4. Generalized Synchronization between ENSO and hydrological variables in Colombia: A recurrence quantification approach

Hernán D. Salas, Germán Poveda, Óscar J. Mesa, and Norbert Marwan
Published in *Frontiers in Applied Mathematics and Statistics*

4.1. Abstract

We use Recurrence Quantification Analysis (RQA) to study features of Generalized Synchronization (GS) between El Niño-Southern Oscillation (ENSO) and monthly hydrological anomalies (HyAns) of rainfall and streamflows in Colombia. To that end, we check the sensitivity of the RQA concerning diverse HyAns estimation methods, which constitutes a fundamental procedure for any climatological analysis at inter-annual timescales. In general, the GS and its sensitivity to HyAns methods are quantified by means of time-lagged joint recurrence analysis. Then, we link the GS results with the dynamics of major physical mechanisms that modulate Colombia's hydroclimatology, including the Caribbean, the CHOCO and the Orinoco Low-Level Jets (LLJs), and the Cross-Equatorial Flow (CEF) over northwestern Amazonia (southern Colombia). Our findings show that RQA exhibits significant differences depending on the HyAns methods. GS results are similar for the HyAns methods with variable annual cycle but the time-lags seem to be sensitive. On the other hand, our results make evident that HyAns in the Pacific, Caribbean and Andean regions of Colombia exhibit strong (weak) GS with the ENSO signal during La Niña (El Niño), when hydrological anomalies are positive (negative). Results from the GS analysis allow us to identify spatial patterns of non-linear dependence between ENSO and the Colombian's climatology. The mentioned moisture transport sources constitute the interdependence mechanism and contribute to explain hydrological anomalies in Colombia during the phases of ENSO. During La Niña (El Niño), GS is strong (weak) for the Caribbean and the CHOCO LLJs whereas GS is moderate (strong) for the Orinoco LLJ. Moreover, moisture advection by

the Caribbean and CHOCO LLJs exhibit synchrony with HyAns at 0 to 2 (2 to 4) month-lags over north-western Colombia and the Orinoco LLJ moisture advection synchonizes with HyAns at similar month-lags over the Amazon region of Colombia. Furthermore, our results suggest a strong (weak) GS between negative (positive) Sea Surface Temperatures (SST) anomalies in the Eastern Pacific and rainfall anomalies in Colombia. In contrast, GS is strong (weak) for positive (negative) SST anomalies in the Central Pacific. Our GS results contribute to advance our understanding on the regional effects of both phases of ENSO in Colombia, whose socio-economical, environmental and ecological impacts cannot be overstated. This work provides a novel approach that reveals new insights into the impact of ENSO on northern South America.

4.2. Introduction

Synchronization is a well-known phenomenon that occurs when different oscillatory systems adjust the time scales of their oscillations due to mutual interaction [*Pikovsky et al.*(2001), *Balanov et al.*(2009)]. Diverse types of synchronization can be determined, such as frequency synchronization (FS), phase synchronization (PS), identical synchronization (IS), generalized synchronization (GS) and lag synchronization (LS) [*Brown and Kocarev*(2000)]. Particularly, GS emerges when the state of the driven system is completely determined by the state of the driving system [*rulkov*(1995)]. Hence, for two signals x and y that have a functional link $y \simeq f(x)$, GS techniques help to identify their synchronization or interrelationships even when their amplitudes behave differently. Diverse methods have been proposed to study synchronization such as those based on recurrences [*Schiff et al.*(1996), *Arnhold et al.*(1999), *Le Van Quyen et al.*(1999), *Quiroga et al.*(2002), *Zbilut et al.*(2006), *Marwan et al.*(2007)], phase differences [*Rosenblum et al.*(1997), *Tass et al.*(1998), *Pikovsky et al.*(2001), *Rosenblum et al.*(2001)], and quasi-simultaneous appearance of events [*Stolbova et al.*(2014), *Malik et al.*(2012), *Rheinwalt et al.*(2016), *Agarwal et al.*(2017)]. In particular, the method based on recurrences (recurrence plots and RQA) is based on properties from the non-linear dynamical systems such as the Poincaré Recurrence Theorem [*Poincaré*(1890)] and the Takens' theorem regarding the artificial reconstruction of strange attractors [*Takens*(1981), *Sauer et al.*(1991)]. In the last two decades, methods based on recurrences have proven to be very useful in different areas such as earth sciences, engineering, and applied physics, among others [*Marwan et al.*(2007), *Marwan*(2008), *Webber*(2014)]. Specifically, *join recurrence plots* (JRPs) have been used to study simultaneous recurrences and the detection of generalized synchronization in different systems [*Romano et al.*(2004)]. Recently, RQA has proved to be useful in diverse areas such as economy, physiology, earth sciences, astrophysics, and engineering [*Marwan et al.*(2007), *Webber*(2014)]. In particular, RQA has been used to investigate complex synchronization scenarios between coupled chaotic oscillators [*Feldhoff et al.*(2013)], emergence of GS in networks of structurally different time-delay systems [*Senthilkumar et al.*(2013), *Suresh et al.*(2016)], and precursors and synchronization of components of the tectonic system [*Hobbs *and Ord*(2018)], among others. The oscillatory nature of diverse hydroclimate phenomena is well known, such as those associated with the annual cycle of insolation (12 months), and the quasi-periodic El Niño-Southern Oscillation (3-4 years), which constitutes the most important modulator of interannual climate variability worldwide [*Sarachik and Cane*(2010)]. In turn, ENSO itself exhibits phase-locking with the annual cycle [*Stein et al.*(2011)], and affects the annual cycle of hydroclimatic variability in tropical South America and elsewhere [*Sarachik and Cane*(2010)]. Thus, GS and RQA provide powerful theoretical and quantitative frameworks to understand the non-linear interaction between ENSO and hydroclimatic processes in the region.

