



Classification of mechanisms, climatic context, areal scaling, and synchronization of floods: the hydroclimatology of floods in the Upper Paraná River basin, Brazil

Carlos H. R. Lima¹, Amir AghaKouchak², and Upmanu Lall³

¹Civil and Environmental Engineering, University of Brasilia, Brasilia, Distrito Federal, Brazil

²Civil and Environmental Engineering, University of California, Irvine, Irvine, California, USA

³Earth and Environmental Engineering, Columbia University, New York, New York, USA

Correspondence to: Carlos H. R. Lima (chrlima@unb.br)

Received: 3 February 2017 – Discussion started: 14 February 2017

Revised: 6 October 2017 – Accepted: 16 October 2017 – Published: 1 December 2017

Abstract. Floods are the main natural disaster in Brazil, causing substantial economic damage and loss of life. Studies suggest that some extreme floods result from a causal climate chain. Exceptional rain and floods are determined by large-scale anomalies and persistent patterns in the atmospheric and oceanic circulations, which influence the magnitude, extent, and duration of these extremes. Moreover, floods can result from different generating mechanisms. These factors contradict the assumptions of homogeneity, and often stationarity, in flood frequency analysis. Here we outline a methodological framework based on clustering using self-organizing maps (SOMs) that allows the linkage of large-scale processes to local-scale observations. The methodology is applied to flood data from several sites in the flood-prone Upper Paraná River basin (UPRB) in southern Brazil. The SOM clustering approach is employed to classify the 6-day rainfall field over the UPRB into four categories, which are then used to classify floods into four types based on the spatiotemporal dynamics of the rainfall field prior to the observed flood events. An analysis of the vertically integrated moisture fluxes, vorticity, and high-level atmospheric circulation revealed that these four clusters are related to known tropical and extratropical processes, including the South American low-level jet (SALLJ); extratropical cyclones; and the South Atlantic Convergence Zone (SACZ). Persistent anomalies in the sea surface temperature fields in the Pacific and Atlantic oceans are also found to be associated with these processes. Floods associated with each cluster present different patterns in terms of frequency, magnitude, spatial variability, scaling, and synchronization of events across the sites and subbasins. These insights suggest new directions for flood risk assessment, forecasting, and management.

1 Introduction

The assumptions of homogeneity, stationarity, and randomness in traditional flood frequency studies have been questioned in numerous studies (e.g., Hirschboeck et al., 2000; Jain and Lall, 2001; Milly et al., 2002; Alila and Mtiraoui, 2002; Kwon et al., 2008; Villarini et al., 2009, 2013; Smith et al., 2011; Westra and Sisson, 2011; Vogel et al., 2011; Neiman et al., 2011; Seo et al., 2012; Merz et al., 2014; Lima et al., 2015). To make progress on understanding and mod-

eling the real world flood process, one needs to better understand how the complex interactions among weather, climate, hydrology, basin attributes, and antecedent conditions evolve over space and time.

Historically, flood studies have followed two distinct research lines: hydrometeorology of floods and flood frequency analysis. Flood hydrometeorology focuses on understanding (i) hydrodynamics of the rainfall–runoff process during flood events, (ii) spatial structure of local rainfall events that are associated with floods, (iii) soil–atmosphere

response, and large-scale circulation patterns associated with the forecast and diagnosis of rainfall events (Maddox, 1983; Kunkel et al., 1994; Pal and Eltahir, 2002; Schumacher and Johnson, 2005; Amengual et al., 2007; Viglione et al., 2010; Li et al., 2013). There is also extensive literature related to the statistical analysis and modeling of flood frequency from local and regional data of rainfall, streamflow, and water basin attributes, including nonstationary approaches (e.g., Thomas and Benson, 1970; Stedinger and Cohn, 1986; Stedinger et al., 1993; Kroll and Stedinger, 1998; Kwon et al., 2008; Lima and Lall, 2010; Cheng et al., 2014; Luke et al., 2017).

In this study, we investigate floods in the Upper Paraná River basin (hereafter, UPRB) in southern Brazil using a hydroclimatology framework to understand the flood-generating mechanisms (Hirschboeck, 1988). The overarching goal is to link frequency of flood events to flood-generating mechanisms to provide a better understanding of the underlying physical processes (Moftakhari et al., 2017). The underlying assumptions in flood frequency studies can be enriched by a formal consideration of the physical mechanisms responsible for the generation of extreme floods. This includes a recognition of the natural climate variability associated with persistence and oscillatory regimes (e.g., El Niño) across different timescales (e.g., interannual, decadal) as well as climatic changes in response to anthropogenic changes in atmosphere, soil, and land use.

Many studies have investigated the interactions between basin attributes and atmospheric circulation leading to extreme or exceptional floods (in the context of this work, it means floods with an exceedance probability of 70 %). However, there is limited knowledge as to how evolving large-scale climate modes on the interannual scale change the occurrence of local precipitation and soil moisture, thus altering the probability distribution and occurrence of floods (Sun et al., 2016). It is argued that the frequency of flood events is very sensitive to modest changes in climate (Knox, 1993; Sun et al., 2016). We explore Hirschboeck's hypothesis (Hirschboeck, 1988) that exceptional floods in basins of all sizes could be related to anomalies in the large-scale atmospheric circulation. This flood hydroclimatological perspective has been applied to identify the moisture transport and large-scale climate patterns associated with floods in the United States (Hirschboeck, 1988; Budikova et al., 2010; Nakamura et al., 2013; Lu and Lall, 2016; Mallakpour and Villarini, 2016), Europe (Jacobeit et al., 2003; Bárdossy and Filiz, 2005; Prudhomme and Genevier, 2010; Lu et al., 2013), and other parts of the world (Kahana et al., 2002). However, such flood studies are rare in South America.

Intuitively, a rainfall system that persists over a given locale with a continuous and sufficient supply of moisture (from advection and recycling) has a high likelihood of generating an exceptional flood. For sufficiently large drainage areas, an extreme flood may require an external flux of advective moisture, i.e., local convective processes may not tend to

produce exceptional floods in these basins. Moreover, such an influx of large-scale advective moisture can lead to an increased potential for large floods as the drainage area and return period increase. Hirschboeck (1988) and Hirschboeck et al. (2000) note that the scale of convective storms that can generate intense short rainfall is typically of $10\text{--}10^2 \text{ km}^2$; therefore, it is unlikely that such convective processes are the main source of exceptional floods over large areas. Conversely, mesoscale convective systems (MCSs), such as a mesoscale convective complex (MCC) and squall lines, tend to cover large areas and persist for several hours and are sources of heavy rainfall in some regions of the USA (Schumacher and Johnson, 2005, 2006) and also Brazil (Zipser et al., 2006; Salio et al., 2007; Durkee and Mote, 2009; Durkee et al., 2009; Marengo et al., 2012), in particular the MCCs to the east of the Andes that impact the La Plata basin. However, there is evidence (Maddox, 1983; Corfidi et al., 1996) that the maintenance and development of such systems is related to large-scale atmospheric circulation features. Thus, tropical and extratropical cyclones and associated fronts become important in the production of extreme rainfall over large areas and are directly related to large-scale atmospheric circulation patterns and to storm paths or well-defined regions of moisture transport in the atmosphere.

We explore extreme floods in the UPRB through a hydroclimatic analysis of flood series across 33 nested-basin sites with drainage areas ranging from 2588 to 823 555 km^2 . The spatiotemporal dynamics of daily rainfall over the basin in the days that preceded the largest flood events is analyzed and classified into clusters of similar patterns based on a self-organizing map (SOM) (Kohonen, 2001) clustering algorithm. This way, we intend to take into account the persistence and alignment of the storm path with the drainage basin that produces a given flood. The associated large-scale atmospheric circulation for each cluster is then analyzed in terms of moisture transport and convergence, high-level circulation, and vorticity. Teleconnections with the Atlantic and Pacific oceans are evaluated using composite analysis of the sea surface temperature (SST) field. For each rainfall cluster, the attributes (frequency, magnitude, scaling, and synchronization) of floods across the UPRB are analyzed in order to produce and characterize a typology for floods in the region according to the dynamics of rainfall patterns and associated atmospheric circulation. Floods generated by snowmelt, tropical cyclones, and storm surges do not affect the UPRB and are thus not investigated in this study. The paper is organized as follows. In the next section we present the region of study and data collection. In Sect. 3 we introduce the clustering algorithm. In Sect. 4 we present the results, and finally in Sect. 5 we offer a summary and discussion.

2 Region of study and hydroclimate data set

2.1 The Upper Paraná River basin, streamflow, and rainfall data set

The UPRB is located in southern Brazil (Fig. 1) and is part of the La Plata basin, which is the second largest basin in South America after the Amazon basin. UPRB concentrates a large population of Brazil and is of utmost importance for the country in terms of flood control, hydropower generation, and agriculture. The rainfall season over UPRB is mostly marked by a peak during the austral summer (summer monsoon system) related to the South American monsoon system (SAMS) and associated South Atlantic Convergence Zone (SACZ; see Barros et al., 2000; Jones and Carvalho, 2002; Berbery and Barros, 2002; Carvalho et al., 2004; Marengo et al., 2012), particularly in the region north of 20° S, where the monsoon system is the dominant forcing (Berbery and Barros, 2002). Rainfall interannual variability has been associated with SST anomalies in the tropical Pacific and South Atlantic oceans (Grimm et al., 1998, 2000; Robertson and Mechoso, 2000; Doyle and Barros, 2002; Grimm, 2003, 2004; Chaves and Nobre, 2004; Cardoso and Dias, 2006; Jorgetti et al., 2014). Intra-seasonal and decadal variability in rainfall and streamflow have also been the subject of many studies (Robertson and Mechoso, 2000; Robertson et al., 2001; Zhou and Lau, 2001; Paegle and Mo, 2002; Carvalho et al., 2004). Most of the moisture that reaches UPRB is from the Amazon region (Drumond et al., 2008; Carvalho et al., 2011), and the rainfall mechanisms are also associated with MCSs along the South American low-level jet (SALLJ; see Velasco and Fritsch, 1987; Marengo et al., 2004; Salio et al., 2007) and transient systems related to extratropical cyclones and cold fronts (Mendes et al., 2007; Silva and Ambrizzi, 2010). El Niño events have also been linked to extreme rainfalls and floods in the UPRB (Camilioni and Barros, 2003; Grimm and Tedeschi, 2009; Muza et al., 2009; Cavalcanti et al., 2015; Antico et al., 2016).

We use mean daily streamflow data from 33 sites in the UPRB (Fig. 1). These sites are located strategically to provide the inflow into the main hydropower reservoirs in the UPRB, which are used not only for generation of electrical energy but also for flood control, water supply, and agriculture. The data set is offered by the Electric System National Operator (ONS), which defines the operational rules of all interconnected hydropower reservoirs in the country. The streamflow data are available from January 1931 to December 2013, but in order to be consistent with the availability of the rainfall data set, we perform all analyses considering the streamflow data restricted to the 1980–2013 period. All series have gone through a consistency process from the ONS and represent naturalized flows from artificial and natural streamflow gauges, which means that any reservoir operation upstream of the streamflow gauge is removed from the original series.

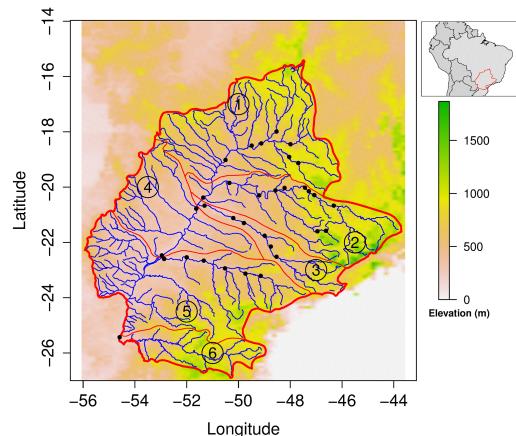


Figure 1. The Paraná River basin (red contour) and streamflow gauges used in this work (black dots). The elevation is in meters and the location of the Paraná River basin within Brazil is shown in the smaller panel in the upper right corner (red line contour). The thinner red line shows the associated subbasins: (1) Paranaíba, (2) Grande, (3) Tietê, (4) Paraná, (5) Paranapanema–Paraná, and (6) Iguaçu.

We limit our analysis to the wet season (November through March), when over 75 % of the floods occur. For each site, we obtain partial duration series of floods by taking the values at which the daily flow exceeds a given threshold. In order to keep a relatively large number of exceptional flood events in each rainfall cluster, we set this threshold as the 70th empirical flood quantile for the wet season. We analyze only independent floods by de-clustering the series (Lang et al., 1999) and taking events with inter-arrival times larger than 15 days, which we believe is a consistent interval to guarantee independence between flood events, considering the different rainfall mechanisms that cause floods in the UPRB. From this procedure, we obtain dates and magnitudes of about 98 flood events (average across all sites, ranging from 76 to 131 events per site) for each of the sites in UPRB analyzed here.

Daily gridded rainfall data ($0.25^\circ \times 0.25^\circ$) for the period 1980–2013 are provided by Xavier et al. (2016). These data consist of interpolated daily rainfall observations from 3625 rainfall gauges and 735 weather stations across Brazil available from different institutions (INMET, ANA, and DAEE). The interpolation schemes and validation procedures are described in Xavier et al. (2016). The rainfall data are delimited by the UPRB boundary as shown in Fig. 1. For each grid point, daily anomalies of rainfall are obtained after removing, from the observed value, the respective long-term monthly mean for that grid point based on the 1980–2013 period.

2.2 Moisture fluxes, vorticity, upper-level winds, and sea surface temperature

Mean daily data of vertically integrated moisture fluxes and the associated divergence field (evaporation – precipitation along an atmospheric column), low- and high-level relative vorticity, and high-level (500 mbar) winds are obtained from the ERA-Interim reanalysis data (Dee et al., 2011). They cover the period from 1980 to 2013 and are retrieved for the region defined by 15° N–60° S and 270–330° W.

