A complex network representation of wind flows

Maximilian Gelbrecht, Niklas Boers, and Jürgen Kurths

Articles you may be interested in

A complex network representation of wind flows

Maximilian Gelbrecht,[1,2,a)] Niklas Boers,[1,3] and Jürgen Kurths[1,2,4,5]
[1]*Potsdam Institute for Climate Impact Research, Potsdam, Germany*
[2]*Department of Physics, Humboldt University, Berlin, Germany*
[3]*Geosciences Department and Laboratoire de Météorologie Dynamique, École Normale Supérieure, Paris, France*
[4]*Department of Control Theory, Nizhny Novgorod State University, Nizhny Novgorod, Russia*
[5]*Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen, United Kingdom*

Climate networks have proven to be a valuable method to investigate spatial connectivity patterns of the climate system. However, so far such networks have mostly been applied to scalar observables. In this study, we propose a new method for constructing networks from atmospheric wind fields on two-dimensional isobaric surfaces. By connecting nodes along a spatial environment based on the local wind flow, we derive a network representation of the low-level circulation that captures its most important characteristics. In our approach, network links are placed according to a suitable statistical null model that takes into account the direction and magnitude of the flow. We compare a simulation-based (numerically costly) and a semi-analytical (numerically cheaper) approach to determine the statistical significance of possible connections, and find that both methods yield qualitatively similar results. As an application, we choose the regional climate system of South America and focus on the monsoon season in austral summer. Monsoon systems are generally characterized by substantial changes in the large-scale wind directions, and therefore provide ideal applications for the proposed wind networks. Based on these networks, we are able to reveal the key features of the low-level circulation of the South American Monsoon System, including the South American Low-Level Jet. Networks of the dry and the wet season are compared with each other and their differences are consistent with the literature on South American climate. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4977699]

It has been shown recently that complex networks provide an important tool for the statistical analysis of the spatial characteristics of climatic variability. They thus constitute important means to improve our understanding of the climate system like, e.g., the impact of the El Niño Southern Oscillation. In general terms, each node of a climate network represents a geographic area and its climatic variability. It is thus a small dynamical subsystem on its own. By various—mostly statistical—means, we can quantify the similarity of the behavior between these subsystems such as, e.g., correlations between time series measured for each of the subsystems. Nodes with strong similarities are then linked to form a network. In this way, networks can be used to represent very complex interactions within the climate system in a mathematical model that is comparably easy to evaluate. Most existing climate network approaches focus on scalar observables like surface temperature, pressure, or precipitation. Here, we propose a new method to construct networks from atmospheric wind fields on isobaric surfaces. We focus on the directed properties of this two-dimensional vector field by connecting nodes along a spatial environment based on the magnitude and direction of the wind. We compare a simulation-based and semi-analytical approach to ensure the statistical significance of these possible connections. Monsoon systems are an ideal field of application for these networks. The large-scale wind direction in a monsoon region changes considerably between the dry and wet season. Therefore, we choose to apply the proposed methodology to investigate South American climate and in particular, the South American Monsoon System (SAMS). Our results show that important features of the South American climate system like the South American Low-Level Jet (SALLJ) are revealed by the wind networks. Furthermore, intraseasonal variations between the dry and wet seasons are accurately and consistently captured by the streamflow networks introduced here.

I. INTRODUCTION

Complex Network have only recently become a method for climate data analysis. Tsonis and Roebber[1] proposed to compute networks of the gridded 500 hPa geopotential height data with the linear Pearson correlation coefficient to define the links. This sparked a lot of interest in the possibilities of applying climate networks to different climatic observables. Networks based on the surface temperature and rainfall have been investigated, as well as networks that used other measures of statistical interdependence like mutual information and rank correlation.[2–4] Complex networks proved to be a suitable tool to uncover and visualize

a)gelbrecht@pik-potsdam.de

phenomena like the El Niño Southern Oscillation.[5,6] Networks representing the synchronizations of extreme rainfall events were able to uncover a prediction mechanism for floods in the South American Andes.[7] Additionally, multiple temperature networks of different geographical regions have been compared,[5] and coupled networks of the geopotential height in different altitudes were used to investigate the vertical structure of the atmosphere.[8] More recently, networks were employed to investigate fluid flows as well. Promising analytic and data-driven approaches to research ocean circulation patterns have been introduced,[9,10] as well as investigations of the dynamics of correlation-based climate networks with underlying flow fields.[11,12]

One of the few climatic observables that have so far not been considered for a complex network analysis is the atmospheric wind field. Wind data provide crucial information about Earth's atmospheric circulation. Climatic phenomena like the monsoons are characterized by large-scale changes in the wind directions. Furthermore, global-scale phenomena like the El Niño Southern Oscillation also directly impact regional wind fields. Wind is, in contrast to other quantities that were previously used to construct climate networks like temperature or rainfall, not a scalar but a three-dimensional observable. Large-scale atmospheric motions are dominated by a balance between the Coriolis and pressure-gradient forces. This *geostrophic equilibrium* causes the wind vectors to be approximately parallel with isobaric surfaces as long as the Coriolis force has a horizontal component, which is the case everywhere except along the equator.[13] For our application, we are interested mainly in advective processes (i.e., horizontal transport of air masses) associated with the South American Monsoon System (SAMS), because these determine their characteristic, large-scale coupling patterns which we intend to analyze with our streamflow network approach. For the case at hand, it is therefore sufficient to investigate the two-dimensional wind on isobaric surfaces, even though heat transfer causes vertical winds, which play a key role for the formation of the large-scale circulation cells of earth's atmosphere. This vertical motion is usually slower than the geostrophic motion, which at 850 hPa is typically of $\mathcal{O}(10 \text{ m s}^{-1})$. The restriction to two-dimensional isobaric surfaces also allows for an easier interpretation of the methods and results: although the three-dimensional field could technically be analyzed with the method introduced below, it would make the results much harder to visualize on two-dimensional geographic maps.

In this study, we introduce a new construction method for climate networks from two-dimensional flows like the low-level atmospheric wind field. The majority of the existing climate network construction methods are based on pairwise correlations of scalar observables. Of course, even from a two-dimensional observable, like the wind field on a given geopotential height, we could derive a network representation solely based on correlations. For example, correlations could be computed between the zonal and meridional components separately, or from the absolute values of the two-dimensional vectors. Several successful

approaches to construct networks from flows based on correlations have been introduced.[11,12] Complementarily, we intend here to introduce a method that directly employs the characteristics of multidimensional vector fields. Several approaches for a direct network representation of flows have been introduced, e.g., by Viebahn and Dijkstra[9] and Ser-Giacomi *et al.*[10] Viebahn and Dijkstra[9] investigated critical transitions of the wind-driven ocean circulation in a model basin by employing complex networks that were constructed with links along or across the streamlines of the flux. Ser-Giacomi *et al.*[10] used Lagrangian simulations to construct complex networks of the surface flow of the Mediterranean sea. Further studies investigated atmospheric blocking with these Lagrangian flow networks as well.[14] Here, we intend to propose a new construction method for the analysis of flow data. We will introduce a simple model of the local flow and propose two different approaches to ensure the significance of the links that we set. We will test and validate our method by applying it to observational wind data of the South American Monsoon System. As large scale changes in the wind field are a substantial feature of every monsoon system, they serve as an ideal application for the proposed methodology. Furthermore, since this climatic subsystem is already quite well-understood, it provides a well-suited real-world example to corroborate our methodology.

In the following, we will first provide a short overview of the South American Monsoon System (SAMS), which is our geographical region of application. Thereafter, we will introduce our methodological framework by giving a brief introduction into basic elements of network theory and the network measures used in this study. Next, different approaches to construct networks from wind data will be proposed and compared. Subsequently, we will show how these networks are applied to the SAMS and interpret our results in a climatological context.

A. The South American Monsoon System

Monsoon systems are usually characterized by a seasonal reversal in the large-scale circulation system that is accompanied by heavy precipitation. However, this change in the prevailing wind direction is not given in South America. This is why for a long time South America was not considered to exhibit a monsoon system. That changed when researchers started to investigate the *Bolivian high*, a high pressure system in the upper troposphere that forms in austral summer over the Altiplano Plateau, a plateau in the Andes of Bolivia and Peru in about 3.75 km height (see Figure 1).[15] In the lower troposphere a thermal low pressure system forms in the *Gran Chaco* region in northern Argentina, Paraguay, and southern Bolivia.