As far as we know, RQA has not been used to test the sensitivity of different methods to estimate hydro-climatic anomalies at interannual timescales (hereafter HyAns). In Chapter 2, we demonstrated that HyAns exhibit significant differences in terms of magnitude, timing, and sign, depending on the estimation method. Furthermore, we also demonstrated that linear and non-linear quantifiers of dependence can be significantly affected by how HyAns are defined (see Chapters 2 and 3). Then, in this work, we test the sensitivity of the GS metrics using the four HyAns methods assessed in Chapters 2 and 3, namely: *(i)* the Traditional Annual Cycle or constant climatology (HyAns–TAC), *(ii)* F-filtering of the Annual cycle by moving averages (HyAns–FAC), *(iii)* Annual cycle extracted by Singular Spectrum Analysis or Singular-Spectrum Annual Cycle (HyAns–SSAC) and, *(iv)* the Modulated Annual Cycle or variable climatology (HyAns–MAC). We also aim to provide advances regarding the robustness of the GS quantifiers based on recurrences for climate analysis at inter-annual timescales.

4.2.1. ENSO effects on hydro-climatology of Colombia

Colombia experiences positive (negative) anomalies in rainfall and streamflows, during the cold (warm) phase of ENSO, whose magnitude and timing vary across the main geographic regions of the country (Caribbean, Andes, Pacific, Orinoco and Amazon) [*Arias et al.*(2015), *Poveda and Mesa*(1997), *Poveda et al.*(2001), *Waylen and Poveda*(2002), *Poveda et al.*(2010), *Bedoya-Soto and Poveda*(2017), *Bedoya-Soto et. al.*(2018)]. Associated with those anomalies, there are extreme hydro-meteorological events of important socio-economical, environmental and ecological impacts. As such, the GS provides a new framework to further understand the impacts on ENSO on the hydroclimatology of Colombia, and the physical mechanisms responsible for such anomalies using a non-linear dynamical systems approach based on recurrences [*Marwan et al.*(2007)]. In general, we investigate GS between hydrological records of monthly series of rainfall and streamflows in Colombia and diverse variables related to moisture transport mechanisms influencing the Colombian hydrology during the extreme phases of ENSO. In particular, our GS-analysis is carried out at inter-annual timescales

considering the extreme phases of ENSO (La Niña and El Niño) separately. Three main moisture-ladden Low-Level Jets are considered as major mechanisms associated with hydroclimatic anomalies in Colombia, namely: *(i)* the Caribbean LLJ [*Wang*(2007)], *(ii)* the **Ch**orro del **O**ccidente **Co**lombiano (hereafter CHOCO LLJ) [*Poveda and Mesa*(2000), *Yepes et al.*(2019)], and *(iii)* the Orinoco LLJ [*Jiménez-Sánchez*(2019)] also known as the Eastern Andean LLJ or Corriente de los Andes Orientales (CAO) [*Montoya*(2001), *Torrealba and Amador*(2010)] (hereafter Orinoco LLJ). Furthermore, we also consider *(iv)* the Cross-Equatorial Flow (CEF) [*Wang and Fu*(2002)] as a possible mechanism of meridional moisture transport from the Amazon River basin to the Colombian territory. Additionally, in order to contrast our analysis with the available literature, we consider the SST gradient between the Pacific Ocean off the Colombian coast and the El Niño 1+2 zone, which makes part of the physical mechanism behind the CHOCO LLJ dynamics [*Poveda and Mesa*(2000)].
Also, we study the emergence of GS between two different types of ENSO and the Colombian hydrology, given that several recent works have shown that major SST warming has occurred frequently in the central Pacific (hereinafter CP El Niño), notably different from the eastern Pacific warming pattern during canonical El Niño events [*Sullivan et al.*(2016), Andreoli et al.(2017), Ashok and Yamagata (2007), Ashok and Yamagata (2009)] (hereinafter EP El Niño). Moreover, the CP El Niño exerts distinct socio-economical, environmental and ecological impacts around the globe [*Sullivan et al.*(2016), Andreoli et al.(2017), Ashok and Yamagata (2007), Ashok and Yamagata (2009)]. [Andreoli et al.(2017)] stated that the different SST anomaly patterns in the tropical Pacific affect the large-scale (Walker circulation and troposferic Rossby waves) and local atmospheric circulation patterns over northern South America producing different anomaly precipitation patterns. In particular, few works have studied the influence of the different types of El Niño events on Colombian's hydroclimatology [*Córdoba-Machado et al.*(2015), *Córdoba-Machado et al.*(2014), *Navarro-Monterroza et al.* (2019), *Serna et al.* (2018)]. Those works indicate that SST anomalies on the Eastern Pacific (EP) and the Central Pacific (CP) ocean are related with positive or negative hydrologic anomalies in the country depending on the location. At this point, there is not a clear consensus on the relationship between the different SST anomalies in the tropical Pacific and its influence on the hydrologic response in Colombia. Recently, [*Sullivan et al.*(2016)] proposed robusts indices for the EP El Niño, CP El Niño, and the Mixed-type El Niño, which are based on the normalized SST anomalies in the Niño3, Niño4, and Niño3.4 areas of the tropical Pacific Ocean, respectively. Therefore, we aim to quantify GS using those indices to infer clues on the influence of the different types of ENSO on Colombia hydrology.

