Ankit Agarwal | Levke Caesar | Norbert Marwan | Rathinasamy Maheswaran | Bruno Merz

Network-based identification and characterization of teleconnections on different scales

SCIENTIFIC REP**O**RTS

OPEN

Network-based identification and characterization of teleconnections on different scales

Ankit Agarwal [1,2,3], Levke Caesar[1,4], Norbert Marwan [1], Rathinasamy Maheswaran[1,5], Bruno Merz [2,3] & Jürgen Kurths[1,2,6]

Sea surface temperature (SST) patterns can – as surface climate forcing – affect weather and climate at large distances. One example is El Niño-Southern Oscillation (ENSO) that causes climate anomalies around the globe via teleconnections. Although several studies identified and characterized these teleconnections, our understanding of climate processes remains incomplete, since interactions and feedbacks are typically exhibited at unique or multiple temporal and spatial scales. This study characterizes the interactions between the cells of a global SST data set at different temporal and spatial scales using climate networks. These networks are constructed using wavelet multi-scale correlation that investigate the correlation between the SST time series at a range of scales allowing instantaneously deeper insights into the correlation patterns compared to traditional methods like empirical orthogonal functions or classical correlation analysis. This allows us to identify and visualise regions of – at a certain timescale – similarly evolving SSTs and distinguish them from those with long-range teleconnections to other ocean regions. Our findings re-confirm accepted knowledge about known highly linked SST patterns like ENSO and the Pacific Decadal Oscillation, but also suggest new insights into the characteristics and origins of long-range teleconnections like the connection between ENSO and Indian Ocean Dipole.

The ocean covers more than two-thirds of the Earth's surface. Given the higher heat capacity of water as compared to air (by approximately a factor of 4) and greater mass in the ocean than in the atmosphere, the ocean can store about 1000 times as much heat as the atmosphere[1]. Although it takes time to transport heat in or out of the ocean, even the ocean mixed-layer has a larger heat capacity than the atmosphere and can, therefore, strongly influence the climatic conditions. As a consequence, the ocean is capable of altering atmospheric conditions on seasonal and longer timescales[2].

As the sea surface forms a direct boundary between ocean and atmosphere, the sea surface temperature (SST) is of particular importance for setting the atmospheric conditions. In the opposite direction, large-scale atmospheric circulation patterns (e.g., quasi-stationary Rossby waves[3]) form atmospheric linkages between different oceanic regions around the globe and control the spatial extent of the interactions among oceans[4]. These long-distance interactions, referred to as *teleconnections*, are marked by a significant correlation of the climatological variables. Teleconnections have periodicities ranging from months to several decades. Teleconnections have been identified to affect temperature, precipitation, storm tracks, severe weather, and droughts[5]. An in-depth knowledge of the dynamics and patterns of the different teleconnections is required for an advanced understanding in climate science which has direct impact on societies, economics and ecosystem.

Extensive research has analysed the teleconnections in the atmosphere and oceans considering a broad range of temporal and spatial scales, in both the tropics and the extratropics[6–8]. Among the most well-known teleconnections are a number of defined patterns that are marked by a strong correlation of the SSTs at different places like the Pacific Decadal Oscillation (PDO) and the North Atlantic Oscillation (NAO), but also the link between El Niño Southern Oscillation (ENSO) and the Asian monsoon system[7,9–12].

[1]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Telegrafenberg, Potsdam, Germany. [2]Institute for Environmental Sciences and Geography, University of Potsdam, Potsdam, Germany. [3]GFZ German Research Centre for Geosciences, Section 4.4: Hydrology, Telegrafenberg, Potsdam, Germany. [4]Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany. [5]MVGR college of Engineering, Vizianagaram, India. [6]Institute of Physics, Humboldt Universität zu Berlin, Berlin, Germany. Correspondence and requests for materials should be addressed to A.A. (email: agarwal@pik-potsdam.de)

Recent studies have raised concern that our understanding of climate processes remains incomplete, because interactions in multi-scale systems, like the climate system, are typically exhibited at unique or at multiple temporal and spatial scales[11,13,14]. Methods are required that help to unveil the interactions in the climate system at different scales, which usually remain hidden when looking at a particular scale only. Also, it is important to quantify the spatial distance of teleconnection variability between remote ocean regions at different timescales which could provide a quantitative understanding of short-range and long-range coupling between different oceanic regions.

To uncover such spatial and temporal variable interactions we use the complex network approach[15], which recently has emerged as a powerful framework in extracting information from large high-dimensional datasets[16,17]. In addition, this non-parametric method allows investigating the topology of local and non-local statistical interrelationships[18]. Recently complex network analysis has facilitated the identification of spatial patterns in various research fields[19]. One such example are climate networks[18,20] that are constructed from climate reanalysis data. Each grid cell of the climate dataset is considered as a network node and links between each pair of grid cells (or nodes) are setup using a similarity measures such as Pearson's correlation coefficient[18,20], event synchronization[21–24], mutual information[14], transfer entropy[25] and multi-scale event synchronization[26]. Only the strongest links, filtered using either a predefined link density or fixed threshold, are retained in the final networks to help uncover the strongest interactions. Several network measures (such as degree, clustering, betweenness, community structures) have been used on the resultant climate network to capture local and global dependent climatic patterns within and among climate variables[23,27–29], like global patterns of extreme rainfall teleconnections[14] and spatial diversity of Indian rainfall teleconnections[30].

To date, the complex network approach for the analysis of climate data has mostly been used to investigate patterns at one reference timescale only. However, the extension of the climate network to multiple timescales can reveal additional and important information as demonstrated recently[26,29,31].

In this study, the wavelet-based correlation method is used by combining the wavelet transform and Pearson correlation[31]. We present a well-accepted climate network constructed from global monthly SST data, with a focus on identifying dependence structures on different timescales over the entire globe. In particular, we construct climate networks on eight different timescales, ranging from the monthly to interdecadal scale, using the wavelet-based correlation measure and interpret the detected short- and long-range links. Our findings re-confirm accepted knowledge about highly linked SST patterns like ENSO and PDO, but also provide new insights into the characteristics and origins of long-range teleconnections.

Data and Methods

In this section, we describe the data characteristics used as well as methodological steps required for the construction of the networks.

Data characteristics. The climate networks constructed are based on the global monthly sea surface temperature (SST) data provided by NOAA's Earth System Research Laboratory (ESRL). The data (Extended reconstructed SST V3b) has a spatial resolution of 2.0-degree latitude x 2.0-degree longitude and is given for the time period 1979–2015. The data is freely available at NOAA's portal (https://esrl.noaa.gov/psd/). As a preprocessing step we have removed 1056 grid points (out of total 10512) with missing values or gaps, hence in total 9456 grid points are considered. To avoid artefacts due to autocorrelation and seasonality, we removed the seasonal cycle and normalized the data. Specifically, we calculate for every month (i.e., separately for all Januaries, Februaries, etc.) the long-term mean and standard deviation. Each data point is then normalized by subtracting the mean and dividing by the standard deviation of the corresponding month at that grid cell. This normalization significantly reduces temporal autocorrelation in the time series[31].

Wavelet multiscale correlation. To construct the SST network, each SST grid cell is considered as a network node, and links between each pair of nodes are set-up based on a statistical relationship between them. The similarity measure used is the wavelet multiscale correlation (WMC) between the monthly anomaly series proposed by[31]. The proposed WMC method is a combination of two well established methods, the maximum overlap discrete wavelet transformation (MODWT) and the Pearson's correlation. All the mathematical details are presented in the Additional information.

