and the *logarithmic root-mean-square factor* F (Golding 1998) of the respective standard deviation field σ . We define the area-weighted versions of E and F as

$$E^*(A, B) = \sqrt{\frac{1}{W} \sum_{i=1}^N w_i (\mu_i^A - \mu_i^B)^2} \tag{9}$$

and

$$F^*(A, B) = \sqrt{\frac{1}{W} \sum_{i=1}^N w_i \left(\log \frac{\sigma_i^A}{\sigma_i^B} \right)^2}, \tag{10}$$

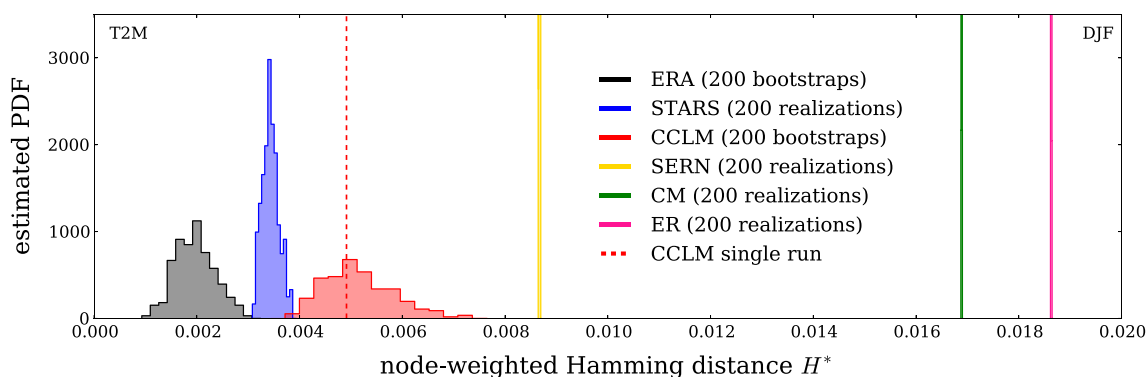


Fig. 4 Estimated probability density functions of the node-weighted Hamming distances between the model networks and the reference network. Time series from T2M in austral summer (DJF). The networks made from resampled reanalysis data (black) bear the closest resemblance to the reference network. STARS (blue) performs better than CCLM (red dashed line). The networks from CCLM

bootstraps (red shaded area) give an impression of the uncertainty of the CCLM modeling, sometimes giving better and sometimes worse results than the actual run. The random models (SERN: yellow, configuration model: green, Erdős–Rényi: pink) perform worse than the RCMs

where A and B denote two distinct datasets to be compared.

Altogether, the measures C^* , H^* , E^* and F^* are comparable in that they are equal to zero in case of perfect agreement and grow with disagreement. Yet they are complementary in that they are based on distinct features of the underlying time series, namely the mean and variance fields in case of E^* and F^* , and the correlation matrix in case of C^* and H^* .

4 Application: regional climate modeling over South America

The climate of South America (SA) is very diverse and includes the humid Amazon rain forest in the central-northern lowlands, which is contrasted by semi-arid regions like the Sertão in northeastern Brazil, deserts like La Guajira in northern Columbia or the Atacama in northern Chile, and permanent ice fields in the south of Patagonia. The Amazon rain forest is also a hotspot of biodiversity and a major carbon sink, whose role as one of the key tipping elements in the global climate system has been well established (Lenton et al. 2008; Boulton et al. 2013).

During austral summer (DJF), the South American Monsoon System is responsible for extensive moisture transport from the southward-shifted Intertropical Convergence Zone (ITCZ) via the Amazon basin and further south towards the extratropics (Vera et al. 2006). The channeling of the easterly trade winds by the Andes and the Brazilian Highlands, also called the South American Low-Level Jet (SALLJ, Marengo and Soares 2004), leads to high precipitation from the Altiplano Plateau to the La Plata basin or, via the South Atlantic Convergence Zone, in southeast SA and the adjacent South Atlantic (Carvalho et al. 2004).

In austral winter (JJA), the ITCZ is shifted northwards. The moisture transport via SALLJ is considerably lower and more moisture from the Atlantic is fed into the then active North American Monsoon System. A major influence on the weather in the southern part of SA is the formation and movement of extratropical cyclones, which during winter tend to be more frequent and of higher complexity (Mendes et al. 2009).

Another prominent influence on the South American climate is the *El Niño* phenomenon (Trenberth 1997), the appearance of a band of unusually warm ocean water in the East Pacific, off the coast of Peru. *El Niño* or, in a wider context, the *El Niño Southern Oscillation* (ENSO), occurs highly erratically on an interannual scale and, depending on its intensity, can impose drastic effects on the SA climate and the ecosystem in general, e.g. by disturbing oceanic food chains, which are sensitive to alterations in the water temperature (Stenseth et al. 2002). Reliable mechanisms for the predictability of ENSO are still highly sought-after in contemporary climate and ocean research (Schneider et al. 2003; Ludescher et al. 2013), and assessing the impacts of climate change on ENSO remains challenging, especially in the long run (Stevenson et al. 2012).

There have been multiple attempts on modeling the regional climate of SA, recently in a coordinated study following the CORDEX conventions (Solman et al. 2013) or with stronger focus on climate change impacts on Amazonia (Cook et al. 2012). Previous works include studies of the long term effects of climate change (Marengo et al. 2009, 2011) and coupled RCM-vegetation modeling (Cook and Vizy 2008). While the agreement of the applied models is often quite high, all of the above studies focus on reproducing seasonal cycles or mean conditions over extended periods, along with variance analyses and observations on the frequency distributions of extremes.

