Manuscript prepared for Nonlin. Processes Geophys.
with version 5.0 of the LaTeX class copernicus.cls.
Date: 10 August 2015

# Review: Visual analytics of climate networks

**Thomas Nocke**[1], **Stefan Buschmann**[2], **Jonathan F. Donges**[1,3], **Norbert Marwan**[1], **Hans-Jörg Schulz**[4], **and Christian Tominski**[5]

[1]Potsdam Institute for Climate Impact Research, Telegrafenberg A31, 14473 Potsdam, Germany
[2]Hasso-Plattner-Institute, Prof.-Dr.-Helmert-Str. 2–3, 14482 Potsdam-Babelsberg, Germany
[3]Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, 114 19 Stockholm, Sweden
[4]Fraunhofer Institute for Computer Graphics Research, Joachim-Jungius-Str. 11, 18059 Rostock, Germany
[5]Institute for Computer Science, University of Rostock, Albert-Einstein-Str. 22, 18059 Rostock, Germany

*Correspondence to:* Thomas Nocke
(nocke@pik-potsdam.de)

**Abstract.** Network analysis has become an important approach in studying complex spatiotemporal behaviour within geophysical observation and simulation data. This new field produces increasing amounts of large geo-referenced networks to be analysed. Particular focus lies currently on the network analysis of the complex statistical interrelationship structure within climatological fields. The standard procedure for such network analyses is the extraction of network measures in combination with static standard visualisation methods. Existing interactive visualisation methods and tools for geo-referenced network exploration are often either not known to the analyst or their potential is not fully exploited. To fill this gap, we illustrate how interactive visual analytics methods in combination with geovisualisation can be tailored for visual climate network investigation. Therefore, the paper provides a problem analysis, relating the multiple visualisation challenges with a survey undertaken with network analysts from the research fields of climate and complex systems science. Then, as an overview for the interested practitioner, we review the state-of-the-art in climate network visualisation and provide an overview of existing tools. As a further contribution, we introduce the visual network analytics tools CGV and GTX, providing tailored solutions for climate network analysis, including alternative geographic projections, edge bundling, and 3D network support. Using these tools, the paper illustrates the application potentials of visual analytics for climate networks based on several use cases including examples from global, regional, and multi-layered climate networks.

## 1 Introduction

Data visualisation created within scientific contexts aims at the provision of meaningful visual representations that support the exchange of working results and provide scientists with appropriate tools to reveal relations and hidden patterns within their data. The advantage of visualisation is that it establishes a direct interface between digital data in a computer and the human perceptual and cognitive abilities, as it compactly and intuitively represents abstract relationships.

Visualisation techniques are available for different data classes – for instance, for 3D scalar data representing three-dimensional phenomena such as CT scans (Zhang et al., 2011), for vector data representing data with a direction such as air flow around an air plane (Brambilla et al., 2012), and multivariate data representing multiple data variables simultaneously (Bürger and Hauser, 2007), such as temperature, humidity, pressure, and wind speed. In scientific contexts, visualisation is often used in a static manner, producing fixed images or animations.

However, nowadays, there is a rising acceptance of interactive visualisation, still mainly for the purpose of presenting scientific findings. Since 1990, significant advances have been made in enhancing visualisation as a flexible, easy-to-use data exploration tool. This includes the possibility to interact directly with a view and having several different linked visualisations that immediately reflect any such interaction or changes of the underlying data, and vice versa. For introductions to interactive visualisation in climate research see (Tominski et al., 2011) and (Wong et al., 2014).

Going even beyond this, the new research field of "visual analytics" has emerged within the last decade (Thomas and

Cook, 2005), based on the idea of coupling human perception abilities with automatised analysis methods and thus allowing new insights into huge complex data sets: "The goal of visual analytics is the creation of tools and techniques to enable people to synthesise information and derive insight from massive, dynamic, ambiguous, and often conflicting data, [to] detect the expected and discover the unexpected, [to] provide timely, defensible, and understandable assessments and [to] communicate assessment effectively for action. [...] The challenge is to identify the best automated algorithm for the analysis task at hand, identify its limits which cannot be further automated, and then develop a tightly integrated solution which adequately integrates the best automated analysis algorithms with appropriate visualisation and interaction techniques." (Keim et al., 2008, 2010). In this context, this paper reviews how the analysis of large geophysical networks such as climate networks, but also earthquakes (Davidsen et al., 2008) and networks of rock fractures or cave passages (Phillips et al., 2015) can strongly benefit from visual analytics.

To provide an example from the field of climate network analysis, similarities of time series from grid or station based climate data can be transferred into a network structure of nodes and edges (see, e.g., Tsonis and Roebber, 2004; Yamasaki et al., 2008; Donges et al., 2009a). Then, various network measures are derived from the network topology and from the often multivariate and typically time-dependent data. Finally, a visualisation step is performed to investigate the spatial or spatiotemporal variability of the network properties and link them to the climate context. By applying the complex network approach to climate data, interesting and new insights into the climate system can be derived – e.g., studying the stability of the global climate with respect to certain climate phenomena (e.g., Berezin et al., 2012), to investigate moisture pathways and propagation of extreme rainfall events (Boers et al., 2013), or even to develop new prediction schemes (Steinhaeuser et al., 2010; Boers et al., 2014; Ludescher et al., 2013, 2014).

With respect to the size of the networks consisting of large numbers of edges on one hand, and due to the restricted availability of suitable visualisation software solutions on the other hand, visualisation used in this context focuses mainly on the static representation of scalar network measures. Figure 1 illustrates typical static representations of mapping node betweenness (Fig. 1(a)) and number of edges (Fig. 1(b)) to a spatial grid by using general purpose tools such as Python or MATLAB.

Such plots provide simplified views of the network data, representing the node information while omitting the structure denoted by the edges. In addition, missing interaction options with such static "overview" images restrict the scientists' investigation options for detecting possibly interesting, partly unknown features in the data in an exploratory sense. Against this background, this article investigates the potentials and challenges arising from interactive visual analytics methods for networks and examines available tools for exploring geo-referenced climate networks, including two tailored solutions developed by the authors. These software solutions tackle major obstacles arising in geo-referenced network visualisation, including "edge clutter" (or "spaghetti plots"), coupled 3D geo-networks, and performance issues with respect to network sizes to be handled interactively.

This article is structured as follows: Sec. 2 provides the background of geophysical climate networks, followed by an in-depth problem analysis including a survey answered by researchers using network visualisation tools in Sec. 3. Thereafter, Sec. 4 discusses suitable visualisation techniques for such networks and Sec. 5 lists software tools in which these techniques are integrated. To illustrate this state-of-the-art, Sec. 6 presents several visualisation examples for three dif-

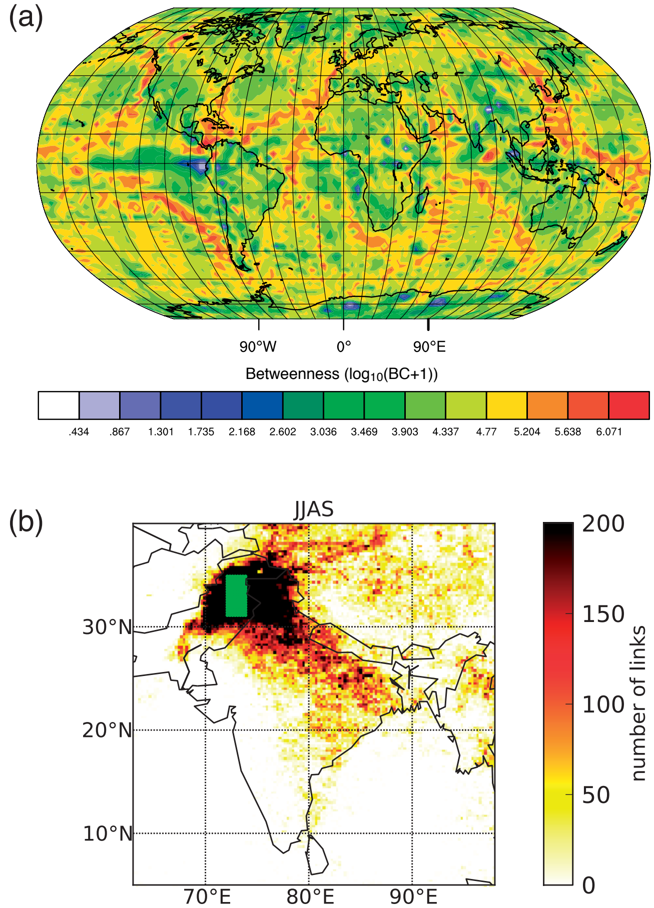



Fig. 1: (a) Network visualisation colour-coding scalar node measures (shortest path betweenness) on the map to study spatiotemporal relations of the surface air temperature field (Python/PyNGL, reprinted from Donges et al. (2009a)). (b) Visualisation highlighting the location of network edges by colouring the number of edges starting in a selected region (green box) of the map (Python/matplotlib.basemap, reprinted from Stolbova et al. (2014)).

ferent types of geophysical climate networks. Finally, Sec. 7 discusses our main findings, provides conclusions, and outlines research challenges for future work.

## 2 Climate networks

Climate researchers investigate the impact of natural phenomena and human society on the Earth's climate and vice versa. These investigations involve a variety of observational data sources as well as complex models, which in turn produce an enormous amount of simulation data. Methods of multivariate statistical analysis, such as empirical orthogonal functions (*principal component analysis*) (Wallace and Gutzler, 1981; von Storch and Zwiers, 1999), are currently the standard means to gain insight into such data.

The analysis of climatological data from the viewpoint of complex network theory (Newman, 2003; Boccaletti et al., 2006) is a recent and versatile method for making sense of the wealth of data that is available to researchers today. The central idea of climate network analysis is to construct a network (or graph) $G = (V, E)$ to represent the structure of particularly strong or significant pairwise statistical relationships that are contained within a spatiotemporally resolved data set (Tsonis and Roebber, 2004; Tsonis et al., 2007, 2008), where $V$ and $E$ denote the sets of nodes and edges in the climate network, respectively. The data sources studied by climate network analysis range from observational data, such as raw meteorological readings collected by the German Weather Service (Rheinwalt et al., 2012), via reanalysis data sets, such as those provided by the NCEP/NCAR Reanalysis 1 project (Kistler et al., 2001), to model simulations, such as those generated by Atmospheric and Oceanic General Circulation Models – e.g., the WCRP CMIP3 multi-model dataset (Meehl et al., 2007).

The nodes $i \in V$ of a climate network represent measurement stations or model grid points, where time series data $x_i(t)$ describing, e.g., temperature or precipitation variability, is available. An edge is introduced between pairs of nodes $(i, j)$ iff the value of a particular measure of statistical association $C_{ij}$ between time series $x_i(t)$ and $x_j(t)$ (e.g., linear Pearson correlation, nonlinear mutual information, or event synchronisation; Donges et al., 2009b; Malik et al., 2012; Runge et al., 2012) exceeds a threshold $T_{ij}$. Accordingly, the network's adjacency matrix $A_{ij}$ is given by

$$A_{ij} = \Theta\left(C_{ij} - T_{ij}\right) - \delta_{ij},$$

where $\Theta(\cdot)$ denotes the Heaviside function and $\delta_{ij}$ Kronecker's delta introduced to remove self-loops. Usually, a global threshold $T$ is prescribed such that $T_{ij} = T$ for all $(i, j)$ (e.g., Donges et al. (2009a,b, 2011); Tsonis and Roebber (2004); Tsonis and Swanson (2008); Yamasaki et al. (2008)), but the threshold may also be chosen adaptively for each pair based on suitable statistical significance tests of time series analysis (e.g., Steinhaeuser et al. (2010); Boers et al. (2013, 2014)).

Such a construction of a climate network, opens up the data to detailed statistical analysis using the tools of complex network theory. While most climate studies focus on standard network measures such as degree or betweenness centralities and their distributions (Newman, 2003; Boccaletti et al., 2006), a number of extensions thereof has been proposed for the specific application to climate network analysis, e.g., for heterogeneous node sizes (fraction of the Earth's surface area a node represented) (Heitzig et al., 2012) in networks of coupled climate networks (Wiedermann et al., 2013) or directed and edge-weighted networks (Zemp et al., 2014a,b).

Climate network analysis has been successfully applied to investigate spatiotemporal climate variability and complex relationships within the climate system and has been shown to provide insights that complement commonly applied methods of eigen analysis of climatological data (Donges et al., 2015). Several stability-focused studies have found evidence that the ENSO phenomenon causes a weakening of spatial statistical interrelationships and thermal stability in the global climate system, as well as reduces predictability (Yamasaki et al., 2008; Tsonis and Swanson, 2008; Berezin et al., 2012). Climate networks have been used to uncover a backbone structure carrying a considerable amount of matter, energy, and dynamical information flow in the global surface air temperature field (Donges et al., 2009a,b) and to unravel subtle shifts in climate subsystems, e.g., a westward propagation of the multidecadal Atlantic oscillation (Feng and Dijkstra, 2014) or a stability change of the Atlantic Meridional Overturning Circulation (van der Mheen et al., 2013). In other studies, the spatial variation of extreme rainfall has been used to uncover typical moisture pathways and extreme rainfall propagation, as well as to investigate involved convergence zones (Malik et al., 2010, 2012; Rheinwalt et al., 2012; Boers et al., 2013, 2014). By introducing new concepts for irregularly sampled time series, palaeo-climate networks have been used to reveal changes of the influences of the Indian Summer Monsoon on the East Asian Summer Monsoon during warm and cold periods (Rehfeld et al., 2012). Among the studies focusing on the El Niño/ Southern Oscillation, teleconnections in general or atmospheric circulation patterns have been subject of interest – e.g., Rossby waves or the Walker circulation (Runge et al., 2012; Wang et al., 2013).

Furthermore, network communities, partitioning, and the network of networks approach have been exploited to identify drivers of the global ocean surface temperature (Tantet and Dijkstra, 2013), to study the interrelationship between North Atlantic and equatorial Pacific (Guez et al., 2012, 2013), to improve statistical predictions of future climate variability (Steinhaeuser et al., 2010; Ludescher et al., 2013; Boers et al., 2014), to perform first attempts in inter-model comparison (Steinhaeuser and Tsonis, 2013; Feldhoff et al., 2014; Lange et al., 2015), and to study the cross-correlation structure between two or more distinct fields of climate variables providing novel insights into the atmosphere's general circulation structure (Donges et al., 2011; Feng et al., 2012; Wiedermann et al., 2015).

Temporal climate networks have been applied to investigate the complex spatio-temporal variability of ENSO teleconnections on regional and global scales relying on standard network measures computed from individual time-slices separately (Yamasaki et al., 2008; Radebach et al., 2013; Ludescher et al., 2013, 2014). More advanced methods and algorithms from the theory of temporal networks promise further deep insights into nonstationary climate system dynamics in the future (Holme and Saramäki, 2012; Iwayama et al., 2012; Lehnertz et al., 2014).

The main aim of climate network analysis is to serve as an explorative technique for investigating the wealth of information contained in the data's spatial correlation structure. Its validity may be confirmed by showing that known statistical relationships and structures are picked up by the method in a way that is consistent with physical expectations and the network theoretical interpretation of specific network measures under study. Moreover, the above cited studies demonstrate that climate network analysis has the potential to uncover previously hidden or unexpected structures in the data, which subsequently have to be put through a process of interpretation and careful analysis using complementary methods to answer relevant questions of interest and to generate new insights into the climate system's functioning (Donges et al., 2015).

## 3 Problem analysis

After the general background motivation, the following Sec. 3.1 analyses the requirements of climate network analysts regarding visualisation and, based on that, Sec. 3.2 will outline the challenges current visualisation tools are facing with respect to these requirements.