The trade winds dominate the low-level wind direction over South America. However, when removing the annual mean from the wind field component-wise, a change in sign between the austral summer and winter is immediately visible (Figure 2). In austral summer the north-easterly trade winds are enhanced. Deflected by the Andes, a strong north-westerly jet forms on the eastern slope of the Andes that then

FIG. 1. Overview of the South American Monsoon System (SAMS): The upper troposphere Bolivian high over the Altiplano Plateau (red area) and the thermal low over the lower troposphere of Gran Chaco (blue area) are key ingredients to the SAMS. The main convection area is the South Atlantic Convergence Zone (SACZ) spanning from the Amazon in south-east direction to the Atlantic. The incoming trade winds are deflected by the Andes and channeled into the South American Low-Level Jet (SALLJ). The SALLJ influences the SACZ and modulates precipitation in Southeastern South America (SESA).
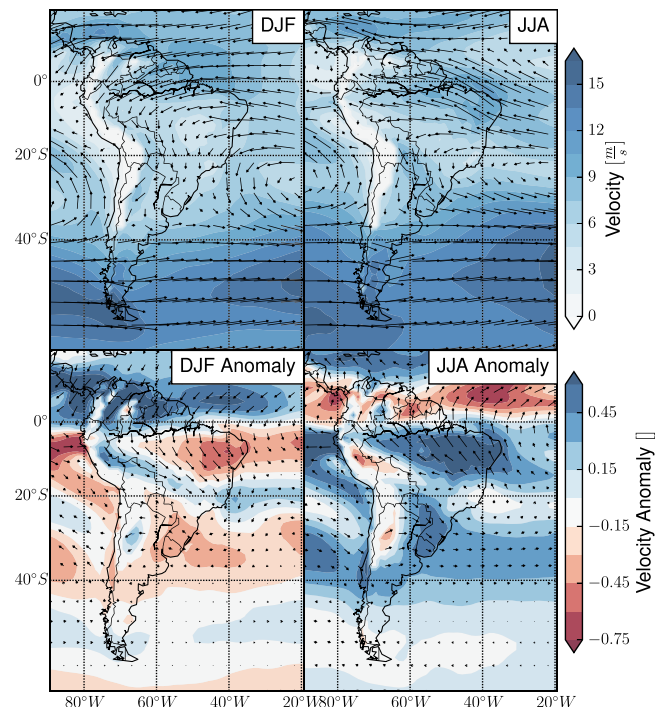

FIG. 2. Wind field of South America on the 850 hPa pressure surface during December, January, and February (DJF) and June, July, and August (JJA). The anomaly values were calculated by removing the annual mean and normalization with the standard deviation. The arrows indicate the average wind direction. Their length is proportional to the absolute value of the wind. During the wet monsoon phase (DJF) there are strong north-easterly trade winds and an enhanced SALLJ along the eastern slopes of the Andes, whereas the SALLJ is suppressed in the dry season. The difference between wet and dry seasons is even more obvious when investigating the anomaly wind field: the prevailing wind direction is reversed.

turns clockwise around the Gran Chaco low. In austral winter this anomalous flow reverses its direction, although the absolute wind fields are still north-westerly.[15]

The total precipitation is similar to other monsoon regions. However, the difference between the wet and the dry seasons is smaller than, e.g., in India or eastern Asia.[16] The largest difference in precipitation between the wet and dry season can be observed in the south-central Amazon basin where the wet season peaks from December to February with an average 300 mm precipitation per month and less than 20 mm per month from June to August.[16] The south-west extension of this area is referred to as *South Atlantic Convergence Zone* (SACZ).[17] Most of the convective activity of the SAMS occurs in this area.[17]

Moisture advected from the Amazon to *Southeastern South America* (SESA) by the *South American Low-Level Jet* (SALLJ) results in the frequent formation of some of the largest Mesoscale Convective Systems on Earth, which can cause heavy thunderstorms, extreme rainfall, and floods.[16,18,19] The SACZ exhibits a dipole-like behavior concerning the precipitation on various time scales.[18,20] In one phase of the dipole the SACZ is enhanced and precipitation in SESA is suppressed. The SALLJ is westward orientated (upper part of the SALLJ-arrow in Fig. 1). In the other phase, the SACZ is suppressed, the southward SALLJ is stronger, and rainfall is enhanced in SESA.[16,20] A strong southward SALLJ, the so-called Gran Chaco Jet, can cause heavy rainfall over central Argentina.[18]

II. DATA

For our calculations we use the MERRA IAU 2D atmospheric single-level diagnostics.[21] It provides us with datasets on a rectangular grid with a latitudinal resolution of $\Delta\lambda = \frac{1}{3}^\circ$ and a longitudinal resolution of $\Delta\Phi = \frac{2}{3}^\circ$. This implies that the grid is denser at higher latitudes due to the longitudes being closer to each other. The wind data comprise two-component time series, a zonal eastward component $u_i(t)$, and a meridional northward component $v_i(t)$. This study uses a dataset spanning from 1985 to 2010, consisting of daily data at a height of 850 hPa which corresponds roughly to 1.47 km (mean of the geopotential height H850 provided by MERRA for South America from 2000 to 2010[21]). All networks were computed with the regular wind time series, but not their anomalies.

All computations are performed with relatively large datasets that cover the geographic region enclosed by 0°W–110°W; 70°S–25°N in order to minimize the influence of possible boundary effects on the constructed networks. The network measures are then cropped to 20°W–90°W; 60°S–20°N for all figures shown. In total, the dataset encompasses $N = 33\,366$ grid cells, each possessing two time series with $N_t = 9496$ time steps. These grid cells will, in the following, be identified as the nodes of the networks. One of the main aims of this work is to identify wind patterns in the

SAMS; therefore, most of the following networks will be constructed with data that are restricted to the monsoon season from December to February (DJF).

III. METHODS

A. Network preliminaries

In this study, we investigate the spatially embedded directed networks $G = (\mathbf{N}, \mathbf{E})$ encompassing a set of N Nodes $\mathbf{N}$ and a set of m directed, unweighted edges $\mathbf{E}$. The adjacency matrix of the network is given by

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge directed from node } j \text{ to } i \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The topological characteristics of the constructed networks encode the spatial coupling structure of the underlying wind data. We evaluate these topological characteristics by means of node-based network measures. The *indegree* and *outdegree* of a node are defined as

$$k_i^{\text{in}} = \sum_{j=1}^{N} A_{ij}, \quad (2)$$

$$k_i^{\text{out}} = \sum_{j=1}^{N} A_{ji}. \quad (3)$$

The proposed networks are spatially embedded, meaning that every node represents a certain area. Because the data are provided on a rectangular geographical grid, the size of this area depends on the latitude of the node. We can correct this for the degree by using the *out-* and *in-area weighted connectivity*,[22] defined as

$$\text{OAWC}_i = \frac{\sum\limits_{j=1}^{N} A_{ji} \cos \lambda_j}{\sum\limits_{j=1}^{N} \cos \lambda_j}, \quad (4)$$

and accordingly for the *in-area weighted connectivity (IAWC)*. The difference between out- and in-area weighted connectivity

$$D_i = \text{OAWC}_i - \text{IAWC}_i, \quad (5)$$

provides us with deeper insights into the network structure, with particular focus on directional asymmetries. By construction, the networks introduced in Section III B tend to form shortest network paths (so-called geodesic paths) along important wind transport routes. The *betweenness centrality*,[23] which is defined by

$$B_i = \sum_{jk} \frac{n_{jk}^{(i)}}{g_{jk}}, \quad (6)$$

is the key measure to evaluate a node's importance for the path structure of the network. In this sum, $n_{jk}^{(i)}$ is 1 if there is

a geodesic path between j and k running through i and 0 otherwise. This number is divided by the total number of geodesic paths between j and k: g_{jk}. Recall that a geodesic path in a network is defined as the shortest sequence of links to connect two given nodes.

B. Streamflow networks

So far, most climate networks focused on statistical co-variability measures to compare two time series, placing a link between the corresponding nodes if their co-variability (quantified, e.g., in terms of the Pearson correlation coefficient) is particularly high. For the case of the two-dimensional isobaric wind field, we propose here an alternative approach that focuses on the directed properties of this two-dimensional vector field: Wind mainly transports air and moisture from one location to another. Given the wind data at a specific location, we can calculate how far the wind transports matter during a specific time and link all nodes reached by the wind. We call networks based on this idea subsequently *streamflow networks*. They will most suitably be analyzed by measures that focus on the paths which these streamflow links create. We will in the following present two ways to construct such networks: One where the corresponding significance tests are based on numerical simulations, and one where they are based on an analytical approximation.

1. Simulation method

To implement the basic idea of networks directly representing the streamflow of the wind fields, we first focus on a single node i and a single step t in the time series. Introducing a fixed travel time T, the distance the wind can reach is $|(u_i(t), v_i(t))| \cdot T$. We add an additional uncertainty $u_{d_i}(\lambda_i)$. This uncertainty is chosen to ensure that two nodes will be linked when the distance the wind can reach passed more than halfway through the distance between two neighboring nodes. This way, we lower the chances of having dead ends in the paths of the network. Furthermore, adding such an uncertainty corresponds to the idea that the nodes represent the entire grid cell area surrounding them, and not just their precise location. For a given rectangular grid, this uncertainty will therefore depend on the latitude λ of node i. We define this uncertainty as the average of the distance to the next neighbor in the meridional direction d_m and the zonal direction d_z

$$u_{d_i}(\lambda_i) = \frac{1}{2}(d_m + d_z(\lambda_i)). \quad (7)$$

The distance in the zonal direction depends on the latitude, whereas the distance in the meridional direction is a constant for a given rectangular grid. The expression

$$d_i(t) = |(u_i(t), v_i(t))| \cdot T + u_{d_i}(\lambda), \quad (8)$$

is thus the distance considered to be covered by the given wind vector. For grids with a very fine resolution, u_{d_i} will be small in comparison with $|(u_i(t), v_i(t))| \cdot T$. Note, however, that it would become more important for coarser grids.
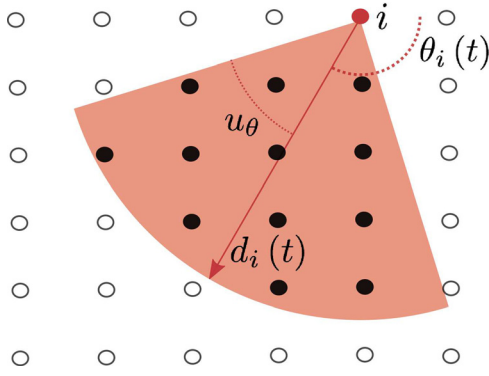
FIG. 3. Basic concept of the streamflow networks: The black filled nodes are considered to be reachable by the wind of the red node. The covered distance $d_i(t)$ is directly proportional to the velocity of the wind; the vector has the same direction as the original wind vector $(u_i(t), v_i(t))$, and the angle uncertainty u_θ is a parameter of the network. The red shaded area indicates the central angle referred to in the text.

