



# Evolution of atmospheric connectivity in the 20th century

F. Arizmendi, A. C. Martí, and M. Barreiro

Instituto de Física, Facultad de Ciencias, Universidad de la República, Igua 4225, Montevideo, Uruguay

Correspondence to: F. Arizmendi (arizmendi.f@gmail.com)

Received: 11 November 2013 – Revised: 24 May 2014 – Accepted: 16 June 2014 – Published: 8 August 2014

**Abstract.** We aim to study the evolution of the upper atmosphere connectivity over the 20th century as well as to distinguish the oceanically forced component from the atmospheric internal variability. For this purpose we build networks from two different reanalysis data sets using both linear and nonlinear statistical similarity measures to determine the existence of links between different regions of the world in the two halves of the last century. We furthermore use symbolic analysis to emphasize intra-seasonal, intra-annual and inter-annual timescales. Both linear and nonlinear networks have similar structures and evolution, showing that the most connected regions are in the tropics over the Pacific Ocean. Also, the Southern Hemisphere extratropics have more connectivity in the first half of the 20th century, particularly on intra-annual and intra-seasonal timescales.

Changes over the Pacific main connectivity regions are analyzed in more detail. Both linear and nonlinear networks show that the central and western Pacific regions have decreasing connectivity from early 1900 up to about 1940, when it starts increasing again until the present. The inter-annual network shows a similar behavior. However, this is not true of other timescales. On intra-annual timescales the minimum connectivity is around 1956, with a negative (positive) trend before (after) that date for both the central and western Pacific. While this is also true of the central Pacific on intra-seasonal timescales, the western Pacific shows a positive trend during the entire 20th century.

In order to separate the internal and forced connectivity networks and to study their evolution through time, an ensemble of atmospheric general circulation model outputs is used. The results suggest that the main connectivity patterns captured in the reanalysis networks are due to the oceanically forced component, particularly on inter-annual timescales. Moreover, the atmospheric internal variability seems to play an important role in determining the intra-seasonal timescale networks.

## 1 Introduction

In the last 15 years, since the pioneering works on complex networks (Watts and Strogatz, 1998; Barabási and Albert, 1999), the use of this statistical methodology of analysis has allowed significant advances in a broad variety of scientific disciplines, including social systems, the internet, neural networks, metabolic networks, and gene expressions, among many others (Albert and Barabási, 2002; Newman, 2010).

The study of the climate system has also benefited from the application of this new theoretical framework (Tsonis et al., 2004, 2006; Donges et al., 2009; Yamasaki et al., 2008; Barreiro et al., 2011), providing valuable insights into climate dynamics. As is well known, the climate system is characterized by phenomena with many different temporal and spatial scales, so that a variety of network approaches have been proposed to study different modes of variability. For example, a large number of articles have been dedicated to the analysis of the El Niño–Southern Oscillation (ENSO) phenomenon (Tsonis et al., 2008; Yamasaki et al., 2008; Martin et al., 2013; Radebach et al., 2013) and how it affects the connectivity of the network (see also Deza et al., 2014 in this volume).

El Niño has been shown to influence the climate of far-away regions through the excitation of anomalous wave trains that propagate longitudinally within the tropics and meridionally to the extratropical regions. The theory that tries to explain these remote connections, or teleconnections, is mainly based on linear wave dispersion of Kelvin and Rossby waves (Held, 1983; James, 1994). Even though there have been several attempts to improve this description, including the influence of El Niño through other mechanisms (Seager et al., 2003), the understanding of atmospheric teleconnections is not yet complete.

Here we aim to deepen the understanding of atmospheric teleconnections, analyzing the 200 mb eddy geopotential

height field from a complex network perspective. This level is chosen because it is the height of the maximum tropical divergence that is crucial for the excitation and propagation of tropical and extratropical anomalies like Rossby wave trains (Trenberth et al., 1998). A second objective of this study is to analyze potential changes in the atmospheric connections during the last century. We focus on the connectivity of the tropical Pacific because it has been shown by several studies (Donges et al., 2009; Barreiro et al., 2011) to be globally the most connected region.

In order to accomplish these objectives, we analyze two atmospheric reanalysis data sets, which represent the best estimate of the historical atmospheric evolution (Broennimann et al., 2009) and an ensemble of experiments performed with an atmospheric general circulation model (AGCM). The networks determining the connectivity among nodes are built in terms of the statistical similarity of the corresponding time series. We use both linear and nonlinear statistical measures (Donges et al., 2009; Barreiro et al., 2011): the linear Pearson correlation (PC) and the nonlinear mutual information (MI). Alternatively, as the PC and the MI obtained from histogram-based probability density functions (PDFs) do not take into account the temporal order of the time series, we introduce the ordinal patterns methodology proposed by Bandt and Pompe (2002) (BP). This method allows us to consider different timescales with the MI as in Barreiro et al. (2011). In general, the networks are represented graphically as two-dimensional maps by plotting the *area-weighted connectivity* AWC, which represents the fraction of the total area of the Earth to which each node is connected.

The total atmospheric variability can be decomposed in a component due to the intrinsic dynamics of the atmosphere that would occur even in the absence of changes to the boundary conditions, and another component due to the forcing of the surface ocean conditions. The latter is potentially more predictable, as it depends on the evolution of the sea surface temperature, which varies on much longer timescales than the atmospheric fields. In some regions, like the tropical Pacific, the ocean–atmosphere interaction allows the prediction of sea surface temperature anomalies (i.e., the development of El Niño) by up to 6–9 months in advance (Neelin et al., 1998). It is thus extremely important for seasonal climate prediction to be able to separate and determine the dynamics that characterizes the intrinsic and forced atmospheric variability. In this study we use the AGCM output to determine if the observed 20th century changes in connectivity are related to the forced or intrinsic dynamics.

Our results show that the most connected regions are in the tropical Pacific, and the overall network structure is broadly similar when using both the linear and nonlinear statistical measures when no timescales are considered. However, using the BP methodology, the nonlinear symbolic analysis reveals several structural changes in the networks. Some regions become more relevant depending on the timescale considered, such as the Southern Hemisphere, relatively more important

on the intra-seasonal timescales, or the northern Pacific, and more relevant on the intra-annual timescales.

Moreover, we show that the evolution of the connectivity in the Pacific depends on the timescale. For example, in the central Pacific the connectivity is smallest around 1956 for intra-seasonal and intra-annual timescales, with negative (positive) trends before (after) that date. On inter-annual timescales, however, the connectivity decreases from early 1900 up to around 1940 and increases afterwards.

This article is organized as follows: Sect. 2 provides a description of the data analyzed and the model used, and a summary of the methodology employed. Section 3 presents the results, and finally in Sect. 4 we present a summary and discuss the main results.

## 2 Data and methodology

### 2.1 Data

We use monthly average geopotential height at 200 mb from the NOAA 20th Century Reanalysis (Compo et al., 2011) and from the NCEP CDAS1 reanalysis data (Kalnay et al., 1996). It is important to keep in mind that a reanalysis is constructed blending observations into an AGCM, and thus the value of the reanalysis compared to a simple model run increases where there are enough observations to constrain the model solution. Prior to 1950, observations were very limited globally, and thus the reanalysis may not represent so much reality as it represents the model solution. In the Southern Hemisphere the situation is even more difficult because there were very few observations prior to the satellite era that began in 1979. Given these limitations the results of this study must be considered as a first approximation of reality based on the best information available to date about the evolution of the 20th century atmospheric circulation.

In addition, we consider the output of the atmospheric general circulation model from the International Center for Theoretical Physics (ICTP-AGCM) forced with global historical sea surface temperature (ERSSTv.2, Smith and Reynolds, 2004). The ICTP-AGCM is a full AGCM with simplified physics and a horizontal resolution of T30 with eight vertical levels (Molteni, 2003; Kucharski et al., 2006). The performance of the ICTP-AGCM in representing the large-scale atmospheric dynamics is comparable to that of state-of-the-art AGCMs and has been used to study global climate variability, including the influence of the oceans on the circulation (Barreiro et al., 2008; Barreiro, 2009; Kucharski et al., 2008, 2009).

Reanalysis and model output are distributed in regular latitude–longitude grids with every point characterized by a time series of the variable of interest. The characteristics of each data set are summarized in Table 1.

The average geopotential height at 200 mb has a meridional structure dominated by relatively large values in the

**Table 1.** Characteristics of data.  $N$ : number of grid points,  $T$ : number of months.

Source	Resolution	Time period	Dimensions	
NCEP CDAS1	2.5°	[1949, 2011]	$N = 10\,224$	$T = 756$
NOAA 20th Century	2°	[1901, 2009]	$N = 16\,020$	$T = 1308$
ICTP AGCM	3.75°	[1901, 2006]	$N = 4980$	$T = 1272$
EN34 index	–	[1901, 2006]	$N = 1$	$T = 1308$

tropical regions that decrease toward the poles. To highlight connections that are deviations from this zonal structure, we consider the eddy geopotential height that is calculated by removing the zonal mean of each latitudinal band for every month. We furthermore consider monthly anomalies; that is, the mean annual cycle is removed prior to analysis, and the time series are linearly detrended at each grid point.