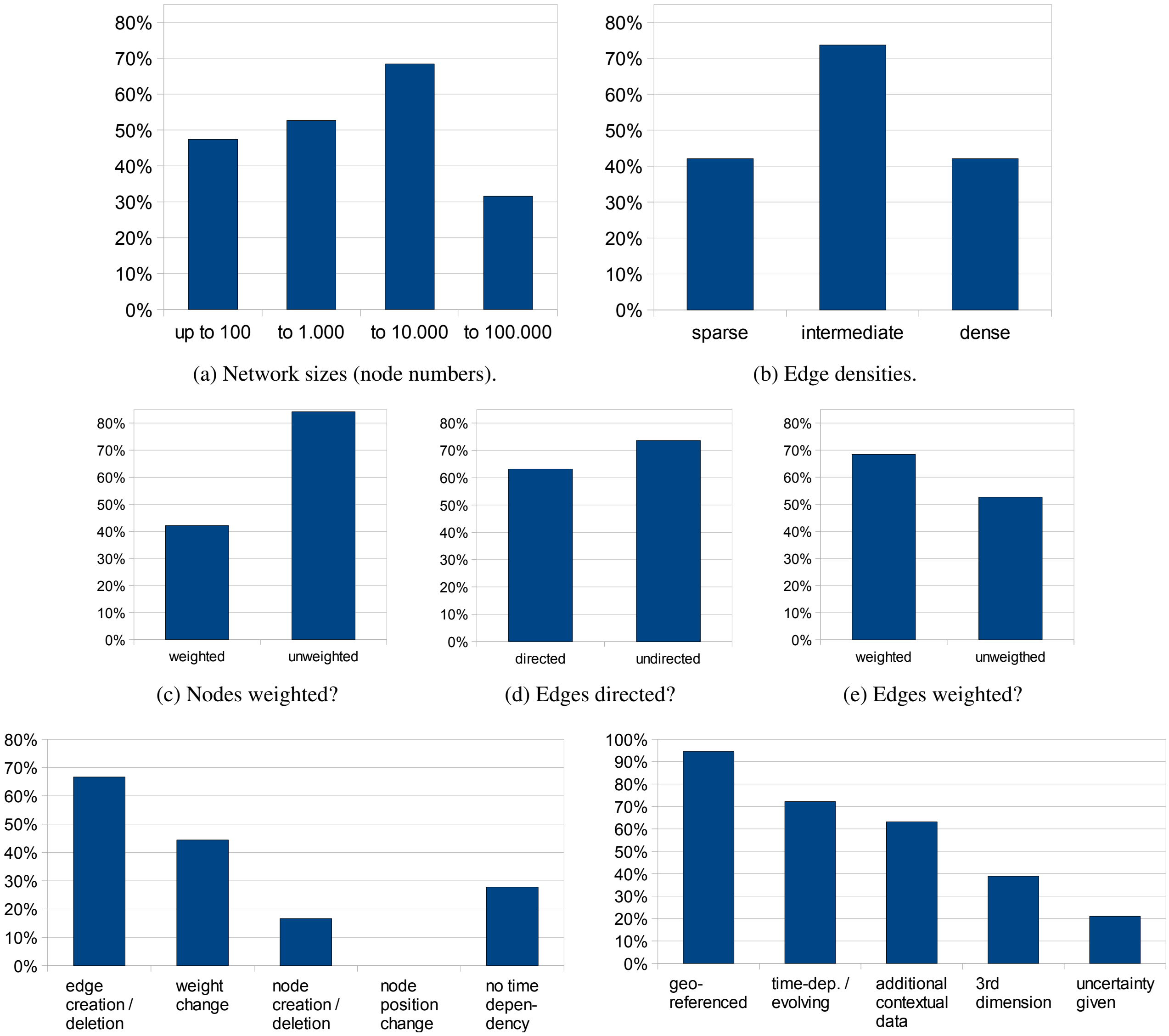
### 3.1 A survey of network analysts' visualisation habits and requirements

Researchers applying climate network analysis have so far mainly relied on static visualisations of statistical results such as degree and edge length distributions (Tsonis and Swanson, 2008), time series of the number of edges $|E(t)|$ for time-dependent climate networks (Yamasaki et al., 2008; Radebach et al., 2013), global maps and scatter plots of local network measures such as degree, closeness, betweenness centrality, and local clustering coefficient (Donges et al., 2009a,b), or line plots showing the evolution of global network measures such as average path length or transitivity with height (Donges et al., 2011). This static approach is not unique to climate network analysis, but appears to be common practice in the modern analysis of general complex networks which is guided by quantitative ideas from physics (most prominently statistical mechanics), mathematics, and social science (Albert and Barabasi, 2002; Newman, 2003; Boccaletti et al., 2006).

However, the sheer number of different metrics in complex network theory complicates the process of gaining an overall picture and, hence, a deeper understanding of climate network structures when following the static approach. This is particularly true since the spatial embedding as well as the possible time dependence of climate networks add additional dimensions to the problem. To get an overview of the situation in geophysical network analysis with respect to visualisation issues, we performed a survey with 19 practitioners within this field at the Potsdam Institute for Climate Impact Research. We asked them, what characteristics their networks have (the data), what the intentions behind the visualisation of such networks are (the tasks), what kind of visualisation they typically apply (techniques and tools), and what their most pressing requirements are with respect to geo-network visualisation.

**Network data.** From the data perspective, geo-referenced networks range from smaller (up to 100 node networks are investigated by 47% of the interviewed) to larger numbers of nodes (up to 10,000 nodes by 68%, and even up to 100,000 nodes by 32%). Edges can be weighted (68%) or unweighted (53%) and directional (63%) or undirectional (74%). With respect to edge density, these investigated networks are sparse (42%), intermediate (74%), or even dense (42%). In general, speed of hardware and software resources is the limiting factor – otherwise even larger networks would be processed and visualised. In addition, in most cases, the used networks are as well geo-referenced (94%), and an additional third dimension (e.g., elevation or atmospheric levels, 39%) may be present. Geophysical networks investigated are often time-dependent / evolving (72%), and associated with nodes and edges are multiple data attributes, which can be derived network measures, or data computed or collected at the corresponding locations (53%). For an overview see Figure 2.

**Visualisation tasks performed on networks.** Interviewed network researchers analyse such networks according to different tasks (see Fig. 3). In general, they are interested in getting familiar with the network's structure (94%), in "finding unknown patterns" (89%), and in presenting results to scientific audiences (89%). Less frequently, scientists perform the general tasks "verify hypotheses" (44%), model validation (33%), and model structure analysis (33%) on their networks. With respect to the analysis of specific structural details, the identification and visual representation of communities/ clusters (72%), the identification of hierarchical structures (44%) and of hubs/ bottlenecks (56%) within the networks is of importance or high importance, whereas loops are not relevant. Addressing general properties of the visualisation, the correct visualisation generation (100%), the aesthetics (83%), the ease of perception (78%), and the com-




(a) Network sizes (node numbers). (b) Edge densities.

(c) Nodes weighted? (d) Edges directed? (e) Edges weighted?

(f) Time dependency properties. (g) Further characteristics.

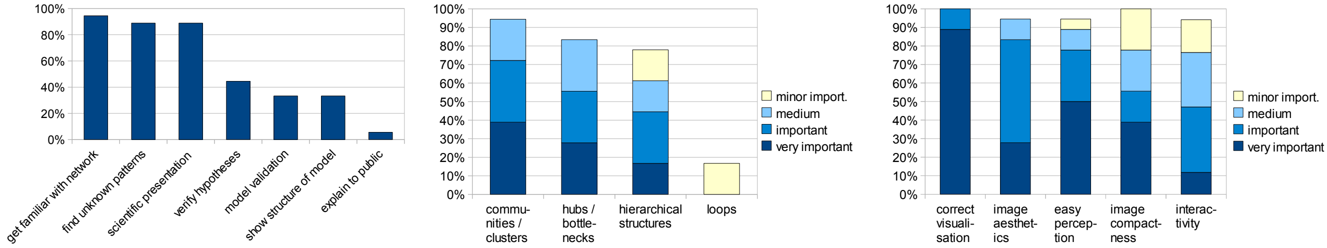
Fig. 2: Network characteristics provided by interviewed network researchers.

pactness of the image (56%) are either “important” or “very important”, whereas the possibility to interact with the image is of importance to only 47% of the interviewed analysts.

**Network visualisation techniques in use.** Beside the explicit scalar representation of network measures, most of the interviewed used node-link diagrams (83%) as network visualisation technique, whereas matrices are used by 28%, and mixed network / tree visualisation techniques are used by one person only (6%).

**Visualisation tools in use.** In addition, interviewed network researchers provided information about which visualisation tools they use for visualising networks (Fig. 4). Most interviewed use (often or sometimes) Python (72%), CGV (28%), MATLAB (22%), Google Earth/ Google Maps (17%), GraphVis (11%), Gephi (11%), or other solutions such as 3Djs (27%). Tools such as Matematica, Network Work Bench, Pajek, GUESS, Tulip and even GIS systems are used only rarely for visual network analytics.

**Further pressing requirements.** Several users asked for new interactive tools, and the main issue is speed-up to represent larger networks interactively. A second issue demanded is the reproducibility of visualisation views, so solutions combining interactivity and script-based steering are requested. A tight integration of automatic network analysis methods with interactive visualisation in the sense of visual analytics is requested by 53% of the interviewed.




(a) General network visualisation tasks. (b) Identification / representation tasks. (c) General visualisation characteristics.

Fig. 3: Tasks performed with network visualisation (a,b) and relevance of general visualisation characteristics (c).

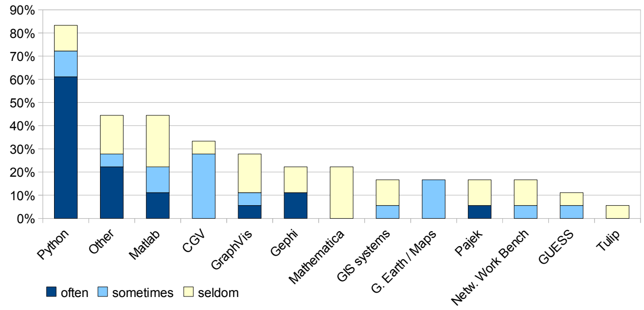



Fig. 4: Tools used to visualise networks.

### 3.2 Visualisation challenges

As described above, geo-referenced networks are often large complex multivariate structures. They typically contain $|V| = \mathcal{O}(10^3 - 10^6)$ nodes and up to $|E| = \mathcal{O}(10^7 - 10^8)$ edges. In case of the larger networks, any attempt to render them for extracting useful information from a direct and unprocessed visualisation (plot) of the network structure is unfeasible due to the following challenges.

**Spatial restrictions/ occlusion:** Representing climate networks on a 3D globe results in occlusion of at least half of the network, which is always hidden on the backside (see Sec. 6 for examples). On the other hand, representing climate networks in a projected 2D space results in distortion of neighbourhoods and clutter. Nodes that are rather close together can end up at opposite sides of the 2D map; edges between such nodes would cross the entire map and give a wrong impression of the actual geo-distance between nodes. Visibility problems aggravate when researchers have to analyse networks with an additional 3rd dimension (see Sec. 6.3).

**Edge clutter:** When the focus lies on the geographic characteristics of the data, node positions need to be fixed according to their geo-position. In such cases, edge clutter becomes a major problem, since large numbers of edges occlude the view. Suitable edge routing or edge bundling algorithms are needed to resolve this issue. However, current algorithms reach their limits in interactive analysis settings, where frequent updates and re-computations are commonplace. More efficient alternatives need to be investigated and developed.

**Multi-faceted analysis:** Climate network data are rich and complex sources of information. They may contain spatial, temporal, structural, and attribute components. It is obvious that such an abundance of information cannot be encoded into a single visual representation. It is rather necessary to use multiple linked views to enable climate researchers to focus on the aspects relevant to the task at hand and to compare and relate different aspects interactively. This requires sophisticated techniques that help the users (1) to navigate and orientate within the visual representations, which is particularly relevant for 3D approaches, (2) to dynamically filter the data for detailed analysis, and (3) to coordinate visualisation and interaction across multiple views and potentially across application boundaries.

While existing network visualisation tools may support the one or the other requirement, they are not tailored to the context of climate network analysis. Given this challenging situation, interactive visualisation promises to provide an intuitive way of combining information from the actual network structure, the network's spatial embedding and several statistical network quantifiers, e.g., degree and shortest-path (edge-) betweenness (Newman, 2003), to generate and test hypotheses ultimately based on the underlying climate data set.

## 4 Techniques for geo-referenced network visualisation

Several overview publications on network visualisation in general and on individual aspects have been published in recent years. von Landesberger et al. (2011) and Hu and Shi (2015) provide overviews on visualisation techniques available for large networks. Hadlak et al. (2015) discuss the visual integration of multiple facets given with the networks, namely hierarchies/clusters on top of the network as well as network attributes, dynamics, and spatialisation, which are all relevant for climate networks, too. In particular for networks with given geo-references, overview articles/books are available from a graph drawing (Wolff, 2013), from an information visualization (Withall et al., 2007; Rozenblat and

Melançon, 2013), and as well as from a cartography perspective (Rodgers, 2005).

Due to the typical size of climate networks, simplification mechanisms are mandatory for reducing the network complexity, else resulting in cluttered “hairball” images. The simplification must be flexible in order to account for various data attributes and analysis tasks, and it must be reproducible (e.g., re-apply stored filters). For an overview of clutter reduction methods in the visualisation field see Ellis and Dix (2007).

This simplification can be performed in each of the three visualisation process steps: (1) filtering/ pre-processing, (2) mapping, and (3) rendering. The filtering/ pre-processing step prepares the data set for visualisation, the mapping step constructs 2D or 3D geometric primitives from the data and parameterises them (for instance, with colour), and the rendering step generates the images from this scene of geometric primitives. In the following, along with these three main visualisation processing steps, we review the state-of-the-art of techniques relevant for the visualisation of climate networks.

### 4.1 Filtering techniques

First, the network itself can be simplified before converting it into a geometric representation. This can be done interactively or automatically. Interactive filtering reduces the number of items rendered by visualising only those nodes and edges that are necessary for the analysis task at hand. Automatic methods reduce the number of nodes and edges based on the network structure or on network measures. The simplification can be done globally for the whole network, or locally for a region of interest, for instance using lens interaction (see, e.g., Fig. 10).

#### 4.1.1 Node filtering

Two general node filtering methods can be distinguished: (1) indiscriminate and (2) selective filtering. **Indiscriminate node reduction methods** sample the nodes of a network, such as the traversal-based sampling which maintains the network connectivity. For an overview of indiscriminate sampling methods see Hu and Lau (2013).

**Selective node filters** reduce the node set based on their properties, either provided with or calculated for the nodes. Typically, “uninteresting” nodes are removed using network measures (e.g., node cardinality or betweenness centrality), but as well by given (climate) parameters provided with the nodes. Also the focus on selected regions or separating land and ocean nodes are one kind of selective node filters (“masking”). A univariate or multivariate node filtering can be applied, filtering out nodes using thresholds or providing a maximum number of nodes to be visualised (e.g., by showing only the $N$ nodes of largest degree).

On the one hand, this can be done interactively by the user, who changes thresholds/ maximum numbers until a suitable, uncluttered image is constructed (see, e.g., Figs. 15 and 18, which are based on node betweenness filters). However, such a choice can lead to arbitrary sub-networks, so a good knowledge of the network and network measure properties is required. Thus, on the other hand, thresholds can be derived based on objective properties of the network structure. An example of such objective properties is to represent only the $n$ most important nodes in terms of betweenness centrality.

To generalise this approach, in the visualisation field the concept of **degree-of-interest functions** (DOI) was introduced (for an overview see Abello et al., 2014), providing flexible means to attach an interest value to a data entity, in this case to nodes (see, e.g., a-priori interestingness and distance to focus by Furnas (1986), user interest by van Ham and Perer (2009), and navigation history by Gladisch et al. (2013)).

Beside the discussed filters, an alternative is the **reduction of redundant or similar nodes**, typically done by a similarity clustering of partial networks (see, e.g., Abello and Pogel, 2006). The result is a network of networks, were – to produce an overview image – individual partial networks can be represented as individual nodes, thus strongly reducing the network complexity. Then, in combination with an adjusted node layout (see Sec. 4.2), they can be unfolded on demand, providing the required degree of network complexity to the user (Hadlak et al., 2011). However, for geo-referenced networks, a geospatial neighbourhood of the clustered nodes is required, otherwise it will be hard to interpret (see Fig. 5).

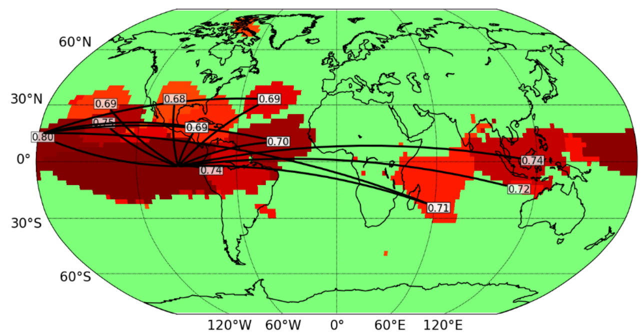



Fig. 5: Climate network visualisation of regional node clusters (reprinted from Hlinka et al., 2014).