4.2.2. Objectives

The objectives of this work are manifold: (i) To quantify non-linear inter-dependencies between the ENSO signal and the anomalies of hydrological variables in Colombia using the GS approach based on RQA; (ii) To assess the sensitivity of the GS-metrics based on

RQA using diverse methods used to define hydro-climatic anomalies; (iii) To evaluate the impact of ENSO on the Colombian hydrology though its effect on the sources of moisture that feed the water cycle dynamics over Colombia; (iv) To calculate GS between different types of El Niño indices (Central Pacific, Eastern Pacific, and Mixed) and the Colombian hydrology; and (v) To examine by means of GS metrics the propagation of the ENSO signals over Colombia.

To achieve the mentioned objectives we revise and characterize hydrological anomalies over Colombia during the phases of ENSO. Besides, a quantification of the linear and non-linear associations among the different variables seems necessary to elucidate new insights and explanations. In particular, we look for evidences about the influence of the aforementioned moisture sources on Colombian hydrology during ENSO. To that end, we consider the behavior of the moisture advection through the Caribbean, CHOCO, and Orinoco LLJs as well as the CEF during the phases of ENSO. Furthermore, our physical interpretation of GS results are based on Gill's theory [Gill (1980)] that states that weakening (strengthening) of the low-level winds is related to the increase (decrease) of the SSTs, which subsequently increase (decrease) of evaporation, increase (decrease) of moisture convergence, and increase (decrease) of rainfall in some regions [*Wang*(2007), *Martin and Schumacher*(2011), *Arias et al.*(2015)].

This work is outlined as follows: Section 4.3 describes the datasets and the region of study. Section 4.4 presents the RQA, the GS-index, the statistical tests for RQA. Results are discussed in Section 4.5 and the final remarks are drawn in Section 4.6.

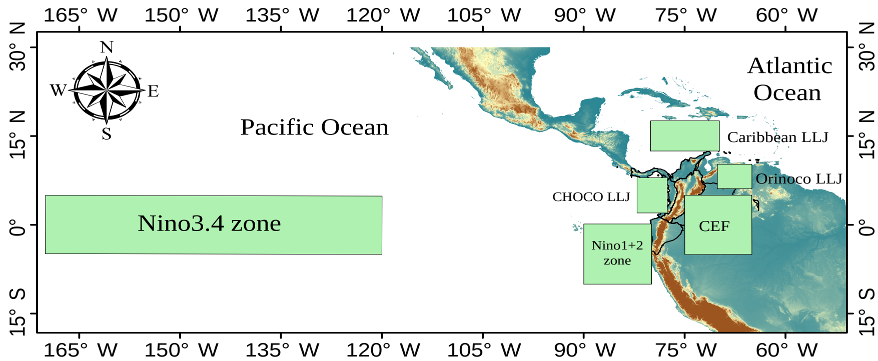
4.3. Data and region of study

We use monthly rainfall and streamflows series from the five major regions of Colombia: Caribbean, Andes, Pacific, Orinoco and Amazon. Fig. **4-1** shows six macro-climatic zones, namely: Nino3.4 and Nino1+2 in the tropical Pacific Ocean, as well as the zones related to the Caribbean, CHOCO, and Orinoco LLJs, and the Cross Equatorial Flow (CEF). Details are provided in sections 4.3.1, 4.3.3, and 4.3.4.

4.3.1. The ENSO signal

Classification of the ENSO events

To classify the ENSO events, we use the monthly Multivariate ENSO Index (MEI) [*Wolter and Timlin (1998)*], which is based on six variables over the tropical Pacific: sea-level pressure, zonal and meridional surface wind components, sea and air surface temperatures, and



Figures 4-1.: Macro-climatic zones associated with the Colombian hydrology.

total cloud fraction. It is considered to be a more reliable estimator of the ENSO state compared to other indexes such as Nino3.4 (based solely on sea surface temperatures) or the Southern Oscillation Index (SOI), which is based on sea-level pressures [Rasmusson and Carpener (1982), Trenberth and Stepaniak (2001)]. Values of the MEI index greater (lesser) than +0.5 (-0.5) define the occurrence of El Niño (La Niña) events, whereas the remaining periods will be referred to as neutral. The information of precipitation and streamflows was classified according to El Niño and La Niña months, following the classification previously described.