Focusing on the actual wind flow, we also have to account for the direction of the wind. The arctangent of the meridional and zonal wind components yields the wind angle $\theta_i(t)$. Next, we also introduce an angle uncertainty u_θ. We define a circle section with the spatial position of node i as its center for every time step. The radius is $d_i(t)$, the central angle is $2 \cdot u_\theta$, and $\theta_i(t)$ its direction (see Fig. 3). All nodes that lie within this circle sector are considered to be reachable by the wind at time t. The angle uncertainty u_θ introduced here remains a parameter to be set. It governs how wide the circle section is and therefore directly influences the degree centrality and link density of the resulting network (see Section III C below regarding suitable choices of this parameter).

Next, we apply this method to the whole time series at node i and count how often each other node j lies within the circle sections determined for each time step t. This provides the hit counts h_{ij}. In order to judge whether these hit counts are significant or not, we compute N_{stat} surrogate hit counts $\eta_{ij}^{(k)}$ that tell us how often node j lies within the circle section spanned by node i in the k-th surrogate. The surrogates are based on uniformly distributed angles as their direction and velocities drawn from the velocity time series of the node. These hit counts are approximately normally distributed

(see Figure 10) and therefore it seems reasonable to calculate the mean $\mu_{\eta,ij}$ and standard deviation $\sigma_{\eta,ij}$ of $\eta_{ij}^{(k)}$ with respect to k to estimate a significance threshold for h_{ij}. Thus, we define the adjacency matrix of the network as

$$A_{ij} = \begin{cases} 1 & \text{if } h_{ij} > \mu_{\eta,ij} + n \cdot \sigma_{\eta,ij} \quad \text{with } n \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The parameter n is either 2 for a 95.5% significance or 3 for a 99.7% significance, assuming a normal distribution. A pseudocode representation of the complete algorithm is shown in the Appendix. It is tailored to be low in memory consumption and flawlessly parallelizable for a fast computation.

The resulting network is directed. Its outdegree thus strongly depends on the absolute value of the wind velocity and does not provide much additional information. A path-based measure like the betweenness centrality seems to be ideal for these kinds of networks. The resulting networks will exhibit strong boundary effects due to the underlying wind fields and the directed nature of the network.

2. Semi-analytic method

The proposed method to calculate the *streamflow networks* has one significant flaw: the computation of the surrogates is very costly and scales with $\mathcal{O}(N_{stat} \cdot N_t \cdot N_{Node})$ in computation time for every node. Using the same basic concept (see Fig. 3), we can, alternatively, analytically estimate how likely it is that a hit count h_{ij} occurs just by chance.

The probability for a single hit from node i to node j can be assumed to be binomially distributed because either a hit occurs with the probability p_{ij} or it does not with probability $1 - p_{ij}$. Hence, we compute the probability for k hits to occur with a cumulative binomial distribution

$$P_{ij}(k \geq h_{ij}) = 1 - P_{ij}(k < h_{ij}), \tag{10}$$

$$= 1 - \sum_{k=0}^{h_{ij}} \binom{N_t}{k} p_{ij}^k (1 - p_{ij})^{N_t - k}. \tag{11}$$

Looking at time series with large N_t, the binomial distribution can be approximated either by a Poisson distribution for small $N_t p_{ij}$ or a normal distribution otherwise

$$P_{ij}(k \geq h_{ij}) = \begin{cases} 1 - \sum_{k=0}^{h_{ij}} \dfrac{(N_t p_{ij})^k \exp(-N_t p_{ij})}{k!} & N_t p_{ij} \text{ small} \\ 1 - \dfrac{1}{\sqrt{2\pi N_t p_{ij}(1 - p_{ij})}} \sum_{k=0}^{h_{ij}} \exp\left(\dfrac{-(k - N_t p_{ij})^2}{2 N_t p_{ij}(1 - p_{ij})}\right) & \text{otherwise.} \end{cases} \tag{12}$$

Both situations will occur when looking at large datasets with big map sections. The approximation with the Poisson distribution is typically used for pairs of nodes with a large spatial distance.

We introduced the streamflow network method by first looking at the absolute value of the velocity and then at its direction, one after the other. It is therefore reasonable to split the elementary probability p_{ij} into a velocity component $p_{|\mathbf{v}|_{ij}}$ and an angular component $p_{\theta_{ij}}$, too. In first order we may approximate these two probabilities as independent from each other. Hence, the elementary probability p_{ij} can be obtained via

$$p_{ij} \approx p_{|\mathbf{v}|_{ij}} \cdot p_{\theta_{ij}}. \tag{13}$$

The velocity probability $p_{|\mathbf{v}|_{ij}}$ is the probability to have a velocity large enough that the node j is reachable for wind from node i. Hence, it is the probability that $|\mathbf{v}|_i > \frac{d_{ij}}{T}$. This can be calculated using the velocity probability density function $p(|\mathbf{v}|)\mathrm{d}v$

$$p_{|\mathbf{v}|_{ij}} = \int_{\frac{d_{ij}}{T}}^{\infty} p(|\mathbf{v}|)\mathrm{d}v. \tag{14}$$

Given a discrete sample time series for $|\mathbf{v}(t)|_i$, this integral can be estimated by sorting the time series and then finding the value closest to $\frac{d_{ij}}{T}$. The relative position of this value in the sorted series is then the estimated value of the integral. This method does not actually try to calculate the probability density function. However, when computing the surrogates, we also do not compute the probability function as we only draw a random value from the time series. So $p_{|\mathbf{v}|_{ij}}$ is actually the probability to draw a velocity from the sample time series large enough for node i and j to be connected. The angular probability $p_{\theta_{ij}}$ is the probability to randomly draw a direction in the interval $[\theta_{ij} - u_\theta, \theta_{ij} + u_\theta]$. Assuming a uniform distribution for this random process as when calculating the surrogates before, this probability is just the fraction of the whole circle which the corresponding circle section occupies. Thus,

$$p_{\theta_{ij}} = \frac{u_\theta}{\pi}. \tag{15}$$

The probability $P_{ij}(k \geq h_{ij})$ can now be calculated using the elementary probability p_{ij}. It tells us how likely it is that h_{ij} (or more) hits occur just by chance. Thresholding this probability, we are able to get the adjacency matrix

$$A_{ij} = \begin{cases} 1 & \text{if } P_{ij}(k \geq h_{ij}) < \psi \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

It seems reasonable to choose the threshold ψ to fit the probability enclosed by $\mu_{\eta,ij} \pm n \cdot \sigma_{\eta,ij}$ when calculating the surrogates, so that both methods yield comparable networks.

This semi-analytic approach is just another way of calculating a streamflow network with similar properties as with the surrogates. It is an unweighted, directed network that emphasizes on the importance of paths through the network. The semi-analytic approach should save us a lot of computation time, though. The only computationally intensive task remaining is the integral for $p_{|\mathbf{v}|_{ij}}$. Solved by sorting

TABLE I. Overview over most used parameter choices for streamflow networks.

Travel time T	1 d
Angle uncertainty u_θ	$\frac{\pi}{6}$
Number of surrogates (simulation method) N_{stat}	200
Significance parameter n (simulation method)	2
Threshold probability ψ (semi-analytic method)	0.05

the time series, with, e.g., quicksort, and then searching for the position in the sorted array with a binary search, its computation time scales with $\mathcal{O}(N_t \cdot \log N_t \cdot \log N_t)$. This is much less than the computation of the surrogates for the simulation method.

C. Parameters used for the network construction

Both methods of calculating streamflow networks have most parameters in common (Table I). The travel time T and angle uncertainty u_θ are crucial for the basic concept of the streamflow networks. The travel time governs the radius of the circle section that is used to calculate the network links. It is therefore directly linked to the outdegree of the nodes and thus the link density of the whole network. Due to the additional distance uncertainty that was introduced, even small travel times T should allow to reasonably compute path-based measures like the betweenness because dead ends in the network are less likely to appear. Figure 4 shows the dependency of the link density on the travel time T. For both methods the link density increases monotonously, in an almost linear way for larger travel times. A travel distance of one day results in a link density large enough for the network to include enough interesting features and is therefore chosen for all following streamflow networks.