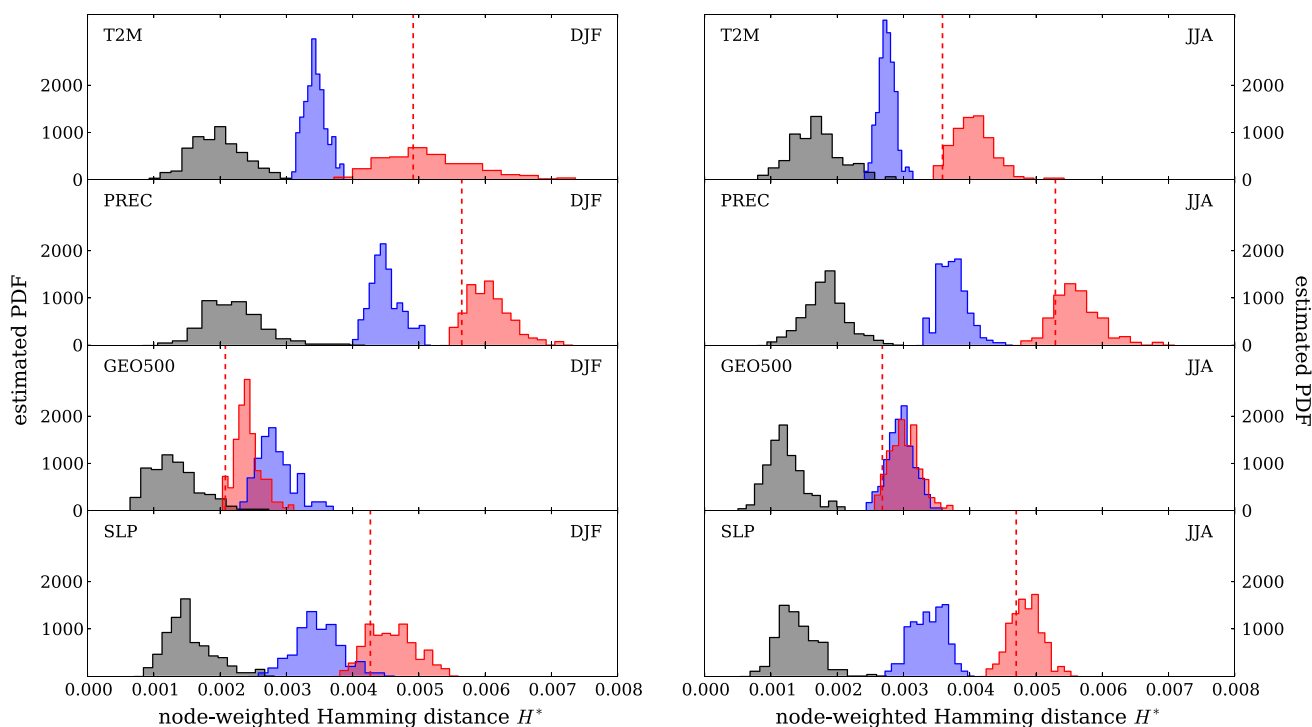


Fig. 5 Node-weighted Hamming distances of networks on modeled and bootstrapped data to the respective ERA-Interim reference network. All variables in austral summer (DJF, left) and winter (JJA, right). The linear network structure is better reproduced by

STARS (blue) than CCLM (bootstraps: red, single run: red dashed line) in both seasons and all variables except GEO500, where CCLM prevails. ERA-bootstraps (black) are always superior to both RCMs. The random network models perform worse and are omitted here

Being based on comparing the network structure, and, hence, spatial correlation patterns of model output and reanalysis data, our methodology is complementary to this established agenda.

4.1 Linear network comparison: H^*

The first step in our analysis concerns the reproduction of the adjacency structure of the reference network by the RCMs and the random models. As an introduction we discuss the probability density functions (PDFs) of the node-weighted Hamming distance H^* (Eq. 8; Fig. 4), derived from networks of T2M time series in austral summer.

As expected, the bootstrapped ERA-Interim data yields networks with the closest resemblance to the reference network and thus the smallest Hamming distance while the random models produce networks with less resemblance, SERN performing best, followed by the configuration model, and Erdős–Rényi being worst. The latter is no surprise due to the complete randomness of the ER model, the expectation value for the unweighted Hamming distance being $2\rho(1 - \rho) = 0.0198$ (Donges et al. 2009a). The slightly better performance of the configuration model is rooted in the model’s conservation of the degree

distribution, giving each node the same degree as the corresponding node in the reference network, thus enhancing the probability of successfully reproducing links. SERN performs much better because it conserves an important link property that is shared by all networks considered here: The probability of finding a link between two nodes is the greater the closer they are to each other geographically. In comparison to ER and CM, this strongly reduces the randomness of link positioning and renders the SERN networks much more similar to the reference network. All random network models produce very narrow distributions due to the rather low link density of 1 % and the high number of 7,000 nodes (law of large numbers). For a better visualization, the cusps of these distributions are cut off in Fig. 4.