Sea Surface Temperatures in the tropical Pacific Ocean – El Niño 3.4 zone (SST3.4)

We use the monthly Extended Reconstructed Sea Surface Temperatures (ERSSTv5) in the El Niño 3.4 region (hereafter SST3.4), in the area 5^o N–5^o S and 170–120^o W. Details in https://www.cpc.ncep.noaa.gov. The mentioned SST3.4 time series was filtered using a 12-month moving average filter to extract the inter-annual variability. Then, we calculate the GS index between SST3.4 and rainfall and streamflows series in Colombia.

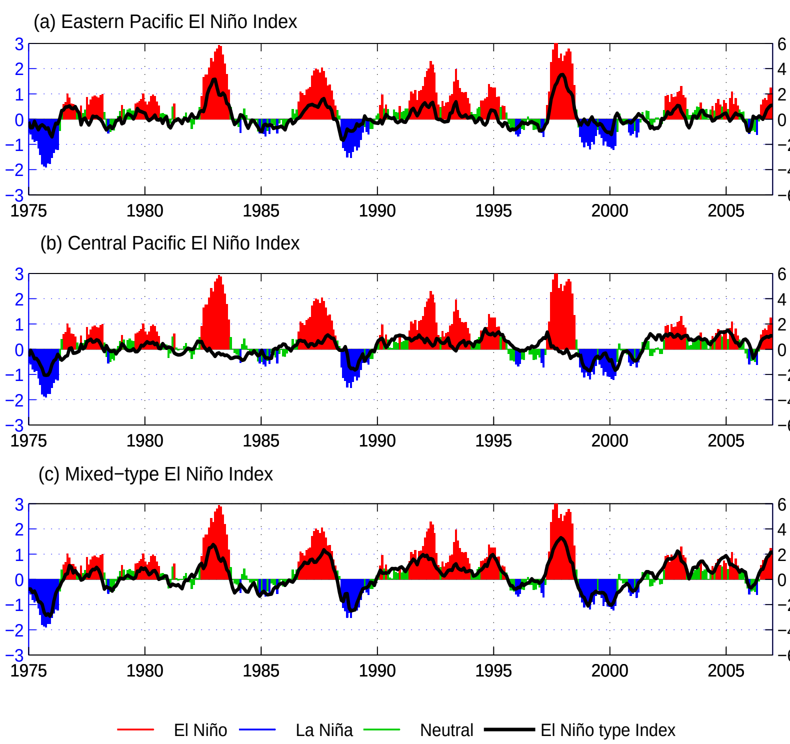
Indices for different types of El Niño

A recent study proposed three simple indices to represent the EP El Niño, CP El Niño, and the mixed type of El Niño (Mixed-type), which can be used for understanding, prediction, and impact assessment [*Sullivan et al.*(2016)]. In general, those indices are based on the SST in the Niño3 and Niño 4 regions, which coincide with the areas that exhibit maximum SST warm anomalies of the EP and CP El Niño, respectively [*Sullivan et al.*(2016)]. Indices for EP type, CP type, and Mixed-type El Niño are calculated as [*Sullivan et al.*(2016)]:

- EP = Niño3normalized - 0.5 * Niño4normalized,
- CP = Niño4normalized - 0.5 * Niño3normalized,

- Mixed-type = EP + CP ≃ Niño3.4normalized,

where Niño3normalized, Niño4normalized, and Niño3.4normalized denote the normalized SSTs anomalies in the Niño3 (5^{o}S–5^{o}N, 150^{o}W–90^{o}W), Niño4 (5^{o}S–5^{o}N, 160^{o}W–150^{o}W), and Niño3.4 (5^{o}S–5^{o}N, 170^{o}W–120^{o}W) regions, respectively [*Sullivan et al.*(2016)].
Fig. **4-2** shows the MEI index and its classification in La Niña, El Niño, and Neutral as explained in section 4.3.1. Furthermore, Fig. **4-2**a shows the EP El Niño index, Fig. **4-2**b shows the CP El Niño index and, Fig. **4-2**c shows the Mixed-type El Niño as defined by [*Sullivan et al.*(2016)].



Figures 4-2.: Indices for different types of El Niño in the period 1975-2006: (left axis) MEI index denoting La Niña (blue), El Niño (red) and Neutral (green) phases of ENSO. (right axis- black line) Eastern Pacific (EP) type (top row), Central Pacific (CP) type (middle row), and Mixed-type(bottom row).

4.3.2. Colombian hydrology

Precipitation

- Dataset 1: We use data from a reanalysis of Colombian rainfall performed by [*Hurtado and Mesa*(2014)], which produced monthly precipitation fields, spanning the period from 1975 to 2006, at a spatial resolution of 5 minutes of arc, in the region 5^{o}S-15^{o}N and 80^{o}W-65^{o}W. Those authors used a set of 2270 rain gauges from diverse institutions including Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia (IDEAM), Empresas Públicas de Medellín (EPM), Centro Nacional de Investigaciones del Café (CENICAFE) and The Global Historical Climatology Network (GHCN). In addition, they also used satellite information from the Global Precipitation Climatology Project(GPCP)-V2, National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR), Tropical Rainfall Measuring Mission (TRMM), and the Geostationary Satellite System (GOES). Finally, the monthly precipitation fields resulted from a combination of the available data and the implementation of the PRISM model over the Colombian territory. The resulting data set is freely available upon request to one the authors (O. J. M) [*Hurtado and Mesa*(2014)].