#### 4.1.2 Edge filtering

Simplification of the node set has an additional advantage: it reduces the edge set as well, as the removal of nodes implies the removal of incident edges. Beyond that, additional methods are available for reducing the edge clutter in a network visualisation. First of all, the most basic edge filter is the thresholding procedure when reconstructing the network from data using correlation measures. An objective criterion can be chosen in such a way that edges represent only significant interrelations of a preselected $p$-value (Donges et al., 2009b; Boers et al., 2014). Next, in the same manner as selective node filters, selective edge filters can be applied, either by (interactively) filtering edges by derived edge measures such as edge shortest-path betweenness (Newman, 2003) or geodesic distance (Donges et al., 2009b) as well as by data given with the edges, or by defining a maximum number of edges to be displayed in combination with a DOI function. In visualisation, such measures have successfully been used to cut down the number of edges – mainly to ensure a proper unfolding of the layout, which tends to become a hairball for small-world (i.e., dense) graphs. For example, van Ham and Wattenberg (2008) use a centrality-based filtering of edges, whereas Nocaj et al. (2014) use a measure of embeddedness. Note that both approaches add the removed edges back in after the layout, as the edge removal is a mere intermediate step to reduce the hairball effect during the computation of the layout. A tailored solution for the specifics of the climate background was introduced by Ebert-Uphoff and Deng (2010), who reduce the edge set based on causal relationships within the climate network.

In addition to these direct edge removal techniques, in the recent decade a new class of methods reducing edge clutter (in particular for cases of a fixed layout) has been developed: edge bundling, which trades edge clutter for overplotting. As edge bundling does change the geometry of the edges and their geometric properties without removing individual edges totally, we discuss it in the following section (see Sec. 4.2.2).

## 4.2 Mapping

Network visualisation techniques applicable for climate networks include network measure charts/ maps, node-link diagrams, and matrix representations. Network measure charts/ maps reduce the problem to the visualisation of scalar data, representing network structure properties instead of the original network structure (and topology). In contrast, the other two classes represent the structure directly – either as node-link diagrams or as matrix representations. Since showing the nodes in columns and rows in matrices does not allow an explicit geo-referencing of the nodes, we argue that node-link diagrams are the most promising (and challenging) class representing climate network structures and thus, will be reviewed in more detail.

**Typical graphical primitives.** For node-link diagrams, network entities are mapped directly to graphical primitives, graphically connecting nodes (e.g., represented as nodes, circles, or spheres) and edges (e.g., represented as straight lines, curves, and cylinders). Depending on the particular type of techniques, network measures and additionally given data at nodes or edges are encoded in visual properties such as colour, size, or thickness. Often, DOI values can be used to steer such visual properties (e.g., mapping it to saturation, transparency, or size).

### 4.2.1 Node layout

**Fixed 2D geo-spatial layouts.** Most often, climate networks are represented on a 2D plane, typically using rectangular or Mercator projection. Edges are drawn as straight or curved lines. Two different kinds of mapping of edges can be observed: (1) edges are drawn directly from one position to the other, independent of their positions (see, e.g., Fig. 15b), or for global climate networks, (2) edges with nodes close to the cylindrical latitudinal cuts of the projection, which would produce a long line through the map, are represented as split lines, ending at or beyond the horizontal map borders (see Fig. 6). In the first case, cluttered images with crossing edges (often in the equatorial regions) of global climate networks are produced, whereas the second case impedes the mental tracking of individual split lines. An example overcoming these drawbacks based on alternative projections is introduced in Sec. 6.1.

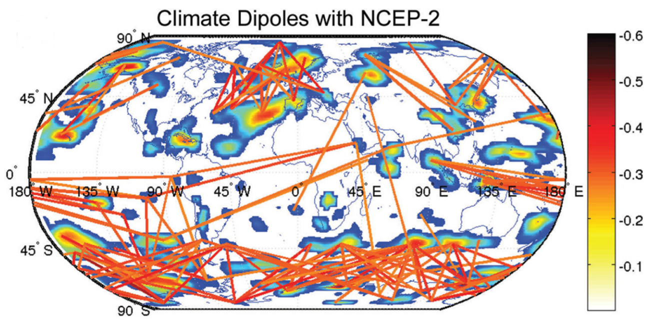
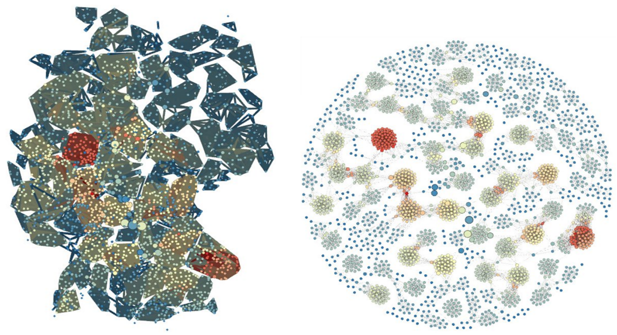



Fig. 6: Node-link diagram with split edges: dipoles extracted from the sea level pressure field (SLP) from NCEP2 reanalysis data, edge shared reciprocal nearest neighbours density mapped to colour (reprinted from Ganguly et al., 2014).

**Node position changing layouts.** To facilitate the network structure perception, avoiding crossing edges and edge clutter, in particular for large networks, layout algorithms changing node positions are applied. Such algorithms try to minimise the number of crossing edges and emphasise the network structure by optimising the spatial alignment of nodes (Díaz et al., 2002). DOI functions can be used to shift clutter from nodes with a higher DOI values to nodes with lower DOI values, for instance adjusting weights in force-based layout algorithms.

An important feature of climate networks, however, is that they are geo-referenced. The position of nodes therefore corresponds to a geographic position on the Earth, which is important to interpret the data, for instance, to relate teleconnections in such networks with actual physical processes. Therefore, graph layout algorithms that change node positions based on the network structure, can be applied only in conjunction with a geo-referenced layout, for instance Fruchterman and Reingold (1991)'s algorithm applied with an initial geo-layout (see Fig. 7).



(a) Geo-layout

(b) Fruchterman-Reingold layout based on initial Geo-layout

Fig. 7: Node-link diagram for a German measurement station based precipitation network (Gephi); node betweenness centrality mapped to circle size, closeness centrality of nodes / edges mapped to circle / line colour (Rheinwalt et al., 2012).

Such readjustments can be categorised in (1) distribution of the nodes (e.g. overlap removal, Fig. 7b), (2) removing of nodes (layout coarsening, see Sec. 4.1.1), and (3) clustering of nodes (to meta-nodes, see Sec. 4.1.1). To provide examples, Gansner et al. (2005) combine removal and distribution of nodes, whereas Hadlak et al. (2011) and Brodkorb et al. (2015) represent sub-networks either as meta-nodes, or integrate alternative visualisation methods or data scales within the geo-referenced node-link layout.

Using smooth, animated transitions between such a geo-layout and a slightly readjusted layout allows the (climate) network analyst to preserve a mental relation between the original geo-location of the nodes within node-position changing layouts (see as well Hadlak et al., 2011). We tested several force-directed layout algorithms, which we initiated with a geo-layout and parameterised them such that they change the node position only locally (e.g. Fruchterman and Reingold (1991) or Jacomy et al. (2014), by reducing the typical parameter strength of these algorithms such as spring forces), see Fig. 7.

**Virtual globes.** In addition to the 2D approaches, interactive 2.5D visualisation techniques can be applied, typically using virtual globes and perspective projections. For example, by mapping the length of an edge to the height of a 3D arch, short and long teleconnection can be visually distinguished more easily. Under certain circumstances, the network can also be used to "imprint" the virtual globe by deforming it accordingly, as it was proposed by (Alper et al., 2007). However, such visualisations can further intensify the edge clutter problem, and generate additional issues such as perspective distortion and occlusion (Elmqvist and Tsigas, 2008). Section 6 discusses two examples from this class.

#### 4.2.2 Edge Bundling

The challenges for visual analysis of large climate networks arise from their geo-reference, which impedes a modification of node positions, because the spatial node embedding is essential for the interpretation of the network. Therefore, standard network layout algorithms, which change the node positions in order to minimise edge overlaps, can be used only in slightly readjusted mode (see above). Thus, because most part of the "edge clutter" cannot be avoided, algorithms to bundle edges with similar properties have been developed. Such edge bundling algorithms reduce the number of individual line segments, and therefore the amount of visual clutter, by performing a spatial clustering and routing of close-by edges through the same path. In that sense, edge bundling can reveal macro-structures of a network, i.e., connections between different subsets of nodes, but at the same time it also conceals direct connections between individual nodes.

Various edge bundling algorithms have been developed over the last years. They differ in what kind of data they are applicable to, their bundling performance, and their visual results, such as strength of bundling or the readability of the resulting bundles. In the following, we compare different edge bundling algorithms with regard to their potential application to climate networks.

Hierarchical edge bundling methods (HEB) (Holten, 2006) use inherent hierarchy information in the data to construct the routing of edges between the levels of the hierarchy. This method needs hierarchical data and is therefore not applicable to general networks, such as in the case of climate networks.

Geometry-based edge bundling (GBEB) (Cui et al., 2008) works on general networks, but needs a so-called "control mesh" that guides the bundling process. This control mesh can either be created manually or it can be derived automatically from the network data by analysing edge patterns. A drawback of this method is that the chosen control mesh, such as a regular grid, has a strong visual influence on the resulting bundling geometry, which can lead to a bundling that may not represent the underlying edge patterns very well and creates visually unpleasing results such as a lot of "zig-zag" edges due to the underlying control mesh.

As an example of edge bundling methods that do not need an additional control geometry, force-directed edge bundling (FDEB) (Holten and van Wijk, 2009) is an algorithm that works on general undirected networks. It uses a physics-based model (Fig. 8a) in which edges attract each other, causing control points to move towards the other edge, while spring forces for each edge act in the opposite direction (keeping the edges intact). By simulating these forces, edges are bundled in a natural looking way and the bundling process can be modified by adjusting the force factors. However, due to the quadratic complexity of the algorithm (the forces have to be simulated for each pair of edges), it is not well suited for large networks, e.g., the authors report that bundling for an example data set of migrations with approx. 10,000 edges took about 80 seconds. This is not suitable for large climate networks, which range between a few hundred to hundreds of thousands of edges.

Using a fast agglomerative bundling approach, the MINGLE algorithm (Gansner et al., 2011) is able to bundle the same migration data set in approx. one second. This algorithm is based on a recursive approach to bundle edges, using an optimisation function based on the principle of "saving ink" (Fig. 8b). It employs spatial data structures and approximative nearest neighbour tests to quickly calculate neighbour graphs for edges and find compatible edges, which are then merged into bundles recursively. In addition, the curvature of bundles can be influenced by setting a maximal turning angle allowed for edges. This approach is scalable for large networks, but offers fewer parameters for controlling the bundling process and produces lesser bundled results as compared to FDEB.

In the search for efficient bundling methods of general networks, image based techniques for edge bundling have been developed, which allow efficient implementations on modern GPU graphics hardware and are, therefore, sufficient for large networks. As a first approach, skeleton-based edge bundling (SBEB) (Ersoy et al., 2011) achieves the bundling effect by calculating the skeleton of edge clusters and attracting the edges towards their centre lines. However, the calculation of 2D skeletons is computationally expensive. Hence, the method was generalised and simplified by using kernel density estimation edge bundling (KDEEB) (Hurter et al., 2012), which computes a density map of the edge drawings using a filter kernel and then moves the graph edges according to the resulting gradients in the density map (Fig. 8c).

While edge bundling helps to reduce edge clutter in large networks, it can also reveal high-level patterns in a data set, in particular, it represents groups of nodes which are connected with other groups of nodes. Detection and analysis of those high-level structures can provide additional insight into a data set. Therefore, challenges lie not only in the calculation of edge bundles on a network, but also in the visual representation of such bundles. For example, visualising the bundling strength of edges can help to visually detect and analyse strong connections between groups of nodes in a network.

Initially, edge bundles are often rendered as curves, e.g., using Bézier curves or B-splines, to emphasise the bundling structure and to generate visually pleasing results by smoothing the direction of edges along their bundles. Yet, this further distorts the actual connections between individual nodes.

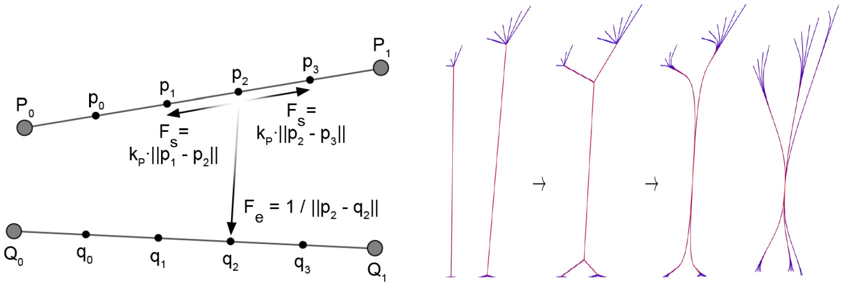
Additionally, a visualisation of the bundling structure itself is desired, e.g., to visually communicate the strength of bundles. This can be achieved by a simple mapping approach, such as depicting bundling strength by colour. Another approach is to improve the visual representation of bundles by means of shading. Lambert et al. (2010) applied a bump mapping approach to enhance the 3D impression of edge bundled networks. This approach influences the colour and brightness of pixels by modifying their surface normals, to create the impression of a "bumpy" 3D surface without modifying the actual geometry. Using this approach, strong bundles appear higher than other ones and therefore stick out visually. Telea and Ersoy (2010) developed an approach for visualising edge bundling layouts by constructing individual shapes for each cluster and rendering these shapes with an image-based technique. In this step, shading is applied to map attributes of the bundles to visual properties, such as colour, luminance, or saturation.

Finally, while edge bundling can reveal high-level structures of a network, it can also be misleading, since it obfuscates the actual connections between individual nodes. Therefore, in addition to analysing high-level patterns of a bundled network, researchers must also be able to access the initial edges of a network without bundling. This can be enabled by interaction techniques that allow a local unbundling of edges, such as brushing and interactive lenses, which reveal the connections of nodes inside the radius of a lens, while the edges outside the lens are bundled. Using this method, the connections of selected nodes can be interactively inspected, while the rest of the view remains uncluttered by the applied edge bundling (Hurter et al., 2011).

Since exploration and analysis of a data set take place interactively, the applied algorithms need to be fast enough to support user interaction, e.g., for adjusting bundling options or to perform fast re-bundling of edges after filtering options have been modified. Therefore, edge bundling does not necessarily need to be performed in real-time, but with a short enough response time for users to support an interactive analysis. Also, the generally high number of edges in climate networks demand for a high bundling performance in order to bundle a few hundred thousand edges in a few seconds. From the described bundling algorithms, MINGLE and KDEEB fulfil these requirements, as they both offer high bundling performance on large data sets. KDEEB, due to its image-based approach, can be easily integrated into a GPU-based rendering pipeline for network visualisation and can be used as a flexible bundling approach, e.g., to enable filter-aware re-bundling of edges. MINGLE, on the other hand, offers the potential to be applied to 3D-edges, e.g., to independently bundle several network layers as well as cross-layer edges in the case of coupled 3D networks (Donges et al., 2011; Feng et al., 2012).