The angle uncertainty u_θ governs the central angle of the circle section and therefore directly affects the outdegree of
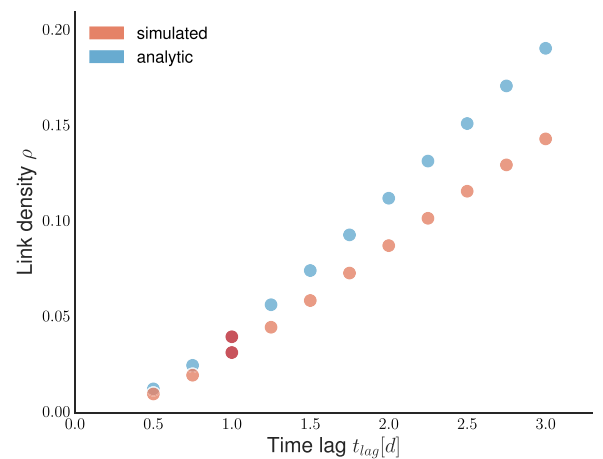

FIG. 4. Dependency of the link density of the delay time T for the simulated and the semi-analytic streamflow network. The calculations were done with the 1985–2010 MERRA dataset for South America and with $n = 2$ respectively $\psi = 0.05$. It shows that the link density increases for the semi-analytic method slightly stronger than for the simulation method. For both methods a delay time of one day (red dot) is picked for further computations.
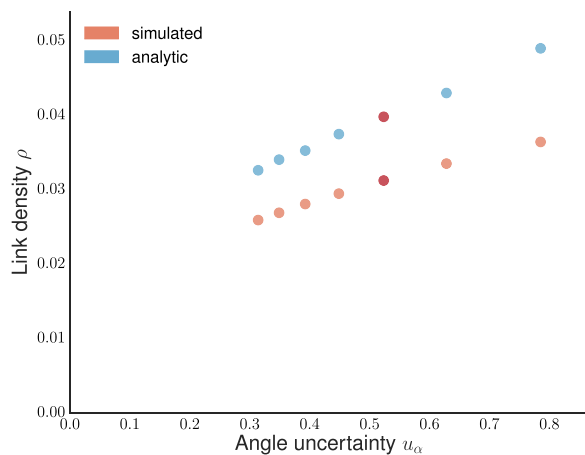
FIG. 5. Dependency of the link density of the angle uncertainty u_θ for the simulated and the semi-analytic streamflow network. The calculations were done with the 1985–2010 MERRA dataset for South America and with $n = 2$ respectively $\psi = 0.05$. It shows that the link density increases in an almost linear way for both methods. The red dots show the angle uncertainty $u_\theta = \pi/6$ picked for further computations.

all nodes as well. However, the angle uncertainty will also affect paths through the network and the overall structure: The extreme case of $u_\theta = \pi$ would, for example, result in an undirected network because the circle section has become a whole circle. We should hence choose an angle uncertainty that is large enough, such that enough connections are possible, but not too large, such that the directed nature of the streamflow networks is not lost. Figure 5 shows that the link density slowly and almost linearly increases with larger u_θ, in a very similar way for both construction methods. For both networks a fixed uncertainty of $u_\theta = \frac{\pi}{6}$ is chosen. Future work could investigate variable angle uncertainties, e.g., by relating them to the variance of the wind direction. However, sensitivity analyses with varying choices of u_θ reveal that the qualitative spatial structure of the networks is very robust to small changes of u_θ (see Figs. 11 and 12), which justifies to choose a fixed value of u_θ as a network parameter. Such a

fixed angle uncertainty leads to a significantly simpler and computationally faster model. Furthermore, the almost linear increase of the link density with larger time lags (Fig. 4) indicates that $u_\theta = \frac{\pi}{6}$ is a reasonable choice given the spatial grid resolution, since this increase would scale quadratically if the angle uncertainty were too large. In different settings and applications, however, we suggest to always explicitly check the robustness of the results to different choices of u_θ, and modify it in accordance to the local dispersions of the wind field if necessary.

Both of the statistical parameters n (simulation method) and ψ (semi-analytic method) will have a direct influence on the link density as well. However, we introduced both parameters as a threshold for statistical significance. Hence, both should be chosen in a way that a certain significance is ensured. It seems therefore convenient to set these parameters such that a 95% confidence level is obtained. An even higher value could result in too low link densities. Here, we choose to set both statistical parameters to a similar level of significance with $n = 2$, which represents 95.45% for normal distributed values, and $\psi = 5\%$. This leads to both methods using a comparable parameter set.

IV. RESULTS

We compute streamflow networks for the core monsoon season in South America (DJF) using the MERRA dataset of wind fields at 850 hPa introduced in Section II. Spatial distributions of network measures, computed for both the analytically and the numerically constructed network, can be seen in Figures 6–8.

In order to reduce boundary effects caused by the spatial embedding, the dataset encompasses $N = 33\,366$ nodes that cover 0°W–110°W; 70°S–25°N. The network measures calculated on this larger dataset are then cropped for the figures to cover 20°W–90°W; 60°S–20°N which corresponds to $N = 17\,066$ nodes.

The OAWC of both network types exhibits the largest values south of 40°S. However, these values are smaller at the
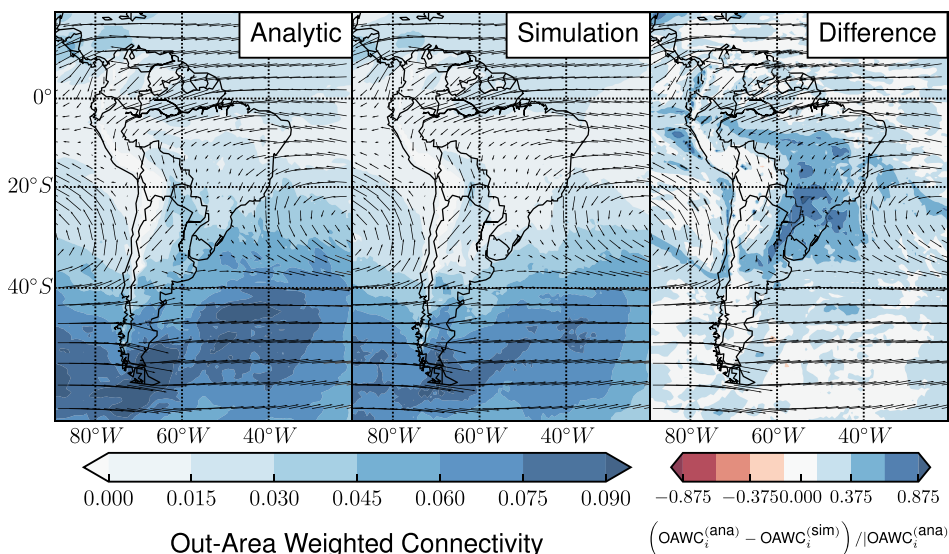

FIG. 6. Out-area weighted connectivity of streamflow networks representing the South American low-level circulation, for both the semi-analytic (a) and the simulation-based (b) network construction method. The relative difference $(\mathrm{OAWC}_i^{(\mathrm{ana})} - \mathrm{OAWC}_i^{(\mathrm{sim})})/|\mathrm{OAWC}_i^{(\mathrm{ana})}|$ between both networks is shown in (c). The strong eastward winds south of 40°S dominate the OAWC. Qualitatively both methods produce networks with similar OAWC. The largest difference can be found in central and southern Brazil.
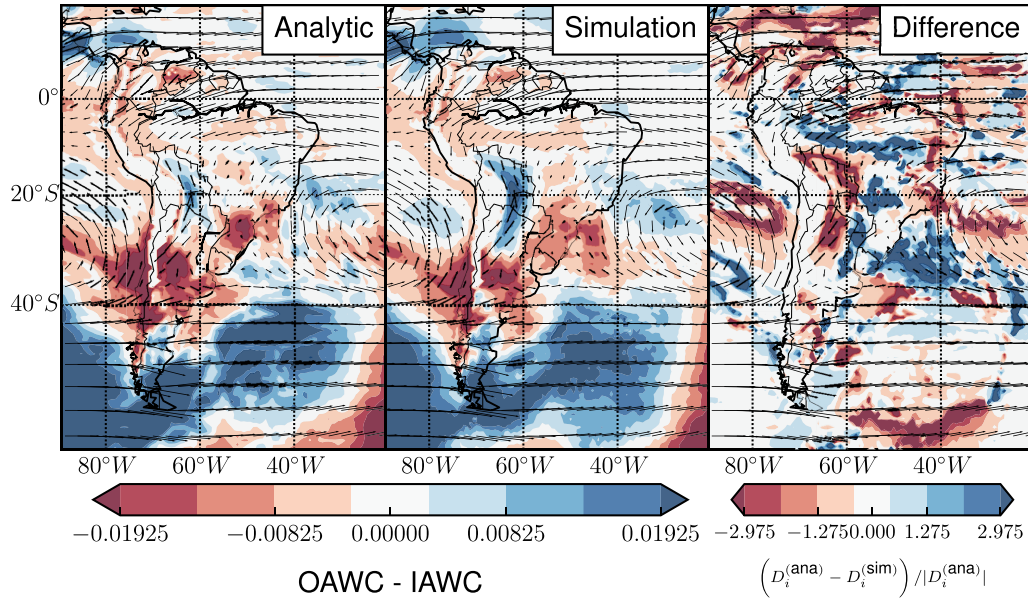
FIG. 7. Difference of out- and in-area weighted connectivity of streamflow networks representing the South American low-level circulation, for both the semi-analytic (a) and the simulation-based (b) network construction method. The relative difference $(D_i^{(ana)} - D^{(sim)}{}_i)/|D^{(ana)}{}_i|$ between both networks is shown in (c). While the high values south of 40°S are caused by the spatial embedding, the large values along the eastern slopes of the Andes in Bolivia and northern Argentina indicate jet-like behavior caused by the South American Low-Level Jet (SALLJ).