Network construction. Formally, a network or graph is defined as an ordered pair G = {N, E}, containing a set of nodes N = {N_1, N_2, …, N_N} together with a set E of edges {i, j} which are 2-element subsets of N. In this work, we consider undirected and unweighted graphs (G), where only one edge can exist between a pair of nodes and self-loops of the type {i,i} are not allowed. Hence, edges simply show connections between nodes, and each edge can be traversed in either direction. This type of graph can be represented by the symmetrical adjacency matrix[22,29]

$$A_{i,j} = \begin{cases} 1, & \{i, j\} \in E \\ 0, & \{i, j\} \notin E. \end{cases}$$

(1)

To construct a climate network (Fig. 1), each grid cell of SST data is considered as a node of the network and undirected and unweighted edges are determined between all possible pairs of nodes based on their similarity. We use the WMC measure to quantify the similarity between the SST time series, yielding a square similarity matrix of 9456 × 9456 (total number of grid cells considered). We generate an adjacency matrix from the similarity matrix by applying a significance-based pruning which is explained in the next section.
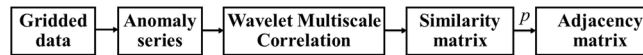
Figure 1. Schematic of network construction. Each grid cell of the SST dataset is considered as node and similarity between each pair of nodes is calculated using the WMC measure. By applying the 95*th* percentile threshold along with multiple testing, a link between each pair of nodes is set up.

Percentile threshold	Original scale	1–2 months	2–4 months	4–8 months	8–16 months	16–32 months	32–64 months	64–128 months
90th	0.32	0.12	0.16	0.24	0.33	0.48	0.55	0.67
95th	0.41	0.17	0.22	0.32	0.44	0.61	0.66	0.76
99th	0.72	0.26	0.32	0.43	0.65	0.86	0.80	0.84
Multiple testing	Original scale	1–2 months	2–4 months	4–8 months	8–16 months	16–32 months	32–64 months	64–128 months
FDR	0.62	0.51	0.56	0.63	0.74	0.84	0.85	0.85

Table 1. Overview of sea surface temperature network threshold values (90*th*−, 95*th*−, 99*th* and FDR) at all timescales. In this study we have selected threshold values obtained using False discovery rates (FDRs) method.

Multiple testing. A number of criteria have been proposed to generate an adjacency matrix from a similarity matrix, such as a fixed amount of link density[22,32] or fixed thresholds[23]. Here, we consider a 5% link density since it is a well accepted criteria globally for the network construction. However, here we combine it with multiple testing attempts to avoid false links by controlling the type I error or adjusting p-values to give only significant links[33,34]. This is a very active area in statistics, and a range of methods have been proposed. Here, we use false discovery rates (FDR) to control p-values (for more details see Benjamini and Yosef[33]). This may seem like a stringent requirement (5% link density combined with multiple testing), but a large number of edges satisfies this criterion and are therefore retained in the network. To give a clear understanding of the final selected threshold, we tabulate all the threshold values (90*th*−, 95*th*−, 99*th* and FDR) in Table 1.

Results

We present the results in three parts, first discussing the scale-specific spatial patterns obtained over different timescales, second investigating the occurrence of short- and long-range linkages between the identified scale-specific spatial patterns over different timescales, and finally providing a 3-D global visualization of the link distributions.

Spatial patterns over different timescales. Using the above-explained similarity measure and pruning method, we construct a separate SST network for each timescale (in total eight scales including the observed timescale, see Fig. 2). The networks reveal, separately for each timescale, how well the temporal evolution of the SSTs of each grid cell is linked to that of the other grid cells, i.e., it shows the number of significantly correlated edges ($\sum A_{ij}$, also called links) that each grid cell (or node) has. This gives us a basis for understanding what kind of teleconnections act at a certain timescale.

Considering only the positively correlated links (Fig. 2), it is obvious that at all scales the majority of the grid cells show a low link number that is larger than zero. We identify this as the background link number at a certain scale that is explained by the near-field spatial correlations of the grid cells[11,35]. In addition, there are distinct regions with a substantially higher number of links.

At the original, i.e., unprocessed scale (Fig. 2a) we observe a highly linked region in the mid to eastern equatorial Pacific accompanied by regions with a medium number of links in the southern Pacific, the western equatorial Pacific and the Indian Ocean. When looking at the scale-specific networks (Fig. 2b–h), it becomes clear that links at the observational scale are aggregated results of the coupling between scale-specific features of global SST data present at different timescales. Moving from lower to higher temporal scales, it becomes apparent that – with the annual cycle removed from the data – there are hardly any significant correlations between the SST evolution in different regions at timescales of less than a year (Fig. 2b–d).

However, at the 8–16 months scale (Fig. 2e), there are two zones with a relatively large number of significant connections in the mid to eastern equatorial Pacific and the southern Pacific. There is also a region with an above global average number of links in the Indian Ocean. While the number of links in this region increases when moving to the larger timescale of 16–32 months (Fig. 2f), the number of links from and to the Southern Ocean decreases.

At the scale of 32–64 months (Fig. 2g), the number of links in the mid to eastern Pacific Ocean decreases and a well-linked region in the form of a tilted horseshoe emerges in the northern Pacific, as well as a region with a medium number of links in the subtropical North Atlantic. On even longer timescales (64–128 months) the highly linked region in the equatorial Pacific re-emerges, as well as several smaller regions with a medium number of links in all three ocean basins (Fig. 2h).

Figure 3 shows the number of negatively correlated links. In contrast to the positive ones (Fig. 2), there is no global background link number due to the fact that the near-field correlation is in general positive. On the original scale, regions whose SST evolutions are negatively correlated to that of other regions can mainly be found in the Pacific Ocean (Fig. 3) centered around the equator as well as in the Southern Ocean. The separation in
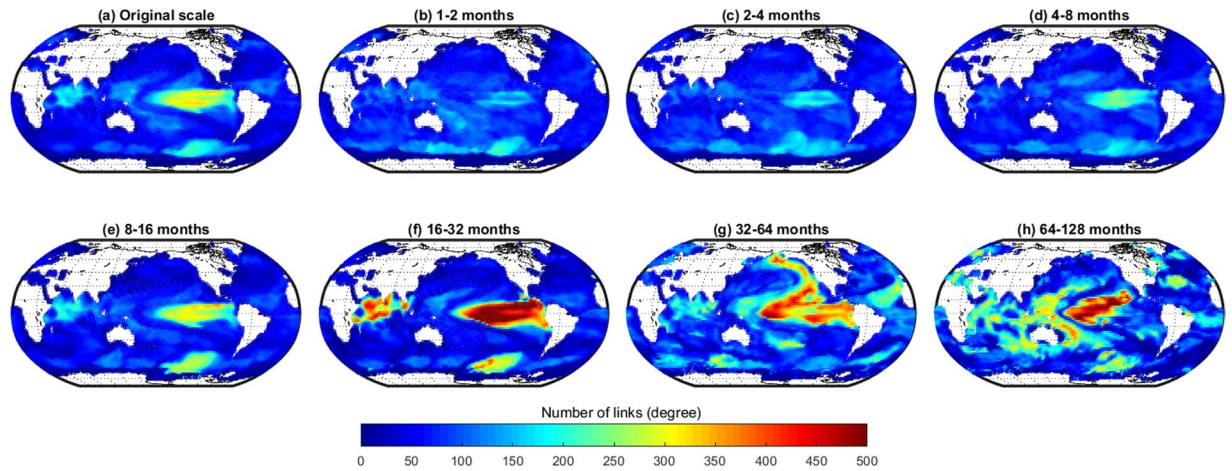
Figure 2. Number of links of each grid cell in a global SST network constructed at different timescales by considering only negative correlations.
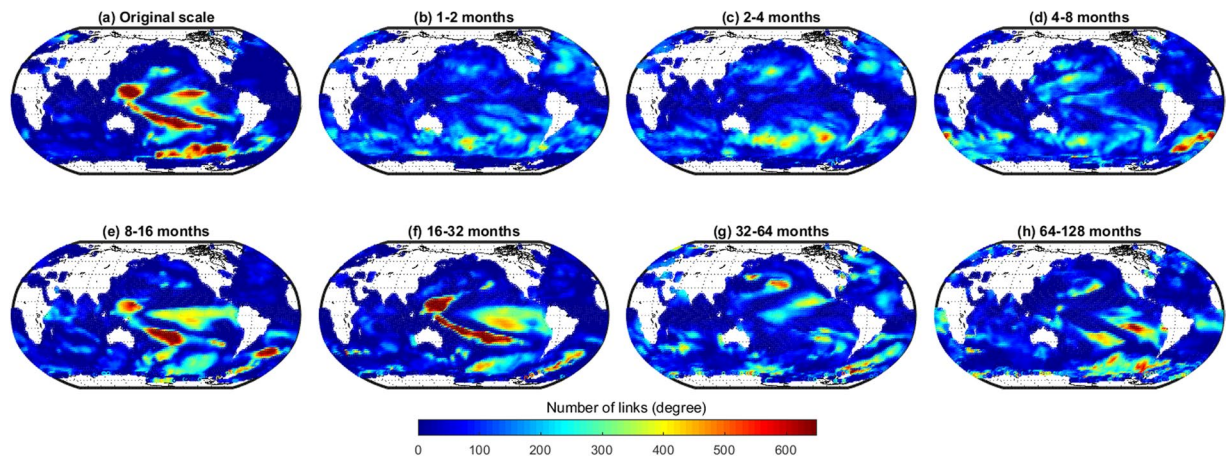