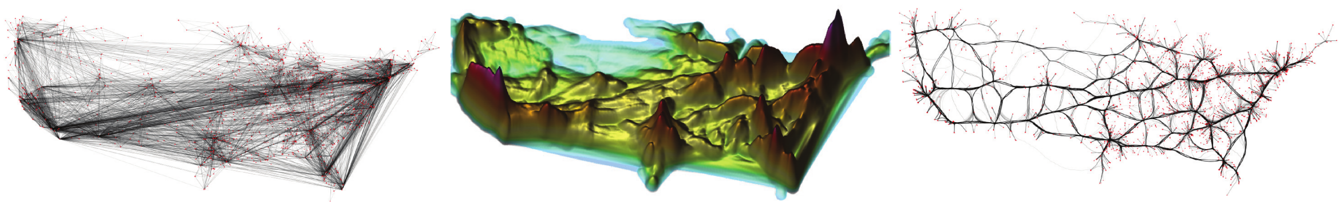
#### 4.2.3 Further mapping aspects

**Temporal / evolving networks.** Typically, time-dependent geo-referenced networks are visualised using animation or space-time cubes (see, e.g., Bach et al., 2013). In the visualisation community widely known techniques such as geographical flow diagrams (e.g., Phan et al. (2005) or Zhu and Guo (2014)) and movement data visualisation approaches (Andrienko et al., 2013) are less relevant for climate net-




(a) Force simulation applied in the FDEB algorithm

(b) Recursive bundling of the MINGLE algorithm



(c) KDEEB application: input graph, density map, and bundled result

Fig. 8: Edge bundling algorithms for geo-referenced networks from the literature. (a) © The Eurographics Association, reprinted from Holten and van Wijk (2009) (b) © 2011 IEEE, reprinted from Gansner et al. (2011) (c) © The Eurographics Association, reprinted from Hurter et al. (2012) .

works representing teleconnections, where no flow is directly associated with the edges. However, if it comes to vulnerability analyses of time-dependent energy networks (e.g., Menck et al. (2014)) or supply chain networks (e.g., Bierkandt et al. (2014)) with physical flows associated, such techniques can be useful. In any case, algorithms preserving frame-to-frame coherence of the network geometry representation such as temporal edge bundling (Hurter et al., 2014) can be beneficial for analysing temporal climate networks, such as those studied by Yamasaki et al. (2008) or Radebach et al. (2013).

**Node labelling.** If climate networks are defined based on measurements, labelling of stations can be relevant to understand local network properties. Then, labels have to be integrated in the occlusion reduction mechanisms.

**3D spatially embedded networks.** If the phenomena represented by the climate network are 3D in longitude, latitude, altitude – such as networks based on 3D atmospheric or oceanic data sets – the occlusion problems are further aggravated. As a typical solution, the visualisation is restricted to two selected layers and their internal and inter-layer edges (see Figs. 9 and 21 for examples). Interactive spatial selection and edge bundling techniques for such real 3D data sets have been developed for neuronal network visualisation (Blaas et al., 2005; Böttger et al., 2014), however, have not yet been applied to 3D climate networks.

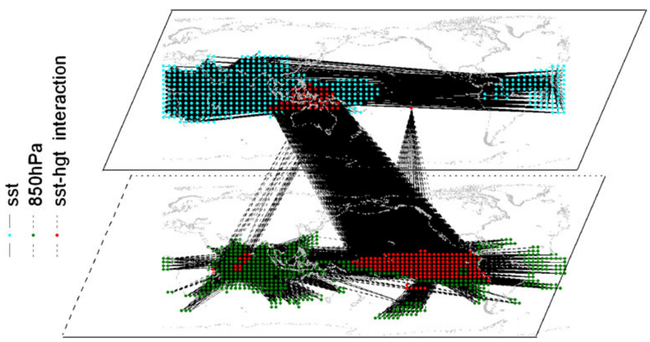



Fig. 9: Two-layered network visualisation; © Springer, reprinted from Feng et al. (2012); original caption: “The graph of bilayer air–sea interaction networks: the dots with olive colour and cyan colour represent the nodes with weighted node degree greater than 0.18 for the lower layer subnetwork and 0.14 for the upper layer subnetwork, respectively. The red dots represent the cross nodes with weighted node degree greater than 0.06. The black dashes or solid lines represent edges”.

### 4.3 Rendering

In the rendering phase, several methods can be used to reduce **visual clutter** and to **highlight structures**. Alpha blending is a common tool for handling overlapping edges, turning regions of high edge density more opaque. More advanced are specific shading techniques, improving the structure perception, e.g., use lightness adaptation for trees (Schulz et al., 2011) and 3D illustrative shading techniques for bundled edges (e.g., in Hurter et al., 2012). Finally, **interactive lenses** have been developed to reduce the local clutter around a user specified focus region (Hurter et al., 2011; Krüger et al., 2013; Tominski et al., 2014). Either, such lenses distort the optics around a network representation region (changing the local rendering properties, so called geometric lenses, thus providing more display space) or change the local primitive mapping (semantic lenses), which can, for instance, regionally show more nodes, change the node layout, hide region crossing edges or highlight edges starting/ ending in the lens region (see Fig. 10).

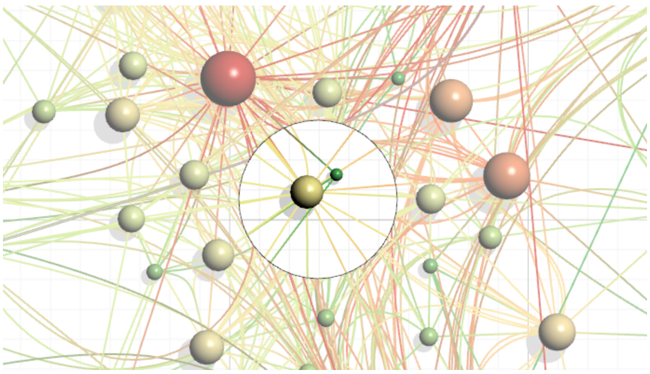


Fig. 10: Circular lens within a network visualisation (CGV): increased saturation and removal of overlapping edges (see as well Tominski et al., 2014).

In particular for climate visualisation on 3D visual globes, **rendering performance** becomes a bottleneck for most network visualisation environments, because geometrically complex representations of nodes (3D spheres) and edges (multiple line segments) result in several million geometric primitives even for medium sized climate networks. However, 3D scene interactions such as zooming, rotating, and panning as well as scene changes by filtering nodes and edges or a modification of data mapping and rendering options must provide interactive feedback, otherwise a visual exploration of such networks is strongly hampered. As a result, the rendering implementation must be highly optimised to support visualisation and analysis of medium-sized to large networks at interactive frame rates.

Techniques for improving the rendering performance for climate networks include both the minimisation of geometric primitives and the optimisation of rendering methods. For example, tessellation and rendering of complex 3D spheres representing the nodes of a network can be supplanted by billboard techniques that render only a simple quad geometry and use GPU fragment shaders to create the visual appearance of a perfect sphere with regard to the current screen resolution. This drastically reduces both geometry size and rendering time (see as well our own solutions in Sec. 5.2). The rendering of polylines, that represent the edges of a network, can be optimised by dynamic (re-)tessellation and level-of-detail techniques: depending on the size of an edge and its distance to the virtual camera, edges which occupy only a small space on the screen can be rendered at lower detail, thus reducing geometric complexity and improving rendering performance.

## 5 Visualisation Systems

To support the analysis of climate networks, tools enabling scientists to visually analyse and present large climate networks are crucial. Relevant features of such tools are sophisticated methods for visualisation and interaction in conjunction with detailed cartographic information.

### 5.1 Tool review

There are a number of relevant graph visualisation tools and systems, including Pajek (de Nooy et al., 2005), GUESS (Adar, 2006), and Gephi (Bastian et al., 2009), summarised in Tab. 1. In the following, we review important properties of these systems with respect to the requirements of climate networks.

**Computational scalability.** The size of a network, in particular the number of edges, is often a major criterion for the general applicability for climate network visualisation. This is influenced by internal data structures, for instance effective handling of sparse matrices, as well as by the effectiveness and efficiency of provided (*layout*) algorithms (e.g., hierarchic network handling and GPU implementations). Tools usable for large networks (more than 100,000 edges) include Gephi, CGV, GTX, and Tulip.

**Interactive network filtering.** Interactive filtering allows the user to specify network parts of interest, e.g., to reduce the displayed network size or to highlight parts of the network, thus avoiding perceptive overload and visual clutter. Whether a certain tool can handle large networks interactively depends strongly on its *computational scalability*. The filtering can be done either interactively, by directly selecting/deselecting nodes or edges of interest in the network structure, or by selections based on additional facets of the network such as node/edge attributes, clusters/hierarchies, time stamps, or spatial regions. CGV, Gephi, GTX, Network Work Bench, GUESS, MATLAB, and Tulip provide interactive network filtering mechanisms.

**Visual scalability.** Even if a large network can be computed and visualised interactively, the resulting visualisation technique itself or its implementation might not be scalable with respect to the display and the human perception system. Typical visualisation techniques support hundreds to several thousand nodes. In particular, if node layout algorithms are not suitable (e.g., because of the geo-reference of the nodes),

Table 1: Overview of existing tools applicable for climate network visualisation.

| Tool name | Availability | Additional properties | Url |
|---|---|---|---|
| 3Djs | open source | multiple additional information visualization techniques | github.com/mbostock/d3 |
| CGV | prototype executable / web service | multiple linked InfoVis techniques; 3D spherical networks | www.informatik.uni-rostock.de/ ct/software/CGV/CGV.html |
| Gephi | open source | network measure calculation | gephi.github.io |
| Google Earth | closed source; web service | multiple GIS functionalities | earth.google.com |
| Graph stream | open source | evolving networks; network measure calculation | graphstream-project.org |
| GraphTool | open source (python) | network measure calculation | graph-tool.skewed.de |
| GraphVis | open source | | www.graphviz.org |
| GTX | prototype executable | alternative geographic projections; 3D spherical networks | www.gtx-vis.org |
| GUESS | open source | development stopped (since 2007) | graphexploration.cond.org |
| igraph | open source (python) | network measure calculation | igraph.org |
| KiNG | open source | development stopped (since 2012) | kinemage.biochem.duke.edu |
| Matlab Graph Visualization | commercial | network measure calculation | www.mathworks.de/products/matlab |
| Network Work Bench | free for non-commercial / closed source | development stopped (since 2009) | nwb.cns.iu.edu |
| NetworkX | open source (python) | network measure calculation | networkx.github.io |
| Node Trix | open source | | www.aviz.fr/Research/Nodetrix |
| Pajek | free for non-commercial / closed source | | pajek.imfm.si |
| Tulip | open source | network measure calculation; multiple linked InfoVis techniques | tulip.labri.fr |

specialised interaction techniques and *edge bundling* become relevant. Interactive lenses are supported by CGV and Tulip. Edge bundling is supported by CGV (FDEB), GTX (MINGLE), GraphViz (MINGLE), and Tulip (based on Lambert et al., 2010). Edge bundling for hierarchical data in circular layouts is supported by the Python package GraphTool and by GraphVis (based on Holten, 2006).

**Layouts.** Node layout algorithms are of minor relevance for visualising climate networks, however for certain questions they can be a supportive feature (see Fig.7). Gephi, the GraphTool package, GraphVis, GUESS, MATLAB, the Network Work Bench, the NetworkX package, Pajek, and Tulip all support a multitude of layout algorithms.

**Geo-embedding.** Functionalities mapping the network within its geographic reference are very important for the interpretation of climate networks. Either planar (CGV, Gephi, GTX, Tulip, Graph Stream) or spherical projections (CGV, GoogleEarth, GTX, Tulip) can be applied to explicitly given node positions in longitude and latitude coordinates. Three different levels of geo-integration can be distinguished:

1. equidistant cylindrical projection without any geographic layers (by scripting GUESS, Network Work Bench, or MATLAB, Graph Stream, and as plugin in Gephi),
2. equidistant cylindrical or spherical projection with several predefined layers, such as topography, land cover, and land use (CGV, Tulip), and in addition
3. multiple flexible projections (GTX) or inclusion of self-defined GIS layers (GoogleEarth).

**Support of 3D spatial networks.** Climate networks can contain nodes at different heights, e.g., to represent several atmospheric layers. To visualise such coupled networks, tools need to support the visualisation of 3D spatial networks. For the interactive visualisation of spatial 3D biological networks such as neuronal networks, several tools and techniques have been developed (Blaas et al., 2005; Böttger et al., 2014). Unfortunately, they are not directly usable for climate networks with an additional level coordinate because of the missing geo-embedding, and to the best of our knowledge, none of the standard tools provide both functionalities. To fill this gap, we integrated a mapping of a third dimension into our own solutions CGV and GTX (see Sec. 6.3). Beyond the node mapping itself, representing 3D spatial networks further increases edge clutter, therefore 3D edge bundling solutions are requested (such as the solution from Böttger et al., 2014). However, we did not find a freely available network visualisation system providing this feature.

**Support of temporal / evolving networks.** Networks that change over time are supported only weakly in most visualisation tools (such as Gephi and Pajek). Events altering the network include node/edge creation, deletion, and attribute change. Graph Stream provides the strongest support for evolving networks, providing a flexible event handling and smooth change of the network layout / representation.

**Application domain.** The majority of tools have been developed domain independently, however, some tools were originally tailored for specific application domains (King for bio-networks, NodeTrix for social networks). In addition, our own solutions CGV und GTX provide functionalities designed for the characteristics of geo-referenced climate networks such as readers supporting the climate network data characteristics, mappings for 3D climate networks (see above), and other GIS related features.

**Scripting.** Most of the presented tools provide a scripting interface, which allows the user to precisely adjust and reproduce network visualisation properties (such as node/edge filters, camera positions) and to hand it to her scientific colleagues. In particular, scripting in the sense of an interaction history storage is essential for building scientists' trust when applying interactively steerable visual analytics tools (GTX, Tulip and Gephi). In addition to that, scripting can also allow scientists to extend the visualisation tool by creating and applying new analysis functions for the data and feeding the results back into the interactive visualisation. This requires the possibility to access and modify network data from within the scripting interface.

**Additional features.** Beyond the discussed characteristics, there are several other features of network visualisation tools potentially relevant for climate networks. In particular, this includes the ability to integrate other facets such as networks that come with an additional hierarchy on top of the network (CGV) and the provision of additional linked visualisation views to display network measures and additional data provided with the nodes and edges (GCV, Tulip, 3Djs). Interactive lenses, which provide local highlighting and reduction of clutter, are provided by Tulip, CGV, and 3Djs. Finally, a very important feature that our survey with climate network analysts revealed was the option to derive network measures on the fly (Gephi, Tulip, Graph Stream, and using the Python packages NetworkX, GraphTool, and igraph).

In summary, there is a large bandwidth of tools available for climate network visualisation, ranging from static, 2D plotting tools with strengths in computational analytics, to highly interactive tools supporting multiple facets. In the class of interactive tools, Tulip and Gephi are most advanced and freely available. Tulip (Fig. 12) is the most sophisticated, providing geo-embedding and combining network views with standard information visualisation techniques. In addition, there is a new category of programmable APIs emerging, which provide easy-to-program graphical user interfaces in combination with full-fledged network and other information visualisation views (Graph Stream, 3Djs, NodeTrix), however, which are typically not scalable for more than 1,000 nodes and for intermediate or dense networks.

### 5.2 Research prototypes

Unfortunately, the existing solutions do not integrate all features required by the interactive visual analytics of climate networks. Either, the spatial reference of the data (e.g., real 3D networks on the globe) and relevant additional cartographic information is not sufficiently supported, or the size of the networks with up to 1,000,000 edges requires efficient data handling in combination with fast GPU-aware rendering and visual clutter reduction mechanisms such as edge bundling. Because of that, we developed two research prototypes to illustrate directions of future visualisation developments for geo-referenced networks.

**CGV**. The CGV system (Tominski et al., 2009) can visualise climate networks in a variety of ways. CGV offers parallel coordinates, geographic 2D map with edge bundling (based on Holten and van Wijk, 2009), 3D globe, and density-based representations. All views are coordinated, meaning highlighting graph entities in one view also highlights the same entities in all other views. This facilitates understanding the different aspects communicated in the different views.

The focus of CGV is on interactive exploration. To this end, CGV integrates several interaction techniques. These include an extended dynamic filtering, elaborate navigation techniques, and graph lenses.

CGV's dynamic filtering mechanism supports the flexible logical combination of threshold filters and interval filters on the attributes of nodes and edges of the climate network. The visual representation reflects the filtering results in two different ways. Either the filtered nodes and edges are omitted or they are dimmed.

To facilitate the navigation of large climate networks, CGV provides a technique called edge-based traveling. Instead of using manual zoom and pan operations, the user can navigate the network by clicking on its edges. A smooth animation will take the user along the clicked edge from one node to the other. Such a navigation aid is particularly useful when working on problems related to paths in the climate network.

CGV integrates interactive lenses to support exploratory analysis tasks. Particularly useful are the edge lens, which reduces edge clutter (see Fig. 10), and the layout lens, which generates local overviews of the connectivity of selected nodes of interest.