We also use daily SST data from the ERA-Interim global SST archive for the 1980–2013 period. Daily SST anomalies for each grid point are calculated by subtracting, from the observed value, the monthly mean for that grid point and related month based on the 1980–2013 period. The SST field is delimited by the region 30° N–80° S and 210° W–20° E.

All data are interpolated for a grid of 2.5° × 2.5°.

3 Technical approach

The spatiotemporal dynamics of daily rainfall over the UPRB in the days that preceded the largest flood events is analyzed and classified into clusters based on a SOM (Kohonen, 2001) clustering algorithm, which is described in the following subsection. By doing this, we consider the persistence and alignment of the storm path with the drainage basin that produces a given flood. For each rainfall cluster, the associated large-scale atmospheric and ocean circulations are evaluated through composite analysis of moisture transport and convergence, high-level circulation, vorticity, and the Atlantic and Pacific SST fields. The attributes (frequency, magnitude, and regional scaling) of floods associated with each rainfall cluster are also analyzed. We then propose a typology for floods in the region according to the rainfall patterns and corresponding atmospheric and ocean circulation. Finally, we employ the ideas of event synchronization and complex networks to explore the spatial dynamics of floods over the UPRB conforming the rainfall clusters. The methodology to evaluate the synchronization of flood events is presented in Sect. 3.2.

3.1 Rainfall clustering

A flood event, defined as a crossing of river stage above its bank height, can vary in duration from a few minutes to months and in spatial extent from a few square kilometers to several millions of square kilometers. A large number of flood studies have focused on understanding physical processes associated with floods in small-scale basins due to the ease of observing critical events in these basins (e.g., Gupta and Dawdy, 1995), while over large areas the focus tends to be on the problem of predicting flood quantiles, with lesser emphasis on understanding the physical mechanisms associated with extreme floods. For instance, the relation of soil moisture and a given rainfall event in producing some floods

over small areas and homogeneous soils is relatively easy to evaluate. Conversely, the problem becomes considerably more complicated as we consider large basins, with drainage areas over 10⁴ km² since (i) the potential of a high heterogeneity in the initial soil moisture field is high and (ii) the location and direction of the storm path along the basin leads to a significant heterogeneity in the spatial and temporal distribution of the rainfall event. Since the influx of large-scale advective moisture may be a particular factor to overlie the initial heterogeneities of the surface conditions for larger basins, we will assume that the spatiotemporal variability (i.e., magnitude, persistence, and alignment of the storm path with the drainage basin) in rainfall is the key factor of producing floods across the UPRB sites evaluated in this work.

Consider that the information regarding the spatiotemporal patterns of rainfall associated with the major flood events is contained in a rainfall data set represented by a matrix $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_T]$, where \mathbf{x}_t is a column vector containing all the relevant information about the spatial variability and persistence of daily rainfall over the UPRB along days $t - \tau, t - \tau - 1, \dots, t$, for some time delay τ . T is the total number of effective days during the austral wet season (November–March) over the 1980–2013 period. Our goal is to extract information about \mathbf{X} through clustering. We use the SOM approach to cluster rainfall information as expressed in \mathbf{X} . SOMs are a particular case of competitive neural networks and were developed by the machine learning community in the 1990s (Kohonen, 2001) for cluster analysis and classification. They have been successfully applied to find clusters in climate systems (e.g., Cavazos, 2000; Hewitson and Crane, 2002; Johnson et al., 2008; Lee and Feldstein, 2013; Bao and Wallace, 2015; Li et al., 2015, 2016; Mioduszewski et al., 2016; Xu et al., 2016). An extended review of applications in climate science is provided by Liu and Weisberg (2011). SOMs are also known as Kohonen neural networks and the basic idea is to obtain a 2-D topology consisting of nodes (or neurons) that are associated with the input space \mathbf{X} , still preserving its topological features.

For the sake of clarity and understanding of the SOM properties and tuned parameters (i.e., parameters that can be subjectively set) used in this work, here we describe the key aspects of SOM. We refer the reader to Kohonen (2001) for more details about SOM. Let us assume that we have K neurons; then initially K representatives (or prototypes, synaptic weight vectors, reference vectors) are randomly chosen from the input space \mathbf{X} and associated with the K neurons. An input vector \mathbf{x}_t is randomly selected from the data set \mathbf{X} and the Euclidean distance between \mathbf{x}_t and each representative $\mathbf{m}_k, k = 1, \dots, K$, is computed. The neuron whose representative yields the smallest distance to \mathbf{x}_t is the winner neuron k^* or best-matching unit (BMU):

$$k^* = \arg \min_k \{||\mathbf{x}_t - \mathbf{m}_k||\}. \quad (1)$$

In the next step, the neurons that are neighbors (neighborhood set) of the winning node k^* are found based on the Euclidean distance and a given threshold c . The representatives corresponding to each grid neighbor of the winning neuron k^* are then updated according to the rule

$$\mathbf{m}_k \leftarrow \mathbf{m}_k + \alpha \cdot (\mathbf{x}_t - \mathbf{m}_k), \quad k \in \mathcal{N}_c(k^*), \quad (2)$$

where α , $0 \leq \alpha \leq 1$ is the so-called learning rate and $\mathcal{N}_c(k^*)$ denotes the set of points in the neighborhood of k^* given the parameter c . The process is then arbitrarily repeated a large number of times (epochs), since there is no explicit error criterion to minimize (Lee and Verleysen, 2007).

Variants of the update rule in Eq. (2) include a time-varying learning rate α and weighted distances based on the proximity of \mathbf{m}_k and the winning neuron \mathbf{m}_{k^*} :

$$\mathbf{m}_k \leftarrow \mathbf{m}_k + \alpha(j) \cdot h(||\mathbf{m}_k - \mathbf{m}_{k^*}||) \cdot (\mathbf{x}_t - \mathbf{m}_k), \quad k \in \mathcal{N}_c(k^*), \quad (3)$$

where $\alpha(j)$ is the learning rate at epoch j and $h(||\cdot||)$ is a neighborhood function around the winner neuron k^* . Common functions for $\alpha(j)$ include the linear, power, and inverse functions with a decrease rate over time. A common function for $h(||\cdot||)$ is the Gaussian kernel

$$h(||\mathbf{m}_k - \mathbf{m}_{k^*}||) = \exp \left\{ -\frac{||\mathbf{m}_k - \mathbf{m}_{k^*}||}{2 \cdot \sigma^2} \right\} I_{k \in \mathcal{N}_c(k^*)},$$

where σ is the width of the kernel (or neighborhood radius) and I the indicator function.

In the batch version of the SOM, instead of presenting a single data vector each time, the entire data set \mathbf{X} is presented to the SOM before any weights are updated and the BMU \mathbf{m}_{k^*} is obtained for each input data \mathbf{x}_t at each epoch, so that each data vector \mathbf{x}_t will belong to a given neuron and the new neurons are updated as

$$\mathbf{m}_k \leftarrow \frac{\sum_{t \in \mathcal{N}_c(k)} w_t \mathbf{x}_t}{\sum_{t \in \mathcal{N}_c(k)} w_t}, \quad (4)$$

where the weight function w_t can be a rectangular function, which is equal to 1 for the neighbors of \mathbf{m}_k and 0 otherwise, or be a smooth function $h(||\mathbf{x}_t - \mathbf{m}_k||)$. In this sense, each new neuron is a weighted average of the data samples that belong to its neighborhood neurons.

For a given number of neurons K , learning rate α , threshold c , and fixed number of epochs, the trained SOM can encode any point \mathbf{x}_t by giving the index k of the closest neuron \mathbf{m}_k , where the distance is computed similarly to Eq. (1). In this way, each data point of the entire data set of rainfall information \mathbf{X} can be assigned (or clustered) into one of the categories $1, \dots, K$.

The final embedding of \mathbf{X} can be evaluated by the mean quantization error (MMQE) of the SOM, which essentially

measures the average distance of each input \mathbf{x}_t to its representative in the output space:

$$\text{MMQE} = \frac{1}{T} \sum_{t=1}^T ||\mathbf{x}_t - \mathbf{m}_{x_t}||, \quad (5)$$

where \mathbf{m}_{x_t} refers to the BMU of the corresponding \mathbf{x}_t .

In order to capture the spatiotemporal dynamics of the rainfall field over the UPRB, including the information of antecedent rainfall for a given day t of the record, we will concatenate the rainfall field over a time window $\tau = 5$ days:

$$\mathbf{x}_t = [\mathbf{r}_{t-5} \ \mathbf{r}_{t-4} \ \mathbf{r}_{t-3} \ \mathbf{r}_{t-2} \ \mathbf{r}_{t-1} \ \mathbf{r}_t]', \quad (6)$$

where \mathbf{r}_t is a row vector representing the observed rainfall field over the Upper Paraná River basin (Fig. 1) at day t , with a dimension of 1178 (number of grid points), so that \mathbf{x}_t has a dimension of 7068.

It is interesting to note that as τ increases, the number of dimensions of \mathbf{x}_t increases as well and the associated rainfall pattern may not necessarily be connected with the flood events. Based on the results discussed in the next section and the lifetime of about 3 days of extratropical cyclones (Simmonds and Keay, 2000) and 3 days of SACZ events (Carvalho et al., 2004), we believe $\tau = 5$ days is an appropriate choice to extract the relevant information regarding the rainfall field during flood events.

To be coherent with the flood data as described in Sect. 2.1, we focus on the November–March daily rainfall. The data set covers the period from 1 January 1980 to 31 December 2013 with a total of 5143 data points. After concatenating the rainfall field as explained in Eq. (6), the number of data points reduces to 5138, starting now in 6 January 1980 and ending in 31 December 2013. This results in a 5138×7068 input data matrix \mathbf{X} to the SOM.

3.2 Flood event synchronization

The spatial dynamics of floods across the UPRB produced by each neuron will be qualitatively explored through the concepts of event synchronization and complex networks, which have been successfully applied in many fields (Quiroga et al., 2002) and also climate science (Malik et al., 2012; Marwan and Kurths, 2015), including for prediction of floods in South America (Boers et al., 2014). Following the nomenclature of Quiroga et al. (2002), let us define the time series of flood event dates (obtained from the partial duration series) for two given streamflow sites x and y as t_i^x and t_j^y , where $i = 1, \dots, m_x$, $j = 1, \dots, m_y$ and m is the total number of events. We define two synchronous flood events whenever the distance between t_i^x and t_j^y is less than a given time lag τ . Let $c^\tau(x|y)$ then be the amount of time in which a flood event in x follows, within the time lag τ , a flood event in y :

$$c^\tau(x|y) = \sum_{i=1}^{m_x} \sum_{j=1}^{m_y} J_{ij}^\tau, \quad (7)$$

where

$$J_{ij}^\tau = \begin{cases} 1 & \text{if } 0 < t_i^x - t_j^y \leq \tau \\ 1/2 & \text{if } t_i^x = t_j^y \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

Similarly, we can calculate $c^\tau(y|x)$. We will then define a measure Q_τ for the event synchronization:

$$Q_\tau = \frac{c^\tau(x|y) + c^\tau(y|x)}{\sqrt{m_x \cdot m_y}}, \quad (9)$$

where $0 \leq Q_\tau \leq 1$, and $Q_\tau = 1$ suggests full synchronization.

The delay behavior (or direction of flow) of the flood events can be measured by

$$q_\tau = \frac{c^\tau(x|y) - c^\tau(y|x)}{\sqrt{m_x \cdot m_y}}, \quad (10)$$

where $-1 \leq q_\tau \leq 1$, and $q_\tau = 1$ implies that flood events in x always precede flood events in y .

When combining all streamflow sites, Q_τ will be the elements of a square symmetric matrix while q_τ will be the entries of a square antisymmetric matrix. The matrix generated from Q_τ can then be converted into a square binary matrix, in which entries will represent only relevant connected sites. This can be accomplished by constructing the adjacency matrix \mathbf{A} :

$$\mathbf{A} == \begin{cases} 1 & \text{if } Q_\tau > T \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where T is a given threshold.

Methods to estimate T usually involve a bootstrap procedure, so that only a certain percentage of the total number of grid points (e.g. 5 %) are connected (Malik et al., 2012; Boers et al., 2014). In the particular case of this work, we are more interested in the gauges that have synchronized flood events with a specific one. Hence, we will define $T = 0.5$ so that we define synchronized gauges when at least 50 % of their flood events occur simultaneously.

The time lag τ should be less than half the minimum inter-event distance so that one single flood event is not synchronized with two events at another site. Based on this, a simple mathematical formulation is presented in Quiroga et al. (2002). In our case, in order to consider independent flood events, we have defined the partial duration series so that flood events are at least 15 days apart. Hence, $\tau = 7$. The average direction in which the flood event propagates will be simply evaluated using the sign of q_τ .

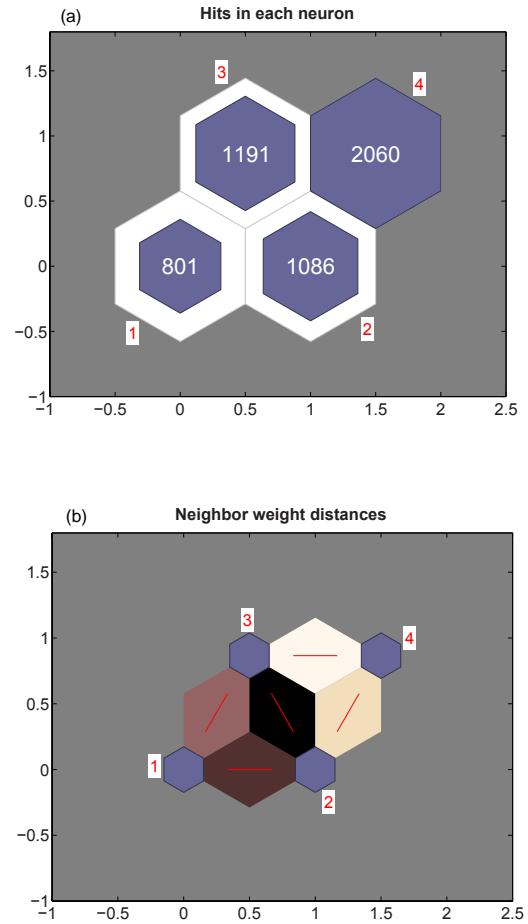


Figure 2. (a) Number of hits in each neuron (blue hexagons). (b) Connecting neighboring neurons (red lines). The colors in the regions containing the red lines indicate the distances between neurons, where darker colors represent larger distances and lighter colors represent smaller distances. The red numbers correspond to the number of the neuron.