Additionally, we find that of the RCMs, STARS outperforms CCLM. There are remarkable differences in the shape of the distributions, those from bootstraps (ERA and CCLM) being wider than the distribution of the STARS ensemble. This can possibly be attributed to the fact that, due to the operating mode of STARS (cf. Sect. 2.1 or Orłowsky et al. 2008), many constraints are imposed on the selection of blocks of consecutive days during the resampling of temperature time series, thus lowering the variability between realizations and resulting in a narrower

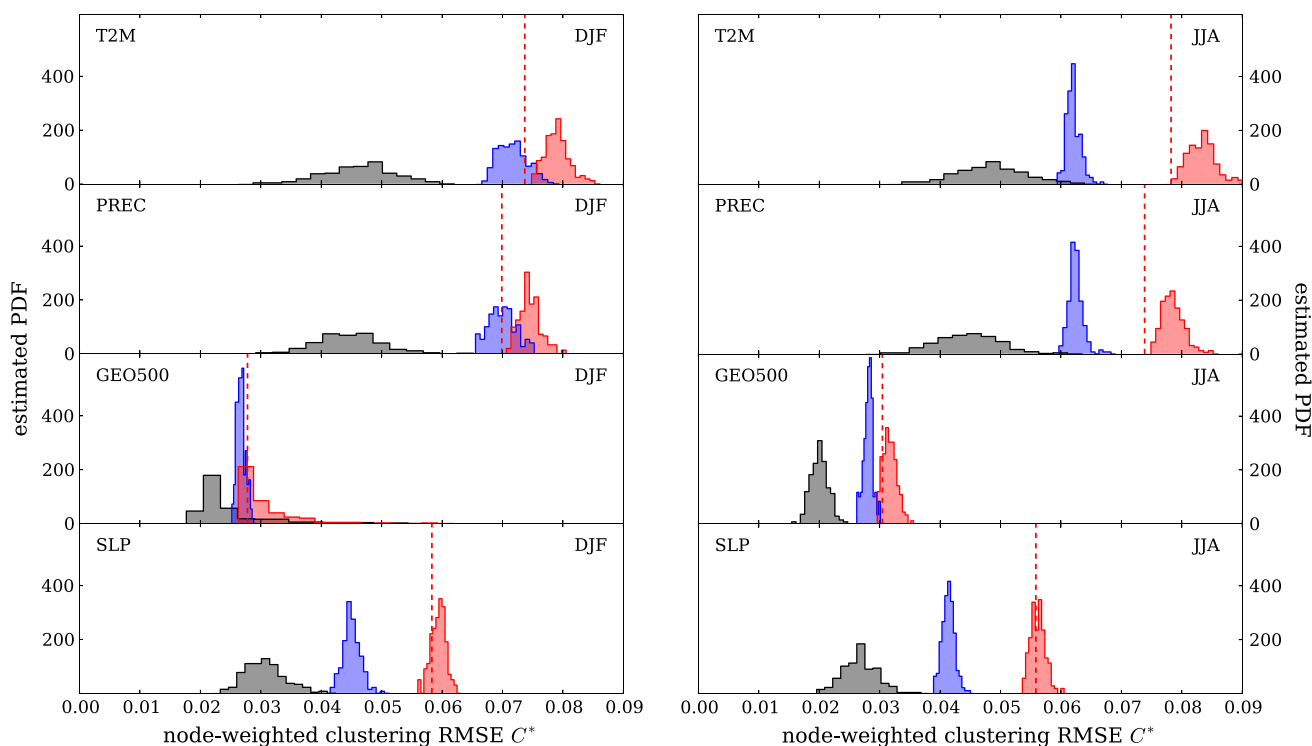


Fig. 6 Node-weighted clustering RMSE of networks on modeled and bootstrapped data with respect to the reference network. All variables in austral summer (DJF, *left*) and winter (JJA, *right*), coloring as in Fig. 5. For the higher-order comparison, the seasonal differences are more pronounced and CCLM performs comparably to STARS in

austral summer (except SLP), while STARS performs better in austral winter for all variables. ERA-bootstraps are always superior to both RCMs. The random network models perform worse and are omitted here

distribution compared to the unbiased bootstrapping procedures applied to ERA and CCLM.

For the other variables (Fig. 5), there is no such clear difference in variability, presumably because these are only indirectly affected by the resampling procedure, via their climatological interrelation to temperature. We have left out the Hamming distance distributions of the random networks here, because their position hardly differs between variables.

Comparing the overall picture, we observe that Hamming distances are greatest for PREC and least for GEO500 across datasets and seasons. This indicates that, out of those variables considered in this study, the dynamics of precipitation are hardest and those of the 500 hPa geopotential are easiest to model. The special position of GEO500 comes as no surprise as it is the only upper level variable, i.e. undisturbed by orographic or other ground-based influences, and since its dynamics have been found relatively easy to model before (Steinhaeuser and Tsonis 2013). Also the outstanding complexity of precipitation dynamics is well-known (Huff and Shipp 1969; Matsoukas et al. 2000; Peters et al. 2001).

Comparing the individual distributions, we find that the resampling-based model STARS has a rather constant relative

distance to the ERA bootstraps (the practical upper bound to model accuracy), which simply reflects the model's functional principle. In contrast, the performance of the dynamical model CCLM varies strongly between variables. Compared to STARS, it generates T2M, PREC, and SLP networks which are less similar and a GEO500 network which is more similar to the respective reanalysis reference. This reflects that model physics differences have a larger impact at the surface than at upper levels where in turn dynamical simulations bear greater similarity to their boundary forcings.