Figure 3. Number of links of each grid cell in a global SST network constructed at different timescales by considering only positive correlations.

the different timescales again reveals that the observed patterns are aggregates of the timescale-specific feature. Again there are only a few highly linked regions on timescales of less than a year (Fig. 3b–d) with the exception of the subtropical South Atlantic (Fig. 3d). On longer timescales of about one to two years the link number in the South Atlantic stays rather stable and highly linked regions in the Pacific emerge (Fig. 3e,f), namely in the mid to eastern equatorial Pacific and the equatorial to south-western Pacific. The latter becomes more linked on the 16–32 months timescale, while the link numbers of the former decrease. Moving to interannual timescales, the regions with a large number of significantly negatively correlated edges globally decrease but a few regions in the South Pacific remain (Fig. 3h).

Occurrence of short- vs. long-range links. We further investigate the length distribution of the significantly correlated edges (both positive and negative) at the different scales by creating histograms showing the number of links plotted over the geographical length link (Fig. 4). The geographical link length (shortest geographical distance between two points on a globe) was computed using Haversine's formula which determines the great circle distance between two spatially embedded points on the sphere[36].

For all scales (Fig. 4a–h) these frequency distributions of the link lengths show that the dominant factor controlling the significance of a correlation between two SST time series is the near-field spatial correlation. This property is embodied by Tobler's First Law of Geography, which states that 'everything is related to everything else, but near things are more related to distant things'[37]. Nevertheless, the number of links decreases again at very short distances. There are two reasons for that: First, because the nodes are evenly spaced angularly on the sphere, neighboring grid points near the equator are much further apart (>500 km) than those near the poles. Second, the maximum number of possible links depends on the geographical link length due to the spherical shape of the Earth and the land ocean distribution on it; in general, the number of possible links increases with the link length until a maximum at a length of about 10,000 km and then decreases again. The second noticeable
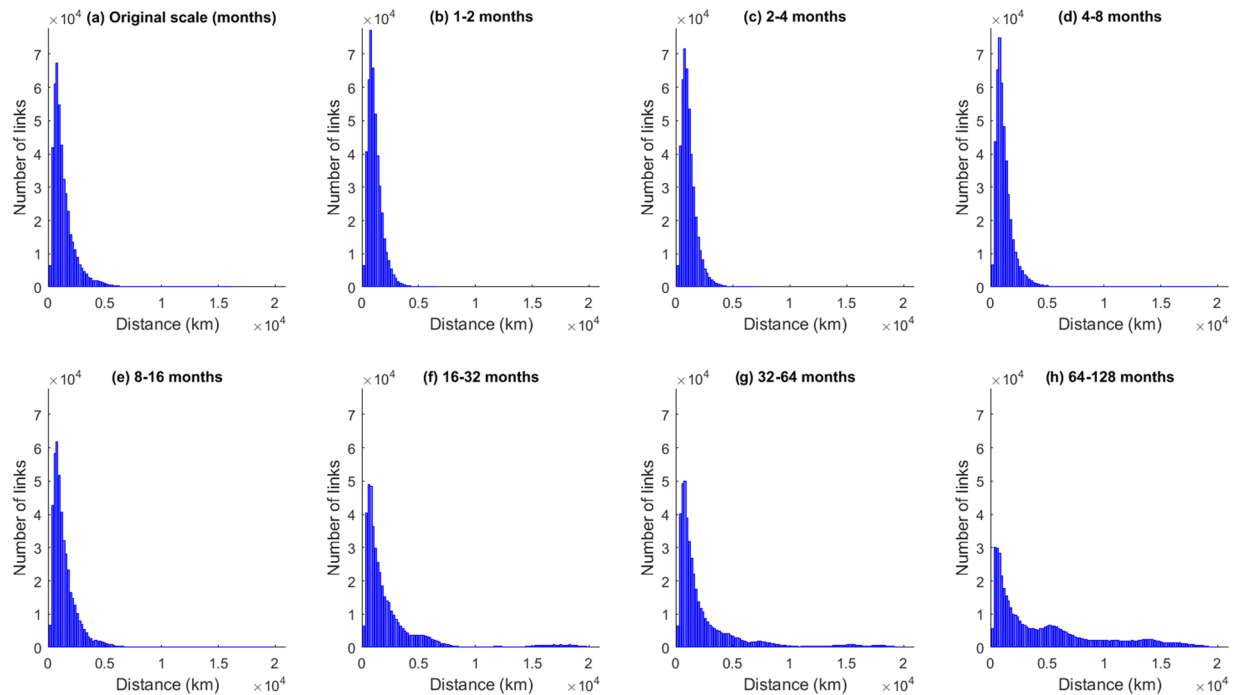
Figure 4. Frequency distribution of the link lengths in SST networks at different timescales. Short-distance links can be interpreted as near-field correlations, while long-distance links suggest long-range spatial dependencies or teleconnections.

feature is a 'fat tail' for the longer timescales (16–32 months and longer, Fig. 4f–h) that suggests the presence of teleconnections, i.e., a similar time-dependent behavior in locations that are geographically separated.

To better understand the nature of these teleconnections, we divide the entire globe into three standard regions, namely the extratropical Northern Hemisphere (NH), the Tropical Hemisphere (TH) and the extratropical Southern Hemisphere (SH), and analyse separately the frequency distributions of the link lengths of edges that connect nodes only within these regions for timescales larger than 16–32 months (Fig. 5).

It can be seen that short-range links appear in all regions on all the considered timescales (i.e. larger than 16 months) with the smallest number of short-range links occurring at the very long timescales (64–128 months) for all three hemispheres. When looking at the long-range links, the three hemispheres show a different behaviour.