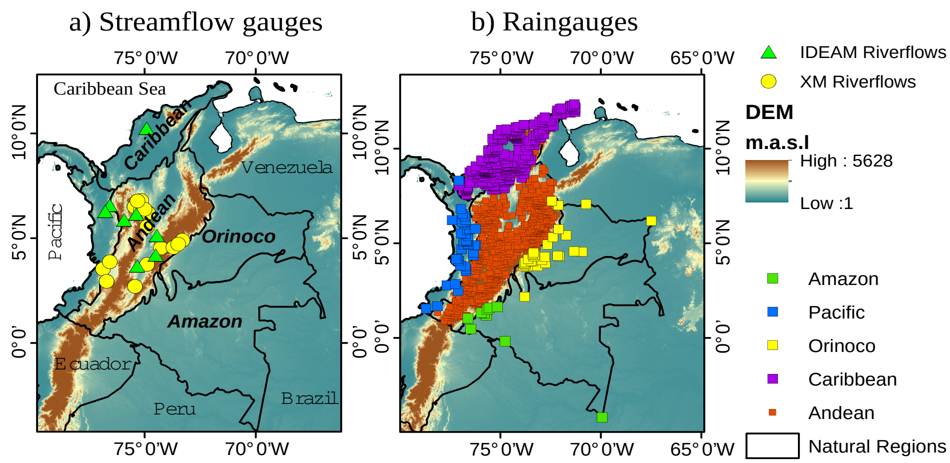
- Dataset 2: We also use data from a set of 937 rain gauges from IDEAM, at monthly resolution, having record lengths of 40 years, for the period 1976-2015, with less than 10% missing records. The geographical distribution of this data set is as follows: 38 time series in the Pacific, 631 in the Andes, 228 in the Caribbean, 30 in the Orinoco, and 10 in the Amazon region. See Fig. **4-3**b.

Streamflows

We use monthly average streamflows data from seven IDEAM's river gauging stations having common record lengths of 38 years, from 1976 to 2013 (see Table S1 in the Supp. Info). Furthermore, we use a set of 18 time series from XM (url: www.xm.com.co), at monthly time-scale, with record lengths 40 years from 1976-2015 (see Table S2 in the Supp. Info). Most river gauging stations are located in the Andean region, whereas long-term data in the Orinoco and Amazon regions are not available. See Fig. **4-3**a.
Table **B-1** details the sample size used to quantify dependency among diverse variables for each dataset. The samples were defined once each month is classified as El Niño, La Niña, and Neutral according to the MEI index (details in section 4.3.1).

Tables 4-1.: Sample size of monthly anomalies according to the phases of ENSO for each dataset.

		Total monthly data points		
Datasets	**Period**	**El Niño**	**La Niña**	**Neutral**
Hurtado & Mesa (2014)	1975-2006	158	65	161
IDEAM–Precipitation	1976-2015	193	85	202
IDEAM–Streamflows	1976-2013	176	85	195
XM–Streamflows	1976-2015	193	85	202



Figures 4-3.: National context including the Colombian Natural Regions as well as the streamflows and rainfall gauges. a) Streamflow gauges. b) Raingauges.

4.3.3. Moisture Advection through Low-Level Jets

In order to investigate the influence of the Low Levels Jets (LLJs) as rainfall-generating moisture transport mechanisms and therefore as physical mechanisms behind the emergence of GS, we use data from the NCEP/NCAR Reanalysis Project [*Kalnay et al.*(1996)]. Moisture advection is calculated as the product between the monthly mean zonal winds and the monthly mean specific humidity at 925 hPa. In particular, moisture advection is calculated over the following geographic zones, which are associated with the three LLJs surrounding Colombia, namely: (i) The Caribbean LLJ moisture advection is calculated over the region 12.5^{o}N–17.5^{o}N, 80^{o}W–70^{o}W. This region is selected based on the work by [*Wang*(2007)], who studied the variability of the Caribbean LLJ; (ii) The CHOCO LLJ moisture advection is quantified over the region 2^{o}N–8^{o}N, 82^{o}W–77.5^{o}W. This region was used in study [*Poveda and Mesa*(2000)], which deals with the characterization and influence of the CHOCO LLJ on ocean-land-atmosphere interactions over the far eastern Pacific; and (iii) The Orinoco LLJ moisture advection is calculated over the region 6^{o}N–10^{o}N, 70^{o}W–65^{o}W. This region is located over the vast Llanos (plains) region over the Colombia-Venezuela border and, selected according to previous works [*Montoya*(2001), *Jiménez-Sánchez*(2019), *Torrealba and Amador*(2010)].

In particular, the CHOCO LLJ dynamics is related to the SST gradient in the far eastern tropical Pacific. That is, the surface air temperature gradient between the Pacific Ocean off the Colombian coast (2^{o}N–8^{o}N, 77.5^{o}W–82^{o}W) and the El Niño 1+2 region (10^{o}S–0^o, 90^{o}W–80^{o}W) from the sea level to 925 hPa [*Poveda and Mesa*(2000)]. Hence, we calculate the SST gradient between both zones and its association with rainfall and streamflows in Colombia using the GS metrics. We will contrast the GS results obtained between ENSO and the Colombian hydrology (rainfall and streamflows) with those obtained between LLJs and the Colombian hydrology to infer similarities and differences.