Performance differences between seasons are smaller than those between models. In austral winter (JJA, Fig. 5, right), the modeling of GEO500 by CCLM is slightly less accurate than in summer (still only 8 % of the STARS realizations perform better than the CCLM run). This might be attributed to a higher complexity of the extratropical cyclogenesis during winter (Mendes et al. 2009) and its relatively greater influence on the South American climate due to the JJA northward displacement of general circulation patterns.

4.2 Nonlinear network comparison: C^*

The comparison of the higher-order structure of the networks, here represented by C^* (Eq. 7; Fig. 6), confirms in

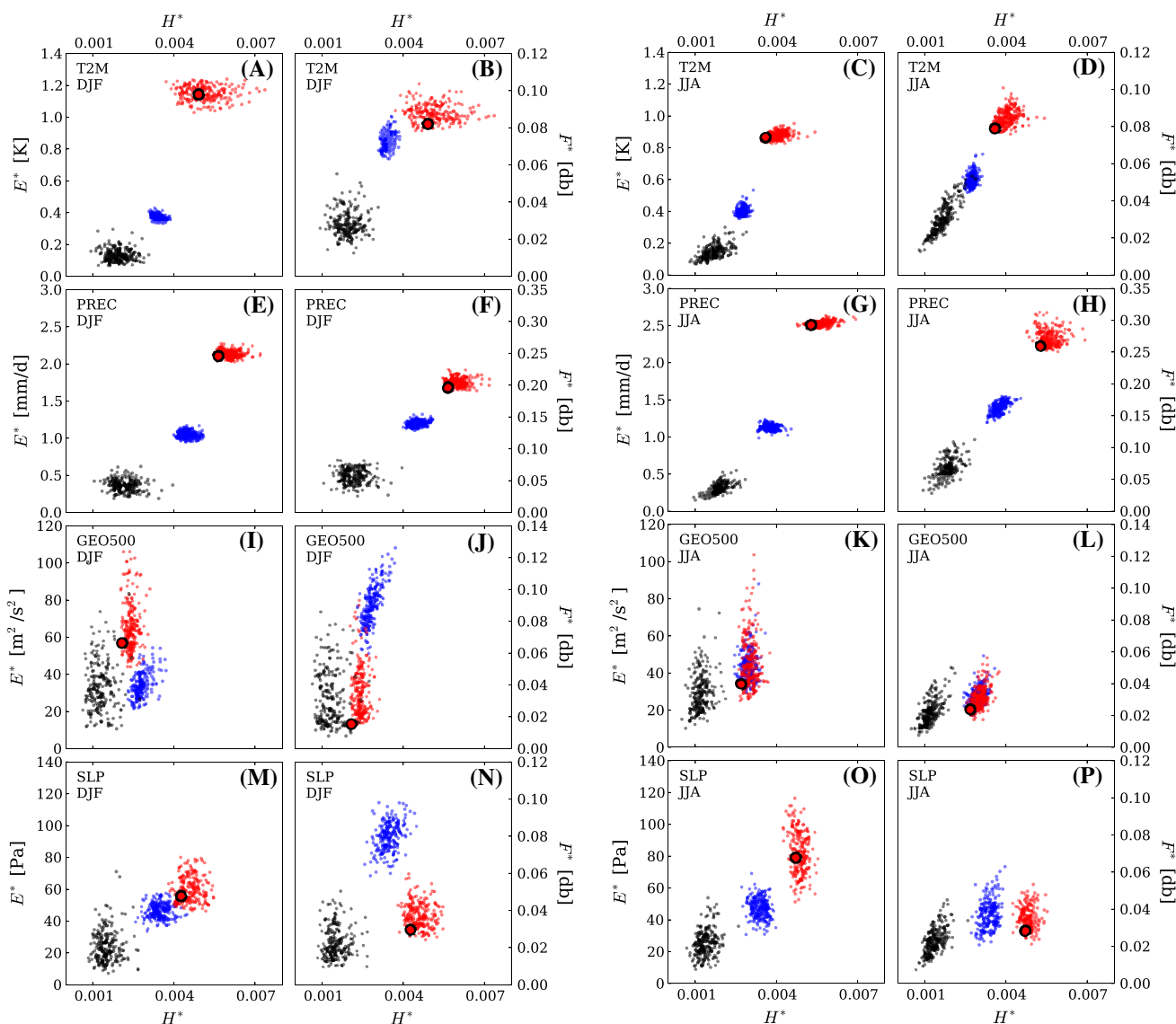


Fig. 7 Mean and variance-based error measures E^* and F^* versus the covariance-based Hamming distance H^* for different variables and seasons as mentioned in the figures. Coloring as in Fig. 5, the CCLM single run is depicted by an accentuated red dot. For T2M and PREC

(panels A–H), STARS performs better with respect to all measures. GEO500 (panels I–L) is dominated by CCLM, and the results are mixed for SLP (panels M–P)

parts the results of the linear comparison: The ERA bootstraps score best, followed by the RCMs and the random models. The latter are omitted in the figures, their ranking being the same as for the linear comparison, with the random networks’ distributions lying even farther out than in Fig. 4. Apparently, reproducing the clustering structure is more challenging for a random model than reproducing the adjacency structure.