4 Results

4.1 Rainfall clustering

We chose a 2×2 hexagonal grid to define the SOM, and the rainfall field is classified into $K = 4$ clusters. This choice is made primarily to associate a relatively large number of flood events in each rainfall cluster. The neighborhood radius c is initially set as 3 and monotonically decreases to 1 (equivalent to six neighbors for a central neuron in an hexagonal grid) when the number of epochs is equal to 100. This is the so-called ordering phase, in which a global order is achieved for the map (Kohonen, 2001). From 100 epochs c is set to 1 (tuning phase). Since the SOM grid consists of four neurons, then only two neighbors will have the size of their neighborhood affected by c (see Fig. 2 and related discussion). The weight function h in Eq. (4) is the rectangular function. The total number of epochs is set to 1000, but we do not ob-

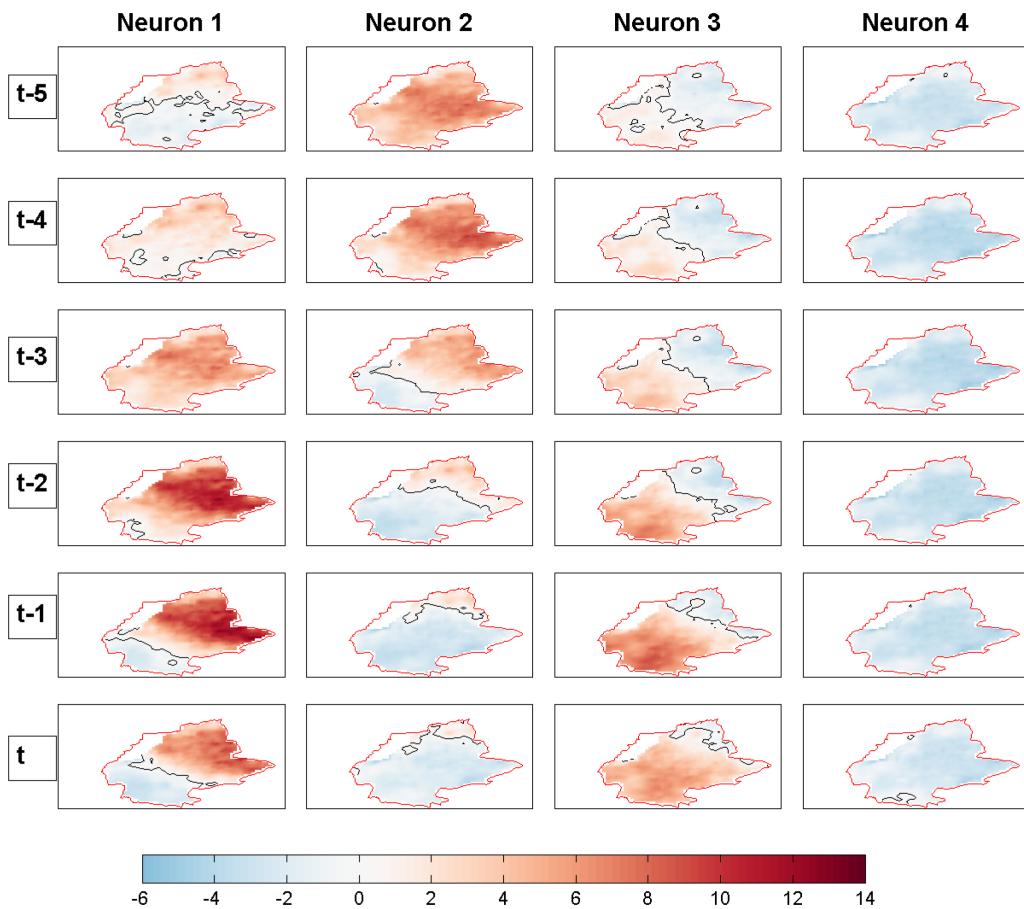


Figure 3. Neuron weights obtained for the SOM. These weights basically represent the rainfall anomalies (in millimeters) over the Upper Paraná River basin from day $t - 5$ (top panels) to day t (bottom panels). The black line shows the zero contour.

serve any significant difference in the MMQE after the first 200 epochs. At 1000 epochs we obtained $\text{MMQE} = 777.69$. We also evaluate MMQE for a 2×3 hexagonal grid and a 3×3 hexagonal grid and observe that the values tend to oscillate around $\text{MMQE} = 777.69$ as a function of the number of epochs, so that any significant differences for the 2×2 grid are observed. The SOM clustering algorithm is obtained using a commercial Neural Network Toolbox (MATLAB, 2014).

Figure 2 shows the final SOM after 1000 epochs in terms of hits in each neuron (panel a) and neighbors and weight distances (panel b). The number of hits is almost evenly distributed among neurons 1, 2, and 3. Neuron 4 has almost the double of hits of the other neurons. Due to the hexagonal grid layout, neurons 2 and 3 are connected to all the remaining neurons, while neurons 1 and 4 are connected only to neurons 3 and 2 (panel b in Fig. 2). The shortest distance is obtained between neurons 3 and 4, followed by the distances between neurons 2 and 4 and neurons 1 and 3.

The analysis above is complemented by looking at the weights of each neuron (Fig. 3), which basically contain the information about the rainfall anomaly field over the UPRB

from day $t - 5$ to day t . Neuron 1 has a north–south seesaw pattern at day $t - 5$ and progressively moves towards a homogeneous field, with a strong rainfall peak at day $t - 3$ centered in the northeastern part of the basin. The north–south dipole structure returns more strongly at day $t - 2$ and persists until day t , but now with a decrease in the rainfall peak. At this point it is worth mentioning that the negative anomalies in the rainfall field do not necessarily imply absence of rainfall, but just that the rainfall at that specific grid point is below its long-term monthly average. Neuron 2 starts with a nearly homogeneous rainfall field at times $t - 5$ and $t - 4$, from which negative rainfall anomalies start in the southern part and cover approximately the entire basin at time t . Neuron 3 starts at time $t - 5$ with a northeast–southwest dipole structure with positive anomalies in the southwest, which progress over time until almost the entire basin is covered by positive anomalies at time t . Neuron 4 has a homogeneous rainfall pattern over the entire basin, with negative anomalies from time $t - 5$ to time t .

Combining the information from Figs. 2 and 3, we observe that the rainfall field represented by neuron 4 is somehow connected to the rainfall patterns expressed by neurons 2 and

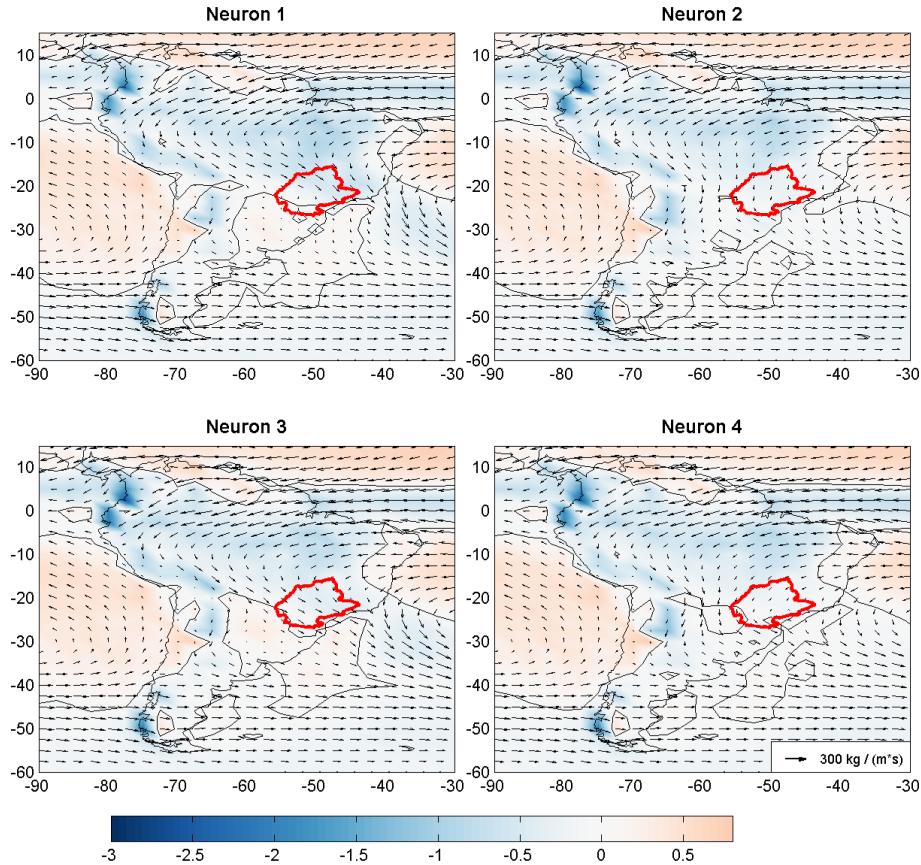


Figure 4. Vertically integrated moisture fluxes ($\text{kg}(\text{m s})^{-1}$, shown by the arrows) and associated divergence field ($10^{-4} \text{ kg}(\text{m}^2 \text{s})^{-1}$) averaged over each neuron class. The red contour line shows the Upper Paraná River basin. The contour for the divergence field equal to zero is also shown.

3 through specific regions of negative anomalies of rainfall. Neurons 2 and 3 also have some connections with neuron 1, expressed mainly by the rainfall field patterns at times $t - 4$ and $t - 3$ for neurons 2 and 1 and at times $t - 4$ to $t - 2$ for neurons 3 and 1, when positive rainfall anomalies occur in the southwestern portion of the basin.

Considering that each neuron represents a given state of the rainfall field during the course of 6 days, we estimate transition probabilities across the states and show them in Table 1. We note that there is a general tendency of the rainfall field to remain in its state (neuron), but the transition probabilities are different among neurons. Neuron 1 is more likely to transition to neuron 2, which is more likely to transition to neuron 4. Neuron 3 has the highest probability to transition to neuron 1, while neuron 4 will more likely stay at its own state, with just a small probability to transition to neuron 3. We further discuss and contextualize these transitions in the next section in which we analyze the atmospheric circulation associated with each neuron.

Table 1. Transition probabilities among neurons.

From neuron	To neuron			
	1	2	3	4
1	0.631	0.347	0.020	0.003
2	0.045	0.621	0.117	0.217
3	0.172	0.067	0.690	0.071
4	0.020	0.026	0.110	0.843

4.2 Atmospheric circulation, moisture transport, and sea surface temperature

The analysis of key atmospheric and ocean variables in each neuron class is conducted here through a composite analysis considering the days corresponding to each neuron class. In this sense, the patterns will reflect the average conditions (climatology) for days t through $t - 5$ as shown in Fig. 3.

The vertically integrated moisture flux and the associated divergence field (evaporation – precipitation along an atmospheric column) averaged over each neuron class are shown in Fig. 4. We can see this as a climatology of the mois-

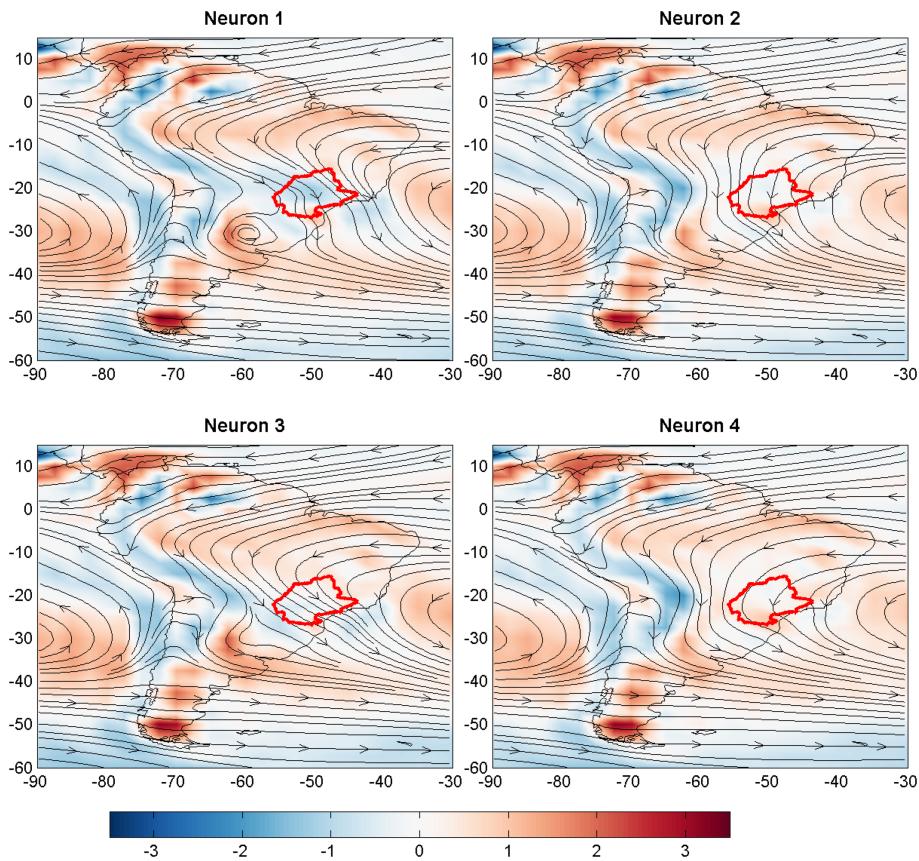


Figure 5. Streamlines for the vertically integrated moisture fluxes and low-level (850 mbar) relative vorticity (10^{-5} s^{-1}) averaged over each neuron class. The red contour line shows the Upper Paraná River basin.

ture transport associated with the rainfall patterns indicated in Fig. 3. Neuron 1 shows an intense moisture transport from the Amazon region, possibly associated with SALLJ episodes (Marengo et al., 2004). The divergence field is negative in the northern portion of the UPRB, suggesting intense rainfall along this region, and slightly positive in the southern part (dry conditions), extending to 50° S. This dipole structure has been reported in several studies (e.g., Nogués-Paegle and Mo, 1997; Díaz and Aceituno, 2003; Liebmann et al., 2004) and is also observed in the rainfall field associated with neuron 1 at time t (bottom panel of Fig. 3). The circulation is similar to the pattern described by Nogués-Paegle and Mo (1997), in which convection in the SACZ is enhanced and more likely to occur during El Niño episodes, while the SALLJ is weak (Liebmann et al., 2004; Silva and Berbery, 2006). However, note that a negative divergence field only indicates that rainfall can potentially occur. This is not a sufficient condition and a negative divergence does not necessarily lead to rainfall.