4.3.4. The Amazonian Cross-Equatorial Flow

The Cross-Equatorial Flow can be related to the hydrological dynamics over eastern and southeastern Colombia (Orinoco and Amazon regions) and the northern-most Amazon basin [*Wang and Fu*(2002)]. Hence, as in the previous section, we estimate the moisture advection as the product between the mean meridional winds and the mean specific humidity at 925 hPa, using data from the NCEP/NCAR Reanalysis Project [*Kalnay et al.*(1996)], over the region 5^{o}S–5^{o}N, 65^{o}W–75^{o}W [*Wang and Fu*(2002)].

4.4. Methods

4.4.1. Generalized synchronization using the Recurrence Quantification Analysis

Recurrence in phase space is a fundamental property of diverse types of dynamical systems [*Marwan et al.*(2007)]. This recurrence behavior can be characterized and quantified by means of the so-called Recurrence Plot (RP), which is a modern tool of non-linear data analysis that allows to visualise and quantify properties of higher-dimensional phase space trajectories [*Eckmann et al.*(1987)]. The construction of the RP from a time series, x_i, requires the representation of the m-dimensional phase space of the system X. For single time series, x_i, it is necessary to reconstruct the dynamics using the time delay embedding technique as [*Packard et al.*(1980), *Takens*(1981)],

$$\vec{x_i} = \left(x_i, x_{i+\omega}, ..., x_{i+\omega(m-1)}\right), \qquad \vec{x_i} \in R^m \tag{4-1}$$

where m is the embedding dimension and ω is the time delay. The false nearest neighbors (FNN) is the most used method to quantify m, whereas the Auto- Information Function (AIF) is commonly used to quantify the time-delay [*Marwan et al.*(2007), *Fraser and Swinney*(1986), *Kantz and Schreiber*(2004), *Marwan*(2011)]. Once m and ω have been estimated, the RP is calculated using the pair-wise proximity test to obtain the recurrence of X for the times i and j as,

$$R_{ij}^X = \Theta(\epsilon - ||\vec{x_i} - \vec{x_j}||), \qquad i, j = 1, ..., N^{'} \tag{4-2}$$

where N is the number of data points, $N^{'} = N - (m-1)\omega$ is the dimension of phase space vectors, ϵ is the threshold to the proximity between the phase space vectors. According to the literature, ϵ must be selected as a few percent (not larger than 10 %) of the maximum phase space diameter [*Webber*(2014), *Mindlin and Gilmore*(1992), *Zbilut and Webber*(1992)]. $||\vec{x}_i - \vec{x}_j||$ is the spatial distance between vectors in phase space. $||\cdot||$ denotes any norm in phase space (e.g., Manhattan, Euclidean, or maximum norm) [*Webber*(2014)], and $\Theta(\cdot)$ is the Heaviside function such as $\Theta(x) = \{1 \mid x > 0; 0 \mid x \leq 0\}$ [*Marwan et al.*(2007)]. The plot of the recurrence binary matrix R^X provides the RP. Then, the probability that one system recurs to a certain state $\vec{x}_i$ is equal to the column-average of the recurrence matrix [*Marwan et al.*(2007)].

$$p(\vec{x}_i) = \frac{1}{N'} \sum_{j=1}^{N'} R^X_{i,j} \tag{4-3}$$

On the other hand, the Joint Recurrence plot (JRP) is used to study the possible influence between two different physical systems, constituting a measure of their simultaneous recurrence. The *joint recurrence matrix* is defined as the Hadamard product of the RPs of the single RPs of systems X and Y [*Marwan et al.*(2007), *Romano et al.*(2004), *Marwan et al.*(2013)],

$$JR^{X,Y}_{i,j} = \Theta(\epsilon - ||\vec{x}_i - \vec{x}_j||) \times \Theta(\epsilon - ||\vec{y}_i - \vec{y}_j||) \tag{4-4}$$

Hence, the probability of finding a simultaneous recurrence at time i in both systems X and Y is equal to the column-average of the JR^{XY} matrix [*Marwan et al.*(2013)],

$$p(\vec{x}_i, \vec{y}_i) = \frac{1}{N'} \sum_{j=1}^{N'} JR^{X,Y}_{i,j} \tag{4-5}$$