### 2.2 Network building

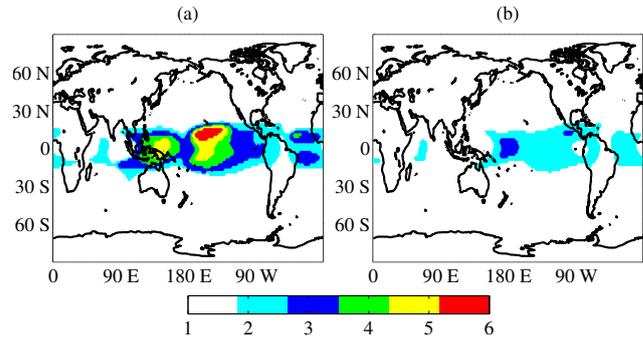
To build the network, every pair of points over the grid is linked whenever a significant statistical similarity between the correspondent time series is detected. Let  $\mathbf{x}(t) = \{x_i(t)\}_{i=1}^N$  with  $t = 1, \dots, T$  be the 200 mb eddy geopotential height anomaly field and  $S(x_i, x_j) = S_{ij}$  the similarity measure between the nodes  $i$  and  $j$ . The elements of the climate network adjacency matrix  $\mathbf{A}$  are then

$$A_{ij} = \Theta(|S_{ij}| - \tau) - \delta_{ij}, \quad (1)$$

where  $\Theta()$  is the Heaviside step function,  $\delta$  is the Kronecker delta and  $\tau$  is a statistical threshold introduced to avoid spurious connections.

In order to set the threshold value, we use the same criterion for every network, randomly relabeling the elements within each time series such that the distributions remain unchanged, and taking a fixed quantile value (99%) of the statistical similarity measure of every possible combination as a threshold. Thus, the null hypothesis of independence can be rejected if the statistical similarity value of the original time series is above this threshold at a 0.01 level of significance.

Recently, Paluš et al. (2011) have shown that two time series that have some autocorrelation, or persistence (defined as the lag time when the autocorrelation drops below  $1/e$ ), induce a bias in the connectivity measure. To avoid this supposed bias, Paluš et al. (2011) suggest considering the persistence of the time series when calculating the connectivity threshold, and show that the network topology differs significantly when this is taken into account. In climate data, the persistence at a particular location is due to local and remote processes. The part of the persistence due to local processes is the one that should be considered in the connectivity threshold because it introduces a bias in the connectivity. However, the persistence due to remote processes is a real connection that occurs through atmospheric circulation anomalies and should be considered when



**Figure 1.** Persistence of the 200 mb eddy geopotential height field of the 20th Century Reanalysis data from NOAA with (a) and without the EN34 component (b).

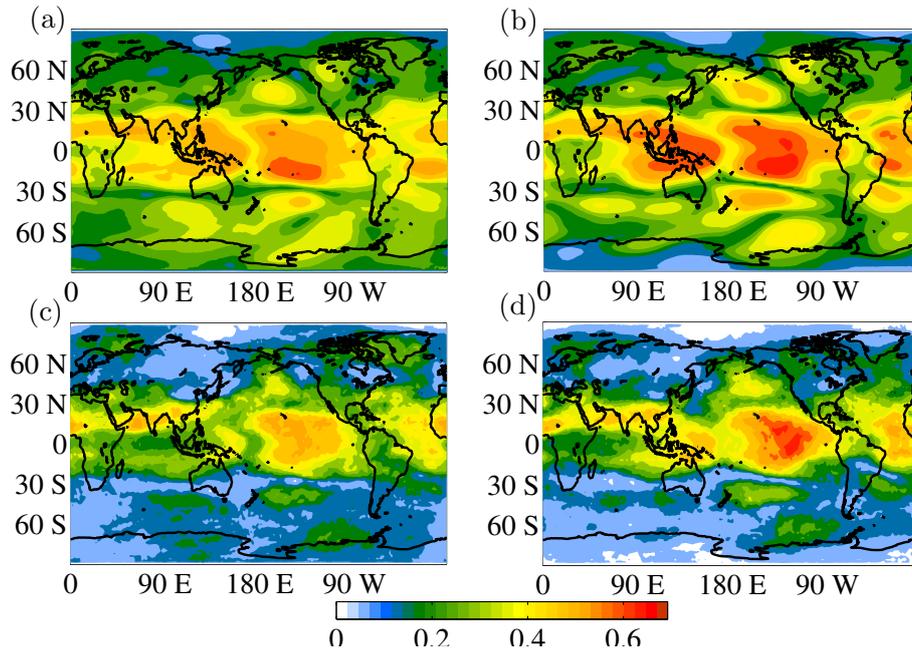
calculating connectivity measures as the area-weighted connectivity used here (see Sect. 2.3).

The persistence of the 200 mb eddy geopotential height of the NOAA 20th Century Reanalysis data is shown in Fig. 1a, where it can be seen that the persistence is 1 month for the extratropics and from 2 to 6 months in the tropical band, with a maximum in the Pacific. This result reinforces the large internal atmospheric variability present in the extratropics and the dominant role of the ocean forcing in the tropical band. Figure 1b shows the persistence after the effect of El Niño (El Niño 3.4 index) has been linearly removed by a regression procedure. Clearly, most of the persistence in the tropical band is due to the El Niño phenomenon that is known to affect the global tropics through the propagation of Kelvin waves (e.g., Lintner and Chiang, 2007). Given that El Niño persists for about 6 months, the connections from the tropical Pacific to the rest of the tropics are also maintained. After El Niño is removed, the persistence in the tropics is about 2 months, which might be because the linear procedure was not enough to remove the El Niño influence completely or due to the existence of local processes that slightly enhance persistence. Nevertheless, it is clear that a persistence of 1 month is characteristic of most of the world, and thus randomly mixing the time series on individual points of the grid provides a good test of the connectivity. This procedure ensures the recognition of the importance of the tropical Pacific in the global connectivity, as clearly seen in Fig. 2.

We constructed networks using linear and nonlinear measures of connectivity. Networks made of purely linear interaction are constructed by evaluating the statistical interdependency between the nodes with the Pearson correlation:

$$\rho_{ij} = \frac{\text{cov}(x_i, x_j)}{\sigma_i \sigma_j} = \frac{E[(x_i - \mu_i)(x_j - \mu_j)]}{\sigma_i \sigma_j}, \quad (2)$$

where  $\mu_{i,j}$  and  $\sigma_{i,j}$  are the mean and standard deviations of the time series  $i, j$ . Since we consider anomalies,  $\mu_i = 0$  for all  $i \in [1, N]$ .



**Figure 2.** AWC maps of the 200 mb eddy geopotential height field with the graph construction methodology applied to the 20th Century Reanalysis data from NOAA. Statistical similarity measures and periods taken: (a) PC (1901–1955), (b) PC (1955–2009), (c) MI with classical histograms (1901–1955) and (d) MI with classical histograms (1955–2009).

Nonlinear connectivity is evaluated using a quantity from information theory, the mutual information:

$$M_{ij} = \sum_{m,n} p_{ij}(m, n) \log \frac{p_{ij}(m, n)}{p_i(m) p_j(n)}, \quad (3)$$

where  $p_i$  and  $p_j$  are the marginal probability density functions of  $x_i$  and  $x_j$ , respectively, and  $p_{ij}$  the joint probability density function. According to information theory, mutual information is a measure of how much information about  $x_i$  is gained by knowing  $x_j$ . It is easy to see that if the time series are independent,  $p_{ij}(m, n) = p_i(m) p_j(n)$ , and therefore  $M_{ij} = 0$ . It is important to notice that both mutual information and linear correlation are symmetric quantities.

The usual way of obtaining the PDF of a time series is by building traditional histograms. Generally, the numbers of bins in the histogram are chosen according to the time series length to ensure good statistics. One important characteristic of the use of histograms is that the time-related information is only considered in the joint probability because, when using histograms to represent the marginal probabilities, all the temporal information is neglected and, as with linear correlation, there is no emphasis on any particular timescale.

To overcome this limitation, the Bandt and Pompe (BP) methodology (Bandt and Pompe, 2002) based on symbolic analysis is employed to separate different timescales. This symbolic analysis consists of building the PDFs by analyzing the patterns formed by consecutive values in the time series. Specifically, each time series  $x_i(t)$  is divided into  $T - D$  overlapping  $D$  dimensional vectors. These vectors

are represented as ordinary patterns (OPs), in the sense that each element is replaced by a number from 0 to  $D - 1$ , depending on its ordinal position (i.e., 0 for the lowest and  $D - 1$  for the largest). Each vector is then assigned one of the  $D!$  possible configurations. For instance, if  $D = 3$ , the possible configurations are (012 021 102 120 201 210), and, for example, the sequence (2.3, 0.7, 8.7) corresponds to the 102 pattern because  $x_i^3 > x_i^1 > x_i^2$ . The BP methodology needs only a weak stationary assumption that the probability for  $x_i^t < x_i^{t+D}$  should not depend on  $t$ . This is satisfied by removing the linear trend in the data. Another condition for calculating the PDFs adequately is that  $T \gg D!$ ; i.e., the number of vectors must be much larger than the number of possible OPs.