The moisture transport in neuron 2 is dominated by a north–south meridional flow crossing the entire basin, with a relatively homogeneous convergence of moisture over the basin, also resembling the rainfall pattern at time t for neu-

ron 2 (bottom panel of Fig. 3). This pattern seems to be associated with a weaker SACZ and stronger SALLJ, as described in Nogués-Paegle and Mo (1997) for positive events.

The moisture transport in neuron 3 is also affected by a strong SACZ and moisture fluxes from the Amazon region, but, when compared with neuron 1 (Fig. 4), the positive divergence (or inhibited precipitation) is far south of the basin and covers a smaller area. The moisture divergence pattern is again similar to the rainfall field at time t for neuron 3, with positive anomalies of rainfall covering almost the entire basin (bottom panel of Fig. 3). Neuron 4 has a moisture transport pattern similar to that of neuron 2, but the origin of the fluxes is more associated with the South Atlantic, with meridional fluxes west of the basin and a less intense moisture convergence. This reflects the rainfall field for neuron 4 (Fig. 3) and is likely associated with the average conditions of moisture transport into the region (Doyle and Barros, 2002; Carvalho et al., 2004).

The dynamics of the moisture transport associated with each neuron class is complemented by analyzing the low-level (850 mbar) relative vorticity (Fig. 5), which can indicate zones of low pressure and cyclonic rotation. A distinguished pattern is found for neuron 1, with negative relative vorticity

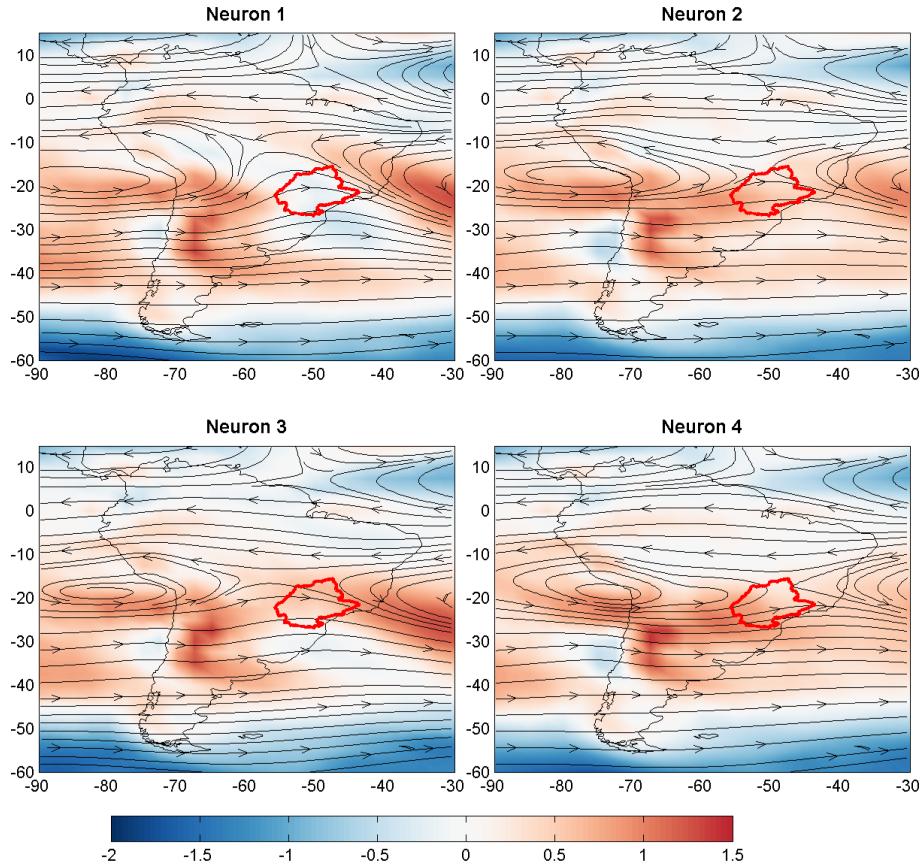


Figure 6. Streamlines for the high-level (500 mbar) wind vector and relative vorticity (10^{-5} s^{-1}) averaged over each neuron class. The red contour line shows the Upper Paraná River basin.

or cyclonic rotation over the entire basin and positive relative vorticity centered around $60^\circ \text{ W } 30^\circ \text{ S}$, which suggests upper-level wave activity and dynamical forcing (divergence in the upper levels) associated with neuron 1. This pattern has been identified in other studies (Liebmann et al., 1999; Robertson and Mechoso, 2000). Neuron 3 also shows cyclonic rotation (negative relative vorticity) in the southern part of the UPRB, extending up to 30° S . Neurons 2 and 4 do not show any sign of intense cyclonic flow over the basin.

The high-level (500 mbar) atmospheric circulation and relative vorticity associated with each neuron class is shown in Fig. 6. Neuron 1 shows a trough in the upper-level circulation that extends to the entire UPRB, with negative vorticity over the entire basin and positive vorticity southwest of it (centered around $45^\circ \text{ W } 30^\circ \text{ S}$). This pattern confirms our hypothesis that this neuron is also associated with upper-level wave activities. Neuron 3 also shows a trough over the basin, but it is weaker and negative vorticity slightly appears only in the south. Neurons 2 and 4 show more of a zonal type of circulation south of 20° S , which resembles the climatology of high-level circulation.

Anomalies in the near-surface air temperature associated with each neuron are shown in Fig. 7. Neurons 1 and 4 have,

respectively, negative and positive anomalies that cover the entire UPRB. Neuron 3 has a contrast of negative anomalies in the south and positive anomalies in the north, suggesting frontal activities. Neuron 2 also has a sharp contrast of anomalies but with opposite signs as compared with neuron 3, and the pattern suggests that it results from the advection of moist and warm air from the Amazon.

Potential SST persistent patterns associated with each neuron are analyzed here by passing a 15-day high-frequency filter on the daily SST anomalies, which are calculated by subtracting, from the daily SST, the average of the corresponding month for the January 1980–December 2013 period. The results are shown in Fig. 8. Neuron 1 and neuron 3 both show positive anomalies in the El Niño region (eastern tropical Pacific), in the central Pacific and tropical Atlantic. A dipole structure appears in both neurons along the southern coast of South America but they are out of phase. The negative SST anomalies off the South American coast associated with neuron 1 have been identified in other studies (Doyle and Barros, 2002) during SACZ activities. It is not clear whether they are a response to the reduced incoming radiation from the intense rainfall band that extends from the Amazon to the South Atlantic or whether they are in fact acting to produce

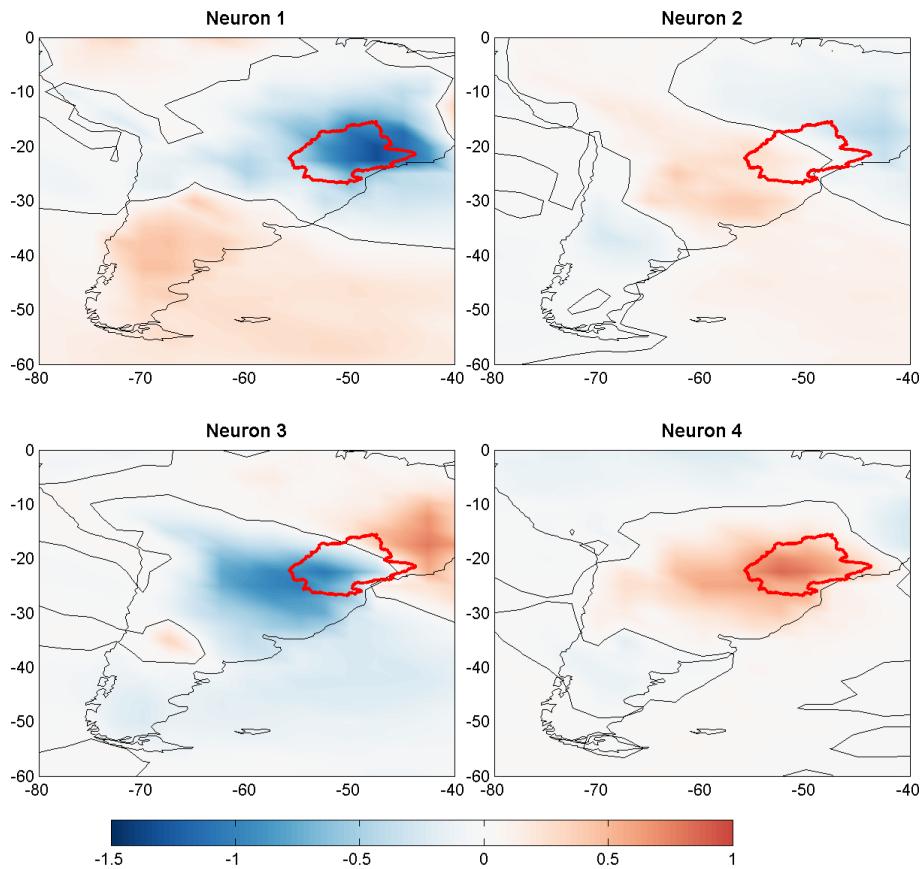


Figure 7. Anomalies ($^{\circ}\text{C}$) in the near-surface air temperature averaged over each neuron class. The red contour line shows the Upper Paraná River basin. The zero contour line is also shown.

the observed circulation pattern. The SST pattern of neuron 3 is similar to that of neuron 1, except that the anomalies off the South American coast near 30° S are positive, which is consistent with the positive rainfall anomalies in the southwestern part of the basin (Fig. 3), as described in Doyle and Barros (2002). Neurons 2 and 4 show a similar pattern along the South Atlantic, with positive anomalies in the equatorial region (up to about 20° S), negative in the subtropic (centered about 30° S), and positive south of 40° S. The SST pattern in the Pacific Ocean for neuron 2 is diffuse, with no remarkable features. Neuron 4 shows positive and negative anomalies that intercalate across the Pacific, with negative anomalies along the El Niño region. The SST anomalies in the Atlantic for neuron 2 are very similar to those observed for neuron 1.

Combining the preceding analyses, we can shed some light on the transition probabilities, hits, and connectivity among neurons as displayed in Fig. 2 and Table 1. Neuron 4 individually has the most hits and likely reflects the average circulation during the wet season, with a strong persistence but reduced SACZ activities. Eventually it precedes neuron 3 (probability = 11 %) and most likely succeeds neuron 2 (probability = 22 %), which is expected given the rainfall pattern, as shown in Fig. 3, and the atmospheric circula-

tion and SST anomalies in Figs. 4 to 8. Neuron 2 also has a slight probability (12 %) to precede neuron 3 and most likely (probability = 35 %) succeeds neuron 1.

When we connect these results with the transition probabilities in Table 1, we can describe the most probable sequence of rainfall states. The dynamical forcing and active SACZ of neuron 1 is most likely preceded by neuron 3 (probability = 17 %), which is marked by active SACZ, high-level waves, and cold fronts, and will most likely be followed by the rainfall pattern of neuron 2 (probability = 35 %), which is coherent with the surface air temperature march as inferred from Fig. 7. Neuron 4 will most likely be followed by neuron 3 (Fig. 3). Neurons 1 and 4 are not connected and the transition probabilities between them are practically zero. In summary, the most likely sequence of neuron transitions, arbitrarily starting at neuron 3, is $3 \rightarrow 1 \rightarrow 2 \rightarrow 4 \rightarrow 3$. However, also note that transition probabilities from one neuron to another are generally smaller than the probabilities to remain in the state (see Table 1).

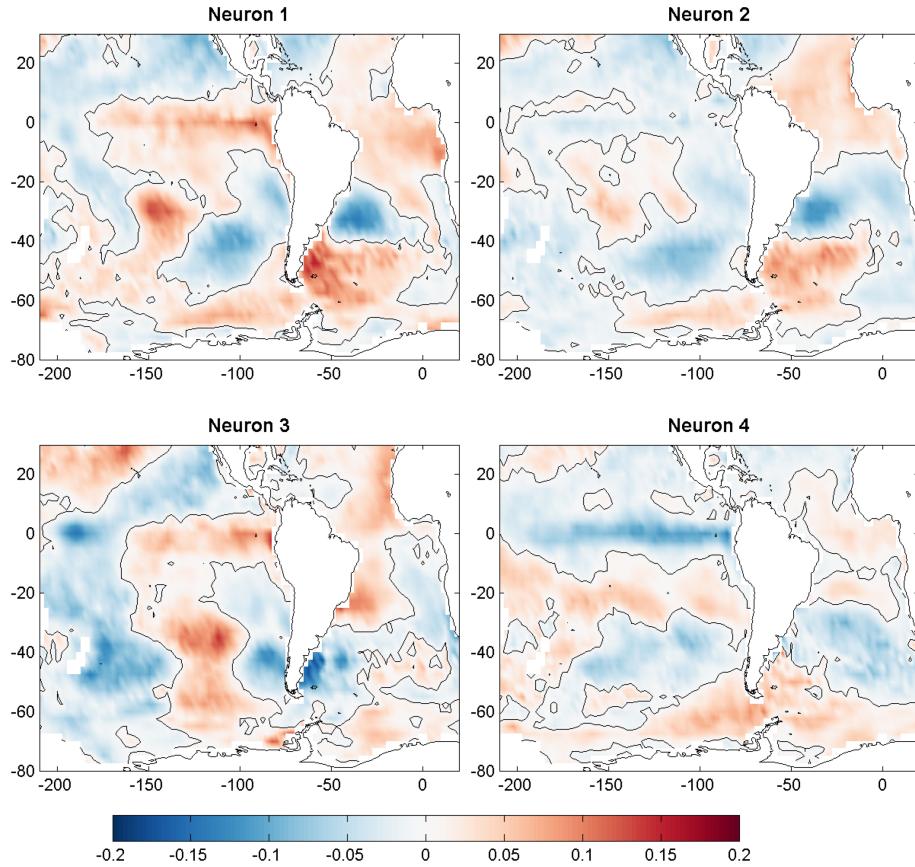


Figure 8. Showing 15-day filtered sea surface temperature (SST) anomalies ($^{\circ}\text{C}$) averaged over each neuron class. The zero contour line is also shown.