The same procedure can be carried out considering the relationship between the system X and the time lagged system $Y(\tau)$. Then, if both systems X and the time lagged $Y(\tau)$ are independent, $p(\vec{x}_i, \vec{y}_{i+\tau}) = p(\vec{x}_i)p(\vec{y}_{i+\tau})$. In contrast, if both systems are in GS, we expect approximately the same recurrences, $p(\vec{x}_i, \vec{y}_{i+\tau}) \approx p(\vec{x}_i) \approx p(\vec{y}_{i+\tau})$.
The computation of the recurrence matrices is most appropriate using a fixed amount of nearest neighbors, N_n, for each column in the matrix than a fixed threshold [*Marwan et al.*(2007), *Arnhold et al.*(1999)]. This corresponds to the original definition of RP [*Eckmann et al.*(1987)]. Therefore, according to [*Marwan et al.*(2007)], $p(\vec{x}_i)$ and $p(\vec{y}_{i+\tau})$ are equal and fixed by N_n. i.e. $p(\vec{x}_i) = p(\vec{y}_{i+\tau}) = N_n/N$. Taking the *recurrence rate* (RR) as $RR = N_n/N$,

$$S(\tau) = \frac{p(\vec{x}_i, \vec{y}_{i+\tau})}{RR} \tag{4-6}$$

where $S(\tau)$ is a time series with values ranging between 0.0 and 1.0, and τ is a time-lag between timeseries x_i and $y_{i+\tau}$ [*Marwan et al.*(2007)]. Consequently, it is possible to define a global index for GS based on the average *joint probability of recurrence* (hereafter GS-index) by choosing the maximum value of $S(\tau)$ and normalizing it [*Marwan et al.*(2007)],

$$GS = max_{\tau} \frac{S(\tau) - RR}{1 - RR} \tag{4-7}$$

where GS ranges from 0 (independent) to 1 (correlated). In order to avoid potential pitfalls in RQA, our procedures are based on pivotal literature on recurrences [*Marwan et al.*(2007), *Marwan*(2011), *Webber*(2014)].

4.4.2. Significance tests

The existence of GS between the systems $X(t)$ and $Y(t+\tau)$ requires a statistical test to verify the significance of the GS-index at time-lag τ. Consequently, it is possible to test the significance of the GS-index based on the joint recurrence matrix between the trajectory in phase space of the original system X (ENSO in our case) and a representative number of phase space trajectories of the system Y (precipitation and streamflows in our case), which are independent copies of the underlying system Y (also called twin surrogates –TS) [*Thiel et. al.*(2006), *Thiel et. al.*(2008)]. This ensures surrogates that are uncorrelated or not synchronized with each other. Then, TS are phase space trajectories sharing the same neighborhood up to the threshold ϵ such that $R^X_{i,k} = R^X_{j,k}$, $k = 1, ..., N'$. In this work, TS are constructed using the algorithm by [*Thiel et. al.*(2006), *Thiel et. al.*(2008)], which is included in the CRP Toolbox for MATLAB provided by TOCSY, available at http://tocsy.agnld.uni-potsdam.de [*Marwan et al.*(2007)].
The null hypothesis is that each TS is an independent trajectory of the system, corresponding to different initial conditions. In this work, we test the significance of the GS-index between X and $Y(\tau)$ using 100 TS of the Y system (hydrological variables). i.e. to calculate the GS-index for each pair of X and the TS and obtain their distribution. Finally, we compare the GS-index for the original time series $x(t)$ and $y(t)$ with the 95th percentile (95% confidence boundary) of the test distribution. The time-lags τ where the GS-index, $GS(\tau)$, takes values out the confidence band imply values for which the null hypothesis is rejected. i.e. the systems X and Y are significantly dependent. In order to test for the null hypothesis, we carry out a multiple testing procedure following the Sidák-Bonferonni equations [Holm (1979), *Abdi*(2007)], and perform 5 independent tests limiting the risk of making at least one Type I error to an overall value of $\alpha[PF] = 0.05$ (PF means *alpha per family of tests*) and then we use $\alpha[PT] \approx 0.01$ (PT means *alpha per test*).

4.4.3. Joint Probability of Recurrences during ENSO

Taking into account that precipitation and streamflows exhibit different kind of anomalies depending on the ENSO phase (warm or cold), we quantify the probability of joint recurrence,

$p(\vec{x}_i, \vec{y}_i)$, at time i (Eq. 4-5) for the cold (La Niña) and warm (El Niño) phases of ENSO separately. After that, we calculate the *joint recurrence matrix* (Eq. 4-4) and then $p(\vec{x}, \vec{y}(\tau))$, where τ is the time-lag. Then, we look for the columns in the JRP that correspond to the classification of ENSO phases as per the MEI index (section 4.3.1). Finally, we calculate $p(\vec{x}_i, \vec{y}_i)$ during La Niña and El Niño conditions.

For the different types of El Niño indices (EP, CP and Mixed-type), we carry out a similar procedure. We quantify the probability of joint recurrence, $p(\vec{x}_i, \vec{y}_i)$, for each index using all data and also periods as per the MEI classification, separately.

4.4.4. Sensitivity of the GS metrics to diverse HyAns methods

Diverse methods commonly used to estimate hydro-climatic anomalies (HyAns) exhibit differences that may induce strong biases and error sources toward the interdependence analysis and modeling of climate time series (see Chapter 2). Such differences may generate ambiguities for the physical interpretation of results (see Chapter 2), and thus the need to consider the uncertainty in HyAns estimation and its consequences for quantification of interdependence using linear and non-linear techniques. In this sense, in Chapter 3 we showed that phase synchronization metrics based on Generalized Phase Differences [*Rosenblum et al.*(2001)] are sensitive to seasonal-residual frequencies in time series of anomalies.