There is an embedding dimension given by the  $D$  parameter of the BP method that fixes the causality timescale. This is the main difference from the classical histogram-based technique used to build the probability density functions. An important advantage is that it is possible to build sequences of patterns not only with consecutive values but with an arbitrary time interval  $\Delta t$  between them. In this study we have chosen to work with  $D = 3$  and  $\Delta t = 1$ -, 4- and 12-month intervals to characterize atmospheric processes on intra-seasonal, intra-annual and inter-annual timescales, respectively.

### 2.3 Area-weighted connectivity

The network representation of a climate field helps us to obtain valuable insight into the climate dynamics; however, it is clear that the specific procedure to build the network is decisive in the information it provides. Moreover, once the network is built, there are different analyses that can be done in order to deepen our understanding of the dynamics that is being analyzed, such as calculating the betweenness centrality for each node (Donges et al., 2009) or detecting the underlying communities in the system (Tsonis et al., 2011).

Following the procedure proposed in previous studies on climate networks (Tsonis et al., 2006; Donges et al., 2009; Barreiro et al., 2011), to analyze the changes in the atmospheric connectivity in the 20th century, we use the *area-weighted connectivity* (AWC) function

$$AWC_i = \frac{\sum_{j=1}^N A_{ij} \cos(\theta_j)}{\sum_{j=1}^N \cos(\theta_j)}, \quad (4)$$

where  $\theta_j$  is the latitude of node  $j$  and the cosine factors are included to take into account the sphericity of the Earth; that is, for a regular planar grid embedded in a spherical surface, the nodes correspond to different areal regions. The AWC provides the fraction of the Earth that each node is connected to, similar to the  $k$  degree usual in complex networks, but area weighted and normalized ( $AWC_i \in [0, 1]$ ).

This study is concerned with analyzing the changes in the atmospheric connectivity during the last hundred years, which, as will be shown below, is dominated by the tropical Pacific. Thus, to further understand the changes in these regions, we calculate the evolution of the AWC in two boxes in the central and western Pacific using the NOAA 20th Century Reanalysis data. Specifically, we considered the nodes in the following regions:

- Western Pacific (WP): [3–7° S, 129–135° E]
- Central Pacific (CP): [3–7° S, 151–157° W].

The evolution is measured by taking 30-year sliding windows in the steps of 1 year and calculating the average AWC for each region.

### 2.4 Intrinsic and forced atmospheric variability

To aid the interpretation of the networks obtained with reanalysis data, we consider the output of the ICTP-AGCM. In particular, as explained in the introduction, we are interested in separating the networks associated with intrinsic (or internal) atmospheric variability and with oceanically forced atmospheric variability, so we constructed an ensemble of 9 ICTP-AGCM runs where the model is forced with global historical sea surface temperatures as boundary conditions and slightly different atmospheric initial conditions.

Given the chaoticity of the atmospheric dynamics, this setup generates  $n = 9$  equiprobable realizations of the atmospheric evolution.

The procedure to separate internal from forced atmospheric variability is as follows. We start with the model outputs, whose characteristics are detailed in Table 1 as  $\mathbf{x}_j^m(t) = \{x_{ji}^m(t)\}_{i=1}^N$ , with  $j = 1, \dots, n$  as the different realizations. The evolution for each realization  $j$  at each grid point can then be separated into two components

$$\mathbf{x}_j^m = \mathbf{x}_F^m + \mathbf{x}_{Ij}^m, \quad (5)$$

where  $\mathbf{x}_F^m$  is the oceanically forced component and  $\mathbf{x}_{Ij}^m$  is the evolution due to atmospheric internal dynamics (independent of the surface boundary conditions). We consider that, to first order, the two components do not interact (Barreiro, 2009).

Thus, assuming that the internal variability is filtered by averaging, the forced component can be estimated by taking the ensemble mean of the runs:

$$\mathbf{x}_F^m \approx \frac{1}{n} \sum_j^n \mathbf{x}_j^m. \quad (6)$$

Once the forced component is estimated, the internal variability in each run can be approximated by applying Eq. (5). To characterize the network of atmospheric internal variability, we calculated the AWC for every  $\mathbf{x}_{Ij}^m$  and then computed the average over the  $n$  AWC values.

### 2.5 Principal component analysis

To aid further understanding of the role of the tropical Pacific in setting the connections, we removed its contribution and constructed the AWC maps again. To do so, we built the AWC maps after removing the first two leading empirical orthogonal functions (EOFs) calculated for the tropical Pacific. The difference between the AWC maps after these EOFs are removed allows us to associate regions of large connectivity with known modes of variability.

We calculated the EOFs of the NOAA 20th Century Reanalysis data in the region [90° E–70° W, 30° N–30° S] and considered the two eigenvectors ( $\mathbf{v}_1, \mathbf{v}_2$ ) with the largest eigenvalues ( $\lambda_1 > \lambda_2 > \dots$ ) of the covariance matrix, which together explain about a quarter of the total variance.

The principal components  $p_i(t)$  are obtained by projecting the original field onto the eigenvectors:

$$p_i(t) = \mathbf{x}'(t) \cdot \mathbf{v}_i, \quad (7)$$

where  $\mathbf{x}'(t)$  is the 200 mb eddy geopotential height field within the tropical Pacific.

To calculate the pattern associated with each EOF over the whole globe, we calculated the linear regression of the original field  $\mathbf{x}(t)$  onto  $p_i(t)$ , such that the parameters  $\hat{a}$  and  $\hat{b}$  that best estimate the relation of time series  $x_j$  with  $p_i$ ,

$$\hat{x}_j = \hat{a}_j p_i + \hat{b}_j, \quad (8)$$

are obtained with the least squares approach. The significance is assessed by evaluating the correlation and comparing the statistics  $T^s = |\rho| \sqrt{T-2} / \sqrt{1-\rho^2}$  with the Student's  $t$  distribution with a two-tailed confidence interval of 95 %, with  $H_0: \rho = 0$  being the null hypothesis. Assuming a red noise type of spectrum, the effective number of degrees of freedom for each node is calculated as  $T^* = (T/2) \log(a)$ , where  $a$  is the correspondent serial correlation value at a lag equal to 1 month.

### 3 Results

#### 3.1 Area-weighted connectivity maps

Being our goal to analyze changes in the structure of atmospheric connectivity on different timescales since the beginning of the 20th century, we divide the NOAA 20th Century Reanalysis data into two equal-length halves: January 1901–December 1955 and January 1955–December 2009. The resulting networks built with Pearson correlation (PC) and mutual information with the probability density functions based on classical histograms (MIH) are represented by the AWC function in Fig. 2.

The spatial structure of the maps obtained with the different measures of connectivity is globally very similar when using both statistical measures. The main difference lies in the number of connections that represent the number of significant links, which is greater in the PC networks than in the MIH networks. The links density  $\delta$ ,

$$\delta = \sum_{ij} \frac{A_{ij}}{N(N-1)} : \delta \in [0, 1], \quad (9)$$

is defined to quantify this property. The networks constructed with PC present  $\delta \approx 0.30$ , while the corresponding value with MIH is  $\delta \approx 0.20$ .

In all the cases, the most connected areas are located in the tropics, particularly in the central Pacific and to a lesser extent over the Atlantic Ocean, southern Asia and over Indonesia between the Indian Ocean and the western Pacific. This larger connectivity in the tropics found with both methodologies agrees with our current understanding of tropical dynamics. Likewise, but weaker in the sense that they vanish for lower threshold values than in the tropics, there are some extratropical regions that show a relatively large number of connections, such as the South and North Pacific. The wavy patterns suggest that they reflect the teleconnection patterns in both hemispheres through the propagation of Rossby waves.

The area of largest connectivity in the central Pacific shows differences according to the methodology: while using PC, the area of maximum connectivity is to the south of the Equator at about 170–130° W; using MIH the (less well-defined) maximum is centered on the Equator further to the east. Nevertheless, both methodologies show that the

central Pacific connectivity is largest in the second half of the century.