eastern than at the western boundary of the map section. Qualitatively, the spatial distribution of OAWC is very similar for both network construction methods. In fact, the relative difference between the semi-analytic and the simulated network is rather small. The semi-analytic network has a 0%–20% larger OAWC (and outdegree) for most parts of the network. The only larger area with an OAWC that is more

than 50% larger for the semi-analytic method can be found in southern Brazil close to the border with Paraguay. This region (10°–30°S, 40°–50°W) exhibits a relatively small average wind velocity of only $2.87\,\mathrm{m\,s^{-1}}$, compared with the average wind velocity of the complete map section, equal to $5.86\,\mathrm{m\,s^{-1}}$. The wind direction is changing in this area as there is an anticyclone east of it and the SALLJ west of it. Possible
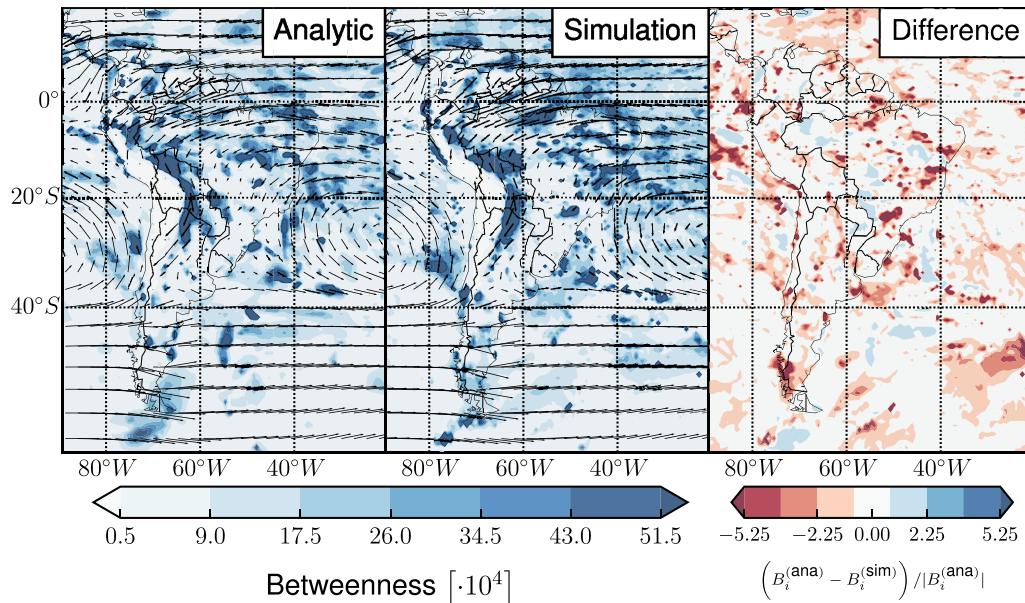

FIG. 8. Betweenness centrality of streamflow networks representing the South American low-level circulation, for both the semi-analytic (a) and the simulation-based (b) network construction method. The relative difference $(B_i^{(ana)} - B_i^{(sim)})/|B_i^{(ana)}|$ between both networks is shown in (c). Both methods exhibit the highest betweenness along the eastern slopes of the Andes. This corresponds to the South American Low-Level Jet (SALLJ), which is a key feature for the moisture transport of the South American Monsoon System. A major difference between both methods is the high betweenness of the simulation network along the Amazon river.
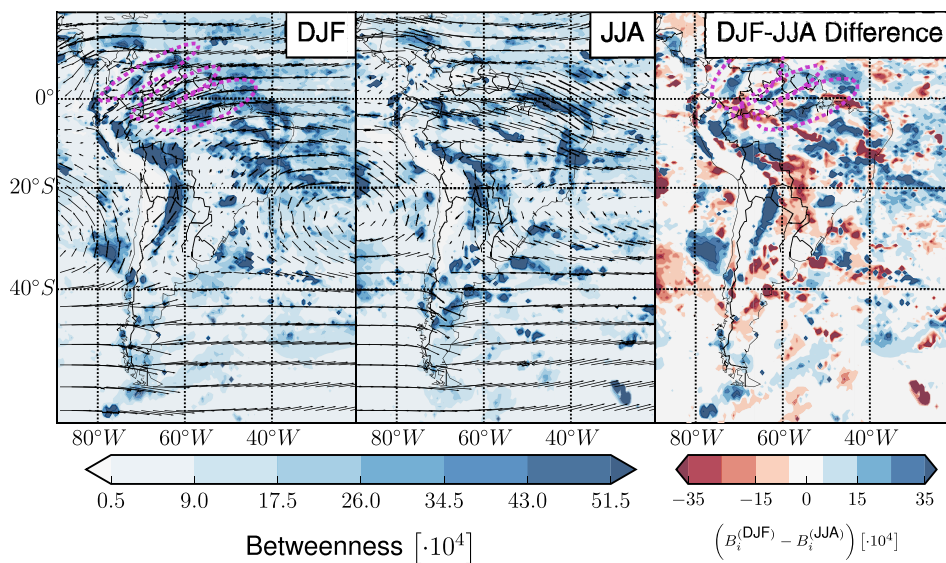
FIG. 9. Difference between wet and dry season of betweenness centrality of streamflow networks of South America, constructed with the numerical simulation method. Note the high betweenness during DJF along the eastern slopes of the Andes and along the three narrow bands marked in magenta in northern South America.

reasons for the relatively large differences between the simulation-based and analytical networks in this area are discussed below.

In order to investigate the directional asymmetries of the streamflow networks, we consider the difference between out- and in-area weighted connectivity. We find that the OAWC of both network types is much larger than the IAWC south-east of 50°S; 30°W, on both sides of the southern Andes in central Argentina and Chile, as well as in SESA. The OAWC is much larger over the Atlantic Ocean south of the continent as well as along the eastern slopes of the Andes in Bolivia, Paraguay, and northern Argentina. The area weighted connectivity difference D_i in this region is larger for the simulation method than for the semi-analytic one. Most part of northern and central South America exhibits a very similar OAWC and IAWC.

The betweenness centrality of both network types is high along the eastern slopes of the Andes as well as in large parts of eastern and northeastern Brazil. Relatively, the simulation-based network shows larger values in these regions and exhibits a more concise pattern. In particular, the simulation-based network has very high betweenness centrality along the Amazon river.

To investigate the seasonal dependence of these spatial patterns, we construct another simulation-based network for the austral winter season (JJA), which is also the dry season in most parts of South America. Figure 9 shows the difference of the betweenness centrality between this JJA-network and the simulation-based DJF-network discussed above. During DJF, the betweenness is much higher along the eastern slopes of the Andes and most parts of northern South America, especially along three narrow bands: one around Los Llanos (0°N–10°N, 60°W–75°W), one south of the Guyana Shield (0°N–5°N, 55°W–70°W), and one along the Amazon river (0°S–5°S, 50°W–65°W) (all are marked in Figure 9). These bands align very well with the average wind field in these areas as well. The JJA-network exhibits higher betweenness centrality in Paraguay and south-western Brazil.

V. DISCUSSION

In the following we will discuss the results of the streamflow networks. While some of the properties of the spatial distribution of the network measures shown before can be attributed to our own method and the boundaries of the map section we choose, other properties unveil important parts of the South American climate.

A. Methodological considerations

The OAWC of the streamflow networks is, by construction, directly related to the wind velocity. In the south of the map section strong eastward winds prevail. That is why the OAWC is the largest in this region. The OAWC of both networks types is overall very similar. The main difference (around 25°S, 55°N) is in a region with very low average wind velocities, east of the SALLJ and west of an anticyclone. It is not clear what exactly causes the difference between both network types in this region, apart from the approximations involved in the semi-analytic method. In particular, the assumption of the semi-analytic method that velocities and directions are independent might not hold true in this case. However, one should expect a low OAWC in this region due to the relatively low wind velocities. This argument favors the simulation-based network.

The streamflow networks exhibit strong boundary effects due to their spatial embedding. These effects are especially visible for the out-area weighted connectivity and area weighted connectivity difference (see Figures 6 and 7): Nodes in the south-eastern corner of the map section show a smaller OAWC than in the south-west, although the average wind velocities are similar in both regions. The wind field at these nodes is predominantly eastward and thus the nodes in the south-east would connect to nodes farther east, which are however not part of the considered map section anymore. For spatially embedded networks based on correlation measures, one can use spatially embedded random networks (SERN) to correct for these boundary effects.[24] However, this method is not applicable to the streamflow networks introduced here:

The nodes of streamflow networks are connected to other nodes in a spatial environment around their position, based on the underlying wind field. SERNs cannot be used as surrogates for networks with this kind of spatial embedding. Fortunately, the boundary effects are mostly restricted to the actual boundary, because the spatial link distance is limited by the wind velocity: Nodes at the edge of the map cannot be linked to nodes in the center or even at the opposite end of the map. This is the reason why we chose a larger spatial area to compute the networks than the map section on which the results are actually shown.