Figure 13 illustrates a CGV session with a climate data set. The figure shows a 2D map representation (centre left) and a 3D globe with an embedded network visualisation (centre right). Additionally, a search box enables label-based node search (top left), a splat view shows the node density of the

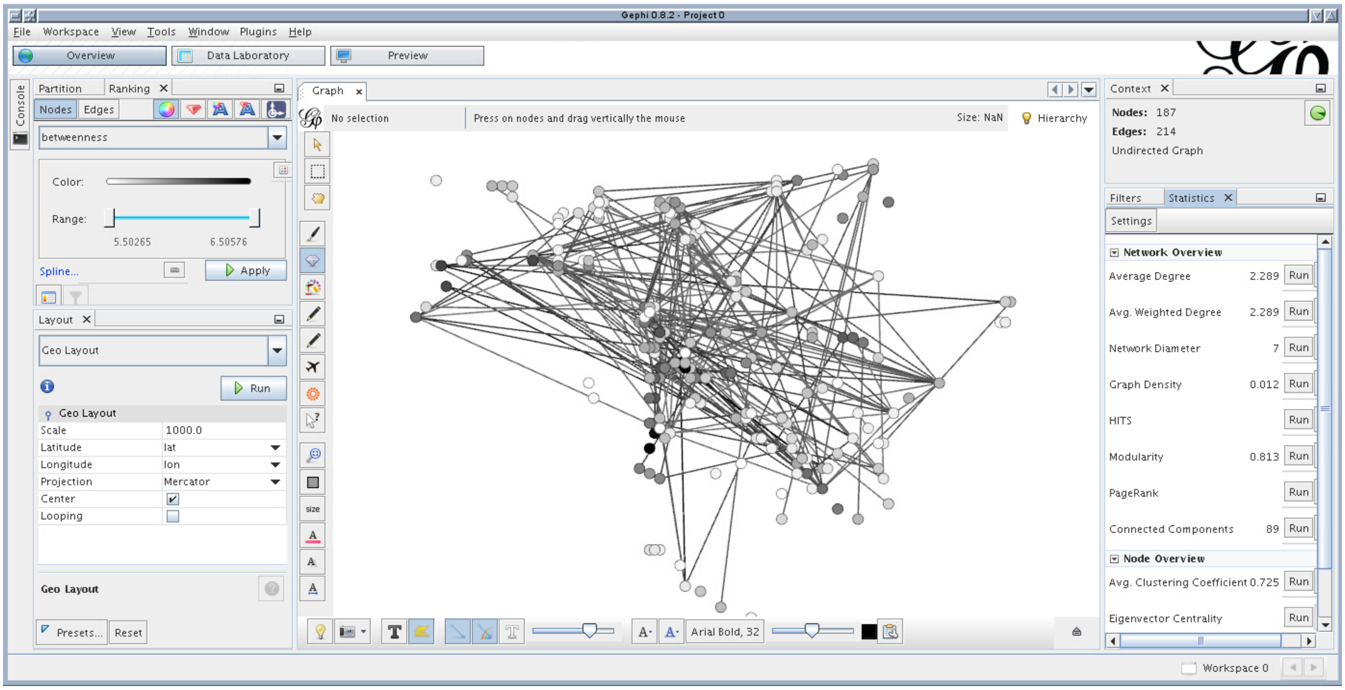



Fig. 11: Gephi screenshot: geo-layout of the African January precipitation network (see Sec. 6.2), filtered by node betweenness ($\geq 5.5$).

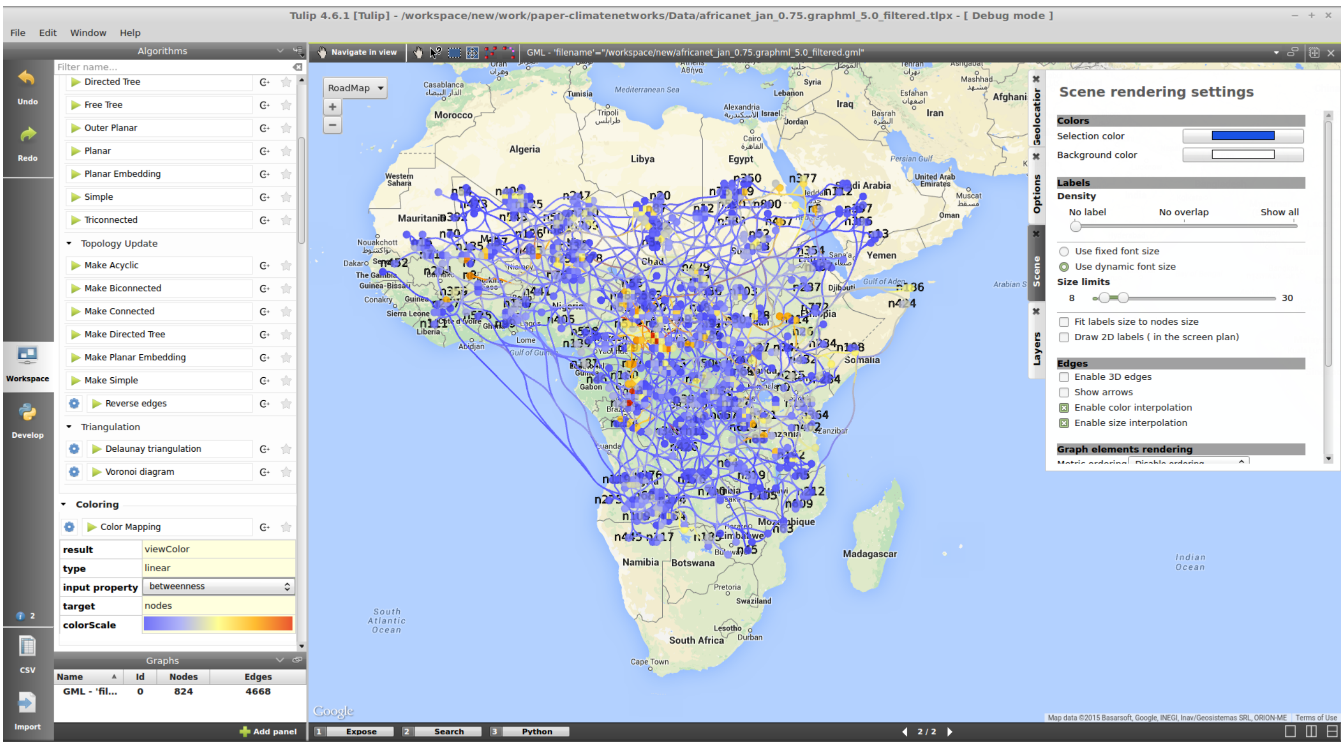



Fig. 12: Tulip screenshot visualising the African January precipitation network (see Sec. 6.2), filtered by node betweenness ($\geq 5.0$) with edge bundling (edge routing) and node betweenness encoded in colour.

network (top middle), and a parallel coordinates view abstractly depicts the node attributes (top right). A dynamic interval filter (bottom) has been applied to filter out nodes with low betweenness values.

**GTX**. While CGV provides a multi-view environment, this comes at additional rendering costs, because each view has to render the climate network. Although CGV can handle larger and/or time-dependent networks, it reaches its limits when huge networks are visualised in multiple views. As an alternative, we are currently developing a single-view tool, named GTX, to provide flexible cartographic information at different levels focused on interactive 3D visualisation of large time-dependent data sets. GTX has been developed both for visual analytics of spatio-temporal trajectory data such as air-traffic movements (Buschmann et al., 2014a) and for interactive visualisation of large geo-referenced (climate-)networks.

GTX is able to process up to 1,000,000 edges at interactive frame rates by combining sophisticated computer graphics and GIS technologies. In particular, data representation and processing is optimised by storing attributed network data in graphics card memory and creating complex geometry on the GPU during rendering, instead of precomputing it on the CPU (Buschmann et al., 2014b). This reduces memory consumption and facilitates interaction by avoiding slow recomputation and updates of geometry data from CPU to GPU memory. As a result, it allows for filtering, mapping, and rendering options to be configured interactively and updated on the fly, thus enabling interactive exploration and analysis by giving direct visual feedback to user interaction. In addition, rendering is optimised by minimising the geometric complexity, using techniques such as billboards to render spheres without tessellation and level-of-detail (LOD) to reduce the visual complexity for edges based on the distance from the virtual camera.

GTX supports both 3D globe representations and 2D geographic projections such as Mercator, transverse Mercator, and circular projection. These representations can be changed on the fly by user interaction, enabling quick comparisons of different projections. Due to the hardware-accelerated implementation, both filtering and mapping options can be configured interactively.

Finally, all of these visualisation options in GTX (i.e., filtering, mapping, geographic projections, and camera options) can be accessed by a scripting interface, using JavaScript as its programming language. This enables scientists to reproduce and share their visualisation configurations, to enhance or implement their own analysis functions on climate network data and visualise their results, and to script reproducible screenshots or even videos for presentation purposes.

Fig. 14 shows an exemplary screenshot of the GTX user interface, visualising a two-layered network (see Sec. 6.3) with height-level depicted by colour and cross-betweenness depicted by node size. A node filter for the attribute "cross-betweenness" has been applied interactively. The tool can be provided to interested researchers on demand.

Further research prototypes potentially relevant for the field of climate network visualisation have been developed for the interactive exploration of spatially-embedded networks (MoleView, Hurter et al. (2011)) and for trajectory visualisation (e.g., Krüger et al. (2013)).

## 6 Examples for visual analytics of climate networks

When analysing climate networks, the geographic nature of those networks plays an important role. Features found in the network structure may correlate with geographic characteristics, indicating a relation between, e.g., topography and the investigated climate phenomena. To detect such interrelationships is an important task of climate network analysis, which should therefore be supported by applied visualisation methods. With our network visualisation solutions GTX and CGV, climate networks are displayed as node-link diagrams, which are embedded in a geographical map, either planar, using a 2D geographical projection (Fig. 15), or as a spherical view on top of a 3D interactive globe (Figs. 16 and 18). The configurable map layer can be used to quickly cross-reference network data with topological or thematic features. Therefore, it provides exchangeable maps for different analysis tasks, e.g., topological or thematic maps.

The node-link diagram is then displayed on top of the map layer. Within this visualisation, nodes are represented as spheres, while 3D polylines depict the edges of the network. It is important to note that, while the positions of nodes correspond to the actual geographical locations of the input data, edges do not necessarily have a geographical meaning. As an edge merely represents a statistical relationship between climate time series at two nodes, the geographical extent of an edge should not be misinterpreted as representing an actual geographical phenomenon. In addition to the network structure, network measures, such as degree and shortest-path betweenness, which can be interpreted as attributes belonging to either nodes or edges, are important when analysing climate networks. To visualise such measures, node and edge attributes can be mapped to visual properties. Node properties include the size and colour of spheres. Edge properties include colour, width, and arc height.

In the following, we provide several examples to illustrate the process of data preprocessing and the subsequent visual analytics for global (Sec. 6.1), regional (Sec. 6.2), and coupled climate networks (Sec. 6.3). In discussing these examples, a case will be made for the added scientific value of advanced computer graphics visualisations of climate networks when compared to commonly used visualisation techniques such as contour plots or coloured maps of scalar fields.

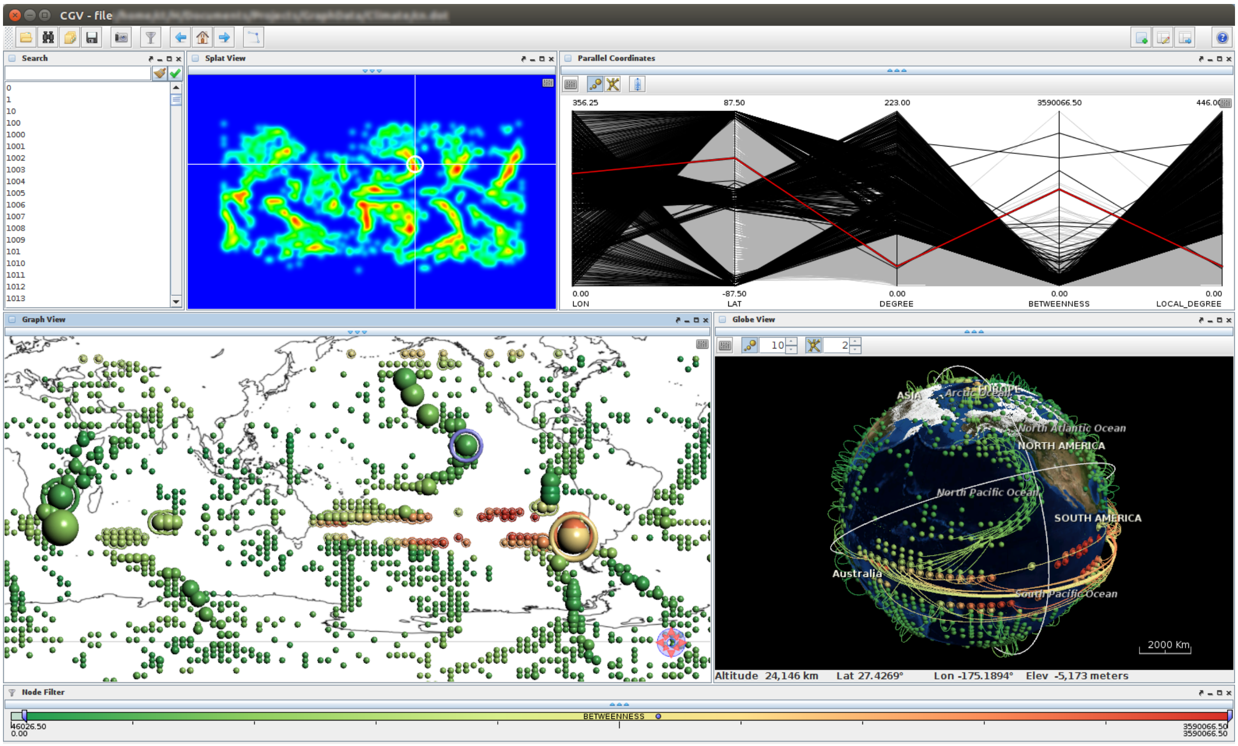



Fig. 13: CGV visualises climate networks in multiple coordinated views and provides advanced interaction.

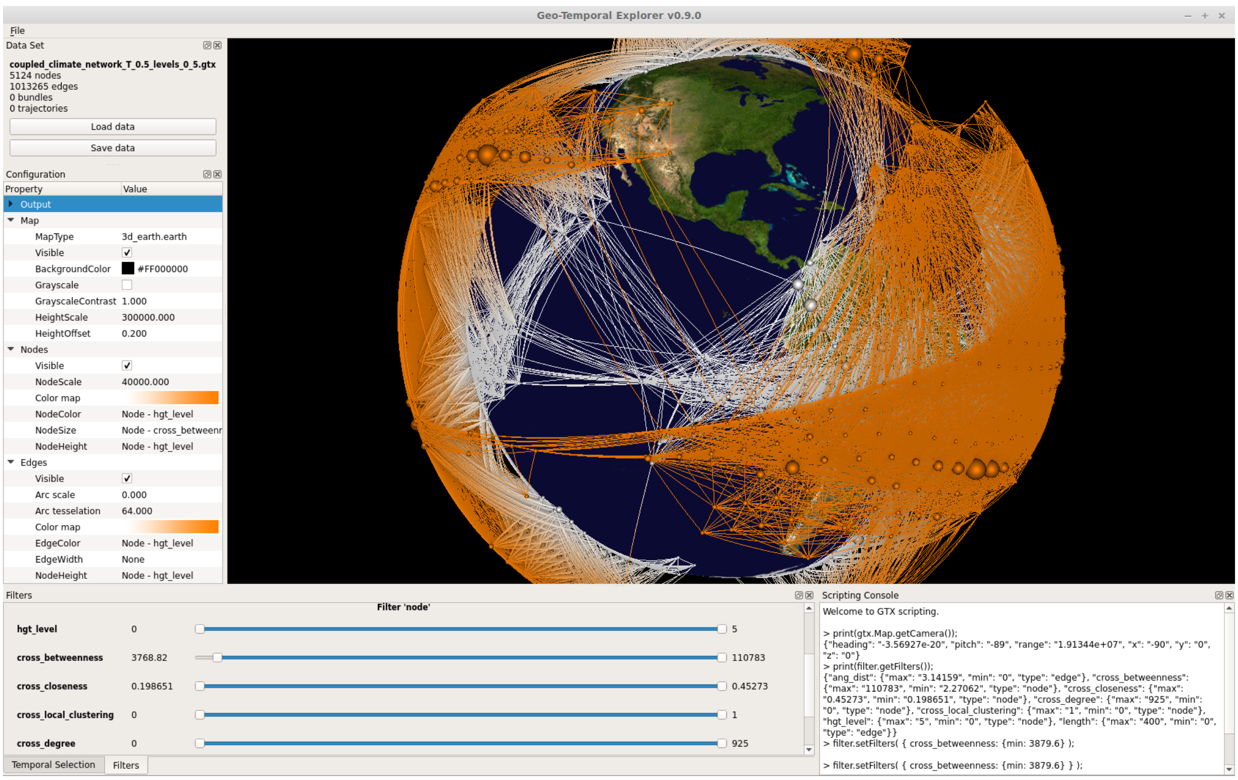
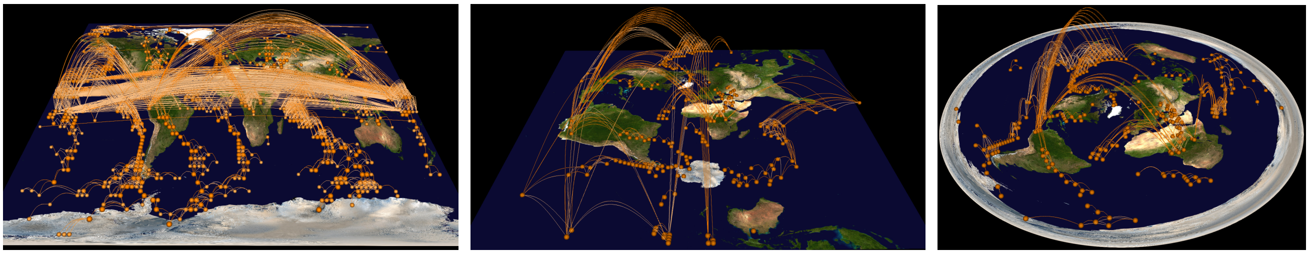



Fig. 14: GTX screenshot showing 3D network visualisation, mapping configuration, interactive filtering, and scripting console.