4.3 Flood response

4.3.1 Frequency and magnitude

The total proportion of flood events in neurons 1 to 4 is equal to 35, 34, 20, and 11 %, respectively. The frequency of floods in each neuron for the streamflow gauges analyzed here is shown in Fig. 9. Neurons 1 and 2 dominate most floods across the UPRB. Neuron 3 dominates the floods along the gauges located in the Paranapanema subbasin (see Fig. 1), while neuron 4 is most associated with floods in the gauges along the Paraná River, particularly with the Itaipu gauge located in the basin outlet, which interestingly is not directly affected by the wave activity of neuron 1 (see following discussion).

The magnitude of floods associated with each neuron class is analyzed by calculating, for each site, the empirical exceedance probability for each data point in the partial duration series, aggregating all estimates across the sites and then estimating the density of such probabilities conditional on the neuron class of the data points. The results are shown in Fig. 10. Neurons 1 and 2 have the peak and largest density in small values of exceedance probability, suggesting that the biggest floods along the UPRB are associated with these pat-

terns of rainfall (Fig. 3) and moisture transport and convergence (Fig. 4). It is worth mentioning that neuron 2 has a rainfall dynamics that is not associated with El Niño events (Fig. 8) but still produces large floods. This highlights the uncertainty and complexity of quantifying the flood hazard during El Niño–Southern Oscillation events, as described in Emerton et al. (2017). The pattern of neuron 3 is more associated with intermediate magnitude flood events while neuron 4 is associated with the smallest flood events, although some large flood events are possible, particularly at the sites where this neuron dominates the frequency of occurrence (Fig. 9).

4.3.2 Spatial scaling

The literature on the scale of flood properties (e.g. quantiles) with drainage area (Gupta and Waymire, 1990; Farquharson et al., 1992; Over and Gupta, 1994; Gupta et al., 1994, 2007; Gupta and Dawdy, 1995; Pandey et al., 1998) suggests that the type of precipitation (e.g. convective versus frontal) and the attributes of the drainage network will jointly determine the different behaviors of the scaling process of flow and drainage area. It is not clear whether such scaling relations

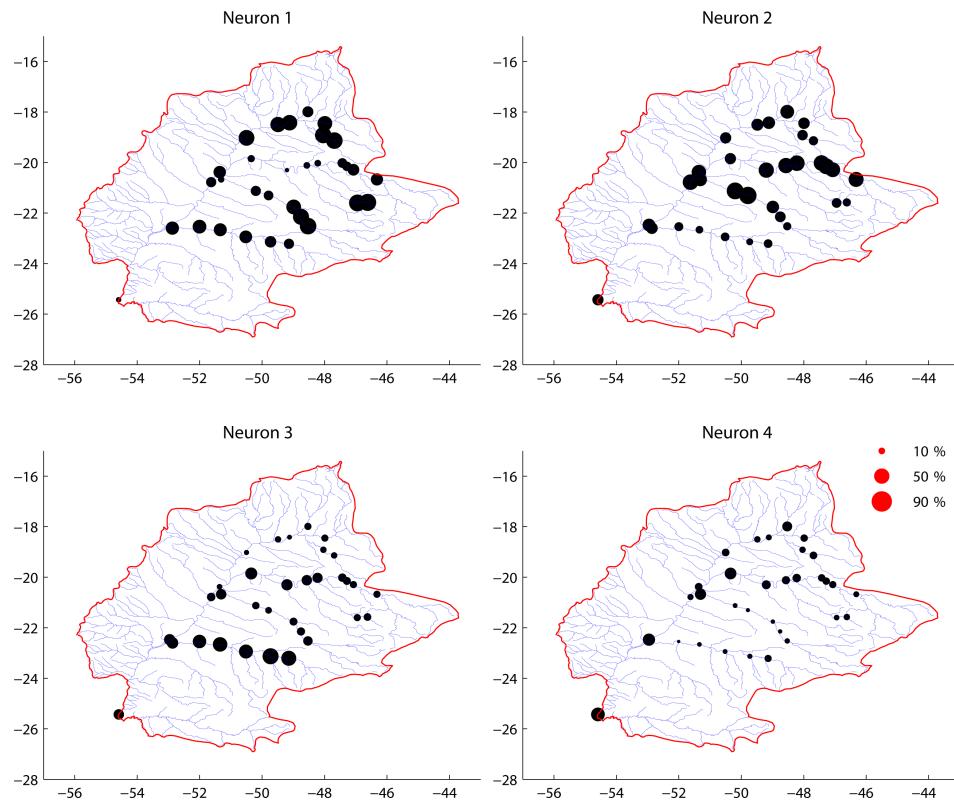


Figure 9. Frequency of flood events in each neuron class for each streamflow gauge. The red dots show the scale for frequencies of 10, 50, and 90 %.

will hold if a mixture of mechanisms can interact to produce large floods. Here we explore the scaling of the first and second sample moments of the flood events with respect to the neuron classes.

Since each flood event at a given site can be assigned to a neuron class, we can easily calculate the sample moments (mean and variance in our case) in each neuron class for each gauge and evaluate how the scaling law of flow moments and drainage area change as a function of the spatiotemporal variability in the rainfall field. Figure 11 shows the scaling of the average flood flow and drainage area for each neuron class. The magnitudes of the slope and intercept coefficients clearly change as a function of the neuron class, but more remarkable differences appear between neurons 1–2 and neurons 3–4. In fact, both slope and intercept estimates of either neurons 1 or 2 are significantly different at the 5 % significance level from the estimates for neurons 3 and 4 using a standard Student's *t* test.

The magnitude of these coefficients also reflects the intensity of rainfall and the spatial pattern associated with each neuron, as shown in Fig. 3. As the rainfall intensity increases, it is expected that the intercept will increase, while the slope is more related to the spatial homogeneity of the rainfall field: as it becomes more homogeneous across the basin, we expect the slope will approach 1. The intercept values as shown in

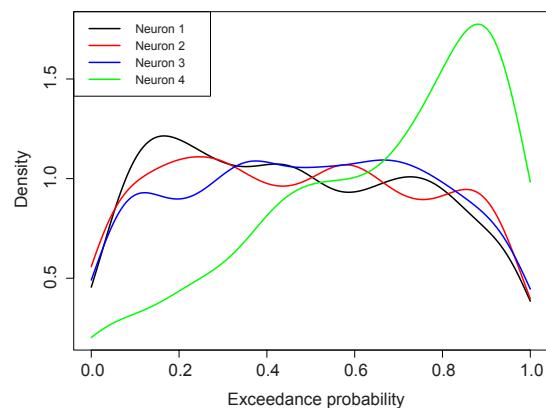


Figure 10. Density of exceedance probabilities in each neuron class.

Fig. 11 increase from neuron 4 to neuron 1, which qualitatively agrees with the rainfall patterns shown in Fig. 3, whose overall magnitude increases from neuron 4 to 1. The slope estimates suggest that the less homogeneous rainfall fields occur in neurons 1 and 2, which is consistent with the pattern displayed in Fig. 3. Neurons 3 and 4 have the largest slope estimates and thus a more homogeneous rainfall field, which is again consistent with the results of Fig. 3.

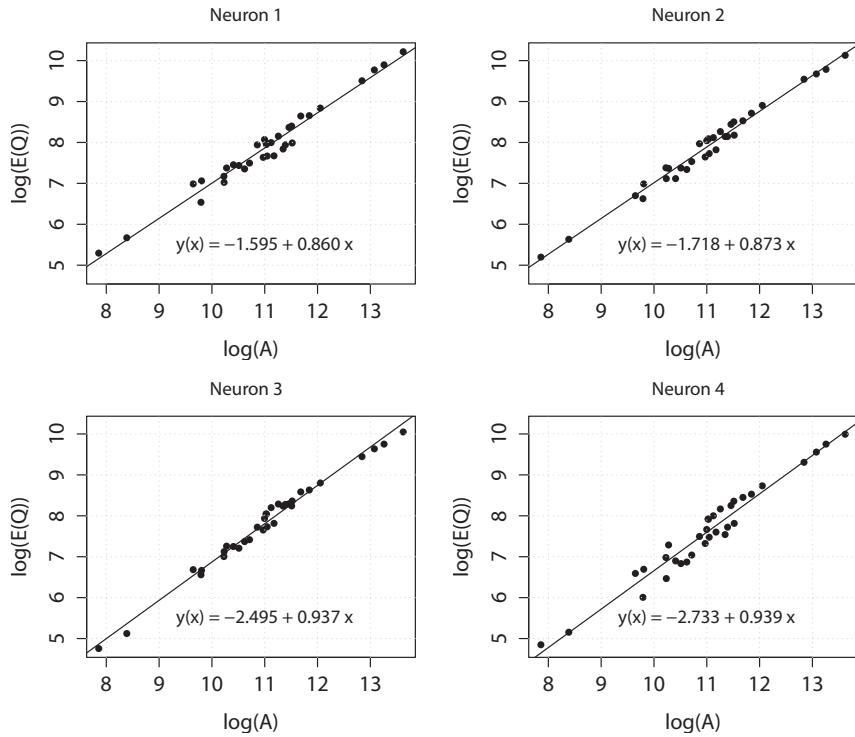


Figure 11. Scaling of average flood flow series in each neuron class. The least-square estimates of intercept and slope are shown in each panel. The black line shows the least-squares regression. Here Q represents the flood series while A represents the drainage area (km^2) of the respective catchment.

The scaling of the sample variance with the drainage area for each neuron class is shown in Fig. 12. As for the average flow scaling, the largest differences among the coefficients are observed between the pair of neurons 1 and 2 and the pair of neurons 3 and 4. Visually, the scaling is clearer for neurons 1 and 2. Neuron 4 shows more dispersed values along the least-squares regression line, suggesting that the mechanisms by which this rainfall pattern produces a given flood across the gauges, particularly for small gauges, are different (see subsequent discussion).

4.3.3 Flood event synchronization

Figure 13 shows a directed network obtained from the adjacency matrix \mathbf{A} and the delay behavior matrix \mathbf{q}_τ considering all flood events across sites and not taking into account the neuron classes. The nodes represent the streamflow gauges in their geographical position while the edges (or branches) represent the existence of synchronization between two sites. The arrow shows the dominant direction of the flood propagation. The flow patterns generally follow the drainage basin direction (Fig. 1): east–west and north–south. However, some exceptions can also be observed, indicating that the size and movement of the storm path may also affect how the sites are synchronized.

If we cluster the flood events into the neuron classes, we can obtain specific adjacency and delay behavior matrices for each neuron. The resulting directed networks are shown in Fig. 14. Now we can observe that the rainfall pattern described by neuron 1 produces the largest synchronization of flood events, given by the number of arrows, including inter- and intra-subbasin connectivity. In general, the cascade of flood events tends to end up in the outlet of the subbasins (see Fig. 1 for the name and location of the subbasins). Neuron 2 has a more intra-subbasin connectivity pattern, which tends to follow the river flow direction and suggest that rainfall upstream of the basin is the more likely cause of floods. The Itaipu site located in the basin outlet is not connected to any site, suggesting that Itaipu floods in this neuron will likely result from the routing flow from upstream sites. Neuron 3 has the northern sites disconnected while a connectivity within and across subbasins is observed. The Tietê subbasin seems to be disconnected from all other subbasins. Finally, neuron 4 shows less connections, and most of them are within the subbasins. The Itaipu site is again completely disconnected; thus, most of its floods associated with neuron 4 are due to routing of upstream flow and floods caused by rainfall of this and other types.

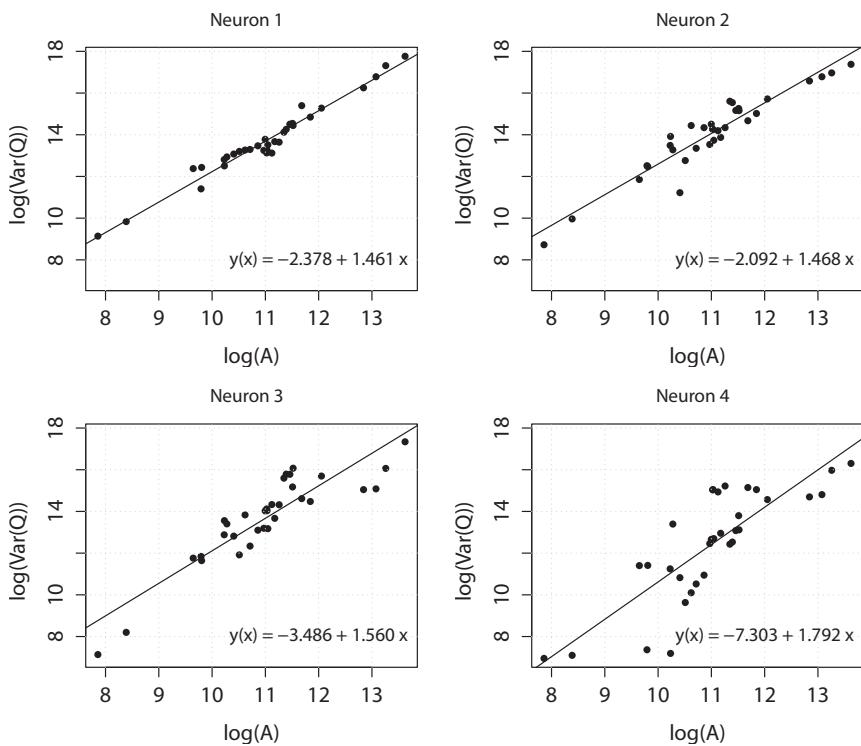


Figure 12. Scaling of variance of flood flow series in each neuron class. The least-square estimates of intercept and slope are shown in each panel. The black line shows the least-squares regression. Here Q represents the flood series while A represents the drainage area (km^2) of the respective catchment.