In the tropics, teleconnections appear on timescales from 16–32 months and on the longer timescales of more than 64 months. At these timescales, there are also long-range links in the two extratropical hemispheres that start developing at the 32–64 months timescale at which the tropics show no long-range links and additionally a notable reduction in short-range links. The geographical lengths of the teleconnections in the tropics can be longer than those in the extratropics due to Earth's spherical shape as grid points in the tropics have larger distances. Hence, we conclude that the length distribution of links varies between the different decadal timescales for the global and the within-region analysis.

As most of the teleconnections that we are interested in act across the three zones, we further analyse the link length frequency distribution from one region (NH, TH, and SH) to the rest of the earth (including within-region links, Fig. 6). This will help us to understand the interactions between tropics and extratropics which cannot be revealed by Fig. 5.

Figure 6 confirms that the extratropics (NH & SH) do not possess teleconnections at the interannual scale (16–32 months) both within the same region (shown in Fig. 5) and with the rest of the earth (Fig. 6). We also notice that the absolute number of short-range connections (Fig. 6) at this scale remains almost unchanged in all three regions except for slight increments in the extratropical SH which depict the near-field interactions between SH and TH. These links connect regions with a distance of about 5,000 km, which is visible in Figs 2 and 3.

The first interesting observation at the intradecadal scale (32–64 months) is the significant increase in short-range connections in the extratropics (NH and SH) in contrast to almost no significant changes in the tropics compared to the interannual scale (Fig. 6). However, Fig. 5e already confirmed that the tropics do not have much within-region short-range connections at the 32–64 months scale compared to the 16–32 month scale (Fig. 5d). Also the increase in short-range connections in NH is mainly credited to internal short-range connections (Fig. 5b). These observations suggest that SH exhibits a significant number of short-range links with the tropics at the intradecadal scale. Moreover, we find that long-range connections appear in the extratropics (NH & SH) which are completely absent at the interannual and other finer scales (Fig. 6a,c) or when looking only at the within-region links (Fig. 5b,h). This results indicates that interactions in the climate system occur at specific timescales but are absent at other timescales, which undermines the necessity of multi-scale analyses.
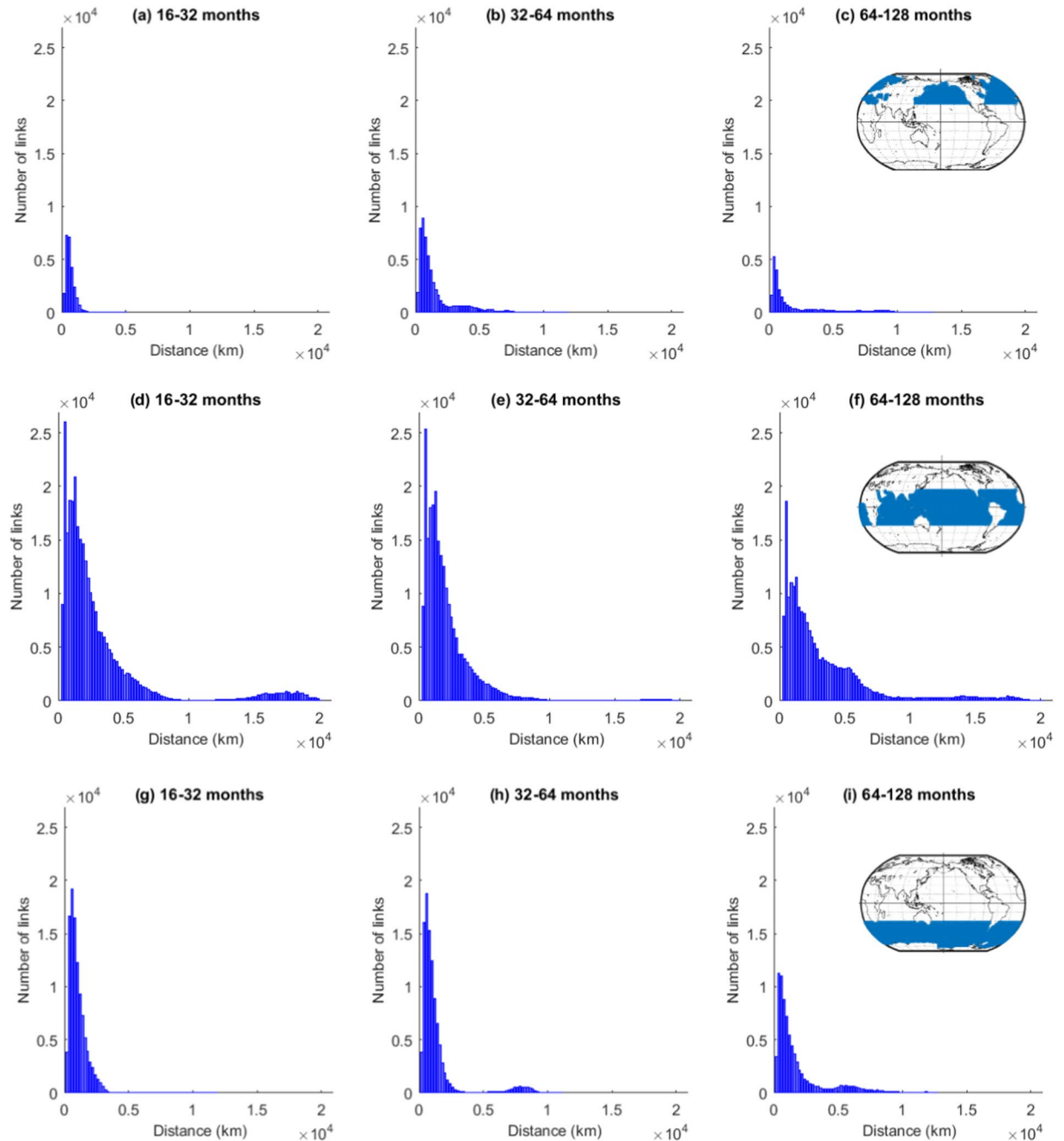
Figure 5. Frequency distribution of the edge lengths in SST networks at different timescales within three standard zones, namely NH, TH and SH (within-region links). Short-distance links show near-field spatial correlation, while long-distance links suggest long-range spatial dependencies or teleconnections within the same zone.

At the decadal scale (64–128 months) and for all three zones, we observe a significant reduction in beyond-region, short-range connections compared to the intradecadal scale. For NH, there is the same reduction as is seen for within-region links, whereas the reduction is larger for SH and TH. This observation indicates a decline in near-field connections within NH and in-between TH and SH. The rise in NH and SH at distances of about 10,000 to 15,000 km shows the significant coupling of the two hemispheres to the other regions. The fact that for both extratropical hemispheres the number of very long links increases, while it stays rather constant for TH, indicates the existence of very long-range links between SH and NH.