Another qualitative similarity is the extensive dominance of the statistical model, most notably in austral winter, where all of the STARS simulations are closer to the reference than the CCLM run. For the summer months, this dominance is equally pronounced

only for SLP, but less pronounced otherwise. In case of T2M, PREC, and GEO500, 80, 49 and 90 % of the STARS ensemble perform better than CCLM, respectively.

Moreover, the clustering RMSE of the CCLM single run is often separated from the bootstrap distribution, the latter scoring worse, most prominently for T2M and PREC. Although visible in the linear difference measure H^* as well (Fig. 5), this feature is more pronounced in the non-linear case (Fig. 6), which implies that the higher-order network structure is more easily disturbed by the bootstrapping procedure than the adjacency structure, as measured by the Hamming distance.

4.3 Conventional measures: E^* and F^*

We now investigate the relation of H^* to the conventional measures E^* and F^* (Eqs. 9, 10; Fig. 7). The measures agree on STARS being better than CCLM at modeling T2M and PREC. In case of GEO500, CCLM performs better according to all three measures except for E^* in DJF (panel I). For SLP, there is no unanimous result: H^* and E^* favor STARS, F^* favors CCLM. Comparing seasons, STARS is generally better in winter than in summer, especially as measured by F^* and H^* . CCLM shows ambiguous interseasonal differences for all three measures.

The complementarity of network-based and conventional difference measures is reflected by the models' ability to simulate SLP and GEO500 in DJF. One of the models may perform better according to both scores (STARS in panel M, CCLM in panel J) or prevail according to H^* but not according to E^* or F^* , respectively (CCLM in panel I, STARS in panel N).

Finally, we observe a consistent difference in the rankings according to the mean- and the variance-based measure. While E^* favors STARS in all cases but one (panel K), the result for F^* is less clear with CCLM even outperforming STARS in both pressure variables. These findings are in line with the statistical model's presumably too low variability of H^* values for T2M as discussed in Sect. 4.1.

4.4 Sensitivity to network construction parameters

We have presented extensive results for only one set of network construction parameters as stated in Sect. 3.2.1, namely $l = 7$ days for the length of the sliding window and $\rho = 1\%$ for the link density of all networks involved. However, calculations were carried out for a wider range of these parameters. We found that varying $l \in [3, 11]$ and $\rho \in [0.5\%, 2\%]$ did not alter the results qualitatively, which confirms the robustness of the demonstrated methodology.

5 Conclusions

In this study we have introduced a novel approach to climate model evaluation based on complex networks. To this end, we have defined two node-weighted difference measures H^* and C^* , which compare adjacency matrices and clustering coefficient fields, respectively. We applied our methodology to evaluate the performance of a statistical vs. a dynamical RCM in simulating the climate of South America.

We have evaluated daily means of 2 m temperature, precipitation, sea level pressure, and geopotential height at 500 hPa, comparing the respective model outputs to our ground truth, the forcing ERA-Interim data. For each variable, climate networks have been constructed based on cross-correlations of the time series, compared using H^* and C^* , and the findings have been related to the classic mean- and variance-based difference measures E^* and F^* .

For the linear network comparison (H^*), we have found that the statistical model STARS is better at reproducing the network structure of the temperature, precipitation, and pressure time series, while the dynamical model CCLM performed better for the geopotential. In the higher-order comparison (C^*), STARS is superior for all variables in austral winter (JJA), while CCLM scores almost comparably in DJF, except for sea level pressure.

While in most cases the conventional difference measures have been in agreement with H^* , there were also cases in which the network structure was better reproduced by a model which was less favored by a conventional measure or vice versa, most notably in the pressure and geopotential variables. Although the construction of climate networks, representing statistical associations within climatological fields, takes more effort than applying rather simple measures like E^* and F^* , these complementary findings demonstrate the novelty and justification for our approach.

The finding of STARS being superior to CCLM for surface variables, not only according to traditional but also to the new network-based difference measures, demonstrates the physical consistency of the statistical model. However, it should be noted that the outcome of this study does not imply a general superiority of statistical to dynamical climate modeling. If, for instance, the reference networks had not been constructed from the driving ERA-Interim reanalysis but from independent observational data, the model ranking might have been different. A study on this matter is underway.

This work was meant to highlight the potential of the methodology. To demonstrate the robustness of our results concerning the comparison of RCM performances, an inclusion of further statistical and dynamical RCMs as well as applications to other regions are required. Such a study could be carried out e.g. in the CORDEX framework. Also the focus on ensemble runs with varying model parameters could be interesting, possibly revealing the influence of specific model components on the network structure of the output and in turn allowing for model improvements.

Further development of the methodology will include the application of additional difference measures, e.g. other metrics on adjacency matrices and the comparison of other network measures like betweenness or closeness to reveal more features of the higher-order network structure

differences. Finally, local versions of network difference measures shall be developed in order to add spatial detail to the comparison.