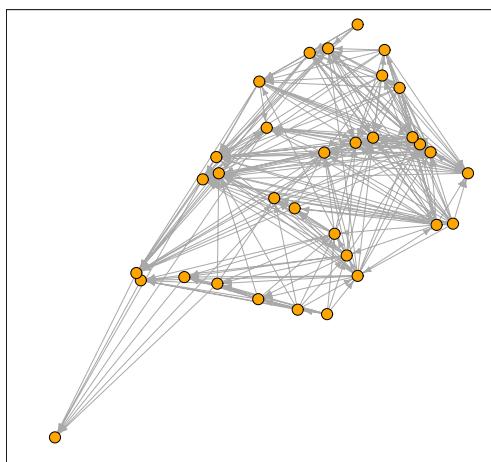


Figure 13. A directed network for the flood events showing synchronization and flow direction (arrows). The dots show the streamflow gauges in their geographical location (see Fig. 1).

5 Summary and conclusions

A general, statistical approach to classifying flood generation mechanisms, the areal scaling of floods, and the synchronization potential of flooding in a large river basin was developed and demonstrated with data from the Upper Paraná

River basin, Brazil. This is the first attempt to describe such floods in a broad, hydroclimatic context. A self-organizing map algorithm was employed to find the spatiotemporal dynamics of the rainfall field over the basin in the days that preceded the major flood events. For each cluster, we analyzed the large-scale moisture transport into the region as well as the upper-level structure and teleconnections associated with SST. The flood response associated with each rainfall pattern was evaluated in terms of magnitude, frequency, spatial scaling, and event synchronization.

Four distinct patterns of rainfall were observed and associated with the atmospheric circulation and moisture transport. The first cluster exhibits strong rainfall concentrated in the northeastern part of the basin, with a peak 2 days before the flood events. It was associated with the moisture transport from the Amazon and intense SACZ, with the presence of cyclones – a pattern that has also been reported in the literature (Liebmann et al., 1999; Robertson and Mechoso, 2000). These events are associated with positive SST anomalies in the tropical Pacific and Atlantic oceans and a dipole structure off the eastern coast of South America, which has also been observed in other studies (e.g., Doyle and Barros, 2002). On average, 35 % of all floods happen during these conditions. The Itaipu streamflow gauge located in the basin outlet is

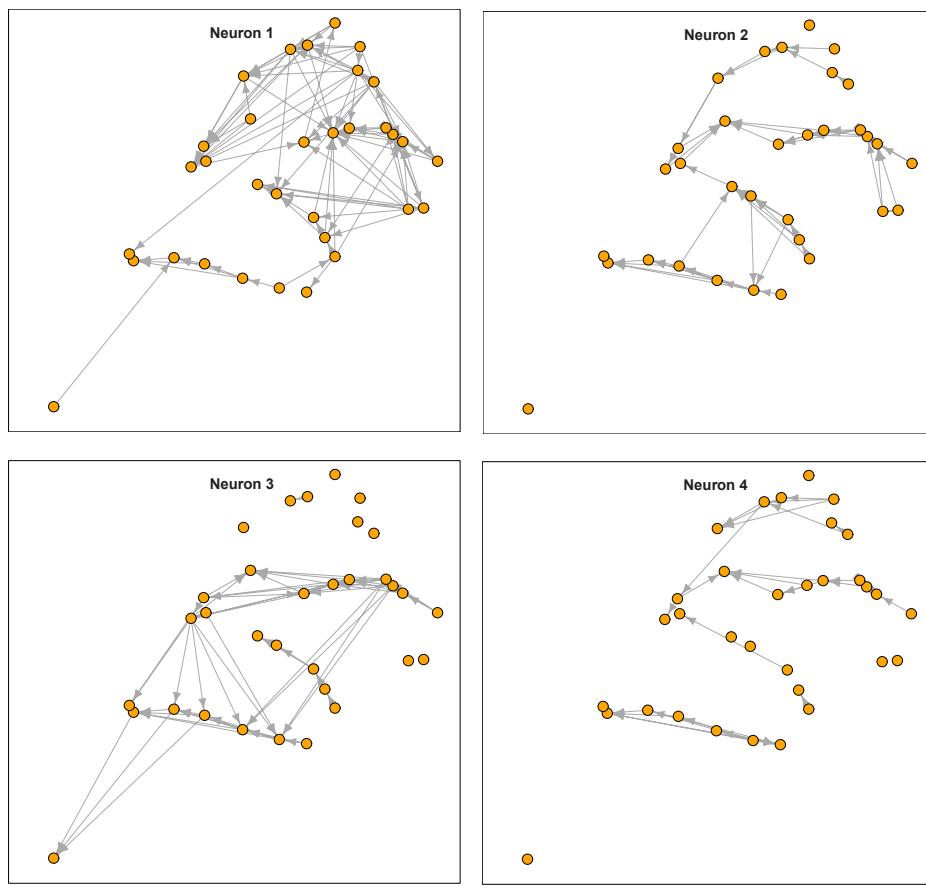


Figure 14. A directed network for the flood events showing synchronization and flow direction (arrows) as a function of neuron class. The dots show the streamflow gauges in their geographical location (see Fig. 1).

less affected, at least directly, by this rainfall pattern. These types of floods are strongly synchronized across all sites.

The third neuron shows features of SACZ episodes associated with extratropical disturbances, possibly fronts and cyclones. The rainfall field is, however, less intense than that of neuron 1 and peaks in the southwestern part of the basin. The composite analysis for the SST field has a pattern similar to that of neuron 1, but the seesaw structure off the eastern South American coast is reversed. On average, 20 % of the floods happen to occur in neuron 3, but this frequency is larger for sites located in the southern part of the basin, particularly in the Paranapanema subbasin. The magnitude of these types of floods is intermediate and there is a synchronization within and across the central and southern subbasins, suggesting connectivity due to the storm track extension and movement and the flood routing along the stream channels. Both neurons 1 and 3 have positive SST anomalies in the El Niño–Southern Oscillation region (eastern tropical Pacific), which has also been associated with extreme rainfall events in the region (Camilloni and Barros, 2003; Grimm and Tedeschi, 2009; Cavalcanti et al., 2015). Therefore, about 55 % of floods (i.e., 35 % in neuron 1 + 20 % in neuron 3)

in the UPRB are linked to El Niño-like SST patterns in the eastern tropical Pacific. Neuron 2 has a rainfall peak in the northeastern part of the basin, between 4 and 5 days before the flood event. The average rainfall field is less intense than neuron 1 but more intense than neuron 3. Neuron 4 shows a SST pattern similar to La Niña-like conditions, with negative SST anomalies in the eastern tropical Pacific, suggesting that about 11 % of floods in the UPRB could be linked with this large-scale phenomenon.

The moisture path shows warm and moist meridional flow across the entire basin, resulting in rainfall possibly due to low-level convergence or eventually frontal activity. The SST field in the Atlantic Ocean is similar to that of neuron 1, but the average conditions in the tropical Pacific are neutral. On average, 34 % of floods are of this type, particularly in the northern subbasins. Together with floods in neuron 1, these are the largest floods in the region. The synchronization of type 3 floods is more intra-subbasin. Finally, type 4 floods are caused by a homogeneous but persistent rainfall field, with most moisture transported from the Atlantic Ocean. There is no evidence of directly extratropical activities and the SST field revealed negative anomalies in the tropical Pacific and

positive in the tropical Atlantic. The near-surface air temperature in this cluster showed positive anomalies, suggesting that local convection might also be an important factor. Although this is the dominant pattern of rainfall, 11 % of the total floods are of these type. These are the less intense floods, with a synchronization that occurs along the main river channels.

The spatial scaling exponents (slope) of floods with drainage area (Figs. 11 and 12) are similar for floods of types 1 and 2, and for types 3 and 4, even though the rainfall mechanisms are different for each pair. The exponent is higher for types 3 and 4, reflecting the higher homogeneity in the rainfall and response pattern. The area exponents for flood variance are considerably higher than those for mean scaling, opening the possibility of a multi-scaling approach. However, once again the exponents are similar for types 1 and 2 and for types 3 and 4. The scaling relationships for variance are not as well constrained for neurons 3 and 4 types of events.

Distinct patterns of flood synchronization and movement are also identified for each neuron. Conditional on the storm track, i.e., large-scale atmospheric flow, these could be further useful to improve analysis and prediction of the potential flood emergence and for the operation of multi-stage flood control systems.

The results obtained in this work are a step forward for flood risk management in the UPRB in two possible ways: flood design and short-term prediction. Local flood frequency analysis could make use of the different flood categories and employ, for instance, a mixture-of-distributions approach (e.g., Alila and Mtiraoui, 2002) for better flood quantile estimates. Regional flood frequency analysis could also consider the different scaling laws and develop a Bayesian approach (as in Lima and Lall, 2010; Cheng et al., 2014; Lima et al., 2016) to better estimate regional parameters.

Finally, the persistent regions with SST anomalies could be used to derive climate predictors for short-term flood risk prediction. The synchronization of the flood events could be explored in more detail to develop short-term flood forecast models conditional on the atmospheric and ocean states and flood situation at nearby sites. Further details of the moisture transport and high-level atmospheric circulation could also be analyzed in order to obtain potential climate predictors for the floods in this region. The proposed method can potentially be used to explore other attributes of floods, the notion of cumulative hazards (Moftakhari et al., 2017), and simultaneous flooding across a basin (Vahedifard et al., 2016), and will be the theme of our future work. The timing of the floods along the wet season and a possible association with the neuron classes can be further explored. In future research, we intend to address some of these topics and also explore how the tools and methodology employed in this work could help evaluate the future flood risk in the UPRB region considering climate changes.

Data availability. The streamflow data for the Upper Paraná River basin are provided by the Brazilian Electric System National Operator (ONS) and can be accessed at <http://www.ons.org.br/>. The rainfall and temperature data are provided by Xavier et al. (2016) and can be accessed at <http://careyking.com/data-download/>. The ERA-Interim global data set (SST, moisture fluxes, divergence field, vorticity, wind field) is available at <http://apps.ecmwf.int/datasets/data/interim-full-mod4/levtype=sfc/>.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Hydro-climate dynamics, analytics and predictability”. It is not associated with a conference.

Acknowledgements. We thank all agencies and authors that provided data sets and codes. The first author acknowledges a postdoctoral fellowship from the Brazilian Government Agency CNPq during part of this work. The second author was partially supported by the NSF award no. EAR-1316536.

Edited by: Julia Hall

Reviewed by: Eduardo Martins, Shaun Harrigan, and one anonymous referee

References

- Alila, Y. and Mtiraoui, A.: Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques, *Hydrol. Process.*, 16, 1065–1084, <https://doi.org/10.1002/hyp.346>, 2002.
- Amengual, A., Romero, R., Gómez, M., Martín, A., and Alonso, S.: A Hydrometeorological Modeling Study of a Flash-Flood Event over Catalonia, Spain, *J. Hydrometeorol.*, 8, 282–303, <https://doi.org/10.1175/JHM577.1>, 2007.
- Antico, A., Torres, M. E., and Diaz, H. F.: Contributions of different time scales to extreme Paraná floods, *Clim. Dynam.*, 46, 3785–3792, <https://doi.org/10.1007/s00382-015-2804-x>, 2016.
- Bao, M. and Wallace, J. M.: Cluster Analysis of Northern Hemisphere Wintertime 500-hPa Flow Regimes during 1920–2014, *J. Atmos. Sci.*, 72, 3597–3608, <https://doi.org/10.1175/JAS-D-15-0001.1>, 2015.
- Bárdossy, A. and Filiz, F.: Identification of flood producing atmospheric circulation patterns, *J. Hydrol.*, 313, 48–57, <https://doi.org/10.1016/j.jhydrol.2005.02.006>, 2005.
- Barros, V., Gonzalez, M., Liebmann, B., and Camilloni, I.: Influence of the South Atlantic convergence zone and South Atlantic Sea surface temperature on interannual summer rainfall variability in Southeastern South America, *Theor. Appl. Climatol.*, 67, 123–133, <https://doi.org/10.1007/s007040070002>, 2000.
- Berbery, E. H. and Barros, V. R.: The Hydrologic Cycle of the La Plata Basin in South America, *J. Hydrometeorol.*, 3, 630–645, [https://doi.org/10.1175/1525-7541\(2002\)003<0630:THCOTL>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0630:THCOTL>2.0.CO;2), 2002.