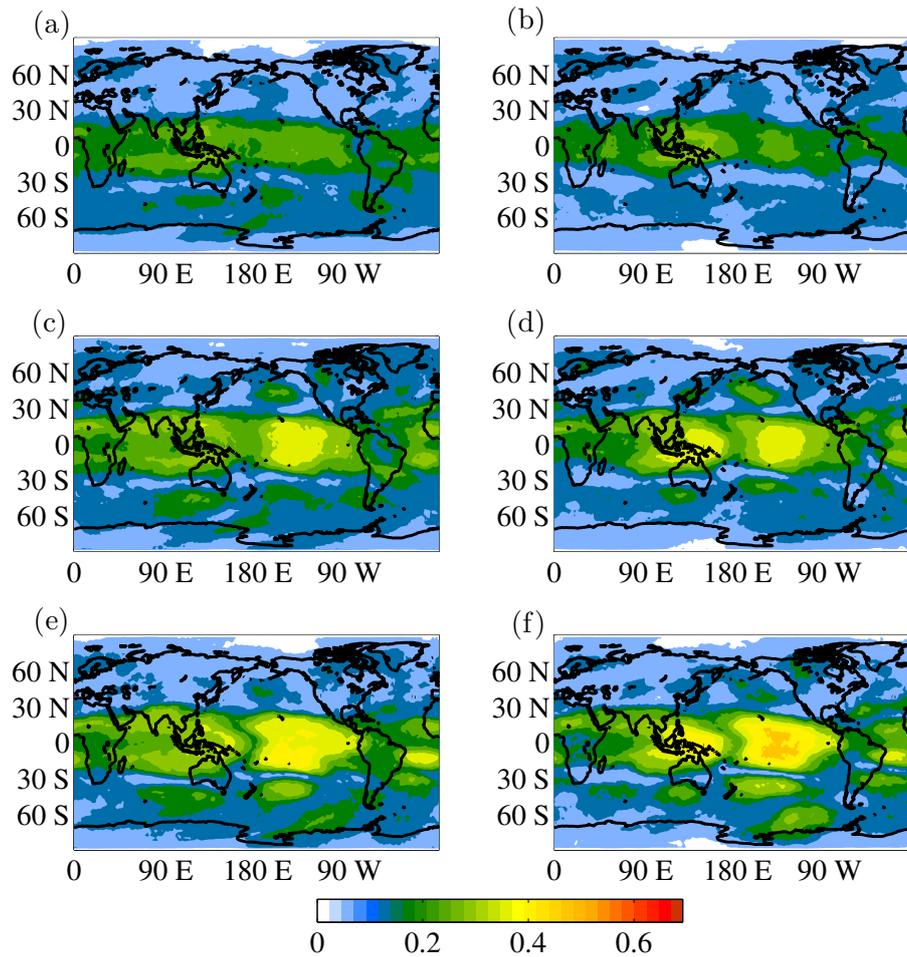
Some additional observations can be made from Fig. 2. Firstly, we note that there are more significant links in the Southern Hemisphere extratropics than in the Northern Hemisphere, especially in the first period. A structure, displaying the signature of Rossby waves propagating from the tropics through the southern Pacific to the South American continent, can be distinguished in all the panels. The increase in the connectivity in the western Pacific and South Asia in the second period can also be distinguished. The opposite occurs in the southern Indian Ocean west of 90° E, which decreases its connectivity, especially with the MIH. These variations suggest that the excitation of Rossby wave trains that propagate from the Indian Ocean sector toward the South Pacific was stronger from 1901 to 1955.

#### 3.2 Ordinal pattern analysis

Figure 3 presents the results of the AWC calculated with the BP methodology for intra-seasonal, intra-annual and inter-annual timescales using  $D=3$  as the embedding dimension. Irrespective of the period considered, the tropics always dominate the connectivity on all timescales. However, the connectivity of the tropical region increases with the timescale, while that of the extratropics does not vary to the same extent. In fact, connectivity in the northern extratropics is largest on intra-annual timescales (particularly in the northern Pacific). In the southern extratropics the regions of maximum connectivity are similar, but increase with the timescale, suggesting a connection with the tropics. Note the existence of a region of large connectivity centered at 60° S only on inter-annual timescales, also seen in the AWC maps calculated with PC and MIH in Fig. 2. This suggests that the symbolic analysis is successful in separating processes according to their timescales and that connectivity in Southern Hemisphere extratropics is strongest on inter-annual timescales. Regarding links density,  $\delta \approx 0.14$  for intra-seasonal networks and  $\delta \approx 0.16$  for intra-annual and inter-annual networks.

Globally, on intra-seasonal timescales,  $\Delta t = 1$  month, the area of larger connectivity in the first period is over the tropical Pacific and between southern Asia and northern Oceania, while in the second half, the global maximum is in the western Pacific and the central Pacific region is less connected. Also, the first period shows larger connectivity areas in the southern Indian Ocean and southern extratropics, with similar differences in structure between the two periods as with PC and MIH (Fig. 2). This behavior is also valid for longer timescales.

In the case of intra-annual timescales,  $\Delta t = 4$  months; there are more links in the Northern Hemisphere, particularly over oceanic areas, and there are also some differences in the connectivity of the tropical Pacific in both time periods. While in the eastern sector during the second period there is lower



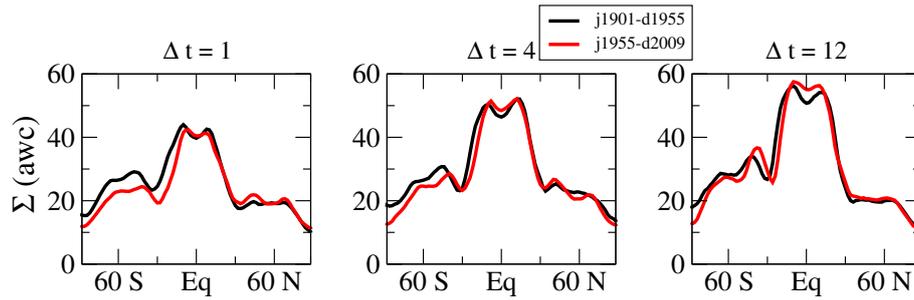
**Figure 3.** AWC maps of the 200 mb eddy geopotential height field obtained from the MI using the BP symbolic analysis for different timescales: intra-seasonal ( $\Delta t = 1$ ) (1901–1955) (a) and (1955–2009) (b), intra-annual ( $\Delta t = 4$ ) (1901–1955) (c) and (1955–2009) (d), and inter-annual ( $\Delta t = 12$ ) (1901–1955) (e) and (1955–2009) (f).

connectivity, the opposite is true in the western region, such as with intra-seasonal networks. On the other hand, both the central and western Pacific are more connected during the second half of the 20th century than during the first half on inter-annual timescales;  $\Delta t = 12$  months.

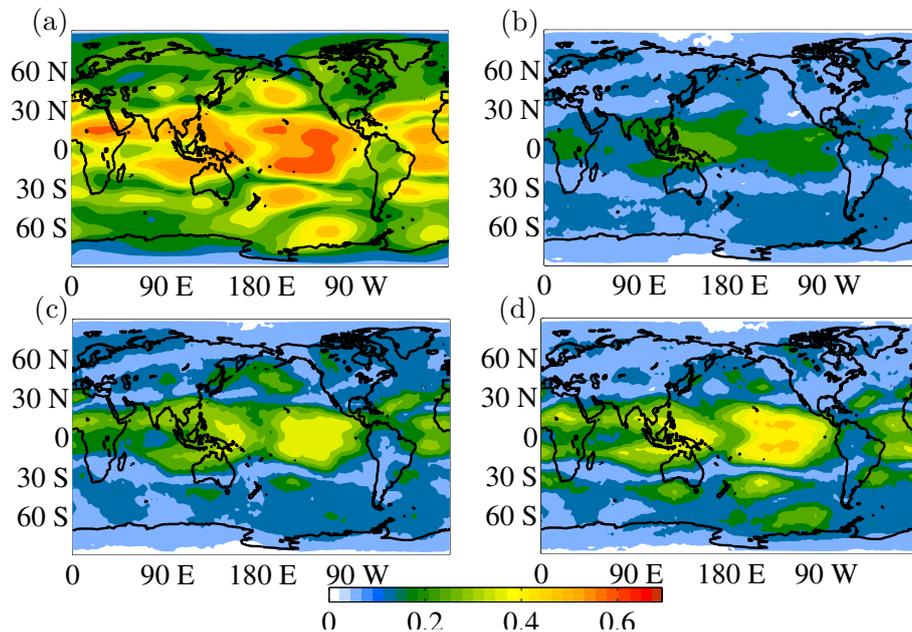
This different behavior in the evolution of the connectivity in certain regions that depend on the timescale can also be noticed over Canada, with larger connectivity in the first period in the intra-annual networks, while in the inter-annual one it is the other way around. Intra-annual and intra-seasonal networks also show larger connectivity in the first half of the century over southern Brazil, contrary to what is seen in the PC and MIH networks in Fig. 2. These results suggest different evolutions of atmospheric processes on different timescales during the period of study, and suggest that the networks from PC and MIH mainly represent inter-annual connections in the tropical band and the Southern Hemisphere.

The hemispheric asymmetries in the AWC are quantified in Fig. 4, where the longitudinal sum of AWC versus latitude is plotted. All the curves present a local minimum on the Equator and two similar maxima near  $10^\circ$  S and  $10^\circ$  N. Clearly, the tropical connectivity increases with the timescale, and is larger in the second half of the century on long timescales. The plots also show that the relatively larger connectivity seen in the Southern Hemisphere during the 1901–1955 period is maximum on intra-seasonal timescales and decreases on longer timescales.

In order to test the robustness of the results, we have also constructed the AWC using the 200 mb eddy geopotential height of the NCEP CDAS1 reanalysis (Kalnay et al., 1996) during the same time period as the second half of the NOAA data. Figure 5 shows the AWC maps derived from the Pearson correlation and the mutual information with the symbolic analysis. The values of  $\delta$  obtained are in good agreement with the maps shown in Figs. 2 and 3. The maps recover the same characteristics obtained for the NOAA 20th Century



**Figure 4.** Zonal sum of the AWC for the NOAA networks built by using the BP methodology with  $\Delta t = 1$ ,  $\Delta t = 4$  and  $\Delta t = 12$ , for both time periods (black: 1901–1955, red: 1955–2009).



**Figure 5.** AWC maps obtained from NCEP/NCAR reanalysis data. The statistical similarity measures used: PC (a) and MI with the BP methodology:  $\Delta t = 1$  (b),  $\Delta t = 4$  (c) and  $\Delta t = 12$  (d).