These analyses uncover that (i) at the interannual scale the tropics show the largest number of significant correlations acting over both short and long distances and the existence of links with a length of about 5,000 km between SH and TH (Fig. 6d,g); (ii) at the intradecadal scale the main characteristics are long-range links between NH and TH with a length of about 5,000 km and 7,000 km (Fig. 6b,e), as well a small number of links between SH
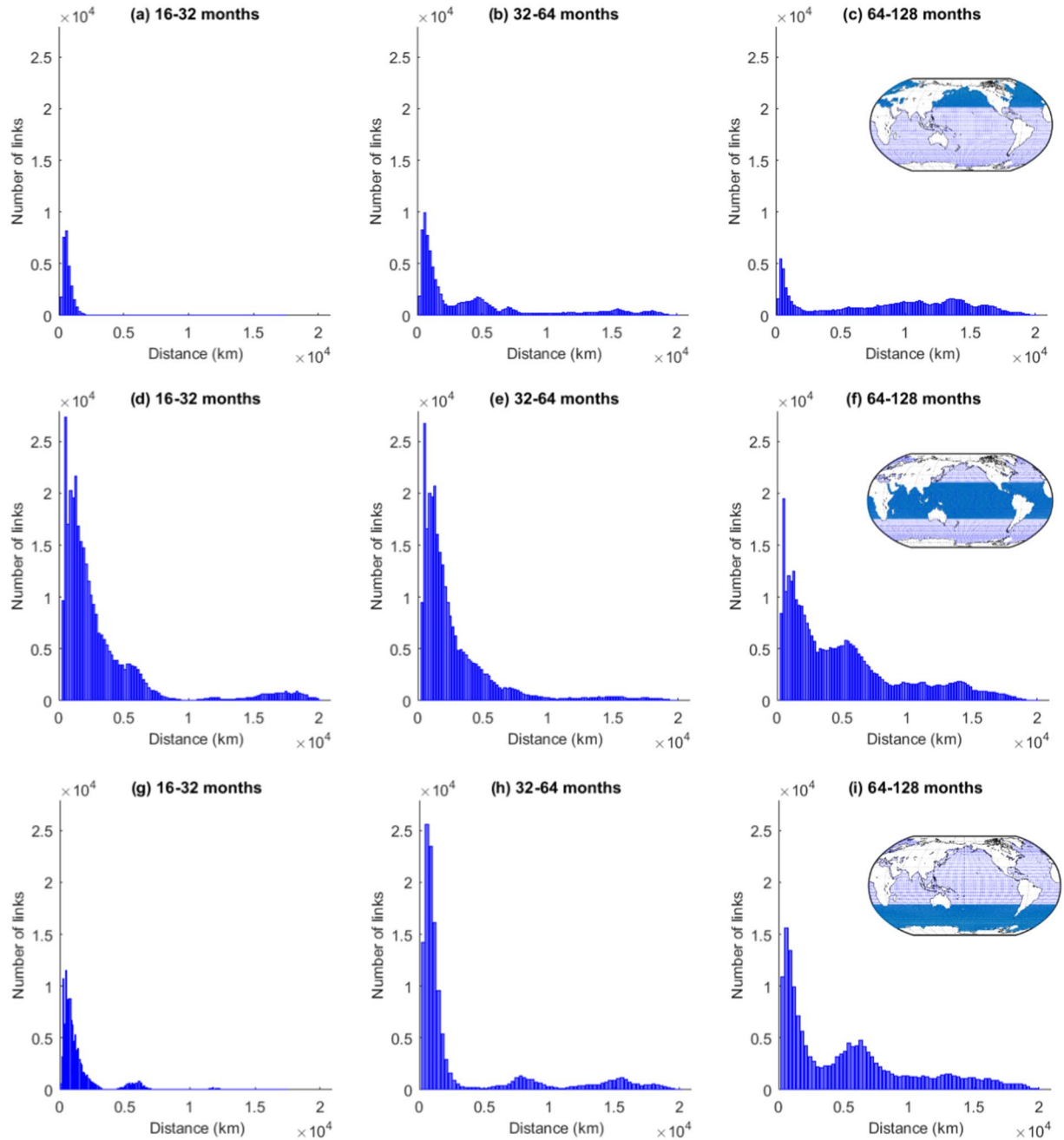
Figure 6. Frequency distribution of the edge lengths in SST networks at different timescales constructed for standard zones, namely NH, TH, and SH, considering the interaction with the whole globe (beyond-region links). Short-distance links show near-field correlation, while long-distance links suggest long-range spatial dependencies or teleconnections.

and TH with a length of about 15,000 km, yet there are no long-range links within TH; (iii) at the decadal scale long-range connections between NH and SH to the tropics as well as interhemispheric (SH to NH) links exist.

3-D visualization of the link distributions. The results listed above identify the ocean regions that are most closely linked when looking at the variability of SST at different timescales going from the intra- to inter-annual and even decadal scale (Figs 2 and 3). The analysis of the frequency distributions of the link lengths gives further insight into the occurrence of teleconnections (Figs 4–6). What is missing is a detailed knowledge about the actual regions that are linked at the different timescales.

We, therefore, visualize the climate networks on spherical 3-D node-link maps (Fig. 7) constructed at different timescales, where the node positions are fixed, and links are plotted using the software GTX tool. To best visualize the structures of the global climate network, especially the connectivity of different regions either via short- or long-range links, we do not plot all links but focus on the regions that were identified as being highly linked. The
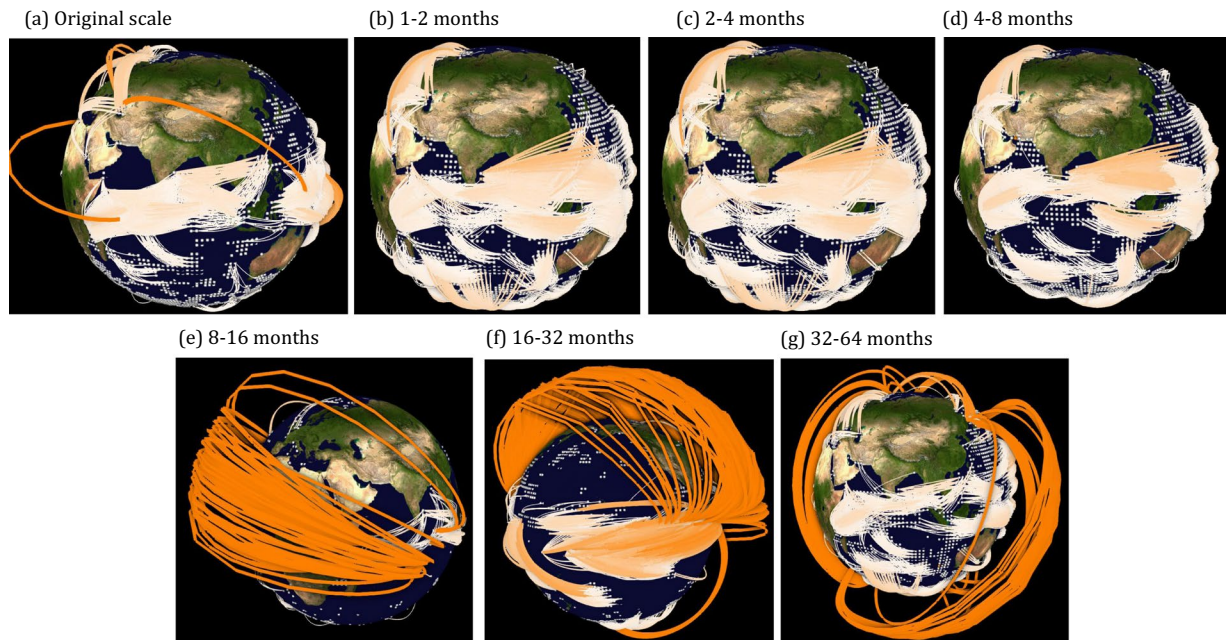
(a) Original scale (b) 1-2 months (c) 2-4 months (d) 4-8 months

(e) 8-16 months (f) 16-32 months (g) 32-64 months

Figure 7. Spherical 3-D globe representation of the short- and long-range connections in SST networks at different timescales. Only nodes beyond a pre-selected betweenness centrality (BC) value are plotted. For instance, in subplot (**a–d**),(**e–f**) and (**g**) nodes are plotted for BC values greater than 90 K, 57 K and 38 K, respectively. Edge color represents the geographical link length.