The difference between out- and in-area weighted connectivity D_i is influenced by these boundary effects in the southern parts of the study area, but it is not only an artifact of the boundary effects: Regions where the winds are slowed down and deflected by orographic barriers like the Andes result in large differences as well. More importantly, the SALLJ region has a much higher out- than in-area weighted connectivity. The winds leading into this region from the Amazon region are deflected by the Andes and are typically quite slow, which results in a low IAWC in the SALLJ region. Additionally, the wind speeds of the SALLJ are higher and the wind direction possesses multiple regimes, so that connections in multiple directions are possible. The observed difference in out- and in-area weighted connectivity can therefore be used to infer such a jet-like behavior. This is also true for the area along the Amazon river and around Los Llanos, which are the only regions in northern South America with a slightly larger out- than in-area weighted connectivity. This effect is present in both methods but it is stronger for the simulation method.

The betweenness is probably the most important measure for the streamflow networks, which by construction call for a path-based analysis of their topology. Due to our choice of a construction method based on the local flows, we associate network paths with paths the wind could take in terms of an uninterrupted transport of air masses. The betweenness centrality, as a measure of the number of shortest paths running through a node, is therefore a key measure to evaluate which areas are the most important ones for the overall path structure. Jet-like structures that interconnect different areas of the network are particularly emphasized by the betweenness centrality.

The difference between the betweenness of both methods is spread over the whole map section. The largest and most important difference is a region of high betweenness along the Amazon river for the simulation method that is only very weakly present for the semi-analytic method.

Both methods produce networks which are overall very similar to each other. Differences appear only in the detail and thus one can safely pick the semi-analytic method when computation time is an issue. However, the simulation method might be able to reveal important wind paths slightly better and is therefore picked to demonstrate the seasonal variability.

B. Climatic interpretation of the results

The streamflow networks provide a network representation of the low-level atmospheric circulation. Many properties of the spatial distribution of the networks measures can be interpreted as key features of the South American climate. The low-level circulation pattern of the SAMS is dominated by the incoming trade winds in northern South America that are, after crossing the Amazon basin, deflected southward by the Andes and channeled into the SALLJ (see Section I A). The SALLJ transports moisture from the Amazon to the subtropics.[25] The here presented streamflow networks are able to unveil these important features: For both methods (simulation-based and semi-analytic), the betweenness shows high values in the SALLJ and the Amazon region, indicating the relevance of these areas in terms of important transport routes of the low-level circulation. The wind paths of the incoming trade winds over the northern Amazon basin are clearly visible in the betweenness in terms of three narrow bands of high betweenness (Figure 8).

Furthermore, South American climate is strongly influenced by frontal systems originating from the southern extra-tropics, which are caused by Rossby waves that propagate across the southern Pacific Ocean.[26–29] In the vicinity of the southern tip of the South American continent, the absolute value of the mean wind field strongly influences the OAWC (Figure 7). However, the spatial pattern of the difference between out- and in-area weighted connectivity is not solely determined by this climatologically westerly wind direction. The large areas of positive differences of out- and in-area weighted connectivity around $50°S$ in combination with the negative differences over southern South America around $40°S$ can be associated with the northward propagating frontal systems, which transport cold, dry air northward at low atmospheric levels. In this sense, the difference between out- and in-area weighted connectivity also gives a notion of sources and sinks of the dominant circulation patterns.

Regarding interseasonal variability, Figure 9 shows that the betweenness centrality is able to capture the activity of the monsoon very well: During the wet season (DJF) the betweenness along the slopes of the Andes and northern South America, where most moisture is transported to southern South America, is much larger than during the dry season (JJA). As mentioned in the previous paragraph, these incoming north-easterly trade winds are visible in the spatial distribution of the betweenness for DJF, but are strongly diminished for JJA.

VI. CONCLUSION

In this study, we have introduced a new method to construct complex networks from two-dimensional wind fields. For this purpose, a simulation-based approach as well as an analytical approximation was compared in detail. Both the simulation-based and the semi-analytic method yield qualitatively similar networks. Networks constructed with the proposed method are able to unveil important characteristics of a regional climate system like the South American monsoon. They thus provide an accurate, coarse-grained representation of the low-level atmospheric circulation that focuses on

the key features thereof. Especially the network measure betweenness centrality captures important wind paths like trade winds and the South American Low-Level Jet very well. Additionally, the spatial pattern of the difference between out- and in-area weighted connectivity provides information about jet-like behavior, as well as about sources and sinks of the dominant circulation patterns. The large difference observed between the wet and dry season networks indicates that the betweenness centrality of the proposed networks can also be an indicator for the overall activity of the monsoon. The enhanced trade winds and South American Low-Level Jet during the wet season are in accordance with the literature about the South American Monsoon System. While we focused on wind flows in Earth's atmosphere in this study, an application of the proposed streamflow network concept to other fluids seems possible as well. Applying the streamflow networks to three dimensional flows is straightforward, too. The availability of suitable three dimensional data could be problematic, though.

ACKNOWLEDGMENTS

APPENDIX: EXTENDED COMPUTATIONAL MATERIAL

1. Simulation method algorithm

ALGORITHM 1. Computation of the edge list of the streamflow network using the simulation method.

$\Delta \leftarrow$ spatial distances of all nodes to each other

$\gamma \leftarrow$ angles between the spatial positions of all nodes to each other

for $i \in$ nodes **do**

$\quad u_d \leftarrow \frac{1}{2}\left(d_m + d_z(\lambda)\right)$ $\qquad\qquad$ ▷ calculate distance uncertainty

$\quad$ **for** $t \in$ time series **do** $\qquad\qquad$ ▷ compute hit counts of actual wind time series

$\qquad d \leftarrow |\mathbf{v}_i(t)| \cdot T + u_d$

$\qquad$ **for** $j \in$ nodes **do** $\qquad\qquad$ ▷ depending on the used grid it is possible to limit the set of nodes j has to loop over based on additional geometrical considerations

$\qquad\quad$ **if** $d > \Delta_{ij}$ **then**

$\qquad\qquad$ **if** $\gamma_{ij} \in [\theta_i - u_\theta, \theta_i + u_\theta]$ **then**

$\qquad\qquad\quad h_{ij} \leftarrow h_{ij} + 1$

$\quad$ **for** $k \in [0, N_{stat}]$ **do** $\qquad\qquad$ ▷ compute hit counts of surrogates

$\qquad$ **for** $t \in$ time series **do**

$\qquad\quad d \leftarrow |\mathbf{v}_i(t)| \cdot T + u_d$

$\qquad\quad$ **for** $j \in$ nodes **do** $\qquad\qquad$ ▷ it is possible to limit the set of nodes j has to loop over (see above)

$\qquad\qquad$ **if** $d > \Delta_{ij}$ **then**

$\qquad\qquad\quad \beta \leftarrow$ random uniformly distributed number $\in [-\pi, \pi]$

$\qquad\qquad\quad$ **if** $\gamma_j \in [\beta - u_\theta, \beta_i + u_\theta]$ **then**

$\qquad\qquad\qquad \eta_{ij}^{(k)} \leftarrow \eta_{ij}^{(k)} + 1$

$\quad$ **for** $j \in$ nodes **do** $\qquad\qquad$ ▷ evaluate surrogates and set links

$\qquad$ **if** $h_{ij} > 0$ **then**

$\qquad\quad \mu \leftarrow$ mean of $\eta_{ij}^{(k)}$ with respect to k

$\qquad\quad \sigma \leftarrow$ standard deviation of $\eta_{ij}^{(k)}$ with respect to k

$\qquad\quad$ **if** $h_{ij} > \mu + n \cdot \sigma$ **then**

$\qquad\qquad$ add (i, j) to edge list **E**

---

2. Selected hit count statistic

TABLE II. Kolmogorov–Smirnov statistic of the distributions of the surrogate hit counts in Figure 10.