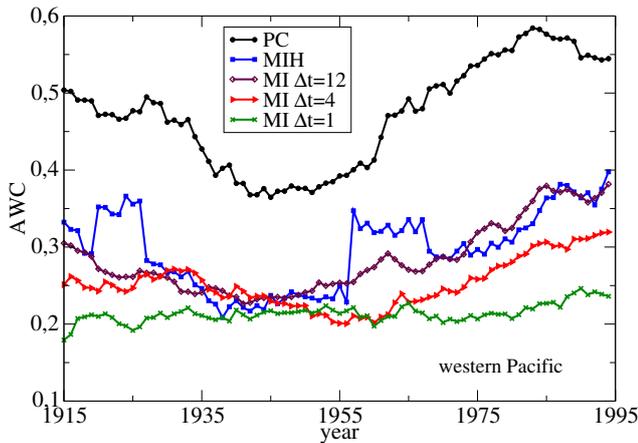
Reanalysis, for example relatively larger connectivity in the northern Pacific on intra-annual timescales, a weakly connected central Pacific on intra-seasonal timescales and a strongly connected Southern Hemisphere extratropical region on inter-annual timescales.

### 3.3 Temporal evolution of the AWC

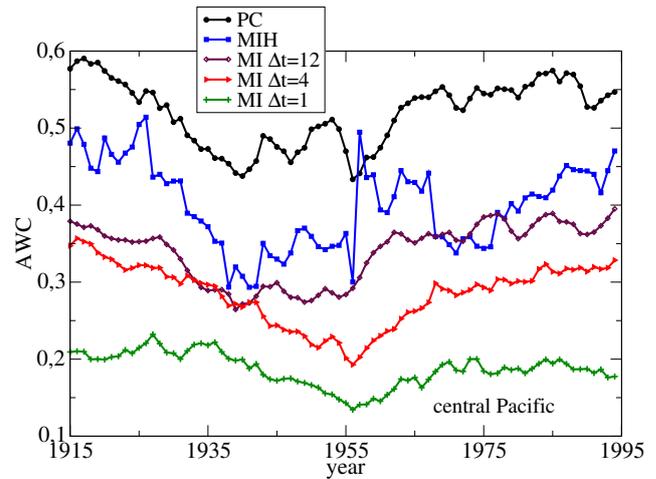
We study further the time evolution of the connectivity for the western Pacific (WP) and central Pacific (CP) regions defined above. To enhance statistics, especially with the mutual information calculated with symbolic analysis, we took 30-year windows such that  $T = 360$  months.

Figures 6 and 7 show plots of AWC versus time for all the different networks for the WP and CP zones, respectively. The behavior of the AWC for both regions depends on the methodology used. Basically, the overall evolution in the PC, MIH and inter-annual networks is very similar, but

intra-annual and intra-seasonal networks behave differently. In the western Pacific, Fig. 6, the mean connectivity of the WP box presents a general decrease for the PC, MIH and intra-seasonal networks from the beginning of the century up to approximately 1940, when the connectivity starts an increasing trend until the end of the century. The connectivity of the intra-annual network shows a small negative trend during the first decades, and from 1956 onwards it starts increasing up to the end of the record. On the other hand, the AWC of the intra-seasonal networks has a small but positive trend during the entire period considered. The PC, MIH and inter-annual networks show similar behavior in the connectivity of the CP box, Fig. 7, starting with a negative trend until 1940 and then an increase in the connectivity to the end of the period. On the other hand, the intra-annual and intra-seasonal networks maintain the negative trend for a few more years, starting the increasing period in 1956, such as the WP box in



**Figure 6.** Temporal evolution of the AWC in the WP box. The networks are built by evaluating 30-year windows of the 200 mb eddy geopotential height field (NOAA) with every methodology used in this work. For each mean AWC value the time periods corresponding to 15 years before and 15 years after were considered.



**Figure 7.** Temporal evolution of the AWC in the CP box. The networks are built by evaluating 30-year windows of the 200 mb eddy geopotential height field (NOAA) with every methodology used in this work. For each mean AWC value the time periods corresponding to 15 years before and 15 years after were considered.

the intra-annual case. Except in the intra-seasonal case, the connectivity values of the central Pacific are larger than in the western Pacific. It is important to mention that the evolution of the connectivity in the MIH networks presents two periods, 1920–1926 and 1956–1967, of anomalous higher values in both regions analyzed. Further analysis indicated that the new connections in both periods are located in the extratropics but have no clear spatial structure, and thus they might be related to noisy behavior.

Figure 8 shows the statistically significant mean Pearson correlation values of the CP and WP boxes for 1915, 1940 and 1995. In agreement with the time evolution shown in Figs. 6 and 7, it can be seen clearly that in the 30 years centered in 1940 the connectivity of these regions is smallest. For the WP the increase in global connectivity seen in the last decades of the century is evident, particularly due to stronger connections with the tropical Atlantic. These maps also show that the CP and WP are not independent of each other, which is expected because phenomena like El Niño connect both regions through changes in the Walker circulation.

### 3.4 Intrinsic and forced variability

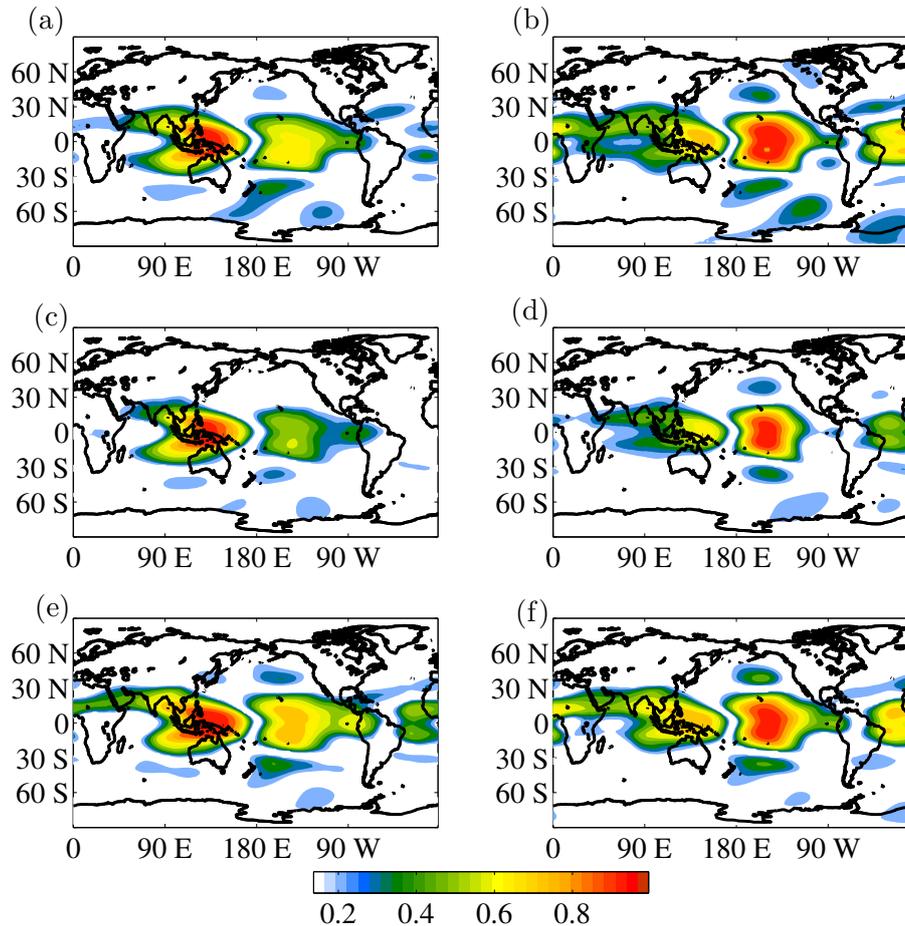
Nine ICTP-AGCM model runs with the same sea surface temperature boundary condition and slightly different atmospheric initial conditions were performed. As is explained in the methodology section, this allows for the separation of the AWC function of the forced and internal variability components. In this case, to continue comparing the behavior in the connectivity of the networks with time, and as the time interval covered by the model output is from 1901 until 2006, the resulting halves are [January 1901, December 1953] and [January 1954, December 2006], with  $T = 636$ . Even though the periods do not match exactly with those of the reanalysis,

they only differ in 2 out of 53 years, and thus the statistics should be fully comparable.

The forced component AWC maps obtained by the BP mutual information methodology are shown in Fig. 9. All timescales present similar spatial structures for both time periods considered, with maxima in the tropics and wavy patterns in the extratropics. The main difference is the intensity of the connectivity, such that during the second period the number of connections increases on all the timescales. Also, on intra-annual timescales Northern Hemisphere regions are more connected than on intra-seasonal timescales and are larger than in the Southern Hemisphere. On inter-annual timescales the connectivity is similar in the northern and southern extratropics, and the model recovers the highly connected region in the South Pacific located at about 35° S, 150° W.

Comparison with reanalysis reveals very similar structures, suggesting that most of the connectivity seen in Fig. 3 is due to oceanically forced variability. The largest differences occur on intra-seasonal timescales in the Indo-Pacific region (particularly in the second half), where the model shows weaker connectivity. Also regarding the reanalysis comparison, it can be noticed that the spatial structure is quite similar except for differences in strength, but there are fewer changes between the first and second halves of the 20th century. This might suggest that the reanalysis differences between both time intervals considered are mostly because of the internal variability.