3-D projection of the networks reveals patterns consistent with the 2-D visualization (Figs 2 and 3). However, it contains additional information about short- and long-range connections, e.g., even though the 2-D visualization at the original scale (Figs 2a and 3a) depicts approximately the same patterns as the 16–32 months scale (Figs 2e and 3e), the 3-D visualizing clarifies that there are hardly any long-range connections present at the original scale (Fig. 7a) compared to the 16–32 month scale (Fig. 7e). This visualization of the connections helps in attributing the interrelationships between selected regions at certain timescales. Thus, the node-link maps add valuable information to the discussion and interpretation of the results obtained in previous Sections.

Discussion

In the following we link our results to known modes of climate variability and teleconnections and highlight the new findings of this study.

The link maps (Figs 2 and 3) clearly reveal the known modes of SST variability on the various time-scales. The highly linked region in the eastern equatorial Pacific can be identified as El Niño/Southern Oscillation (ENSO), a periodic fluctuation in SST (accompanied by a fluctuation in air pressure) across the equatorial Pacific, that occurs with an average time between peak events of about four years[38] and can therefore be seen best at the 16–32 month scale (Fig. 3f). At the same timescale, the region of variability of the Indian Ocean Dipole (IOD) shows a large link number, as IOD events acts on about the same timescale as ENSO[39]. ENSO and IOD are known to impact each other via the atmosphere[40–42], and the node-link map reveals that the links between the two modes exist mainly between the northern part of the ENSO tongue and the equatorial Indian Ocean (Fig. 7e,f). Furthermore, the analysis exhibits that the arrow shaped pattern of negatively correlated links in the western equatorial Pacific (Fig. 3f) that occurs at the same timescale as El Niño is neither linked to ENSO nor to IOD.

The last prominent feature at the intraannual timescale is the highly linked area in the Southern Ocean (e.g., Figs 2f and 3f). Our 2-D visualization confirms here what has just been analysed by Ferster *et al.*[43]: a link between SST in the Southern Ocean to ENSO events via the Southern Annular Mode (SAM), i.e. the north–south movement of the westerly wind belt that circles Antarctica. Positive states of SAM correlate with a cooling of SSTs in the high-latitude Southern Ocean, and a warming within the Southern Hemisphere sub-tropics and mid-latitudes[43]; this negative correlation can be seen in in our 2-D visualization (Fig. 3f). Our 3-D visualization further uncover a teleconnection between the very eastern part of ENSO and the region in the Southern Ocean near Western Antarctica (Fig. 7). From the 2-D maps it follows that this teleconnection has to be positively correlated (Fig. 2f).

Going to longer timescales (32–64 months) we find the majority of the highly linked regions in the northern Pacific. The form and location of these regions resemble strongly the pattern of the Pacific Decadal Oscillation (PDO), a pattern in SST and sea-level pressure that varies between positive and negative phases on periods of 20–30 years. Yet we find high link numbers at the interannual scale. A detailed look shows that this is not a contradiction to the known facts: Viewing the 3-D visualization (Fig. 7g) we find that most of the links are not within the PDO pattern (these indeed should act on a 20–30 year scale), but are teleconnections to the Southern Ocean. These can act on shorter timescales, likely via a western Pacific ocean-atmosphere pathway, whereby SST anomalies can propagate from the southern to the northern hemisphere[5].

But not all teleconnections work via the atmosphere. The 3-D visualization (Fig. 7g) shows a number of links going from the North Atlantic to the South Atlantic (this can also be seen in Fig. 3f). This negative correlation likely shows the so-called see-saw response[44] that is due to the transport of heat from the Southern Ocean to the North Atlantic via the Atlantic Meridional Overturning Circulation (AMOC). If the AMOC is stronger (as it has been in the period after year 2000 compared to the years before[45]), more heat is transported towards the north which leads to a cooling in the Southern Ocean and a warming in the subpolar North Atlantic. (The link diagram for the next '64–128' months scale is not available, since the network size has grown to an extent which was unable to process in the used software.)

More general information on where and at what distances the teleconnections act is given by the histograms of the frequency distributions of the link lengths. Within one region we find teleconnections mainly in the tropics (Fig. 5d) acting on a length scale of about 18,000 km. There are also a few long-range links in both the extratropics NH and SH, but here on all timescales the short-range links dominate. This is different when looking at the links from one region to the others (Fig. 6). Here a number of long-range links are detected, especially on the 64–128 month scale on spatial scales of about 6,000–7,000 km as well as about 10,000–15,000 km. This is likely due to the fact that teleconnections often work via Rossby waves that link the tropics to the extratropics. The wavelengths of these Rossby waves depend on the strength of the westerly winds, but are typical about 7,000 km for weak westerly flow in the mid-latitudes and increase for stronger winds and more polar regions[46].

In general, we can conclude that our method enables us to distinguish between regions that are highly linked through significantly correlated short-range edges and those that have long-range teleconnections. While the former are typical for proximity-based correlation on very short timescales as well as for correlations within the region where a climate mode takes place (like El Niño or the Indian Ocean dipole) and are mainly due to the advection of heat by the oceanic surface, the long-range links correspond to teleconnection patterns that either act via the atmosphere, e.g., through Rossby waves, or through the major open circulation systems like the AMOC. The latter warrant even more investigation as they are of particular importance for the understanding and forecasting of lasting weather events like droughts, flooding and temperature extremes.

Conclusions

Ocean-atmospheric teleconnections are intriguing phenomena in the climate system as they connect very remote regions. The thorough diagnosis of their characteristics will improve our understanding of the climate system with large potential benefits, such as improvement in seasonal forecasting of climate variables and weather extremes. Here, we propose a unified framework to investigate and unravel short-range and long-range teleconnections at different timescales and spatial distances based on wavelet multi-scale correlation and give a first insight into the potential usage of this method. The study detects the highly linked regions in the oceans at different scales along with the identification of the link type, i.e. spatial proximity connections, short-or long-range teleconnection. In addition, the proposed method allows identifying the regions that are highly linked within themselves regarding their SST evolution and distinguish them from regions that have long-range teleconnections to other ocean regions. It unravels, for example, which parts of the well-known ENSO SST pattern in the equatorial Pacific are intralinked and from which regions the teleconnections with other climatic modes originate. This adds valuable information to the understanding of the functioning of teleconnections. The method has further potential to understand how far a particular teleconnection has an atmospheric or oceanic origin. This could be worthwhile for regional climate modeling since it highlights the regions that are necessary to include when modeling climate variability at a certain scale.

Wavelet multi-scale correlation. MODWT decomposes the time series into different time scales or frequency components. The wavelet decomposition is realized using the two basis functions known as father wavelet ($\phi(t)$) and mother wavelet ($\psi(t)$). The general admissibility conditions for ψ to be called a wavelet function are

$$\int_{-\infty}^{\infty} \psi(t)\mathrm{d}t = 0 \tag{2}$$

$$\int_{-\infty}^{\infty} \psi(t)^2\mathrm{d}t = 1 \tag{3}$$

Any function $f(t)$ can be expressed through these basis functions and their scaled and translated versions are

$$
\begin{aligned}
f(t) = {} & \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\Phi_{J,k}(t) \\
& + \sum_k d_{J-1,k}\Phi_{J-1,k}(t) + \cdots + \sum_k d_{1,k}\Phi_{1,k}(t)
\end{aligned}
\tag{4}
$$

where J is the total number of scales to be analyzed, and k is in the range of 1 to l (length of the time series). The coefficients $S_{J,k}$, are the approximation coefficients and $d_{J,k}, \ldots, d_{1,k}$ are the wavelet transform coefficients at scales J to 1, while the functions $\phi_{J,k}(t)$ and $\{\Psi_{j,k}| j = 1, \ldots, J - 1, J\}$ are the basis functions which are obtained through translation and dilation of the father of father ($\phi(t)$) and mother ($\psi(t)$) wavelets function respectively.