Acknowledgments This paper was developed within the scope of the IRTG 1740/TRP 2011/50151-0, funded by the DFG/FAPESP. Furthermore, this work has been financially supported by the Leibniz Society (project ECONS), and the Stordalen Foundation (JFD). For certain calculations, the software packages `pyunicorn` (Donges et al. 2013a) and `igraph` (Csárdi and Nepusz 2006) were used. The authors would like to thank Manoel F. Cardoso, Niklas Boers, and the reviewers for helpful comments on the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Baker IT, Prihodko L, Denning AS, Goulden M, Miller S, da Rocha HR (2008) Seasonal drought stress in the Amazon: reconciling models and observations. *J Geophys Res* 113(G1):G00B01. doi:10.1029/2007JG000644
- Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T (2011) Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Mon Weather Rev* 139(12):3887–3905. doi:10.1175/MWR-D-10-05013.1
- Barnett L, Di Paolo E, Bullock S (2007) Spatially embedded random networks. *Phys Rev E* 76(5):056115. doi:10.1103/PhysRevE.76.056115
- Bechtold P, Köhler M, Jung T, Doblas-Reyes F, Leutbecher M, Rodwell MJ, Vitart F, Balsamo G (2008) Advances in simulating atmospheric variability with the ECMWF model: from synoptic to decadal time-scales. *Q J R Meteorol Soc* 134(634):1337–1351. doi:10.1002/qj.289
- Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D (2006) Complex networks: structure and dynamics. *Phys Rep* 424(4–5):175–308. doi:10.1016/j.physrep.2005.10.009
- 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. doi:10.1002/grl.50681
- Böhm U, Gerstengarbe FW, Hauffe D, Kücken M, Österle H, Werner PC (2003) Dynamic regional climate modeling and sensitivity experiments for the northeast of Brazil. In: Gaiser T, Krol M, Frischkorn H, de Araújo JC (eds) *Global change and regional impacts*. Springer, Berlin, pp 153–170
- Bott A (1989) A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes. *Mon Weather Rev* 117(5):1006–1016
- Boulton CA, Good P, Lenton TM (2013) Early warning signals of simulated Amazon rainforest dieback. *Theor Ecol* 6(3):373–384. doi:10.1007/s12080-013-0191-7
- Carvalho L, Jones C, Liebmann B (2004) The South Atlantic convergence zone: intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *J Clim* 17:88–108
- Ciliberti S, Martin OC, Wagner A (2007) Innovation and robustness in complex regulatory gene networks. *Proc Natl Acad Sci* 104(34):13591–13596. doi:10.1073/pnas.0705396104
- Cohen R, Havlin S (2008) *Complex networks: structure, stability and function*. Cambridge University Press, Cambridge
- Cook B, Zeng N, Yoon JH (2012) Will Amazonia dry out? Magnitude and causes of change from IPCC climate model projections. *Earth Interact* 16(3):1–27. doi:10.1175/2011EI398.1
- Cook KH, Vizy EK (2008) Effects of twenty-first-century climate change on the Amazon rain forest. *J Clim* 21(3):542–560. doi:10.1175/2007JCLI1838.1
- Csárdi G, Nepusz T (2006) The `igraph` software package for complex network research. *InterJ Complex Syst* 1695
- Davin EL, Seneviratne SI (2011) Role of land surface processes and diffuse/direct radiation partitioning in simulating the European climate. *Biogeosci Discuss* 8(6):11601–11630. doi:10.5194/bgd-8-11601-2011
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen I, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553–597. doi:10.1002/qj.828
- Dobler A, Ahrens B (2010) Analysis of the Indian summer monsoon system in the regional climate model COSMO-CLM. *J Geophys Res* 115(D16):101. doi:10.1029/2009JD013497
- Donges J, Heitzig J, Runge J, Schultz HCH, Wiedermann M, Zech A, Feldhoff JH, Rheinwalt A, Kutza H, Radebach A, Marwan N, Kurths J (2013a) Advanced functional network analysis in the geosciences: the `pyunicorn` package. In: *EGU General Assembly*, vol 15, p 3558
- Donges JF, Zou Y, Marwan N, Kurths J (2009a) Complex networks in climate dynamics. *Eur Phys J Spec Top* 174(1):157–179. doi:10.1140/epjst/e2009-01098-2
- Donges JF, Zou Y, Marwan N, Kurths J (2009b) The backbone of the climate network. *Europhys Lett* 87(4):48007. doi:10.1209/0295-5075/87/48007
- Donges JF, Petrova I, Loew A, Marwan N, Kurths J (2013b) Relationships between eigen and complex network techniques for the statistical analysis of climate data. *Rev arxiv13056634 [physicsdata-an]* arXiv:1305.6634v1
- Efron B (1979) Bootstrap methods: another look at the jackknife. *Ann Stat* 7(1):1–26
- Erdős P, Rényi A (1959) On random graphs I. *Publ Math Debrecen* 6:290–297
- Fischer T, Menz C, Su B, Scholten T (2013) Simulated and projected climate extremes in the Zhujiang river basin, South China, using the regional climate model COSMO-CLM. *Int J Climatol*. doi:10.1002/joc.3643
- Fountalis I, Bracco A, Dovrolis C (2013) Spatio-temporal network analysis for studying climate patterns. *Clim Dyn*. doi:10.1007/s00382-013-1729-5
- Giorgi F, Jones C, Asrar GR (2009) Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bull* 58(3):175–183
- Golding BW (1998) Nimrod: a system for generating automated very short range forecasts. *Meteorol Appl* 5(1):1–16. doi:10.1017/S1350482798000577
- Gozolchiani A, Havlin S, Yamasaki K (2011) Emergence of El Niño as an autonomous component in the climate network. *Phys Rev Lett* 107(14):148501. doi:10.1103/PhysRevLett.107.148501
- Gutsch M, Lasch P, Suckow F, Reyer C (2011) Management of mixed oak-pine forests under climate scenario uncertainty. *For Syst* 20(3):453–463
- Hamming RW (1950) Error detecting and error correcting codes. *Bell Syst Tech J* 29(2):147–160