- Boers, N., Bookhagen, B., Barbosa, H. M. J., Marwan, N., Kurths, J., and Marengo, J. A.: Prediction of extreme floods in the eastern Central Andes based on a complex networks approach, *Nat. Commun.*, 5, 5199, <https://doi.org/10.1038/ncomms6199>, 2014.
- Budikova, D., Coleman, J. S. M., Strope, S. A., and Austin, A.: Hydroclimatology of the 2008 Midwest floods, *Water Resour. Res.*, 46, w12524, <https://doi.org/10.1029/2010WR009206>, 2010.
- Camilloni, I. A. and Barros, V. R.: Extreme discharge events in the Paraná River and their climate forcing, *J. Hydrol.*, 278, 94–106, [https://doi.org/10.1016/S0022-1694\(03\)00133-1](https://doi.org/10.1016/S0022-1694(03)00133-1), 2003.
- Cardoso, A. O. and Silva Dias, P. L.: The relationship between ENSO and Paraná River flow, *Adv. Geosci.*, 6, 189–193, <https://doi.org/10.5194/adgeo-6-189-2006>, 2006.
- Carvalho, L., Jones, C., and Liebmann, B.: The south atlantic convergence zone: intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall, *J. Climate*, 17, 88–108, [https://doi.org/10.1175/1520-0442\(2004\)017<0088:TSACZI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0088:TSACZI>2.0.CO;2), 2004.
- Carvalho, L. M. V., Silva, A. E., Jones, C., Liebmann, B., Dias, P. L. S., and Rocha, H. R.: Moisture transport and intraseasonal variability in the South America monsoon system, *Clim. Dynam.*, 36, 1865–1880, <https://doi.org/10.1007/s00382-010-0806-2>, 2011.
- Cavalcanti, I., Carril, A., Penalba, O., Grimm, A., Menéndez, C., Sanchez, E., Cherchi, A., Sorensson, A., Robledo, F., Rivera, J., Pántano, V., Bettolli, L., Zaninelli, P., Zamboni, L., Tedeschi, R., Dominguez, M., Ruscica, R., and Flach, R.: Precipitation extremes over La Plata Basin – Review and new results from observations and climate simulations, *J. Hydrol.*, 523, 211–230, <https://doi.org/10.1016/j.jhydrol.2015.01.028>, 2015.
- Cavazos, T.: Using Self-Organizing Maps to Investigate Extreme Climate Events: An Application to Wintertime Precipitation in the Balkans, *J. Climate*, 13, 1718–1732, [https://doi.org/10.1175/1520-0442\(2000\)013<1718:USOMTI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<1718:USOMTI>2.0.CO;2), 2000.
- Chaves, R. R. and Nobre, P.: Interactions between sea surface temperature over the South Atlantic Ocean and the South Atlantic Convergence Zone, *Geophys. Res. Lett.*, 31, l03204, <https://doi.org/10.1029/2003GL018647>, 2004.
- Cheng, L., AghaKouchak, A., Gilletland, E., and Katz, R. W.: Non-stationary extreme value analysis in a changing climate, *Climatic Change*, 127, 353–369, <https://doi.org/10.1007/s10584-014-1254-5>, 2014.
- Corfidi, S. F., Merritt, J. H., and Fritsch, J. M.: Predicting the movement of mesoscale convective complexes, *Weather Forecast.*, 11, 41–46, 1996.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- Díaz, A. and Aceituno, P.: Atmospheric Circulation Anomalies during Episodes of Enhanced and Reduced Convective Cloudiness over Uruguay, *J. Climate*, 16, 3171–3185, [https://doi.org/10.1175/1520-0442\(2003\)016<3171:ACADEO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<3171:ACADEO>2.0.CO;2), 2003.
- Doyle, M. E. and Barros, V. R.: Midsummer Low-Level Circulation and Precipitation in Subtropical South America and Related Sea Surface Temperature Anomalies in the South Atlantic, *J. Climate*, 15, 3394–3410, [https://doi.org/10.1175/1520-0442\(2002\)015<3394:MLLCAP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3394:MLLCAP>2.0.CO;2), 2002.
- Drumond, A., Nieto, R., Gimeno, L., and Ambrizzi, T.: A Lagrangian identification of major sources of moisture over Central Brazil and La Plata Basin, *J. Geophys. Res.*, 113, D14128, <https://doi.org/10.1029/2007JD009547>, 2008.
- Durkee, J. D. and Mote, T. L.: A climatology of warm-season mesoscale convective complexes in subtropical South America, *Int. J. Climatol.*, 30, 418–431, <https://doi.org/10.1002/joc.1893>, 2009.
- Durkee, J. D., Mote, T. L., and Shepherd, M.: The Contribution of Mesoscale Convective Complexes to Rainfall across Subtropical South America, *J. Climate*, 22, 4590–4605, <https://doi.org/10.1175/2009JCLI2858.1>, 2009.
- Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J., and Pappenberger, F.: Complex picture for likelihood of ENSO-driven flood hazard, *Nat. Commun.*, 8, 14796, <https://doi.org/10.1038/ncomms14796>, 2017.
- Farquharson, F. A. K., Meigh, J. R., and Sutcliffe, J. V.: Regional Flood Frequency-Analysis in Arid and Semiarid Areas, *J. Hydrol.*, 138, 487–501, 1992.
- Grimm, A.: The El Niño impact on the summer monsoon in Brazil: regional processes versus remote influences, *J. Climate*, 16, 263–280, [https://doi.org/10.1175/1520-0442\(2003\)016<0263:TENIOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0263:TENIOT>2.0.CO;2), 2003.
- Grimm, A.: How do La Niña events disturb the summer monsoon system in Brazil?, *Clim. Dynam.*, 22, 123–138, <https://doi.org/10.1007/s00382-003-0368-7>, 2004.
- Grimm, A., Ferraz, S. E. T., and Gomes, J.: Precipitation Anomalies in Southern Brazil Associated with El Niño and La Niña Events, *J. Climate*, 11, 2863–2880, [https://doi.org/10.1175/1520-0442\(1998\)011<2863:PAISBA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2863:PAISBA>2.0.CO;2), 1998.
- Grimm, A. M. and Tedeschi, R. G.: ENSO and Extreme Rainfall Events in South America, *J. Climate*, 22, 1589–1609, <https://doi.org/10.1175/2008JCLI2429.1>, 2009.
- Grimm, A. M., Barros, V. R., and Doyle, M. E.: Climate Variability in Southern South America Associated with El Niño and La Niña Events, *J. Climate*, 13, 35–58, [https://doi.org/10.1175/1520-0442\(2000\)013<0035:CVISSA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0035:CVISSA>2.0.CO;2), 2000.
- Gupta, V. K. and Dawdy, D. R.: Physical interpretations of regional variations in the scaling exponents of flood quantiles, *Hydrolog. Process.*, 9, 347–361, <https://doi.org/10.1002/hyp.3360090309>, 1995.
- Gupta, V. K. and Waymire, E.: Multiscaling Properties of Spatial Rainfall and River Flow Distributions, *J. Geophys. Res.*, 95, 1999–2009, <https://doi.org/10.1029/JD095iD03p01999>, 1990.
- Gupta, V. K., Mesa, O. J., and Dawdy, D. R.: Multiscaling theory of flood peaks: Regional quantile analysis, *Water Resour. Res.*, 30, 3405–3421, <https://doi.org/10.1029/94WR01791>, 1994.
- Gupta, V. K., Troutman, B. M., and Dawdy, D. R.: Towards a Nonlinear Geophysical Theory of Floods in River Networks: An Overview of 20 Years of Progress, in: *Nonlinear dynamics*

- in Geosciences, edited by: Tsonis, A. and Elsner, J., 121–151, Springer, New York, USA, 2007.
- Hewitson, B. C. and Crane, R. G.: Self-organizing maps: applications to synoptic climatology, *Clim. Res.*, 22, 13–26, <https://doi.org/10.3354/cr022013>, 2002.
- Hirschboeck, K. K.: Flood Hydroclimatology, in: Flood Geomorphology, edited by: Baker, V., Kochel, R. C., and Patton, P. C., John Wiley & Sons, Inc, 1988.
- Hirschboeck, K. K., Ely, L., and Maddox, R. A.: Hydroclimatology of meteorologic floods, in: Inland Flood Hazards: Human, Riparian and Aquatic Communities, edited by: Wohl, E., Cambridge University Press, Cambridge, UK, 2000.
- Jacobbeit, J., Glaser, R., Luterbacher, J., and Wanner, H.: Links between flood events in central Europe since AD 1500 and large-scale atmospheric circulation modes, *Geophys. Res. Lett.*, 30, 1172, <https://doi.org/10.1029/2002GL016433>, 2003.
- Jain, S. and Lall, U.: Floods in a changing climate: Does the past represent the future?, *Water Resour. Res.*, 37, 3193–3205, <https://doi.org/10.1029/2001WR000495>, 2001.
- Johnson, N. C., Feldstein, S. B., and Tremblay, B.: The Continuum of Northern Hemisphere Teleconnection Patterns and a Description of the NAO Shift with the Use of Self-Organizing Maps, *J. Climate*, 21, 6354–6371, <https://doi.org/10.1175/2008JCLI2380.1>, 2008.
- Jones, C. and Carvalho, L. M. V.: Active and Break Phases in the South American Monsoon System, *J. Climate*, 15, 905–914, [https://doi.org/10.1175/1520-0442\(2002\)015<0905:AABPIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0905:AABPIT>2.0.CO;2), 2002.
- Jorgetti, T., da Silva Dias, P. L., and de Freitas, E. D.: The relationship between South Atlantic SST and SACZ intensity and positioning, *Clim. Dynam.*, 42, 3077–3086, <https://doi.org/10.1007/s00382-013-1998-z>, 2014.
- Kahana, R., Ziv, B., Enzel, Y., and Dayan, U.: Synoptic climatology of major floods in the Negev Desert, Israel, *Int. J. Climatol.*, 22, 867–882, <https://doi.org/10.1002/joc.766>, 2002.
- Knox, J. C.: Large Increases in Flood Magnitude in Response to Modest Changes in Climate, *Nature*, 361, 430–432, 1993.
- Kohonen, T.: Self-Organizing Maps, Vol. 30, Springer-Verlag Berlin Heidelberg, 3rd Edn., <https://doi.org/10.1007/978-3-642-56927-2>, 2001.
- Kroll, C. N. and Stedinger, J. R.: Regional Hydrologic Analysis: Ordinary and generalized least squares revisited, *Water Resour. Res.*, 34, 121–128, <https://doi.org/10.1029/97WR02685>, 1998.
- Kunkel, K. E., Changon, S. A., and Angel, J. R.: Climatic Aspects of the 1993 Upper Mississippi River Basin Flood, *B. Am. Meteorol. Soc.*, 75, 811–822, [https://doi.org/10.1175/1520-0477\(1994\)075<0811:CAOTUM>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<0811:CAOTUM>2.0.CO;2), 1994.
- Kwon, H.-H., Brown, C., and Lall, U.: Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling, *Geophys. Res. Lett.*, 35, L05404, <https://doi.org/10.1029/2007GL032220>, 2008.
- Lang, M., Ouarda, T. B. M. J., and Bobée, B.: Towards operational guidelines for over-threshold modeling, *J. Hydrol.*, 225, 103–117, [https://doi.org/10.1016/S0022-1694\(99\)00167-5](https://doi.org/10.1016/S0022-1694(99)00167-5), 1999.
- Lee, J. A. and Verleysen, M.: Nonlinear Dimensionality Reduction, Springer-Verlag New York, <https://doi.org/10.1007/978-0-387-39351-3>, 2007.
- Lee, S. and Feldstein, S. B.: Detecting ozone- and greenhouse gas-driven wind trends with observational data, *Science*, 339, 563–567, <https://doi.org/10.1126/science.1225154>, 2013.
- Li, D., Bou-Zed, E., Baeck, M. L., Jessup, S. M., and Smith, J. A.: Modeling Land Surface Processes and Heavy Rainfall in Urban Environments: Sensitivity to Urban Surface Representations, *J. Hydrometeorol.*, 14, 1098–1118, <https://doi.org/10.1175/JHM-D-12-0154.1>, 2013.
- Li, J., Sharma, A., Evans, J., and Johnson, F.: Addressing the mischaracterization of extreme rainfall in regional climate model simulations – A synoptic pattern based bias correction approach, *J. Hydrol.*, <https://doi.org/10.1016/j.jhydrol.2016.04.070>, online first, 2016.
- Li, X., Li, C., Ling, J., and Tan, Y.: The Relationship between Contiguous El Niño and La Niña Revealed by Self-Organizing Maps, *J. Climate*, 28, 8118–8134, <https://doi.org/10.1175/JCLI-D-15-0123.1>, 2015.
- Liebmann, B., Kiladis, G. N., Marengo, J., Ambrizzi, T., and Glick, J. D.: Submonthly Convective Variability over South America and the South Atlantic Convergence Zone, *J. Climate*, 12, 1877–1891, [https://doi.org/10.1175/1520-0442\(1999\)012<1877:SCVOSA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1877:SCVOSA>2.0.CO;2), 1999.
- Liebmann, B., Kiladis, G. N., Vera, C. S., Saulo, A. C., and Carvalho, L. M. V.: Subseasonal Variations of Rainfall in South America in the Vicinity of the Low-Level Jet East of the Andes and Comparison to Those in the South Atlantic Convergence Zone, *J. Climate*, 17, 3829–3842, [https://doi.org/10.1175/1520-0442\(2004\)017<3829:SVORIS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<3829:SVORIS>2.0.CO;2), 2004.
- Lima, C. H., Lall, U., Troy, T., and Devineni, N.: A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates, *J. Hydrol.*, 541, 816–823, <https://doi.org/10.1016/j.jhydrol.2016.07.042>, 2016.
- Lima, C. H. R. and Lall, U.: Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow, *J. Hydrol.*, 383, 307–318, <https://doi.org/10.1016/j.jhydrol.2009.12.045>, 2010.
- Lima, C. H. R., Lall, U., Troy, T. T., and Devineni, N.: A climate informed model for nonstationary flood risk prediction: application to Negro River at Manaus, Amazonia, *J. Hydrol.*, 522, 594–602, <https://doi.org/10.1016/j.jhydrol.2015.01.009>, 2015.
- Liu, Y. and Weisberg, R. H.: A Review of Self-Organizing Map Applications in Meteorology and Oceanography, in: Self Organizing Maps – Applications and Novel Algorithm Design, InTech, <https://doi.org/10.5772/13146>, 2011.
- Lu, M. and Lall, U.: Tropical Moisture Exports, Extreme Precipitation and Floods in Northeast US, *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2016-403>, 2016.
- Lu, M., Lall, U., Schwartz, A., and Kwon, H.: Precipitation predictability associated with tropical moisture exports and circulation patterns for a major flood in France in 1995, *Water Resour. Res.*, 49, 6381–6392, <https://doi.org/10.1002/wrcr.20512>, 2013.
- Luke, A., Vrugt, J. A., AghaKouchak, A., Matthew, R., and Sanders, B. F.: Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States, *Water Resour. Res.*, 53, 5469–5494, <https://doi.org/10.1002/2016WR019676>, 2017.
- Maddox, R. A.: Large-Scale Meteorological Conditions Associated with Midlatitude, Mesoscale Convective Complexes, *Mon.*