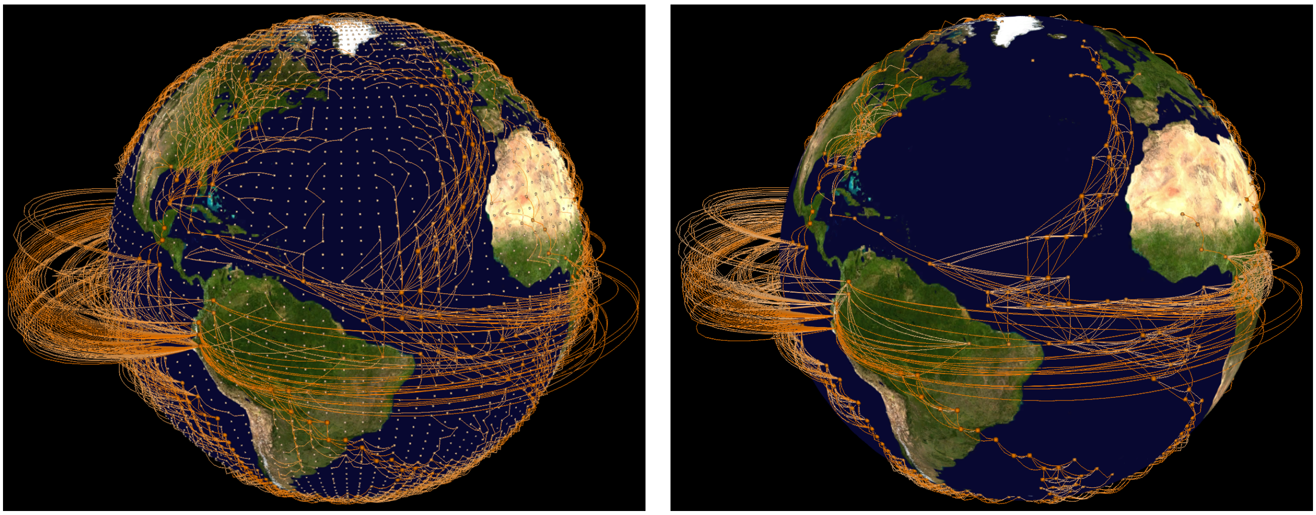


(a) Mercator projection, network filtered by node-betweenness ($\geq 80000$).

(b) Transverse Mercator projection, network filtered by node-betweenness ($\geq 195000$).

(c) 360 degree spherical projection, network filtered by node-betweenness ($\geq 170000$).

Fig. 15: Visualisation of a global surface air temperature climate network using different 2D map projections (GTX). Filtering has been applied to highlight nodes with large shortest-path betweenness.



(a) Global network, filtered by edge-betweenness ($\geq 10000$).

(b) Global network, filtered by node-betweenness ($\geq 80000$).

Fig. 16: Visualisation of a global surface air temperature climate network embedded on the surface of a 3D virtual planet using two different filtering criteria (GTX).

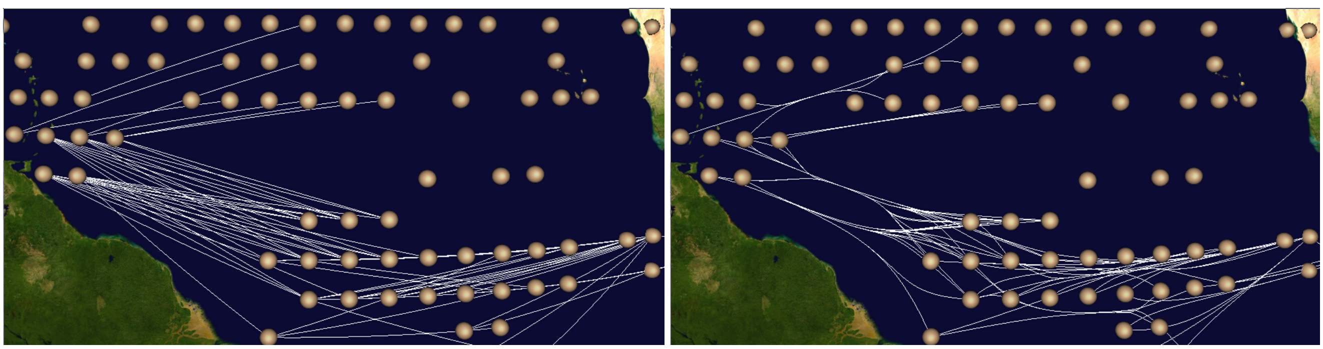


Fig. 17: Surface air temperature network (GTX) zoomed to the equatorial Atlantic ocean based on a node betweenness filtering ($\geq 100.000$) and an edge angular distance filtering ($\geq 0.26$), illustrating the effect of edge bundling : straight edges (left), bundled edges (right).

### 6.1 Global climate networks

To illustrate the visual analytics of global climate networks, we investigate the correlation structure of the monthly averaged global surface air temperature (SAT) field taken from a 20th century reference run (20c3m, as defined in the IPCC AR4 (Solomon et al., 2007)) by the HadCM3 model (Meehl et al., 2007) covering the time span from January 1860 through December 1999. Following the protocol of Donges et al. (2009a,b), a global threshold $T$ is chosen such that $0.5\%$ of all possible edges associated to the largest values of Pearson correlation (without lag) between SAT time series are included in the climate network. The resulting network contains approximately 6,000 nodes and 115,000 edges.

This SAT climate network is visualised as a node-link diagram, where the positions of nodes are fixed at the geographical locations of the corresponding model grid points (Fig. 16). Using dynamic filtering facilities as described in Section 5.1, we can interactively filter for nodes and edges with large shortest-path betweenness, therefore highlighting structures of particular importance for matter and energy flow in the climate system that tend to preferentially follow shortest paths in the network (Molkenthin et al., 2014; Tupikina et al., 2014).

Two-dimensional latitude-longitude projections of the filtered network (Fig. 15) reveal patterns that are in accordance with the backbone structure of significantly increased node betweenness discussed in Donges et al. (2009a) (compare Fig. 1a). The transverse Mercator (Fig. 15b) and spherical projections (Fig. 15c) avoid most of the mainly tropical edge clutter induced by the commonly used standard Mercator projection (Fig. 15a). Specifically, the transverse Mercator appears particularly useful for climate network analysis of global climatological fields, because it allows viewing the whole network and at the same time avoids strong geometric distortions near the poles. For example, in the SAT network, this projection provides a clearly arranged overview of the chains of high betweenness nodes emerging around the coasts of Antarctica and the Arctic as well as their global connectivity (Fig. 15b).

Compared to commonly used visual representations of climate network properties, such as the shortest-path node betweenness maps shown in Figs. 1a and 19, a particular advantage of the visualisations presented here is that information on node and edge attributes can be viewed simultaneously and intuitively, while in classical tools, edge-based network properties are usually displayed in colour-coded matrix views, see, e.g., Donner et al. (2010, Fig. 13)). In the global SAT climate network, this shows that edges with large shortest-path betweenness tend to fall into one of two categories: very short or very long edges. This observation is most clearly pronounced in spherical representations of the filtered climate network (Figs. 16, 18) which show less visual clutter than the two-dimensional projections (Fig. 1), but restrict the view to one visible hemisphere only.

Based on both views, hypotheses on some of these short and long range edges with large shortest-path betweenness can be formulated. For example, high betweenness short range edges as part of the substrate lattice (Radebach et al., 2013) of the SAT field may represent advection of heat by strong surface ocean currents such as the Canary current along the West coasts of Europe and North Africa (Fig. 16) or the Peru (Fig. 18) and California currents along the West coasts of the Americas (Fig. 16). Our visualisations furthermore reveal that short high betweenness edges often connect high betweenness nodes with comparably low degree (see, e.g., the structure resembling parts of the California current along the West coast of North America in Fig. 18). Hence, these nodes can indeed be considered as parts of critical bottleneck or backbone structures in the global surface air temperature network that channel shortest-path connections between a large fraction of pairs of regions on the Earth's surface while bearing few direct connections themselves (Donges et al., 2009a; Molkenthin et al., 2014; Tupikina et al., 2014). Long range edges with large shortest-path betweenness correspond to teleconnection patterns such as the El Niño-Southern Oscillation (ENSO) region in the tropical Pacific (Radebach et al., 2013) or the tropical Walker circulation with edges connecting regions over the Pacific, Atlantic, and Indian Oceans along the equator (Figs. 15, 16, and 18). Visual analytics also reveals further unexpected structures such as the chain of high betweenness nodes and edges crossing South America roughly from East to West (Figs. 15 and 16) that call for further research to understand the underlying processes and their potential significance for climate dynamics.

In the case of longer distant edges the visual clutter can be reduced using edge bundling. Figure 17 illustrates how a bundling can be used for the SAT network. Zooming into a certain region, bundling allows for a compact representation of major network structures by simplifying dense crossing edge regions. Thus agglomerate directions of teleconnections within the network can be better perceived.

Finally it should be noted that interactively varying the filter criteria is useful to evaluate the robustness of such patterns when testing hypotheses on underlying climatic processes and selecting regions or network substructures for further, more detailed analysis. In this spirit, Figs. 15, 16, and 18 have been generated applying differing thresholds for the node and edge shortest-path betweenness filters, respectively.

### 6.2 Regional climate networks

Regional climate networks are a special case of spatially embedded networks, because they are confined by an artificial boundary. The boundary cuts potential network edges and can, therefore, influence the network statistics and requires specific correction schemes (Rheinwalt et al., 2012). Nevertheless, the visualisation of the edges of the regional network is useful and allows for investigations of climate interactions.

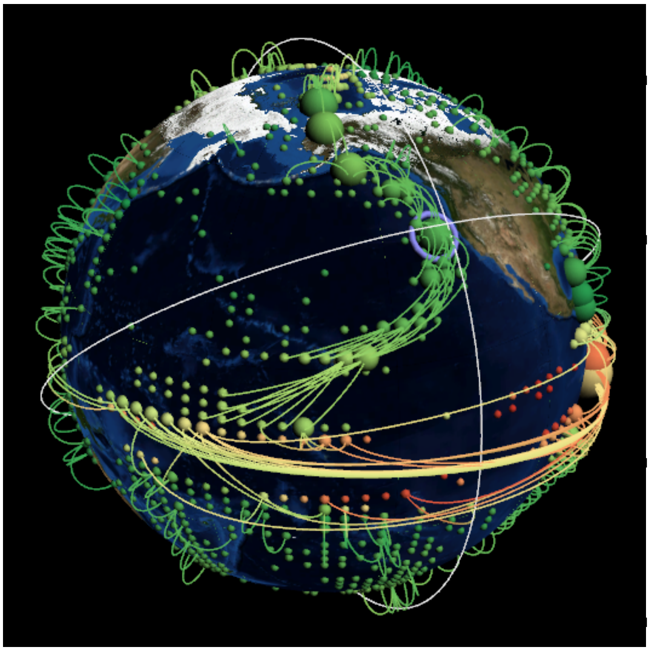


Fig. 18: Spherical three-dimensional globe representation of a surface air temperature network (CGV). Node colour (green for small values, red for large values) and size encode the node attributes degree and shortest-path betweenness, respectively.

As an example, we consider the climate network of interrelations between local rainfall variation over Africa. Here, the boundaries are mainly defined by the African coast. Instead of using continuous data, such as air temperature, we are interested in the rainfall variation. On a daily scale, rainfall occurs as events, therefore, Pearson correlation cannot be applied and other methods for studying interrelations are more appropriate, e.g., event synchronisation (Malik et al., 2010; Boers et al., 2013; Stolbova et al., 2014). In our example, we focus on a monthly scale and, therefore, will be able to use Pearson correlation. The data covers the period between 1901 and 2003 and is equally spaced on a regular grid of size $0.5° \times 0.5°$ (source CRU; New et al., 2002). We reconstruct climate networks for the months July and January separately using a fixed threshold for the correlation coefficient of 0.75 (Fig. 19).

For a planar visualisation of node attributes, we show an exemplary representation of the node betweenness prepared with MATLAB (Fig. 19). High values of betweenness characterise regions with high convection activity. These regions correspond to the Intertropical Convergence Zone (ITCZ). The seasonally separated view highlights the seasonal movement of the ITZC from its northern position in summer to a more southern position in winter. Moreover, it indicates further convergence zones, such as the Congo air boundary, crossing Africa from the Red Sea/ Gulf of Aden towards and along the Congo river (Nicholson, 2000).

A CGV based network visualisation (Fig. 20) allows for a combined view of the network edges together with node degree (node size) and betweenness (node colour). Applying a filtering procedure hiding all nodes of low betweenness can unveil such regions that correspond to known convergence zones.

### 6.3 Coupled climate networks

Coupled climate networks (CCNs) have been developed as an extension of climate network analysis that aims at systematically studying the complex structure of statistical interrelationships between different climatological fields such as surface air temperature and pressure (Donges et al., 2011). In this way, CCNs complement classical methods such as coupled pattern or maximum covariance analysis that are frequently used for the statistical analysis of multiple climatological fields (Petrova, 2012; Donges et al., 2015). CCN analysis has been applied to study vertical wind field interactions between different isobaric surfaces in the atmosphere (Donges et al., 2011), ocean-atmosphere coupling in the tropics (Feng et al., 2012) and the northern hemisphere (Wiedermann et al., 2015), interrelationships between precipitation and evaporation fields (Donges et al., 2015), as well as for predicting El Niño events (Ludescher et al., 2013, 2014). Coupled climate networks typically call for three-dimensional geographically embedded network visualisations using either available latitude, longitude, and height information for each node (e.g., for the three-dimensional geopotential height field studied in Donges et al. (2011)) or displaying overlapping surface fields as two or more stacked layers of nodes (e.g., for the evaporation and precipitation fields analysed in Petrova (2012); Donges et al. (2015)).

Here, we reconsider vertical interactions in global atmospheric geostrophic wind dynamics as studied in Donges et al. (2011). The coupled climate network is constructed from the monthly averaged and vertically resolved geopotential height field from Reanalysis 1 data provided by the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) (Kistler et al., 2001), covering the time span from January 1948 through February 2009. To study the vertical interaction structure between near surface and upper tropospheric geostrophic wind, we select data on isobaric surfaces at 1,000 mbar and 600 mbar, respectively, and interpolate to an icosahedral grid for each layer separately, resulting in a total of $2 \times 2,562 = 5,124$ nodes in the CCN. All nodes in each layer are assigned to one of two subnetworks needed for statistically analysing the interaction structure between the two isobaric surfaces (Donges et al., 2011). Finally, an edge is added for each pair of nodes within and between isobaric surfaces where the Pearson correlation (without lag) of the corresponding geopotential height time series is $\geq 0.5$.