Figure	$d_{crit}(\alpha = 0.001)$[a]	d_{max}
(a)	0.1949	0.1024
(b)	0.0872	0.0680
(c)	0.0872	0.0657
(d)	0.0872	0.0703

[a]Calculated according to Ref. 30 with $d_{crit}(\alpha = 0.001) = \frac{1.949}{\sqrt{n}}$.
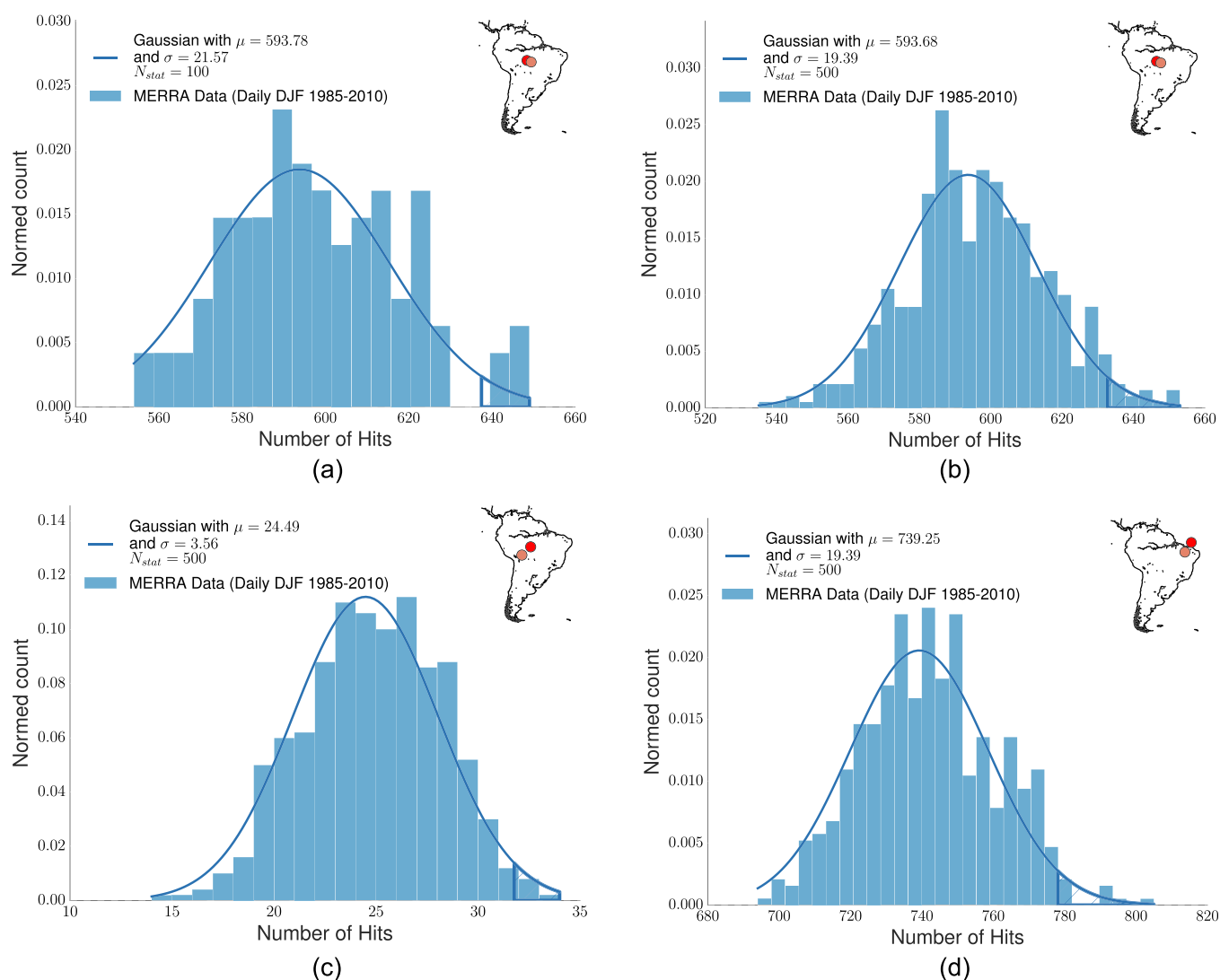

FIG. 10. Histograms of the distribution of the hit count surrogates for different locations in South America. The distribution is approximately normally distributed when there are only few hits when looking at larger distances (bottom-left panel) and also for large hit counts when smaller distances are chosen (bottom-right panel). The results of a Kolmogorov–Smirnov test are shown in Table II. Both top-row panels show the surrogate distribution between the same locations, only with a different number of surrogates. Apparently already as few as 100 surrogates can be enough to get a good estimate for a significance threshold. The significance threshold $\mu_{\eta,ij} + n \cdot \sigma_{\eta,ij}$ itself is marked in every panel with a dark blue line.

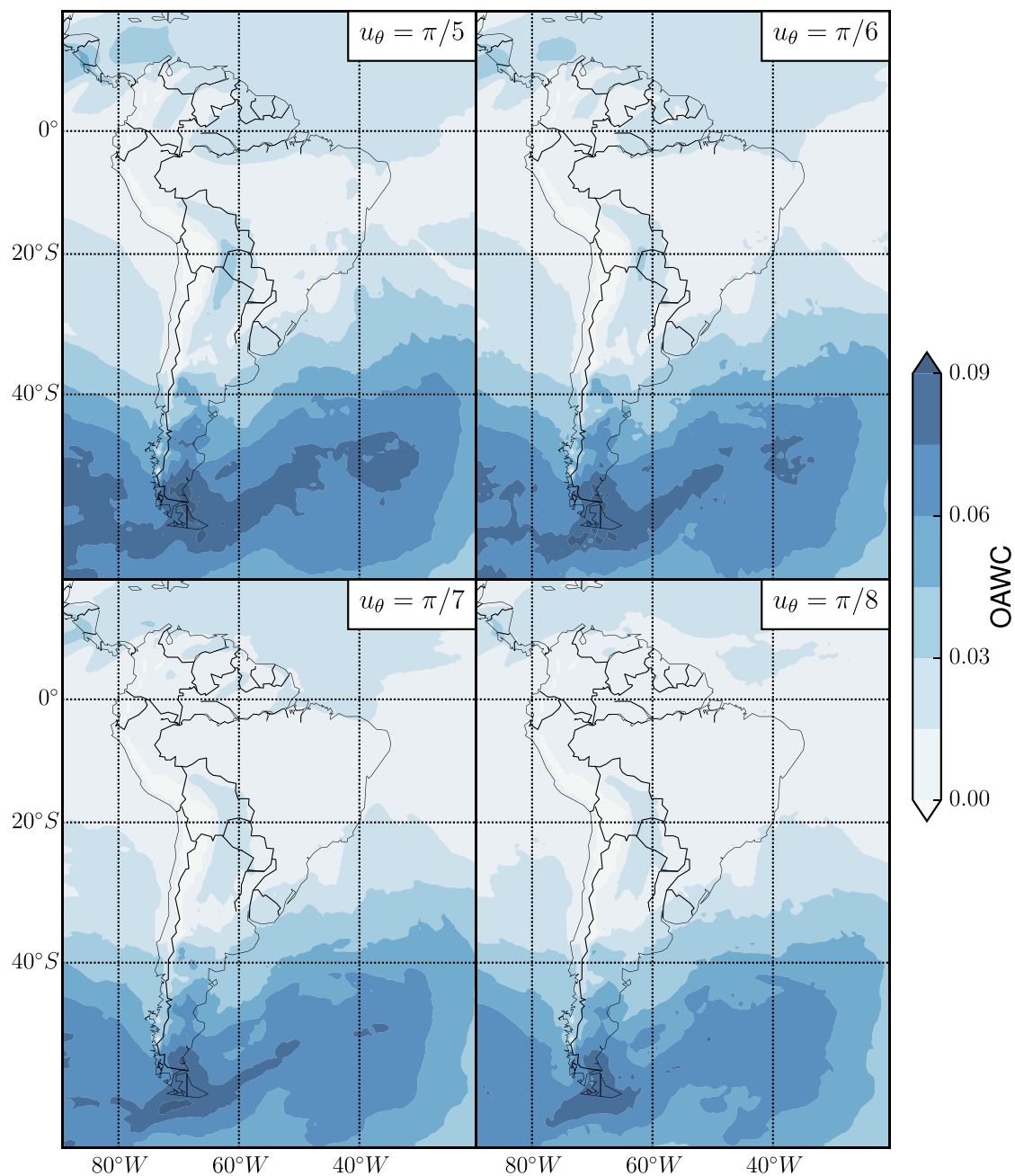## 3. Robustness of network measures to changes of the angle uncertainty