The mother and father wavelet is scaled (or dilated) by a factor j and translated (or shifted) by a factor k to give

$$\phi_{j,k}(t) = 2^{\frac{-j}{2}}\phi(2^{-j}t - k) \tag{5}$$

$$\Phi_{j,k}(t) = 2^{\frac{-j}{2}}\Phi(2^{-j}t - k) \tag{6}$$

Further, the values of the wavelet transform coefficients at each of the scale and the approximation coefficients at scale J are estimated by:

$$d_{j,k} \approx \int \Phi_{j,k}(t)f(t)dt, \quad j = 1, \ldots, J - 1, J \tag{7}$$

and

$$s_{j,k} \approx \int \phi_{J,k}(t)f(t)dt \tag{8}$$

where the scaling coefficients $s_{J,k}$ capture the smooth trend of the time series at the coarse scale 2^J, which are also called smooth coefficients; and the wavelet coefficients $d_{j,k}$, also known as detail coefficients can detect deviations from the coarsest scale to the finest scale.

The original series $f(t)$ can be reconstructed by summing up the detailed components and the smooth components:

$$f(t) = S_{J,k} + D_{J,k} + D_{J-1,k} + \cdots + D_{1,k} \tag{9}$$

where $S_{J,k} = \sum_k s_{J,k}\phi_k(t)$; $D_{J,k} = \sum_k d_{J,k}\Phi_{J,k}(t)$; $D_{1,k} = \sum_k d_{1,k}\Phi_{1,k}(t)$.

Further, consider two time series $\{X(t)\}$ and $\{Y(t)\}$ with the same length T. The wavelet multi-scale correlation (WMC) measure between both time series can be estimated as[47,48]

$$WMC = \rho_{X,Y}^{l_j} \equiv \frac{COV_{X,Y}^{l_j}}{Var_Y^{l_j}Var_X^{l_j}} \tag{10}$$

where

$$Var_X^{l_j} = \frac{1}{T_j}\sum_{t=M_j-1}^{T-1}[d_{j,t}^X]^2 \tag{11}$$

$$Var_Y^{l_j} = \frac{1}{T_j}\sum_{t=M_j-1}^{T-1}[d_{j,t}^Y]^2 \tag{12}$$

$$COV_{X,Y}^{l_j} = \frac{1}{T_j}\sum_{t=M_j-1}^{T-1}d_{j,t}^X d_{j,t}^Y \tag{13}$$

$d_{j,t}^{(\cdot)}$ denotes the MODWT wavelet coefficient of variables $\{X, Y\}$ at scale l_j; $T'_j = T - M_j + 1$ stands for the number of coefficients unaffected by the boundary; $M_j = (2^j - 1)(M - 1) + 1$ represents the length of the scale l_j wavelet filter, and M is the width of the wavelet filter. The multi-scale correlation measure $\rho_{X,Y}^{l_j}$ denotes the scale-wise correlation between X and Y at different l_j scales. Like the Pearson correlation coefficient (PCC), the value of $\rho_{X,Y}^{l_j}$ ranges between -1 and 1. $\rho_{X,Y}^{l_j} = 0$ implies the variables X and Y are not correlated at scale l_j. $\rho_{X,Y}^{l_j} = 1$ and $\rho_{X,Y}^{l_j} = -1$ indicate that the variables are perfectly correlated and anti-correlated, respectively. The values of $\rho_{X,Y}^{1l_j} \ldots \rho_{X,Y}^{l_j}$ indicate the strength of the relation between X and Y at different temporal scale bands.

References

1. Solomon, S., Qin, D. & Manning, M. *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 New York, 2007).
2. Bigg, G. R., Jickells, T. D., Liss, P. S. & Osborn, T. J. The role of the oceans in climate. *International Journal of Climatology* **23**, 1127–1159, https://doi.org/10.1002/joc.926 (2003).
3. Kornhuber, K. *et al.* Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters* **14**, 054002 (2019).
4. Griggs, D. J. & Noguer, M. Climate change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. *Weather* **57**, 267–269, https://doi.org/10.1256/004316502320517344 (2002).
5. Mamalakis, A., Yu, J.-Y., Randerson, J. T., AghaKouchak, A. & Foufoula-Georgiou, E. A new interhemispheric teleconnection increases predictability of winter precipitation in southwestern US. *Nature Communications* **9**, https://doi.org/10.1038/s41467-018-04722-7 (2018).
6. Fovell, R. G., Tung, W.-W. & American Meteorological Society. *Multiscale convection-coupled systems in the tropics: a tribute to Dr. Michio Yanai*, OCLC: 1024071001 (2017).
7. Kucharski, F., Kang, I.-S., Straus, D. & King, M. P. Teleconnections in the Atmosphere and Oceans. *Bulletin of the American Meteorological Society* **91**, 381–383, https://doi.org/10.1175/2009BAMS2834.1 (2010).
8. Ramirez, E., da Silva Dias, P. L. & Raupp, C. F. M. Multiscale Atmosphere–Ocean Interactions and the Low-Frequency Variability in the Equatorial Region. *Journal of the Atmospheric Sciences* **74**, 2503–2523, https://doi.org/10.1175/JAS-D-15-0325.1 (2017).
9. Azad, S. & Rajeevan, M. Possible shift in the ENSO-Indian monsoon rainfall relationship under future global warming. *Scientific Reports* **6**, https://doi.org/10.1038/srep20145 (2016).

10. Behera, S. K., Krishnan, R. & Yamagata, T. Unusual ocean-atmosphere conditions in the tropical Indian Ocean during 1994. *Geophysical Research Letters* **26**, 3001–3004, https://doi.org/10.1029/1999GL010434 (1999).

11. Steinhaeuser, K., Ganguly, A. R. & Chawla, N. V. Multivariate and multiscale dependence in the global climate system revealed through complex networks. *Climate Dynamics* **39**, 889–895, https://doi.org/10.1007/s00382-011-1135-9 (2012).

12. Zhou, D., Gozolchiani, A., Ashkenazy, Y. & Havlin, S. Teleconnection Paths via Climate Network Direct Link Detection. *Physical Review Letters* **115**, https://doi.org/10.1103/PhysRevLett.115.268501 (2015).

13. Casagrande, E., Mueller, B., Miralles, D. G., Entekhabi, D. & Molini, A. Wavelet correlations to reveal multiscale coupling in geophysical systems. *Journal of Geophysical Research: Atmospheres* **120**, 7555–7572, https://doi.org/10.1002/2015JD023265 (2015).

14. Paluš, M. Linked by Dynamics: Wavelet-Based Mutual Information Rate as a Connectivity Measure and Scale-Specific Networks. In Tsonis, A. A. (ed.) *Advances in Nonlinear Geosciences*, 427–463, https://doi.org/10.1007/978-3-319-58895-7_21 (Springer International Publishing, Cham, 2018).

15. Newman, M. E. J. The Structure and Function of Complex Networks. *SIAM Review* **45**, 167–256, https://doi.org/10.1137/S003614450342480 (2003).

16. Mitra, C., Kittel, T., Choudhary, A., Kurths, J. & Donner, R. V. Recovery time after localized perturbations in complex dynamical networks. *New Journal of Physics* **19**, 103004, https://doi.org/10.1088/1367-2630/aa7fab (2017).