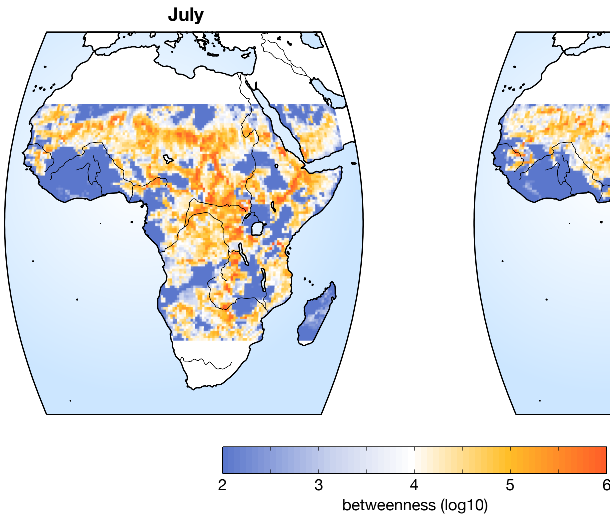
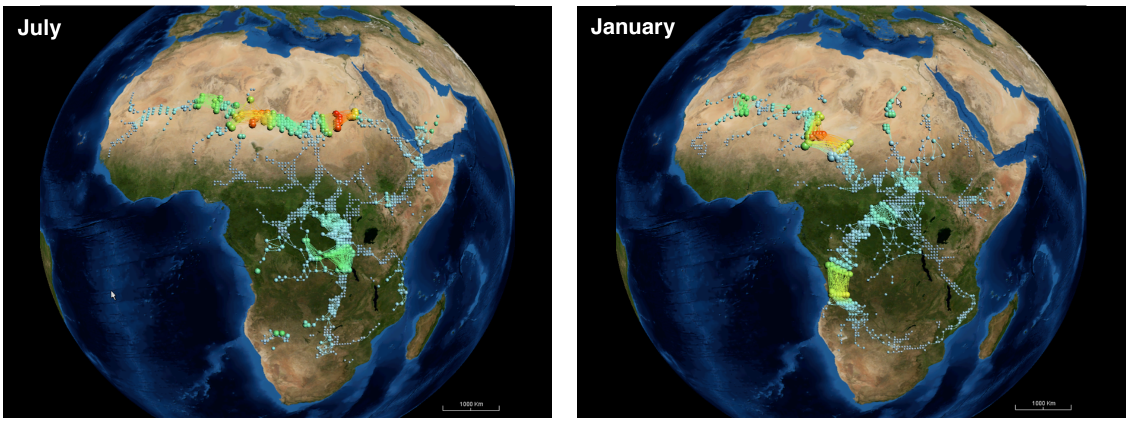



Fig. 19: Shortest-path betweenness of regional African rainfall networks for July and January (Matlab).




(a) Network visualisation July.

(b) Network visualisation January.

Fig. 20: Rainfall network of Africa for July and January unveiling regions that correspond to known convergence zones (CGV). The network visualisation (bottom row) has been filtered to hide nodes with low shortest-path betweenness to highlight regions corresponding to important moisture convergence zones. Additionally, the node attributes degree (node size) and shortest-path betweenness (node colour: blue shades $\rightarrow$ small values, red shades $\rightarrow$ large values) are displayed.

The resulting coupled climate network is visualised on the sphere in a three-dimensional view in GTX (Fig. 21). Here, the vertical dimension is strongly exaggerated, because the average vertical distance between the two isobaric surfaces ($O(10^1)$ km) is much smaller than the typical horizontal separation of nodes ($O(10^3)$ km). The network was filtered for nodes with large cross-shortest path betweenness (Donges et al., 2011) to highlight regions that are potentially particularly important for channelling vertical interactions between geostrophic wind field dynamics on both isobaric surfaces. Specifically, on the surface layer, we observe a circumpolar band of nodes with high cross-betweenness over the Arctic that is particularly pronounced over the Pacific side of the Arctic Ocean (white nodes in Fig. 21). Nodes in this circumpolar band are connected to high cross-betweenness nodes in the upper troposphere (red nodes in Fig. 21) over Europe, the Pacific Ocean, and above the East coast of North America. These observations indicate that the Arctic vortex is particularly important for mediating vertical interactions between near-surface and upper tropospheric atmospheric dynamics. While the Arctic circumpolar band of high betweenness nodes was already evident in the classical two-dimensional contour plots analysed in Donges et al. (2011), being able to intuitively display the connectivity of this region with nodes in the upper troposphere presents a large added value of the visual analytics approach in this particular use case.

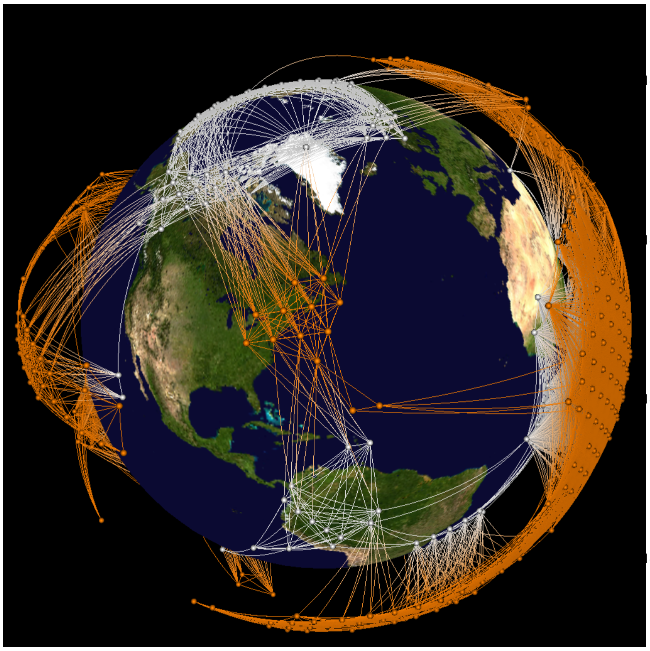


Fig. 21: Coupled climate network constructed from the geopotential height field on two isobaric surfaces (GTX): near-surface (1000 mbar, white nodes) and upper troposphere (600 mbar, red nodes). The network visualisation has been filtered to highlight nodes with large cross-betweenness indicating regions that are potentially important for mediating vertical interactions in the atmosphere's geostrophic wind dynamics, e.g., the circumpolar band of high cross-betweenness nodes above the Arctic.

From a visualisation point-of-view, it is desirable to be able to interactively visualise and study the full three-dimensional CNN including all 17 layers of nodes of the geopotential height field provided by the NCEP/NCAR reanalysis data set, to use three-dimensional edge bundling and related techniques to reduce clutter, and to also overlay this network visualisation with additional data sources such as vector-valued wind direction and speed or atmospheric moisture content.

## 7 Conclusion

In summary, the visualisation of climate data is an important means to gain insights in climate and climate related science and to communicate those insights. However, most frequently, climate data is processed using conventional statistical methods such as empirical orthogonal function analysis, and visualisation is often used for producing a final static image. This is appropriate for presentation purposes (such as in the IPCC reports), however, it does not exploit the power of the human visual system in combination with the strengths of computer-based automatic data analysis. Such an in-depth climate data analysis – tightly coupling statistics and visualisation – is subject of ongoing research.

In general, interactive visual analytics of large, time-dependent, geo-referenced climate networks is still a challenging problem. The combined application of interactive methods provided by visualisation and geographic information systems and of non-linear analysis methods is still hampered for climate science users. Hence, we strive to integrate existing approaches and to develop novel concepts for practical solutions for climate scientists.

In particular, there is the pressing issue of time-dependency of climate networks (Yamasaki et al., 2008; Radebach et al., 2013). Time-dependency implies additional conceptual and technical challenges, because the dimension of time can be structured in a number of different ways and because the data size is multiplied by the number of time steps (Aigner et al., 2011). Up to now, in most visualisation systems, individual time steps have to be loaded separately, which hinders the exploration of temporal trends and patterns in the data. New visualisation views have to be integrated to address this problem.

Additionally, uncertainty of model structure and hence of the generated data will play an increasingly important role. As a result, we have to consider the 3D visualisation of uncertain network structures with uncertain attributes, which we think is a formidable challenge.

A further not yet solved problem with the interactive visualisation approach is that filter settings are not derived from quantitative criteria, thereby rendering the results can be arbitrary to some degree. Thus, a direction for future research in visual climate networks analytics will be to identify objective filter thresholds (e.g., based on the network stability) and to provide these thresholds as reference filter values to network analysts within a visualisation session. It should be noted that visualisations such as the one presented in Fig. 18 have already proven highly valuable and successful in visually exploring large climate networks, as well as intuitively conveying the basic ideas and results of climate network analysis to scientific audiences at international conferences (see, e.g., Zou et al., 2011).

*Acknowledgements.* The authors thank Nora Molkenthin for support designing the questionnaire and executing parts of the expert interviews and all interviewed for their valuable feedback. This work was funded by the Federal Ministry for Education and Research (BMBF), Germany within the Potsdam Research Cluster for Georisk Analysis, Environmental Change, and Sustainability (PROGRESS) and project GLUES as well as the Stordalen Foundation. Further support by the federal state of Mecklenburg-Vorpommern and EFRE within the project "Basic and Applied Research in Interactive Document Engineering and Maritime Graphics" is gratefully acknowledged. We thank Dorota Andrich and Wieland Hagen for implementation support, as well as Steffen Hadlak for his helpful comments on Section 5. Climate network construction and quantitative analysis has been performed using the

Python package `pyunicorn` (Donges et al., 2013) available at `http://tocsy.pik-potsdam.de/pyunicorn.php`.

*Image copyrights.*



## References

Abello, J. and Pogel, A.: Graph Partitions and Concept Lattices, Discrete Methods in Epidemiology, AMS-DIMACS Series, 70, 115–138, 2006.

Abello, J., Hadlak, S., Schumann, H., and Schulz, H.-J.: A Modular Degree-of-Interest Specification for the Visual Analysis of Large Dynamic Networks, IEEE Transactions on Visualization and Computer Graphics, 20, 337–350, 2014.

Adar, E.: GUESS: A Language and Interface for Graph Exploration, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), ACM, 2006.

Aigner, W., Miksch, S., Schumann, H., and Tominski, C.: Visualization of Time-Oriented Data, Springer, 2011.

Albert, R. and Barabasi, A. L.: Statistical Mechanics of Complex Networks, Reviews of Modern Physics, 74, 47–97, 2002.

Alper, B., Sümengen, S., and Balcisoy, S.: Dynamic visualization of geographic networks using surface deformations with constraints, in: Proc. of the Computer Graphics International Conference (CGI), 2007.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., and Wrobel, S.: Visual Analytics of Movement, Springer, http://www.springer.com/computer/database+management+\%26+information+retrieval/book/978-3-642-37582-8, 2013.

Bach, B., Pietriga, E., and Fekete, J.-D.: GraphDiaries: Animated Transitions and Temporal Navigation for Dynamic Networks, Transactions on Visualization and Computer Graphics (TVCG), 20, 740–754, 2013.

Bastian, M., Heymann, S., and Jacomy, M.: Gephi: An Open Source Software for Exploring and Manipulating Networks, in: International AAAI Conference on Weblogs and Social Media, 2009.

Berezin, Y., Gozolchiani, A., Guez, O., and Havlin, S.: Stability of climate networks with time., Scientific reports, 2, 666, 2012.

Bierkandt, R., Wenz, L., Willner, S. N., and Levermann, A.: Acclimate—a model for economic damage propagation. Part 1: basic formulation of damage transfer within a global supply network and damage conserving dynamics, Environment Systems and Decisions, 34, 507–524, 2014.

Blaas, J., Botha, C., Peters, B., Vos, F., and Post, F.: Fast and reproducible fiber bundle selection in DTI visualization, in: IEEE Visualization'05, pp. 59–64, 2005.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U.: Complex Networks: Structure and Dynamics, Physics Reports, 424, 175–308, 2006.

Boers, N., Bookhagen, B., Marwan, N., Kurths, J., and Marengo, J.: Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System, Geophysical Research Letters, 40, 4386–4392, 2013.

Boers, N., Bookhagen, B., Barbosa, H. M. J., Marwan, N., Kurths, J., and Marengo, J. A.: Prediction of extreme floods in the eastern Central Andes based on a complex networks approach, Nature communications, 5, 5199, 2014.

Böttger, J., Schafer, A., Lohmann, G., Villringer, A., and Margulies, D. S.: Three-Dimensional Mean-Shift Edge Bundling for the Visualization of Functional Connectivity in the Brain, IEEE Transactions on Visualization and Computer Graphics, 20, 471–480, 2014.

Brambilla, A., Carnecky, R., Peikert, R., Viola, I., and Hauser, H.: Illustrative Flow Visualization: State of the Art, Trends and Challenges, in: EuroGraphics 2012 State of the Art Reports (STARs), pp. 75–94, http://diglib.eg.org/EG/DL/conf/EG2012/stars/075-094.pdf, 2012.

Brodkorb, F., Kuijper, A., Andrienko, G., Andrienko, N., and von Landesberger, T.: Overview with Details for Exploring Geolocated Graphs on Maps, Information Visualization, to appear, 2015.

Bürger, R. and Hauser, H.: Visualization of Multi-variate Scientific Data, in: EuroGraphics 2007 State of the Art Reports (STARs), pp. 117–134, http://www.cg.tuwien.ac.at/research/publications/2007/buerger-2007-star/, 2007.

Buschmann, S., Trapp, M., and Döllner, J.: Real-Time Animated Visualization of Massive Air-Traffic Trajectories, in: Proceedings of CyberWorlds 2014, pp. 172–181, IEEE Computer Society, 2014a.

Buschmann, S., Trapp, M., Lühne, P., and Döllner, J.: Hardware-accelerated attribute mapping for interactive visualization of complex 3D trajectories, in: Proc. of International Conference on Information Visualization Theory and Applications, pp. 355–363, 2014b.

Cui, W., Zhou, H., Qu, H., Wong, P. C., and Li, X.: Geometry-based edge clustering for graph visualization, Visualization and Computer Graphics, IEEE Transactions on, 14, 1277–1284, 2008.

Davidsen, J., Grassberger, P., and Paczuski, M.: Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity, Physical Review E, 77, 066 104, 2008.

de Nooy, W., Mrvar, A., and Batagelj, V.: Exploratory Social Network Analysis with Pajek, Cambridge University Press, 2005.

Díaz, J., Petit, J., and Serna, M.: A Survey of Graph Layout Problems, ACM Comput. Surv., 34, 313–356, 2002.

Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: The Backbone of the Climate Network, Europhysics Letters, 87, 48 007, 2009a.

Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Complex Networks in Climate Dynamics, European Physical Journal – Special Topics, 174, 157–179, 2009b.

Donges, J. F., Schultz, H. C. H., Marwan, N., Zou, Y., and Kurths, J.: Investigating the topology of interacting networks: theory and application to coupled climate subnetworks, The European Physical Journal B, 84, 635–651, 2011.

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

Donges, J. F., Petrova, I., Loew, A., Marwan, N., and Kurths, J.: How complex climate networks complement eigen techniques for the statistical analysis of climatological data, Climate Dynamics (online first), 2015.

Donner, R. V., Zou, Y., Donges, J. F., Marwan, N., and Kurths, J.: Recurrence networks – A novel paradigm for nonlinear time series analysis, New Journal of Physics, 12, 033 205, 2010.

Ebert-Uphoff, I. and Deng, Y.: Causal Discovery Methods for Climate Networks, Tech. Rep. GT-ME-2010-001, Georgia Institute of Technology, 2010.

Ellis, G. and Dix, A.: A Taxonomy of Clutter Reduction for Information Visualisation, IEEE Transactions on Visualization and Computer Graphics, 13, 1216–1223, 2007.

Elmqvist, N. and Tsigas, P.: A Taxonomy of 3D Occlusion Management for Visualization, IEEE Transactions on Visualization and Computer Graphics, 14, 1095–1109, 2008.

Ersoy, O., Hurter, C., Paulovich, F. V., Cantareiro, G., and Telea, A.: Skeleton-based edge bundling for graph visualization, Visualization and Computer Graphics, IEEE Transactions on, 17, 2364–2373, 2011.

Feldhoff, J. H., Lange, S., Volkholz, J., Donges, J. F., Kurths, J., and Gerstengarbe, F.-W.: Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate, Clim. Dynam. (online first), 2014.

Feng, A., Gong, Z., Wang, Q., and Feng, G.: Three-dimensional air–sea interactions investigated with bilayer networks, Theor. Appl. Climatol., 109, 635–643, 2012.