- Handorf D, Dethloff K (2012) How well do state-of-the-art atmosphere-ocean general circulation models reproduce atmospheric teleconnection patterns? *Tellus A* 1:1–27
- 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
- Hohenegger C, Brockhaus P, Bretherton CS, Schär C (2009) The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *J Clim* 22(19):5003–5020
- Huff FA, Shipp WL (1969) Spatial correlations of storm, monthly and seasonal precipitation. *J Appl Meteorol* 8(4):542–550
- Jaeger EB, Anders I, Luthi D, Rockel B, Schar C, Seneviratne SI (2008) Analysis of ERA40-driven CLM simulations for Europe. *Meteorol Z* 17(4):349–367. doi:10.1127/0941-2948/2008/0301
- Jones P (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon Weather Rev* 127(3):2204–2210
- Lange S, Rockel B, Volkholz J, Bookhagen B (2014) Regional climate model sensitivities to parametrizations of convection and non-precipitating subgrid-scale clouds over South America. *Clim Dyn*. doi:10.1007/s00382-014-2199-0
- Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ (2008) Tipping elements in the earth's climate system. *Proc Natl Acad Sci* 105(6):1786–1793. doi:10.1073/pnas.0705414105
- Li Z, Brissette F, Chen J (2012) Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in Nordic watersheds. *Hydrol Process*. doi:10.1002/hyp.9499
- Liu XD, Osher S, Chan T (1994) Weighted essentially non-oscillatory schemes. *J Comput Phys* 115(1):200–212. doi:10.1006/jcph.1994.1187
- Ludescher J, Gozolchiani A, Bogachev MI, Bunde A, Havlin S, Schellnhuber HJ (2013) Improved El Niño forecasting by cooperativity detection. *Proc Natl Acad Sci*. doi:10.1073/pnas.1309353110
- Lutz J, Volkholz J, Gerstengarbe FW (2013) Climate projections for southern Africa using complementary methods. *Int J Clim Chang Strateg Manag* 5(2):130–151. doi:10.1108/17568691311327550
- Malik N, Bookhagen B, Marwan N, Kurths J (2011) Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Clim Dyn* 39(3–4):971–987. doi:10.1007/s00382-011-1156-4
- Marengo J, Soares W (2004) Climatology of the low-level jet east of the Andes as derived from the NCEP-NCAR reanalyses: Characteristics and temporal variability. *J Clim* 17(12):2261–2280
- Marengo JA, Ambrizzi T, da Rocha RP, Alves LM, Cuadra SV, Valverde MC, Torres RR, Santos DC, Ferraz SET (2009) Future change of climate in South America in the late twenty-first century: intercomparison of scenarios from three regional climate models. *Clim Dyn* 35(6):1073–1097. doi:10.1007/s00382-009-0721-6
- Marengo JA, Chou SC, Kay G, Alves LM, Pesquero JF, Soares WR, Santos DC, Sueiro G, Betts R, Chagas DJ, Gomes JL, Bustamante JF, Tavares P (2011) Development of regional future climate change scenarios in South America using the Eta CPTec/HadCM3 climate change projections: climatology and regional analyses for the Amazon, São Francisco and the Paraná River basins. *Clim Dyn* 38(9–10):1829–1848. doi:10.1007/s00382-011-1155-5
- Matsoukas C, Islam S, Rodriguez-Iturbe I (2000) Detrended fluctuation analysis of rainfall and streamflow time series. *J Geophys Res* 105(D23):29165–29172. doi:10.1029/2000JD900419
- Mendes D, Souza EP, Ja Marengo, Mendes MCD (2009) Climatology of extratropical cyclones over the South American-southern oceans sector. *Theor Appl Climatol* 100(3–4):239–250. doi:10.1007/s00704-009-0161-6
- Newman M (2009) *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
- Nikulin G, Jones C, Giorgi F, Asrar G, Büchner M, Cerezo-Mota R, Christensen OB, Déqué M, Fernandez J, Hänsler A, van Meijgaard E, Samuelsson P, Sylla MB, Sushama L (2012) Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *J Clim* 25(18):6057–6078
- Orlowsky B, Fraedrich K (2008) Upscaling European surface temperatures to North Atlantic circulation-pattern statistics. *Int J Climatol*. doi:10.1002/joc
- Orlowsky B, Gerstengarbe FW, Werner PC (2008) A resampling scheme for regional climate simulations and its performance compared to a dynamical RCM. *Theor Appl Climatol* 92(3–4):209–223. doi:10.1007/s00704-007-0352-y
- Orlowsky B, Bothe O, Fraedrich K, Gerstengarbe FW, Zhu X (2010) Future climates from bias-bootstrapped weather analogs: an application to the Yangtze River Basin. *J Clim* 23(13):3509–3524. doi:10.1175/2010JCLI3271.1
- Panitz HJ, Dosio A, Büchner M, Lüthi D, Keuler K (2013) COSMO-CLM (CCLM) climate simulations over CORDEX-Africa domain: analysis of the EAR-Interim driven simulations at 0.44° and 0.22° resolution. *Clim Dyn* 1–24. doi:10.1007/s00382-013-1834-5
- Peters O, Hertlein C, Christensen K (2001) A complexity view of rainfall. *Phys Rev Lett* 88:018701. doi:10.1103/PhysRevLett.88.018701
- Reyer C, Lasch-Born P, Suckow F, Gutsch M, Murawski A, Pilz T (2013) Projections of regional changes in forest net primary productivity for different tree species in Europe driven by climate change and carbon dioxide. *Ann For Sci*. doi:10.1007/s13595-013-0306-8
- Rheinwält A, Marwan N, Kurths J, Werner P, Gerstengarbe FW (2012) Boundary effects in network measures of spatially embedded networks. *Europhys Lett* 100(2):28002. doi:10.1209/0295-5075/100/280021
- Rockel B, Geyer B (2008) The performance of the regional climate model CLM in different climate regions, based on the example of precipitation. *Meteorol Z* 17(4):487–498
- Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). *Meteorol Z* 17(4):347–348. doi:10.1127/0941-2948/2008/0309
- Schneider EK, DeWitt DG, Rosati A, Kirtman BP, Ji L, Tribbia JJ (2003) Retrospective ENSO forecasts: sensitivity to atmospheric model and ocean resolution. *Mon Weather Rev* 131(12):3038–3060
- Solman SA, Sanchez E, Samuelsson P, Rocha RP, Li L, Marengo J, Pessacg NL, Remedio aRC, Chou SC, Berbery H, Treut H, Castro M, Jacob D (2013) Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-Interim reanalysis: model performance and uncertainties. *Clim Dyn*. doi:10.1007/s00382-013-1667-2
- Sommeria G, Deardorff J (1977) Subgrid-scale condensation in models of nonprecipitating clouds. *J Atmos Sci* 34(2):344–355
- Steinhaeuser K, Tsonis AA (2013) A climate model intercomparison at the dynamics level. *Clim Dyn*. doi:10.1007/s00382-013-1761-5
- Stenseth NC, Mysterud A, Ottersen G, Hurrell JW, Chan KS, Lima M (2002) Ecological effects of climate fluctuations. *Science* 297(5585):1292–6. doi:10.1126/science.1071281
- Stevenson S, Fox-Kemper B, Jochum M, Neale R, Deser C, Meehl G (2012) Will there be a significant change to El Niño in the