17. Mitra, C., Choudhary, A., Sinha, S., Kurths, J. & Donner, R. V. Multiple-node basin stability in complex dynamical networks. *Physical Review E* **95**, 032317 (2017).

18. Donges, J. F., Zou, Y., Marwan, N. & Kurths, J. The backbone of the climate network. *EPL (Europhysics Letters)* **87**, 48007, https://doi.org/10.1209/0295-5075/87/48007 (2009).

19. Gu, J. *et al.* Recent Progress in Some Active Topics on Complex Networks. *Journal of Physics: Conference Series* **604**, 012007, https://doi.org/10.1088/1742-6596/604/1/012007 (2015).

20. Tsonis, A. A., Swanson, K. L. & Roebber, P. J. What Do Networks Have to Do with Climate? *Bulletin of the American Meteorological Society* **87**, 585–595, https://doi.org/10.1175/BAMS-87-5-585 (2006).

21. Boers, N. *et al.* The South American rainfall dipole: A complex network analysis of extreme events: BOERS ET AL. *Geophysical Research Letters* **41**, 7397–7405, https://doi.org/10.1002/2014GL061829 (2014).

22. Stolbova, V., Martin, P., Bookhagen, B., Marwan, N. & 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, https://doi.org/10.5194/npg-21-901-2014 (2014).

23. Donges, J. F., Zou, Y., Marwan, N. & Kurths, J. Complex networks in climate dynamics: Comparing linear and nonlinear network construction methods. *The European Physical Journal Special Topics* **174**, 157–179, https://doi.org/10.1140/epjst/e2009-01098-2 (2009).

24. Rheinwalt, A. *et al.* Non-linear time series analysis of precipitation events using regional climate networks for Germany. *Climate Dynamics* **46**, 1065–1074, https://doi.org/10.1007/s00382-015-2632-z (2016).

25. Runge, J. *et al.* Identifying causal gateways and mediators in complex spatio-temporal systems. *Nature Communications* **6**, 8502, https://doi.org/10.1038/ncomms9502 (2015).

26. Agarwal, A., Marwan, N., Rathinasamy, M., Merz, B. & Kurths, J. Multi-scale event synchronization analysis for unravelling climate processes: a wavelet-based approach. *Nonlinear Processes in Geophysics* **24**, 599–611, https://doi.org/10.5194/npg-24-599-2017 (2017).

27. Marwan, N., Donges, J. F., Zou, Y., Donner, R. V. & Kurths, J. Complex network approach for recurrence analysis of time series. *Physics Letters A* **373**, 4246–4254, https://doi.org/10.1016/j.physleta.2009.09.042 (2009).

28. Mitra, C., Kurths, J. & Donner, R. V. Rewiring hierarchical scale-free networks: Influence on synchronizability and topology. *EPL (Europhysics Letters)* **119**, 30002 (2017).

29. Agarwal, A. *Unraveling spatio-temporal climatic patterns via multi-scale complex networks*. Ph.D. thesis, https://doi.org/10.25932/publishup-42395 (Universität Potsdam, 2019).

30. Kurths, J. *et al.* Unraveling the spatial diversity of Indian precipitation teleconnections via nonlinear multi-scale approach. *Nonlinear Processes in Geophysics Discussions* 1–19, https://doi.org/10.5194/npg-2019-20 (2019).

31. Agarwal, A., Maheswaran, R., Marwan, N., Caesar, L. & Kurths, J. Wavelet-based multiscale similarity measure for complex networks. *The European Physical Journal B* **91**, https://doi.org/10.1140/epjb/e2018-90460-6 (2018).

32. Agarwal, A., Marwan, N., Maheswaran, R., Merz, B. & Kurths, J. Quantifying the roles of single stations within homogeneous regions using complex network analysis. *Journal of Hydrology*, https://doi.org/10.1016/j.jhydrol.2018.06.050 (2018).

33. Benjamini, Y. & Yosef, H. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* **57**, 289–300 (1995).

34. Reiner, A., Yekutieli, D. & Benjamini, Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* **19**, 368–375, https://doi.org/10.1093/bioinformatics/btf877 (2003).

35. Scarsoglio, S., Laio, F. & Ridolfi, L. Climate Dynamics: A Network-Based Approach for the Analysis of Global Precipitation. *PLoS One* **8**, e71129, https://doi.org/10.1371/journal.pone.0071129 (2013).

36. Van Brummelen, G. *Heavenly mathematics: the forgotten art of spherical trigonometry*. (Princeton University Press, Princeton; Oxford, 2013).

37. Tobler, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* **46**, 234, https://doi.org/10.2307/143141 (1970).

38. Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E. & Huang, D. Predictability of El Niño over the past 148 years. *Nature* **428**, 733–736, https://doi.org/10.1038/nature02439 (2004).

39. Feng, M. & Meyers, G. Interannual variability in the tropical Indian Ocean: a two-year time-scale of Indian Ocean Dipole. *Deep Sea Research Part II: Topical Studies in Oceanography* **50**, 2263–2284, https://doi.org/10.1016/S0967-0645(03)00056-0 (2003).

40. Luo, J.-J. *et al.* Interaction between El Niño and Extreme Indian Ocean Dipole. *Journal of Climate* **23**, 726–742, https://doi.org/10.1175/2009JCLI3104.1 (2010).

41. Stuecker, M. F. *et al.* Revisiting ENSO/Indian Ocean Dipole phase relationships. *Geophysical Research Letters* **44**, 2481–2492, https://doi.org/10.1002/2016GL072308 (2017).

42. Yamagata, T. *et al.* Coupled Ocean-Atmosphere Variability in the Tropical Indian Ocean. In Wang, C., Xie, S. & Carton, J. (eds) *Geophysical Monograph Series*, 189–211, https://doi.org/10.1029/147GM12 (American Geophysical Union, Washington, D. C., 2013).

43. Ferster, B., Subrahmanyam, B. & Macdonald, A. Confirmation of ENSO-Southern Ocean Teleconnections Using Satellite-Derived SST. *Remote Sensing* **10**, 331, https://doi.org/10.3390/rs10020331 (2018).

44. Rahmstorf, S. Ocean circulation and climate during the past 120,000 years. *Nature* **419**, 207–214, https://doi.org/10.1038/nature01090 (2002).

45. Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G. & Saba, V. Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature* **556**, 191–196, https://doi.org/10.1038/s41586-018-0006-5 (2018).

46. Hoskins, B. J. & Karoly, D. J. The steady linear response of a spherical atmosphere to thermal and orographic forcing. *Journal of the Atmospheric Sciences* **38**, 1179–1196 (1981).

47. Lark, R. M. *et al.* Scale- and location-dependent correlation of nitrous oxide emissions with soil properties: an analysis using wavelets. *European Journal of Soil Science* **55**, 611–627, https://doi.org/10.1111/j.1365-2389.2004.00620.x (2004).

48. Fernández-Macho, J. Wavelet multiple correlation and cross-correlation: A multiscale analysis of Eurozone stock markets. *Physica A: Statistical Mechanics and its Applications* **391**, 1097–1104, https://doi.org/10.1016/j.physa.2011.11.002 (2012).

Acknowledgements

Author Contributions

A.A. substantially contributed to conception, design, acquisition of data, analysis and interpretation of data. A.A. plotted all the figures and took the lead in writing the manuscript. L.C. participated in drafting the article and revising it critically for important intellectual content. B.M. & N.M. were involved in the extensive discussion and designing the content. R.M. contributed in formulating method. B.M. & J.K. supervised the work and gave final approval of the version to be submitted.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.