Figure 10 shows the longitudinal sum of AWC versus latitude for different timescales. Compared with the reanalysis, the curves obtained are less smooth, but the general shapes are similar. The values of AWC are, as expected, substantially



**Figure 8.** Maps of the cross-correlation using the NOAA data. The color represents the mean value of the PC of each location, with the WP and CP boxes (defined in the text) taking 30-year windows centered at 1915, 1940 and 1995. (a) and (b) correspond to 1915, WP and CP, respectively, (c) and (d) to 1940, WP and CP, respectively, and (e) and (f) to 1995, WP and CP, respectively.

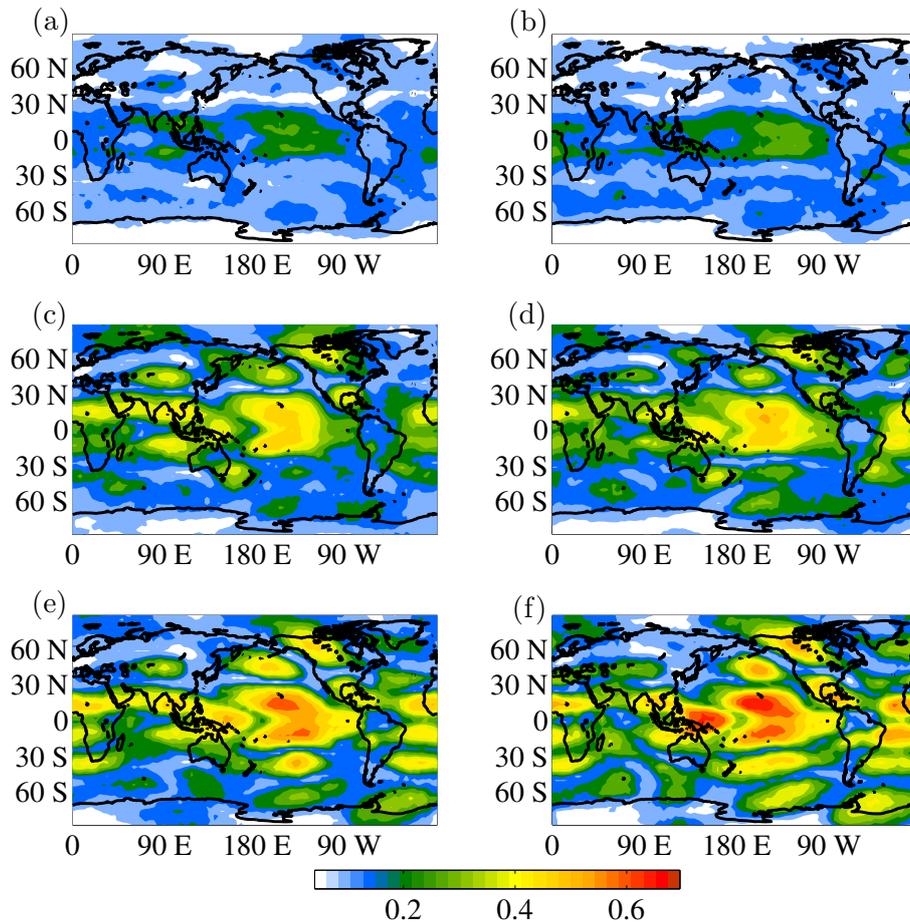
smaller than in the reanalysis, given that internal variability has been filtered. The differences in magnitudes are smaller on longer timescales, indicating the dominant role of the oceans in the inter-annual global connectivity.

As in the reanalysis data, the second period shows enhanced connectivity in the tropics for long timescales. There are asymmetries between the two maxima near the Equator that are, however, not present in the reanalysis data networks. Moreover, in the Southern Hemisphere extratropics, the intra-seasonal networks show largest connectivity in the second period, contrary to the results obtained in the reanalysis. This suggests that atmospheric internal variability plays an important role in the connectivity of the southern extratropics on intra-seasonal timescales.

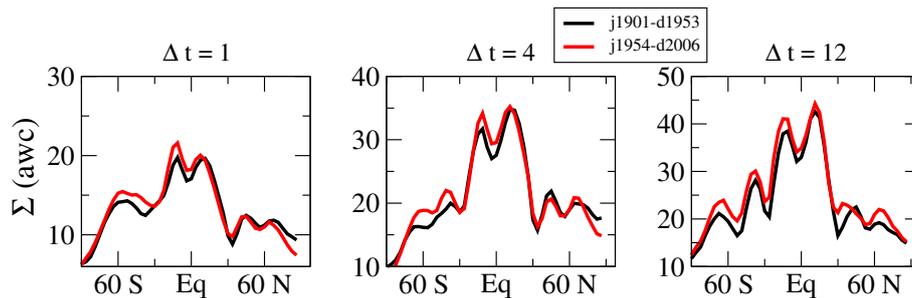
The AWC maps of internal atmospheric variability are shown in Fig. 11. Both Pearson correlation and mutual information with  $\Delta t = 1$  and 4 have mostly the same spatial structure but differ where global maxima are located. AWC maps are also strongly similar between periods.

On intra-seasonal timescales the regions of maximum connectivity are in the western Pacific and in the southern extratropics. There are clear connections in the Southern Hemisphere between the South Pacific, South America and the South Atlantic. There is also a small difference between the two periods: the connectivity in the western Pacific region increases, and it decreases in the central Pacific, similar to the NOAA 20th Century Reanalysis data. This result, together with the AWC maps of the forced atmospheric variability, suggests that the changes in connectivity observed in the reanalysis during the 20th century are due to changes in both forced and intrinsic components.

On intra-annual timescales, on the other hand, the AWC maps present the same well-connected spots as the intra-seasonal case, but also additional regions mainly in the North Pacific and the North Atlantic. The wavy patterns suggest the existence of wave train propagation all over the globe. The Pearson correlation network has different link density, but the structure is quite similar to the intra-annual case.



**Figure 9.** AWC maps of the ICTP-AGCM oceanically forced component. The networks are built by evaluating the mutual information with the BP methodology:  $\Delta t = 1$ : (1901–1953) (a) and (1954–2006) (b),  $\Delta t = 4$ : (1901–1953) (c) and (1954–2006) (d), and  $\Delta t = 12$ : (1901–1953) (e) and (1954–2006) (f).



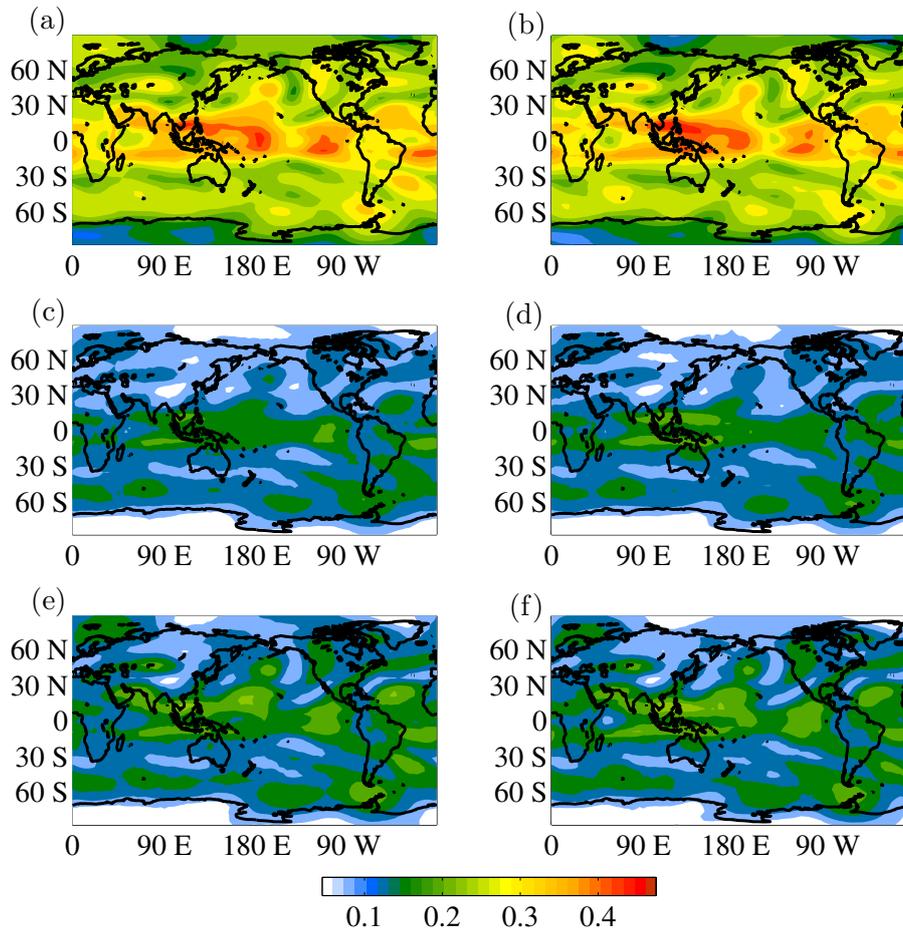
**Figure 10.** Zonal sum of the AWC function for the ICTP-AGCM networks built by using the BP methodology with  $\Delta t = 1$ ,  $\Delta t = 4$  and  $\Delta t = 12$ , for both time periods (black: 1901–1953, red: 1954–2006). It is a measure of how many connections each latitude has.