- twenty-first century? *J Clim* 25(6):2129–2145. doi:[10.1175/JCLI-D-11-00252.1](https://doi.org/10.1175/JCLI-D-11-00252.1)
- Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (2013) IPCC, 2013: summary for policymakers. In: climate change 2013: the physical science basis. Contribution of working group I to the 5th assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge (in press)
- Stoner AMK, Hayhoe K, Wuebbles DJ (2009) Assessing general circulation model simulations of atmospheric teleconnection patterns. *J Clim* 22(16):4348–4372. doi:[10.1175/2009JCLI2577.1](https://doi.org/10.1175/2009JCLI2577.1)
- Strogatz SH (2001) Exploring complex networks. *Nature* 410(6825):268–276
- Trenberth K (1997) The definition of El Niño. *Bull Am Meteorol Soc* 78(12):2771–2777
- Tsonis AA, Roebber PJ (2004) The architecture of the climate network. *Phys A Stat Mech Appl* 333:497–504. doi:[10.1016/j.physa.2003.10.045](https://doi.org/10.1016/j.physa.2003.10.045)
- Tsonis AA, Swanson K, Kravtsov S (2007) A new dynamical mechanism for major climate shifts. *Geophys Res Lett* 34(13):L13705. doi:[10.1029/2007GL030288](https://doi.org/10.1029/2007GL030288)
- van der Mheen M, Ha Dijkstra, Gozolchiani A, den 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](https://doi.org/10.1002/grl.50515)
- Vera C, Higgins W, Amador J (2006) Toward a unified view of the American monsoon systems. *J Clim* 19:4977–5000
- Wagner S, Fast I, Kaspar F (2011) Climatic changes between 20th century and pre-industrial times over South America in regional model simulations. *Clim Past Discuss* 7(5):2981–3022. doi:[10.5194/cpd-7-2981-2011](https://doi.org/10.5194/cpd-7-2981-2011)
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393(6684):440–2. doi:[10.1038/30918](https://doi.org/10.1038/30918)
- Werner P, Gerstengarbe F (1997) Proposal for the development of climate scenarios. *Clim Res* 8:171–182. doi:[10.3354/cr008171](https://doi.org/10.3354/cr008171)
- 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(2):28007. doi:[10.1209/0295-5075/102/28007](https://doi.org/10.1209/0295-5075/102/28007)
- Yamasaki K, Gozolchiani A, Havlin S (2008) Climate networks around the globe are significantly affected by El Niño. *Phys Rev Lett* 100(22):1–4. doi:[10.1103/PhysRevLett.100.228501](https://doi.org/10.1103/PhysRevLett.100.228501)
- Zahn M, von Storch H (2008) A long-term climatology of North Atlantic polar lows. *Geophys Res Lett* 35(22). doi:[10.1029/2008GL035769](https://doi.org/10.1029/2008GL035769)
- Zhou C, Zemanová L, Zamora G, Hilgetag C, Kurths J (2006) Hierarchical organization unveiled by functional connectivity in complex brain networks. *Phys Rev Lett* 97(23):238103. doi:[10.1103/PhysRevLett.97.238103](https://doi.org/10.1103/PhysRevLett.97.238103)