FIG. 11. Out-area weighted connectivity of streamflow networks using the simulation method. All networks were computed using DJF data and the same parameter except for the angle uncertainty u_θ. u_θ was varied to investigate the qualitative behaviour of the OAWC with changing u_θ. The OAWC is very robust to small changes of u_θ.
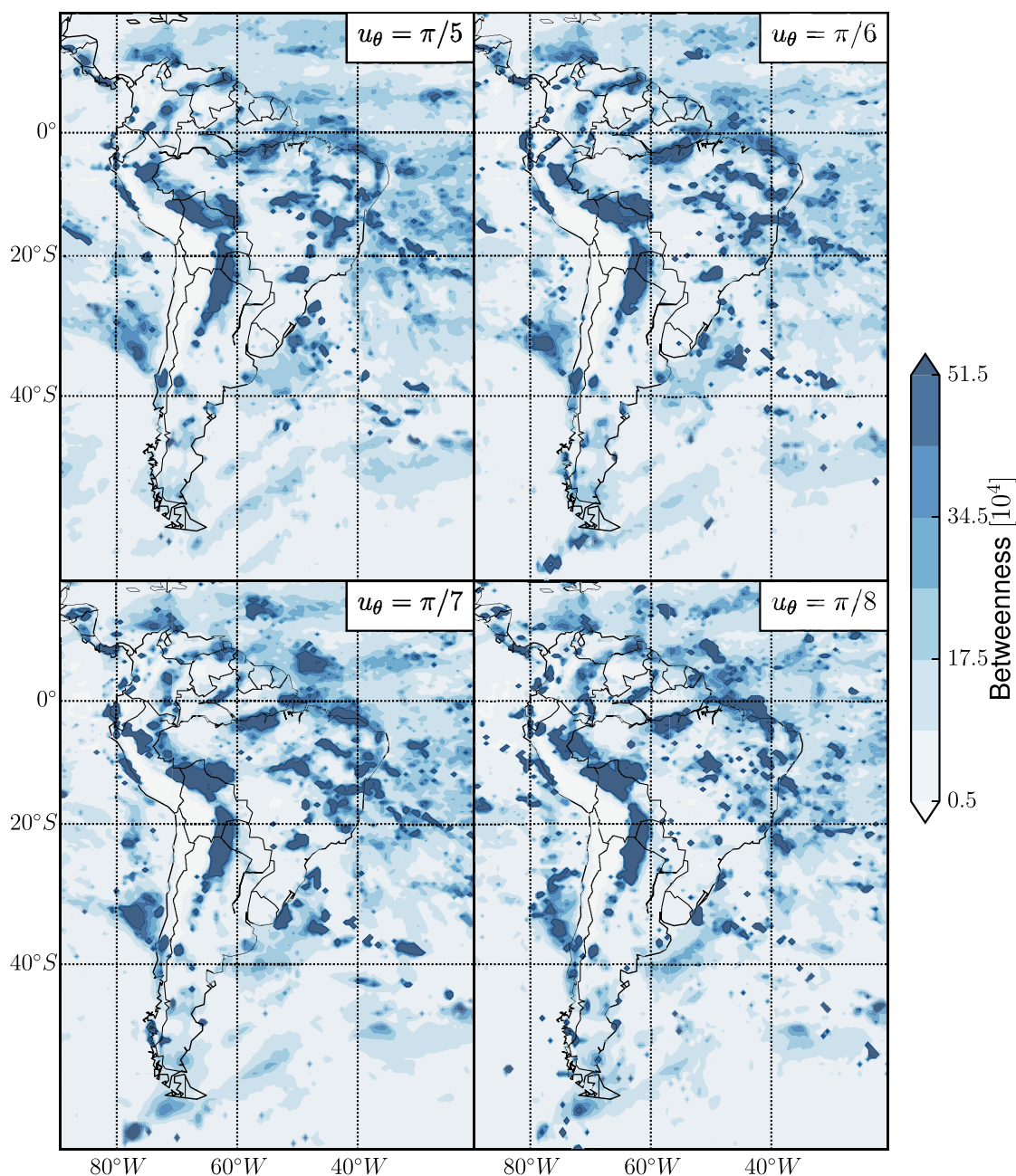
FIG. 12. Betweenness of streamflow networks for different values of u_θ similar to Fig. 11. Note that also for this measure, variations around the value $\pi/6$ do not affect the qualitative characteristics of the resulting betweenness centrality distribution.

[1]A. Tsonis and P. Roebber, "The architecture of the climate network," Physica A **333**, 497–504 (2004).

[2]F. J. Donges, Y. Zou, N. Marwan, and J. Kurths, "Complex networks in climate dynamics," Eur. Phys. J.: Spec. Top. **174**, 157–179 (2009).

[3]N. Malik, B. Bookhagen, N. Marwan, and J. Kurths, "Analysis of spatial and temporal extreme monsoonal rainfall over south Asia using complex networks," Clim. Dyn. **39**, 971–987 (2012).

[4]N. Boers, B. Bookhagen, N. Marwan, J. Kurths, and J. Marengo, "Complex networks identify spatial patterns of extreme rainfall events of the south American monsoon system," Geophys. Res. Lett. **40**, 4386–4392, doi:10.1002/grl.50681 (2013).

[5]K. Yamasaki, A. Gozolchiani, and S. Havlin, "Climate networks around the globe are significantly affected by El Niño," Phys. Rev. Lett. **100**, 228501 (2008).

[6]A. A. Tsonis and K. L. Swanson, "Topology and predictability of el niño and la niña networks," Phys. Rev. Lett. **100**, 228502 (2008).

[7]N. Boers, B. Bookhagen, H. Barbosa, N. Marwan, J. Kurths, and J. Marengo, "Prediction of extreme floods in the eastern central andes based on a complex networks approach," Nat. Commun. **5**, 5199 (2014).

[8]J. F. Donges, H. C. H. Schultz, N. Marwan, Y. Zou, and J. Kurths, "Investigating the topology of interacting networks," Eur. Phys. J. B **84**, 635–651 (2011).

[9]J. Viebahn and H. A. Dijkstra, "Critical transition analysis of the deterministic wind-driven ocean circulation – a flux-based network approach," Int. J. Bifurcation Chaos **24**, 1430007 (2014).

[10]E. Ser-Giacomi, V. Rossi, C. López, and E. Hernández-García, "Flow networks: A characterization of geophysical fluid transport," Chaos **25**, 036404 (2015).

[11]N. Molkenthin, K. Rehfeld, N. Marwan, and J. Kurths, "Networks from flows - from dynamics to topology," Sci. Rep. **4**, 4119 (2014).

[12]L. Tupikina, N. Molkenthin, C. López, E. Hernández-García, N. Marwan, and J. Kurths, "Correlation networks from flows: The case of forced and time-dependent advection-diffusion dynamics," PLoS One **11**, 1–12 (2016).

[13]M. L. Salby, *Physics of the Atmosphere and Climate*, 2nd ed. (Cambridge University Press, Cambridge, 2012).

[14]E. Ser-Giacomi, R. Vasile, I. Recuerda, E. Hernández-García, and C. López, "Dominant transport pathways in an atmospheric blocking event," Chaos **25**, 087413 (2015).

[15]J. Zhou and K.-M. Lau, "Does a monsoon climate exist over south America?," J. Clim. **11**, 1020–1040 (1998).

[16]C. R. M. Brant Liebmann, "The south American monsoon system," in *The Global Monsoon System: Research and Forecast*, edited by C.-P. Chang, Y. Ding, N.-C. Lau, R. H. Johnson, B. Wanga, and T. Yasunari (World Scientific, 2011), Chap. 9, pp. 137–157.

[17]L. M. V. Carvalho, C. Jones, and B. Liebmann, "The south Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall," J. Clim. **17**, 88–108 (2004).

[18]J. A. Marengo, B. Liebmann, A. M. Grimm, V. Misra, P. L. Silva Dias, I. F. A. Cavalcanti, L. M. V. Carvalho, E. H. Berbery, T. Ambrizzi, C. S. Vera, A. C. Saulo, J. Nogues-Paegle, E. Zipser, A. Seth, and L. M. Alves, "Recent developments on the south American monsoon system," Int. J. Climatol. **32**, 1–21 (2012).

[19]C. Vera, J. Baez, M. Douglas, C. B. Emmanuel, J. Marengo, J. Meitin, M. Nicolini, J. Nogues-Paegle, J. Paegle, O. Penalba, P. Salio, C. Saulo, M. A. S. Dias, P. S. Dias, and E. Zipser, "The south American low-level jet experiment," Bull. Am. Meteorol. Soc. **87**, 63–77 (2006).

[20]C. Vera, W. Higgins, J. Amador, T. Ambrizzi, R. Garreaud, D. Gochis, D. Gutzler, D. Lettenmaier, J. Marengo, C. R. Mechoso, J. Nogues-Paegle, P. L. S. Dias, and C. Zhang, "Toward a unified view of the American monsoon systems," J. Clim. **19**, 4977–5000 (2006).

[21]M. M. Rienecker, M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W. Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C. R. Redder, R.

Reichle, F. R. Robertson, A. G. Ruddick, M. Sienkiewicz, and J. Woollen, "MERRA: NASA's modern-era retrospective analysis for research and applications," J. Clim. **24**, 3624–3648 (2011).

[22]A. A. Tsonis, K. L. Swanson, and P. J. Roebber, "What do networks have to do with climate?," Bull. Am. Meteorol. Soc. **87**, 585–595 (2006).

[23]M. Newman, *Networks: An Introduction* (OUP, Oxford, 2010).

[24]A. Rheinwalt, N. Marwan, J. Kurths, P. Werner, and F.-W. Gerstengarbe, "Boundary effects in network measures of spatially embedded networks," Europhys. Lett. **100**, 28002 (2012).

[25]J. A. Marengo, W. R. Soares, C. Saulo, and M. Nicolini, "Climatology of the low-level jet east of the Andes as derived from the NCEP-NCAR reanalyses: Characteristics and temporal variability," J. Clim. **17**, 2261–2280 (2004).

[26]J. R. Siqueira and L. A. T. Machado, "Influence of the frontal systems on the day-to-day convection variability over south America," J. Clim. **17**, 1754–1766 (2004).

[27]J. R. Siqueira, W. B. Rossow, L. A. T. Machado, and C. Pearl, "Structural characteristics of convective systems over south America related to cold-frontal incursions," Mon. Weather Rev. **133**, 1045–1064 (2005).

[28]N. Boers, A. Rheinwalt, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. Marengo, and J. Kurths, "The south American rainfall dipole: A complex network analysis of extreme events," Geophys. Res. Lett. **41**, 7397–7405, doi:10.1002/2014GL061829 (2014).

[29]N. Boers, H. M. J. Barbosa, B. Bookhagen, J. A. Marengo, N. Marwan, and J. Kurths, "Propagation of strong rainfall events from southeastern south America to the central Andes," J. Clim. **28**, 7641–7658 (2015).

[30]J. H. Lothar Sachs, *Angewandte Statistik*, 13th ed. (Springer Verlag, Berlin, Heidelberg, 2009).