### 3.5 Principal component analysis

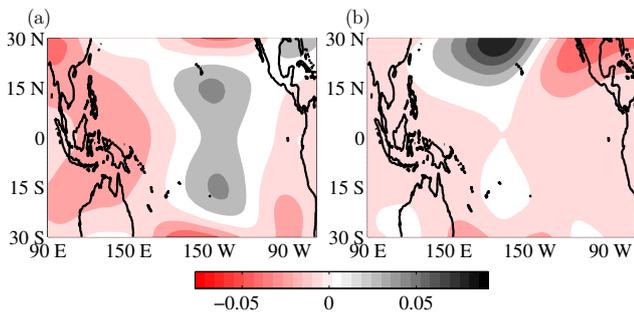
Considering the influence, or connectivity, of the tropical Pacific, a principal component analysis was done using the NOAA 20th Century Reanalysis data (Compo et al., 2011) in this region. In particular, the time series within the following coordinates were considered: [90° E, 110° W–30° S, 30° N]. As in the sections above, we worked with anomalies of the

200 mb eddy geopotential height and considered the data in halves: [January 1901–December 1955] and [January 1955–December 2009].

Figure 12 shows the first two eigenvectors of the 1955–2009 period, rotated following the simplicity (varimax) criterion. The percentages of explained variance are 14.9 and 9.1, respectively. It is worth mentioning that the leading eigenvector has the same structure as the connectivity degree (Donges



**Figure 11.** AWC maps of the ICTP-AGCM internal variability. The networks are built with PC (1901–1953) (a) and (1954–2006) (b), BP methodology with  $\Delta t = 1$  (1901–1953) (c) and (1954–2006) (d), and BP with  $\Delta t = 4$  (1901–1953) (e) and (1954–2006) (f).



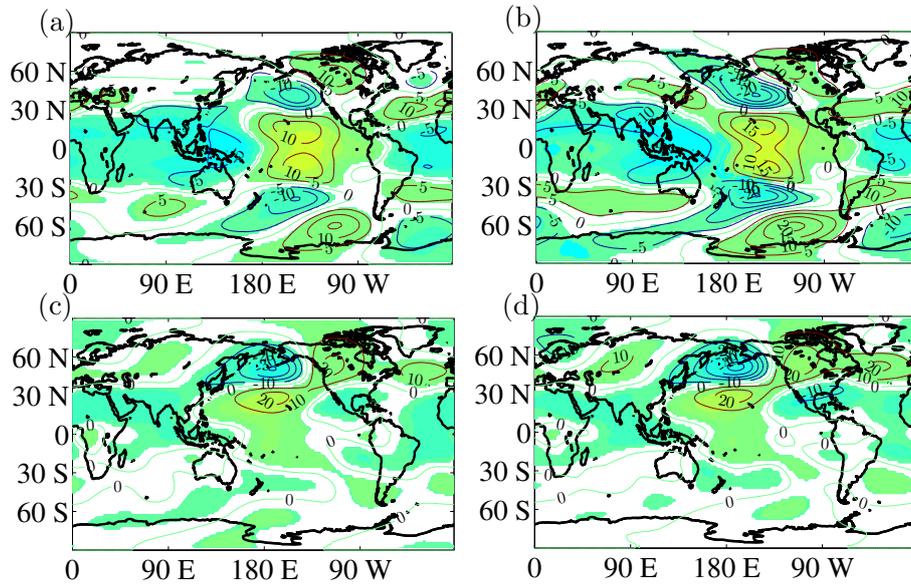
**Figure 12.** The first two eigenvectors of the covariance matrix (EOFs) of the NOAA reanalysis data considering the time period (1955–2009) rotated by the simplicity criterion. Variance explained: 14.9 % (ENSO pattern related) (a), 9.1 % (PNA pattern related) (b).

et al., 2013). The first EOF is consistently strongly related to the ENSO pattern.

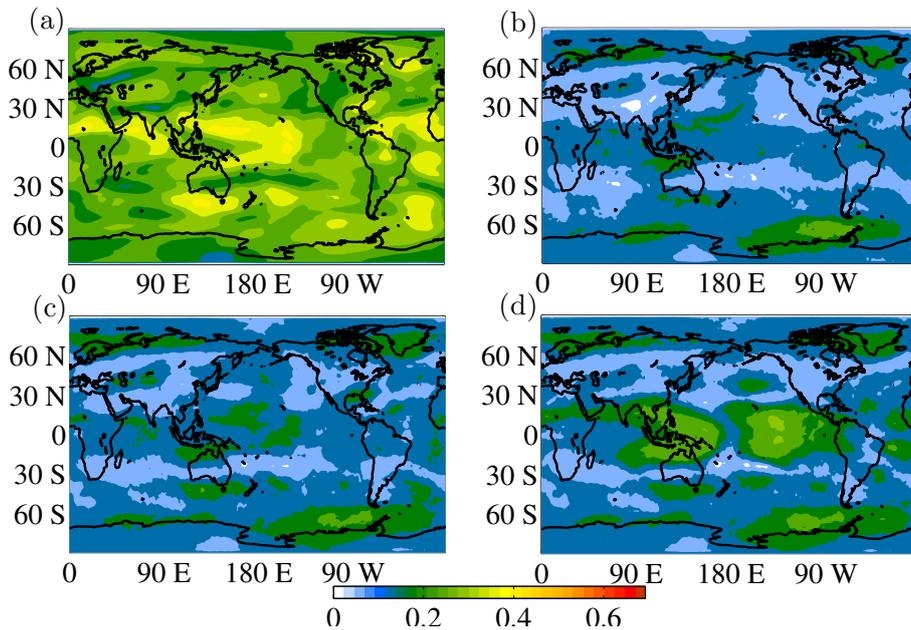
The maps in Fig. 13 show the regressions of global eddy geopotential height anomalies onto the principal components of the first two EOFs. Clearly, the regression with the leading

principal component has a very similar structure to the map of the AWC of the ICTP-AGCM forced component, indicating that the ocean forcing is to a large extent from the tropical Pacific. The second eigenvector is also related in both periods to the Pacific–North American (PNA) pattern. Note that for the leading EOF in the second period, the extratropical wave train that propagates from the central Pacific is much stronger than in the first period.

Taking into account that the first two patterns are related to well-known modes of variability such as ENSO and PNA, we build the AWC maps for the original data without these principal components (Fig. 14). The AWC function of the Pearson correlation network without the ENSO-related component (Fig. 14a) shows an important decrease mainly in the tropical regions (compare it with Fig. 2b). The overall appearance of the resulting structure is similar to the one obtained for the atmospheric internal variability of the ICTP-AGCM, showing again the important influence of ENSO on the atmospheric dynamics. On the other hand, Fig. 14b shows the AWC map of the mutual information, with BP symbolic analysis on the intra-seasonal scale ( $\Delta t = 1$ ) without



**Figure 13.** Regressions of the two principal components obtained with the global eddy geopotential height field where the correlation is statistically significant (these areas are colored for visualization only). (a) and (c) from 1901 to 1955; (b) and (d) from 1955 to 2009.



**Figure 14.** AWC maps of the NOAA reanalysis using data from the second period (1955–2009). The networks are built after removing one of the two principal components related to ENSO and PNA patterns. PC without ENSO-related component (a), MI with intra-seasonal timescale without ENSO-related component (b), and MI with intra-annual timescale without the ENSO and PNA components ((c) and (d), respectively).

the ENSO-related component revealing that the connectivity in the tropical Pacific decreases, but to a far lesser extent than in Fig. 14a. The general structure in the tropical connectivity is maintained, and some subpolar regions become more connected. This suggests that this network connectivity is less ENSO dependent, and thus less influenced by the

oceanically forced component than the Pearson correlation network. In the bottom panels the AWC maps of the mutual information on the intra-annual scale network (BP with  $\Delta t = 4$ ) without the ENSO-related component (Fig. 14c) and without the PNA-related component (Fig. 14d) are shown. Comparison reveals that in the northern Pacific, the region

where both EOFs have weight, the maximum connectivity region seen in Fig. 3d is composed of a center due to the ENSO component located at 40° N, 160° W and a broadly connected region to the northwest related to the PNA.

#### 4 Conclusions

We have studied the variability of the 200 mb eddy geopotential height anomaly field by means of networks built using linear and nonlinear statistical similarity measures. The analysis considered two reanalysis data sets (NOAA 20th Century and NCEP CDAS1) and the output of an atmospheric general circulation model (ICTP-AGCM). This combination allowed us to study the evolution of the upper atmosphere connectivity over the 20th century as well as to disentangle the networks due to the forced and internal components.

Overall, the resulting connectivity structures built using Pearson correlation (linear) or the mutual information (non-linear) have similar spatial features dominated by high connectivity in the tropics and smaller connectivity centers shaped as wave trains in the extratropics.