Feng, Q. Y. and Dijkstra, H.: Are North Atlantic multidecadal SST anomalies westward propagating?, Geophysical Research Letters, pp. 541–546, 2014.

Fruchterman, T. M. J. and Reingold, E. M.: Graph drawing by force-directed placement, Software: Practice and Experience, 21, 1129–1164, 1991.

Furnas, G. W.: Generalized Fisheye Views, SIGCHI Bull., 17, 16–23, 1986.

Ganguly, A. R., Kodra, E. A., Agrawal, A., Banerjee, A., Boriah, S., Chatterjee, S., Chatterjee, S., Choudhary, A., Das, D., Faghmous, J., Ganguli, P., Ghosh, S., Hayhoe, K., Hays, C., Hendrix, W., Fu, Q., Kawale, J., Kumar, D., Kumar, V., Liao, W., Liess, S., Mawalagedara, R., Mithal, V., Oglesby, R., Salvi, K., Snyder, P. K., Steinhaeuser, K., Wang, D., and Wuebbles, D.: Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques, Nonlinear Processes in Geophysics, 21, 777–795, 2014.

Gansner, E., Koren, Y., and North, S.: Topological fisheye views for visualizing large graphs, IEEE Transactions on Visualization and Computer Graphics, 11, 457–468, 2005.

Gansner, E. R., Hu, Y., North, S., and Scheidegger, C.: Multilevel agglomerative edge bundling for visualizing large graphs, in: Pacific Visualization Symposium (PacificVis), 2011 IEEE, pp. 187–194, IEEE, 2011.

Gladisch, S., Schumann, H., and Tominski, C.: Navigation Recommendations for Exploring Hierarchical Graphs, in: Advances in Visual Computing, edited by Bebis, G., Boyle, R., Parvin, B., Koracin, D., Li, B., Porikli, F., Zordan, V., Klosowski, J., Coquillart, S., Luo, X., Chen, M., and Gotz, D., vol. 8034 of *Lecture Notes in Computer Science*, pp. 36–47, Springer Berlin Heidelberg, 2013.

Guez, O., Gozolchiani, A., Berezin, Y., Brenner, S., and Havlin, S.: Climate network structure evolves with North Atlantic Oscillation phases, EPL (Europhysics Letters), 98, 38 006, 2012.

Guez, O., Gozolchiani, A., Berezin, Y., Wang, Y., and Havlin, S.: Global climate network evolves with North Atlantic Oscillation phases: Coupling to Southern Pacific Ocean, EPL (Europhysics Letters), 103, 68 006, 2013.

Hadlak, S., Schulz, H.-J., and Schumann, H.: In Situ Exploration of Large Dynamic Networks, IEEE Transactions on Visualization and Computer Graphics, 17, 2334–2343, 2011.

Hadlak, S., Schumann, H., and Schulz, H.-J.: A Survey of Multifaceted Graph Visualization, State-of-the-Art Report at EuroVis'15, pp. 1–20, 2015.

Heitzig, J., Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes, Eur. Phys. J. B, 85, 38, 2012.

Hlinka, J., Hartman, D., Jajcay, N., Vejmelka, M., Donner, R., Marwan, N., Kurths, J., and Paluš, M.: Regional and inter-regional effects in evolving climate networks, Nonlinear Processes in Geophysics, 21, 451–462, 2014.

Holme, P. and Saramäki, J.: Temporal networks, Physics Reports, 519, 97 – 125, http://www.sciencedirect.com/science/article/pii/S0370157312000841, temporal Networks, 2012.

Holten, D.: Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data, Visualization and Computer Graphics, IEEE Transactions on, 12, 741–748, 2006.

Holten, D. and van Wijk, J. J.: Force-Directed Edge Bundling for Graph Visualization, Computer Graphics Forum, 28, 983–990, 2009.

Hu, P. and Lau, W. C.: A Survey and Taxonomy of Graph Sampling, CoRR, abs/1308.5865, 2013.

Hu, Y. and Shi, L.: Visualizing large graphs, Wiley Interdisciplinary Reviews: Computational Statistics, 7, 115–136, 2015.

Hurter, C., Telea, A., and Ersoy, O.: Moleview: An attribute and structure-based semantic lens for large element-based plots, Visualization and Computer Graphics, IEEE Transactions on, 17, 2600–2609, 2011.

Hurter, C., Ersoy, O., and Telea, A.: Graph bundling by kernel density estimation, Computer Graphics Forum, 31, 865–874, 2012.

Hurter, C., Ersoy, O., Fabrikant, S., Klein, T., and Telea, A.: Bundled Visualization of DynamicGraph and Trail Data, Visualization and Computer Graphics, IEEE Transactions on, 20, 1141–1157, 2014.

Iwayama, K., Hirata, Y., Takahashi, K., Watanabe, K., Aihara, K., and Suzuki, H.: Characterizing global evolutions of complex systems via intermediate network representations, Scientific Reports, 2, 423, 2012.

Jacomy, M., Venturini, T., Heymann, S., and Bastian, M.: ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software, PLoS ONE, 9, e98 679, 2014.

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., and Melancon, G.: Visual Analytics: Definition, Process, and Challenges, in: Information Visualization, edited by Kerren, A., Stasko, J., Fekete, J.-D., and North, C., vol. 4950 of *Lecture Notes in Computer Science*, pp. 154–175, Springer Berlin Heidelberg, 2008.

Keim, D. A., Kohlhammer, J., Ellis, G., and Mansmann, F., eds.: Mastering the Information Age – Solving Problems with Visual Analytics, Eurographics, 2010.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., Dool, H. V. D., Jenne, R., and Fiorino, M.: The NCEP–NCAR 50-Year Reanalysis: Monthly Means CD–ROM and Documentation, Bulletin of the American Meteorological Society, 82, 247–268, 2001.

Krüger, R., Thom, D., Wörner, M., Bosch, H., and Ertl, T.: TrajectoryLenses–A Set-based Filtering and Exploration Technique for Long-term Trajectory Data, Computer Graphics Forum, 32, 451–460, 2013.

Lambert, A., Bourqui, R., and Auber, D.: Winding roads: Routing edges into bundles, Computer Graphics Forum, 29, 853–862, 2010.

Lange, S., Donges, J. F., Volkholz, J., and Kurths, J.: Local difference measures between complex networks for dynamical system model evaluation, PLoS ONE (in press), 2015.

Lehnertz, K., Ansmann, G., Bialonski, S., Dickten, H., Geier, C., and Porz, S.: Evolving networks in the human epileptic brain, Physica D: Nonlinear Phenomena, 267, 7 – 15, http://www.sciencedirect.com/science/article/pii/S0167278913001826, evolving Dynamical Networks, 2014.

Ludescher, J., Gozolchiani, A., Bogachev, M. I., Bunde, A., Havlin, S., and Schellnhuber, H. J.: Improved El Niño forecasting by cooperativity detection, Proc. Natl. Acad. Sci. USA, 110, 11 742–11 745, 2013.

Ludescher, J., Gozolchiani, A., Bogachev, M. I., Bunde, A., Havlin, S., and Schellnhuber, H. J.: Very early warning of next El Niño, Proc. Natl. Acad. Sci. USA, 111, 2064–2066, 2014.

Malik, N., Marwan, N., and Kurths, J.: Spatial structures and directionalities in Monsoonal precipitation over South Asia, Nonlinear Processes in Geophysics, 17, 371–381, 2010.

Malik, N., Bookhagen, B., Marwan, N., and Kurths, J.: Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks, Climate Dynamics, 39, 971–987, 2012.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E.: THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research, Bulletin of the American Meteorological Society, 88, 1383, 2007.

Menck, P. J., Heitzig, J., Kurths, J., and Schellnhuber, H. J.: How dead ends undermine power grid stability, Nature communications, 5, 2014.

Molkenthin, N., Rehfeld, K., Marwan, N., and Kurths, J.: Networks from flows – from dynamics to topology, Scientific Reports, 4, 4119, 2014.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Climate Research, 21, 1–25, 2002.

Newman, M. E. J.: The Structure and Function of Complex Networks, SIAM Review, 45, 167–256, 2003.

Nicholson, S.: The nature of rainfall variability over Africa on time scales of decades to millenia, Global and Planetary Change, 26, 137–158, 2000.

Nocaj, A., Ortmann, M., and Brandes, U.: Untangling Hairballs, in: Graph Drawing, edited by Duncan, C. and Symvonis, A., vol. 8871 of *Lecture Notes in Computer Science*, pp. 101–112, Springer, 2014.

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

Phan, D., Xiao, L., Yeh, R., Hanrahan, P., and Winograd, T.: Flow Map Layout, in: Proc. of IEEE Symposium on Information Visualization, edited by Stasko, J. and Ward, M. O., pp. 219–224, IEEE Computer Society, 2005.

Phillips, J. D., Schwanghart, W., and Heckmann, T.: Graph theory in the geosciences, Earth-Science Reviews, 143, 147–160, 2015.

Radebach, A., Donner, R. V., Runge, J., Donges, J. F., and Kurths, J.: Disentangling different types of El Niño episodes by evolving climate network analysis, Phys. Rev. E, 88, 052 807, 2013.

Rehfeld, K., Marwan, N., Breitenbach, S. F. M., and Kurths, J.: Late Holocene Asian summer monsoon dynamics from small but complex networks of paleoclimate data, Climate Dynamics, 41, 3–19, 2012.

Rheinwalt, A., Marwan, N., Kurths, J., Werner, P., and Gerstengarbe, F.-W.: Boundary effects in network measures of spatially embedded networks, EPL (Europhysics Letters), 100, 28 002, 2012.

Rodgers, P.: Graph Drawing Techniques for Geographic Visualization, in: Exploring Geovisualization, edited by Dykes, J., MacEachren, A. M., and Kraak, M.-J., pp. 143–158, Elsevier, 2005.

Rozenblat, C. and Melançon, G., eds.: Methods for Multilevel Analysis and Visualisation of Geographical Networks, vol. 11 of *Methodos Series*, Springer, 2013.

Runge, J., Heitzig, J., Petoukhov, V., and Kurths, J.: Escaping the Curse of Dimensionality in Estimating Multivariate Transfer Entropy, Physical Review Letters, 108, 258 701, 2012.

Schulz, H.-J., Hadlak, S., and Schumann, H.: Point-Based Visualization for Large Hierarchies, IEEE Transactions on Visualization and Computer Graphics, 17, 598–611, 2011.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., M.Tignor, and Miller, H., eds.: IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Steinhaeuser, K. and Tsonis, A. A.: A climate model intercomparison at the dynamics level, Climate Dynamics, 42, 1665–1670, 2013.

Steinhaeuser, K., Chawla, N. V., and Ganguly, A. R.: Complex Networks as a Unified Framework for Descriptive Analysis and Predictive Modeling in Climate Science, Science And Technology, 2010.

Stolbova, V., Martin, P., Bookhagen, B., Marwan, N., and Kurths, J.: Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka, Nonlinear Processes in Geophysics, 21, 901–917, 2014.

Tantet, A. and Dijkstra, H. A.: An interaction network perspective on the relation between patterns of sea surface temperature variability and global mean surface temperature, Earth System Dynamics Discussions, 4, 743–783, 2013.

Telea, A. and Ersoy, O.: Image-Based Edge Bundles: Simplified Visualization of Large Graphs, Computer Graphics Forum, 29, 843–852, 2010.

Thomas, J. J. and Cook, K. A.: Illuminating the path: The research and development agenda for visual analytics, IEEE Computer Society Press, 2005.

Tominski, C., Abello, J., and Schumann, H.: CGV – An Interactive Graph Visualization System, Computers & Graphics, 33, 660–678, 2009.

Tominski, C., Donges, J. F., and Nocke, T.: Information Visualization in Climate Research, in: Proceedings of the International Conference Information Visualisation (IV), pp. 298–305, IEEE Computer Society, 2011.

Tominski, C., Gladisch, S., Kister, U., Dachselt, R., and Schumann, H.: A Survey on Interactive Lenses in Visualization, in: EuroVis State-of-the-Art Reports, pp. 43–62, Eurographics Association, 2014.

Tsonis, A. and Swanson, K.: Topology and Predictability of El Niño and La Niña Networks, Phys. Rev. Lett., 100, 228 502, 2008.

Tsonis, A. A. and Roebber, P. J.: The Architecture of the Climate Network, Physica A, 333, 497–504, 2004.

Tsonis, A. A., Swanson, K., and Kravtsov, S.: A new dynamical mechanism for major climate shifts, Geophysical Research Letters, 34, n/a–n/a, l13705, 2007.

Tsonis, A. A., Swanson, K. L., and Wang, G.: On the Role of Atmospheric Teleconnections in Climate, Journal of Climate, 21, 2990 – 3001, 2008.

Tupikina, L., Rehfeld, K., Molkenthin, N., Stolbova, V., Marwan, N., and Kurths, J.: Characterizing the evolution of climate networks, Nonlinear Processes in Geophysics, 21, 705–711, 2014.

van der Mheen, M., Dijkstra, H. A., Gozolchiani, A., den Toom, M., Feng, Q., Kurths, J., and Hernandez-Garcia, E.: Interaction network based early warning indicators for the Atlantic MOC collapse, Geophysical Research Letters, 40, 2714–2719, 2013.

van Ham, F. and Perer, A.: "Search, Show Context, Expand on Demand": Supporting Large Graph Exploration with Degree-of-Interest, IEEE Transactions on Visualization and Computer Graphics, 15, 953–960, 2009.

van Ham, F. and Wattenberg, M.: Centrality Based Visualization of Small World Graphs, Computer Graphics Forum, 27, 975–982, 2008.

von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J., Fekete, J.-D., and Fellner, D.: Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges, Computer Graphics Forum, 30, 1719–1749, 2011.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, 1999.

Wallace, J. and Gutzler, D.: Teleconnections in the Geopotential Height Field During the Northern Hemisphere Winter, Monthly Weather Review, 109, 784–812, 1981.

Wang, Y., Gozolchiani, A., Ashkenazy, Y., Berezin, Y., Guez, O., and Havlin, S.: Dominant Imprint of Rossby Waves in the Climate Network, Physical Review Letters, 111, 138 501, 2013.

Wiedermann, M., Donges, J. F., Heitzig, J., and Kurths, J.: Node-weighted interacting network measures improve the representation of real-world complex systems, Europhys. Lett., 102, 28 007, 2013.

Wiedermann, M., Donges, J. F., Handorf, D., Kurths, J., and Donner, R. V.: Hierarchical structures in Northern Hemispheric extratropical winter ocean-atmosphere interactions, arXiv preprint arXiv:1506.06634, 2015.

Withall, M., Phillips, I., and Parish, D.: Network visualisation: a review, IET Communications, 1, 365–372(7), 2007.

Wolff, A.: Graph Drawing and Cartography, in: Handbook of Graph Drawing and Visualization, edited by Tamassia, R., pp. 697–736, CRC Press, 2013.

Wong, P. C., Shen, H.-W., Leung, R., Hagos, S., Lee, T.-Y., Tong, X., and Lu, K.: Visual analytics of large-scale climate model data, in: Large Data Analysis and Visualization (LDAV), 2014 IEEE 4th Symposium on, pp. 85–92, 2014.

Yamasaki, K., Gozolchiani, A., and Havlin, S.: Climate Networks Around the Globe are Significantly Affected by El Niño, Physical Review Letters, 100, 228 501, 2008.

Zemp, D., Schleussner, C.-F., Barbosa, H., Van der Ent, R., Donges, J., Heinke, J., Sampaio, G., and Rammig, A.: On the importance of cascading moisture recycling in South America, Atmospheric Chemistry and Physics, 14, 13 337–13 359, 2014a.

Zemp, D., Wiedermann, M., Kurths, J., Rammig, A., and Donges, J. F.: Node-weighted measures for complex networks with directed and weighted edges for studying continental moisture recycling, EPL (Europhysics Letters), 107, 58 005, 2014b.

Zhang, Q., Eagleson, R., and Peters, T. M.: Volume Visualization: A Technical Overview with a Focus on Medical Applications, J. Digital Imaging, 24, 640–664, 2011.

Zhu, X. and Guo, D.: Mapping Large Spatial Flow Data with Hierarchical Clustering, Transactions in GIS, 18, 421–435, 2014.

Zou, Y., Donges, J. F., and Kurths, J.: Recent Advances in Complex Climate Network Analysis, Complex Systems and Complexity Science, 8, 27–38, in Chinese., 2011.