- Weather Rev., 111, 1475–1493, [https://doi.org/10.1175/1520-0493\(1983\)111<1475:LSMCAW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<1475:LSMCAW>2.0.CO;2), 1983.
- Malik, N., Bookhagen, B., Marwan, N., and Kurths, J.: Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks, *Clim. Dynam.*, 39, 971–987, <https://doi.org/10.1007/s00382-011-1156-4>, 2012.
- Mallakpour, I. and Villarini, G.: Investigating the relationship between the frequency of flooding over the central United States and large-scale climate, *Adv. Water Resour.*, 92, 159–171, <https://doi.org/10.1016/j.advwatres.2016.04.008>, 2016.
- Marengo, J. A., Soares, W. R., Saulo, C., and Nicolini, M.: Climatology of the Low-Level Jet East of the Andes as Derived from the NCEP–NCAR Reanalyses: Characteristics and Temporal Variability, *J. Climate*, 17, 2261–2280, [https://doi.org/10.1175/1520-0442\(2004\)017<2261:COTLJE>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2261:COTLJE>2.0.CO;2), 2004.
- Marengo, J. A., Liebmann, B., Grimm, A. M., Misra, V., Silva Dias, P. L., Cavalcanti, I. F. A., Carvalho, L. M. V., Berbery, E. H., Ambrizzi, T., Vera, C. S., Saulo, A. C., Nogues-Paegle, J., Zipser, E., Seth, A., and Alves, L. M.: Recent developments on the South American monsoon system, *Int. J. Climatol.*, 32, 1–21, <https://doi.org/10.1002/joc.2254>, 2012.
- Marwan, N. and Kurths, J.: Complex network based techniques to identify extreme events and (sudden) transitions in spatio-temporal systems, *Chaos*, 25, 097609, <https://doi.org/10.1063/1.4916924>, 2015.
- MATLAB: version 8.3.0.532 (R2014a), The MathWorks Inc., Natick, Massachusetts, 2014.
- Mendes, D., Souza, E., Trigo, I., and Miranda, P.: On precursors of South American cyclogenesis, *Tellus A*, 59, 114–121, <https://doi.org/10.1111/j.1600-0870.2006.00215.x>, 2007.
- Merz, B., Aerts, J., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Blöschl, G., Bouwer, L. M., Brauer, A., Cioffi, F., Delgado, J. M., Gocht, M., Guzzetti, F., Harrigan, S., Hirschboeck, K., Kilsby, C., Kron, W., Kwon, H.-H., Lall, U., Merz, R., Nissen, K., Salvatti, P., Swierczynski, T., Ulbrich, U., Viglione, A., Ward, P. J., Weiler, M., Wilhelm, B., and Nied, M.: Floods and climate: emerging perspectives for flood risk assessment and management, *Nat. Hazards Earth Syst. Sci.*, 14, 1921–1942, <https://doi.org/10.5194/nhess-14-1921-2014>, 2014.
- Milly, P. C. D., Wetherald, R. T., Dunne, K. A., and Delworth, T. L.: Increasing risk of great floods in a changing climate, *Nature*, 415, 514–517, <https://doi.org/10.1038/415514a>, 2002.
- Mioduszewski, J. R., Rennermalm, A. K., Hammann, A., Tedesco, M., Noble, E. U., Stroeve, J. C., and Mote, T. L.: Atmospheric drivers of Greenland surface melt revealed by self-organizing maps, *J. Geophys. Res.-Atmos.*, 121, 5095–5114, <https://doi.org/10.1002/2015JD024550>, 2016.
- Moftakhari, H. R., AghaKouchak, A., Sanders, B. F., and Matthew, R. A.: Cumulative hazard: The case of nuisance flooding, *Earth's Future*, 5, 214–223, 2017.
- Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., and Matthew, R. A.: Compounding effects of sea level rise and fluvial flooding, *P. Natl. Acad. Sci. USA*, 114, 9785–9790, 2017.
- Muza, M. N., Carvalho, L. M. V., Jones, C., and Liebmann, B.: Intraseasonal and Interannual Variability of Extreme Dry and Wet Events over Southeastern South America and the Subtropical Atlantic during Austral Summer, *J. Climate*, 22, 1682–1699, <https://doi.org/10.1175/2008JCLI2257.1>, 2009.
- Nakamura, J., Lall, U., Kushnir, Y., Robertson, A. W., and Seager, R.: Dynamical Structure of Extreme Floods in the U.S. Midwest and the United Kingdom, *J. Hydrometeorol.*, 14, 485–504, <https://doi.org/10.1175/JHM-D-12-059.1>, 2013.
- Neiman, P. J., Schick, L. J., Ralph, F. M., Hughes, M., and Wick, G. A.: Flooding in Western Washington: The Connection to Atmospheric Rivers, *J. Hydrometeorol.*, 12, 1337–1358, <https://doi.org/10.1175/2011JHM1358.1>, 2011.
- Nogués-Paegle, J. and Mo, K. C.: Alternating Wet and Dry Conditions over South America during Summer, *Mon. Weather Rev.*, 125, 279–291, [https://doi.org/10.1175/1520-0493\(1997\)125<0279:AWADCO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0279:AWADCO>2.0.CO;2), 1997.
- Over, T. M. and Gupta, V. K.: Statistical-Analysis of Mesoscale Rainfall – Dependence of a Random Cascade Generator on Large-Scale Forcing, *J. Appl. Meteorol.*, 33, 1526–1542, [https://doi.org/10.1175/1520-0450\(1994\)033<1526:SAOMRD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<1526:SAOMRD>2.0.CO;2), 1994.
- Paegle, J. N. and Mo, K. C.: Linkages between Summer Rainfall Variability over South America and Sea Surface Temperature Anomalies, *J. Climate*, 15, 1389–1407, [https://doi.org/10.1175/1520-0442\(2002\)015<1389:LBSRVO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1389:LBSRVO>2.0.CO;2), 2002.
- Pal, J. S. and Eltahir, E. A. B.: Teleconnections of soil moisture and rainfall during the 1993 midwest summer flood, *Geophys. Res. Lett.*, 29, 12-1–12-4, <https://doi.org/10.1029/2002GL014815>, 2002.
- Pandey, G., Lovejoy, S., and Schertzer, D.: Multifractal analysis of daily river flows including extremes for basins of five to two million square kilometres, one day to 75 years, *J. Hydrol.*, 208, 62–81, [https://doi.org/10.1016/S0022-1694\(98\)00148-6](https://doi.org/10.1016/S0022-1694(98)00148-6), 1998.
- Prudhomme, C. and Genevier, M.: Can Atmospheric circulation be linked to flooding in Europe, *Hydrol. Process.*, 25, 1180–1190, <https://doi.org/10.1002/hyp.7879>, 2010.
- Quiroga, R., Kreuz, T., and Grassberger, P.: Event synchronization: a simple and fast method to measure synchronicity and time delay patterns, *Phys. Rev. E*, 66, 041904, <https://doi.org/10.1103/PhysRevE.66.041904>, 2002.
- Robertson, A. W. and Mechoso, C. R.: Interannual and interdecadal variability of the South Atlantic convergence zone, *Mon. Weather Rev.*, 128, 2947–2957, [https://doi.org/10.1175/1520-0493\(2000\)128<2947:IAIVOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2947:IAIVOT>2.0.CO;2), 2000.
- Robertson, A. W., Mechoso, C. R., and Garcia, N. O.: Interannual prediction of the Paraná River, *Geophys. Res. Lett.*, 28, 4235–4238, <https://doi.org/10.1029/2000GL012197>, 2001.
- Salio, P., Nicolini, M., and Zipser, E. J.: Mesoscale Convective Systems over Southeastern South America and Their Relationship with the South American Low-Level Jet, *Mon. Weather Rev.*, 135, 1290–1309, <https://doi.org/10.1175/MWR3305.1>, 2007.
- Schumacher, R. S. and Johnson, R. H.: Organization and environmental properties of extreme-rain-producing mesoscale convective systems, *Mon. Weather Rev.*, 133, 961–976, <https://doi.org/10.1175/MWR2899.1>, 2005.
- Schumacher, R. S. and Johnson, R. H.: Characteristics of US extreme rain events during 1999–2003, *Weather Forecast.*, 21, 69–85, <https://doi.org/10.1175/WAF900.1>, 2006.
- Seo, Y., Schmidt, A. R., and Sivapalan, M.: Effect of storm movement on flood peaks: Analysis framework based on characteristic timescales, *Water Resour. Res.*, 48, W05532, <https://doi.org/10.1029/2011WR011761>, 2012.

- Silva, G. A. M. D. and Ambrizzi, T.: Summertime moisture transport over Southeastern South America and extratropical cyclones behavior during inter-El Niño events, *Theor. Appl. Climatol.*, 101, 303–310, <https://doi.org/10.1007/s00704-009-0218-6>, 2010.
- Silva, V. B. S. and Berbery, E. H.: Intense Rainfall Events Affecting the La Plata Basin, *J. Hydrometeorol.*, 7, 769–787, <https://doi.org/10.1175/JHM520.1>, 2006.
- Simmonds, I. and Keay, K.: Mean Southern Hemisphere extratropical cyclone behavior in the 40-year NCEP–NCAR reanalysis, *J. Climate*, 13, 873–885, [https://doi.org/10.1175/1520-0442\(2000\)013<0873:MSHECB>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0873:MSHECB>2.0.CO;2), 2000.
- Smith, J. A., Villarini, G., and Baeck, M. L.: Mixture Distributions and the Hydroclimatology of Extreme Rainfall and Flooding in the Eastern United States, *J. Hydrometeorol.*, 12, 294–309, <https://doi.org/10.1175/2010JHM1242.1>, 2011.
- Stedinger, J. R. and Cohn, T. A.: Flood Frequency Analysis with Historical and Paleoflood Information, *Water Resour. Res.*, 22, 785–793, <https://doi.org/10.1029/WR022i005p00785>, 1986.
- Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency Analysis of Extreme Events, in: *Handbook of Hydrology*, edited by: Maidment, D. R., 18, 18.1–18.66, McGRAW-HILL, INC., 1993.
- Sun, Q., Miao, C., AghaKouchak, A., and Duan, Q.: Century-scale causal relationships between global dry/wet conditions and the state of the Pacific and Atlantic Oceans, *Geophys. Res. Lett.*, 43, 6528–6537, 2016.
- Thomas, D. M. and Benson, M. A.: Generalization of Streamflow Characteristics From Drainage-Basin Characteristics, Tech. rep., U.S. Geol. Surv. Water Supply Paper 1975, 55 pp., 1970.
- Vahedifard, F., AghaKouchak, A., and Jafari, N. H.: Compound hazards yield Louisiana flood, *Science*, 353, 1374–1374, 2016.
- Velasco, I. and Fritsch, J. M.: Mesoscale convective complexes in the Americas, *J. Geophys. Res.-Atmos.*, 92, 9591–9613, <https://doi.org/10.1029/JD092iD08p09591>, 1987.
- Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M., and Blöschl, G.: Quantifying space-time dynamics of flood event types, *J. Hydrol.*, 394, 213–229, <https://doi.org/10.1016/j.jhydrol.2010.05.041>, 2010.
- Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D., and Krajewski, W. F.: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, *Adv. Water Resour.*, 32, 1255–1266, <https://doi.org/10.1016/j.advwatres.2009.05.003>, 2009.
- Villarini, G., Smith, J. A., Vitolo, R., and Stephenson, D. B.: On the temporal clustering of US floods and its relationship to climate teleconnection patterns, *Int. J. Climatol.*, 33, 629–640, <https://doi.org/10.1002/joc.3458>, 2013.
- Vogel, R. M., Yaindl, C., and Walter, M.: Nonstationarity: Flood Magnification and Recurrence Reduction Factors in the United States, *J. Am. Water Resour. As.*, 47, 464–474, <https://doi.org/10.1111/j.1752-1688.2011.00541.x>, 2011.
- Westra, S. and Sisson, S. A.: Detection of non-stationarity in precipitation extremes using a max-stable process model, *J. Hydrol.*, 406, 119–128, <https://doi.org/10.1016/j.jhydrol.2011.06.014>, 2011.
- Xavier, A. C., King, C. W., and Scanlon, B. R.: Daily gridded meteorological variables in Brazil (1980–2013), *Int. J. Climatol.*, 36, 2644–2659, <https://doi.org/10.1002/joc.4518>, 2016.
- Xu, G., Osborn, T. J., Matthews, A. J., and Joshi, M. M.: Different atmospheric moisture divergence responses to extreme and moderate El Niños, *Clim. Dynam.*, 47, 393–410, <https://doi.org/10.1007/s00382-015-2844-2>, 2016.
- Zhou, J. and Lau, K.-M.: Principal Modes of Interannual and Decadal Variability of Summer Rainfall over South America, *Int. J. Climatol.*, 21, 1623–1644, <https://doi.org/10.1002/joc.700>, 2001.
- Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., and Yorty, D. P.: Where are the most intense thunderstorms on earth?, *B. Am. Meteorol. Soc.*, 87, 1057–1071, <https://doi.org/10.1175/BAMS-87-8-1057>, 2006.