The use of mutual information calculated with the BP symbolic analysis allows for the separation of the connectivity, depending on the timescale. On intra-seasonal scales ( $\Delta t = 1$ ), the connectivity of the central and western Pacific regions shows significant changes between both halves of the 20th century: while in the first half the tropical Pacific is fully connected, as with the Pearson correlation or mutual information with histograms, in the second period the eastern area appears less connected and the global maximum is in the western Pacific.

On the intra-annual timescale ( $\Delta t = 4$ ), the main feature is the maximum in connection with the northern Pacific, which is larger than on the other two timescales considered. Finally, for the inter-annual timescale ( $\Delta t = 12$ ), the southern subtropical Pacific arises as a connectivity hub, also seen in the AWC constructed with the Pearson correlation and mutual information with histograms, but which is not so clear in the other timescale networks. In other regions such as Canada, the connectivity decreases (increases) in the second period on intra-annual (inter-annual) timescales. These examples indicate that the BP methodology is useful in separating different timescale processes.

Furthermore, in order to deepen our understanding of the changes in the connectivity during the 20th century, we have studied the AWC time evolution by considering a 30-year sliding window of two main boxes located in the western Pacific (WP) and the central Pacific (CP). The temporal evolution in the connectivity of these regions indicates a decreasing behavior at the beginning of the century until 1940, when it starts increasing until the end of the century for the PC, MIH and inter-annual networks. Intra-annual networks, on the other hand, present a negative trend during the first decades, and a positive trend from 1956 onwards.

The connectivity thus evolves similarly in WP and CP for the PC, MIH, inter-annual and intra-annual networks. On the other hand, the connectivity of intra-seasonal networks is very similar to the intra-annual networks in CP, but shows a positive trend in WP during the entire 20th century. Even though is not trivial to interpret these results, we might assume that when the connectivity of these main regions decreases (increases) in a particular timescale network, the teleconnections associated with these regions and with this particular timescale are weakened (strengthened) in the temporal period considered. In this case, both the mid-century minima of all the networks' connectivity in the central Pacific, as well as the continuously increasing connectivity in the intra-seasonal timescale networks in the western Pacific, deserve further investigation.

Finally, using the ICTP-AGCM we were able to separate the oceanically forced component from the atmospheric internal variability and calculate the respective connectivity networks. The results suggest that the main patterns of connectivity captured by the reanalysis networks are due to the oceanically forced component, particularly on inter-annual timescales. On intra-seasonal timescales the atmospheric internal variability seems to play an important role in determining the network structure.

This study showed a great richness in upper atmospheric connectivity during the 20th century in time and space, revealed through the use of linear and nonlinear similarity measures and a technique to separate timescales. We focused on the tropical Pacific due to its global relevance. Subsequent studies should focus on extratropical regions, where this technique can prove useful for improving climate predictions.

*Acknowledgements.* The authors would like to acknowledge the support of the LINC project (no. 289447) funded by EC's Marie Curie ITN program (FP7-PEOPLE-2011-ITN).

Edited by: J. Donges

Reviewed by: two anonymous referees

#### References

- Albert, R. and Barabási, A. L.: Statistical mechanics of complex networks, *Rev. Mod. Phys.*, 74, 47–97, 2002.
- Bandt, C. and Pompe, B.: Permutation entropy: A natural complexity measure for time series, *Phys. Rev. Lett.*, 88, 174102, 2002.
- Barabási, A. L. and Albert, R.: Emergence of scaling in random networks, *Science*, 286, 509–512, 1999.
- Barreiro, M.: Influence of ENSO and the South Atlantic Ocean on climate predictability over Southeastern South America, *Clim. Dynam.*, 35, 1493–1508, 2010.
- Barreiro, M., Fedorov, A., Pacanowski, R., and Philander, S. G.: Abrupt Climate Changes: How Freshening of the Northern Atlantic Affects the Thermohaline and Wind-Driven Oceanic Circulations, *Annu. Rev. Earth Planet. Sci.*, 36, 33–58, 2008.

- Barreiro, M., Marti, A. C., and Masoller, C.: Inferring long memory processes in the climate network via ordinal pattern analysis, *Chaos*, 21, 013101, doi:10.1063/1.3545273, 2011.
- Broennimann, S., Stickler, A., Griesser, T., Fischer, A. M., Grant, A., Ewen, T., Zhou, T., Scharner, M., Rozanov, E., and Peter, T.: Variability of large-scale atmospheric circulation indices, *Meteorol. Z.*, 18, 379–396, 2009.
- Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, A., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, *Q. J. Roy. Meteorol. Soc.*, 137, 1–28, doi:10.1002/qj.776, 2011.
- Deza, J. I., Masoller, C., and Barreiro, M.: Distinguishing the effects of internal and forced atmospheric variability in climate networks, *Nonlin. Processes Geophys.*, 21, 617–631, doi:10.5194/npg-21-617-2014, 2014.
- Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Complex networks in climate dynamics, *Eur. Phys. J.-Spec. Top.*, 174, 157–179, doi:10.1140/epjst/e2009-01098-2, 2009.
- Donges, J. F., Petrova, I., Loew, A., Marwan, N., and Kurths, J.: Relationships between eigen and complex network techniques for the statistical analysis of climate data, *Rev arxiv:1305.6634v1*, [physicsdata-an] arXiv:1305.6634v1, 2013.
- Held, I. M.: Stationary and quasi-stationary eddies in the extratropical troposphere: Theory. Large-scale Processes in the Atmosphere, edited by: Hoskins, B. J. and Pearce, R. P., Academic Press, 127–168, 1983.
- James, I. N.: Introduction to circulating atmospheres, Cambridge University Press, Cambridge, UK, 1994.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, 77, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.
- Kucharski, F., Molteni, F., and Bracco, A.: Decadal interactions between the western tropical Pacific and the North Atlantic Oscillation, *Clim. Dynam.*, 26 79–91, 2006.
- Kucharski, F., Bracco, A., Yoo, J. H., and Molteni, F.: Atlantic forced component of the Indian monsoon interannual variability, *Geophys. Res. Lett.*, 35, L04706, doi:10.1029/2007GL033037, 2008.
- Kucharski, F., Bracco, A., Yoo, J. H., Tompkins, A., Feudale, L., Ruti, P., and Dell’Aquila, A.: A Gill-Matsun-type mechanism explains the Tropical Atlantic influence on African and Indian Monsoon Rainfall, *Q. J. Roy. Meteorol. Soc.*, 135, 569–579, doi:10.1002/qj.406, 2009.
- Lintner, B. R. and Chiang, J. C. H.: Adjustment of remote tropical climate of El Niño conditions, *J. Climate*, 20, 2544–2557, 2006.
- Martin, E. A., Paczuski, M., and Davidsen, F.: Interpretation of link fluctuations in climate networks during El Niño periods, *Eur. Phys. Lett.*, 102, 48003, 2013.
- Molteni, F.: Atmospheric simulations using a GCM with simplified physical parametrizations, I. Model climatology and variability in multi-decadal experiments, *Clim. Dynam.*, 20, 175–191, 2003.
- Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F. F., Wakata, Y., Yamagata, T., and Zebiak, S. E.: ENSO theory, *J. Geophys. Res.-Oceans*, 103, 14261–14290, 1998.
- Newman, M. E. J.: *Networks: An Introduction*, Oxford University Press, Oxford, 2010.
- Paluš, M., Hartman, D., Hlinka, J., and Vejmelka, M.: Discerning connectivity from dynamics in climate networks, *Nonlin. Processes Geophys.*, 18, 751–763, doi:10.5194/npg-18-751-2011, 2011.
- 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, 052807, doi:10.1103/PhysRevE.88.052807, 2013.
- Seager, R., Harnik, N., Kushnir, Y., Robinson, W., and Miller, J.: Mechanisms of hemispherically symmetric climate variability, *J. Climate*, 16, 2960–2978, 2013.
- Smith, T. M. and Reynolds, R. W.: Improved extended reconstruction of SST (1854–1997), *J. Climate*, 17, 2466–2477, 2004.
- Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-C., and Ropelewski, C.: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures, *J. Geophys. Res.*, 103, 14291–14324, 1998.
- Tsonis, A. A. and Roebber, P. J.: The architecture of climate network, *Physica A*, 333, 497–504, 2004.
- Tsonis, A. A. and Swanson, K. L.: Topology and predictability of El Niño and La Niña networks, *Phys. Rev. Lett.*, 100, 228502, doi:10.1103/PhysRevLett.100.228502, 2008.
- Tsonis, A. A., Swanson, K. L., and Roebber, P. J.: What do networks have to do with climate?, *Bull. Amer. Meteorol. Soc.*, 87, 585–596, doi:10.1175/BAMS-87-5-585, 2006.
- Tsonis, A. A., Wang, G., Swanson, K. L., Rodrigues, F. A., and da Fontoura Costa, L.: Community structure and dynamics in climate networks, *Clim. Dynam.*, 37, 933–940, 2011.
- Watts, D. J. and Strogatz, S. H.: Collective dynamics of ‘small-world’ networks, *Nature*, 393, 440–442, doi:10.1038/30918, 1998.
- Yamasaki, K., Gozolchiani, A., and Havlin, S.: Climate Networks around the Globe are significantly affected by El Niño, *Phys. Rev. Lett.*, 100, 228501, 2008.