Hence, we verify the sensitivity of GS metrics based on RQA using the four HyAns methods assessed in Chapters 2 and 3, namely: *(i)* the Traditional Annual Cycle or constant climatology (HyAns–TAC), *(ii)* **F**-filtering of the Annual cycle by moving average (HyAns–FAC), *(iii)* Annual cycle extracted by Singular Spectrum Analysis or Singular-Spectrum Annual Cycle (HyAns–SSAC) and, *(iv)* the Modulated Annual Cycle or variable climatology. In this way, we evaluate the robustness of our GS approach in relation to the diverse HyAns estimation methods.

As mentioned in Chapter 2, the HyAns–FAC method is appropriate for climatological purposes owing to the following properties: (i) a firm mathematical foundation, (ii) a varying climatology (annual cycle), (iii) an adequate separation of noise, intra-annual, seasonal and inter-annual components, since (iv) it does not leave significant residual seasonal frequencies in HyAns, (v) its extreme values are in a middle range in comparison with the HyAns–SSAC and HyAns–MAC methods, (vi) its autocorrelation function is comparable with HyAns–SSAC and HyAns–MAC, and (vii) its computational cost is one-tenth of HyAns–SSAC and hundredths times less than HyAns–MAC.

4.5. Results and discussion

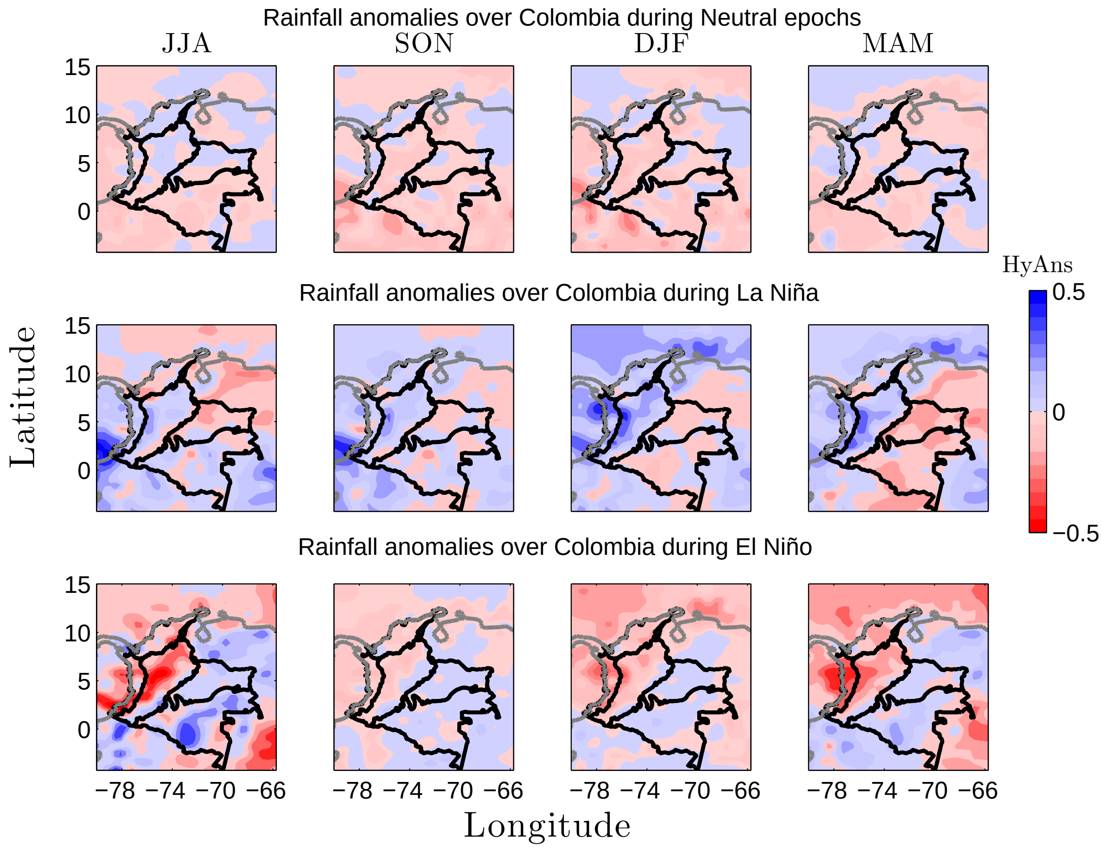
4.5.1. Rainfall anomalies in Colombia and their association with ENSO by means of Pearson correlations

First, we quantify rainfall anomalies over Colombia considering both phases of ENSO (El Niño and La Niña), as explained in section 4.3.1. Rainfall anomalies were calculated using a F-Filtering of the annual cycle by moving averages, which filters out adequately the seasonal signal [Douglass (2011), Douglass (2015)]. Fig. **C-11** shows that during La Niña, Colombia experiences wetter (drier) than normal periods in the Pacific, Andean and Caribbean (Orinoco and Amazon) regions. In contrast, during El Niño, Colombia evidences drier (wetter) than normal periods in the Pacific, Andean and Caribbean (Orinoco and Amazon) regions. On the other hand, Fig. **C-11** shows that, during La Niña (El Niño), the Orinoco and Amazon regions exhibit negative (positive) rainfall anomalies throughout the year. It is worth noting that the influence of both phases of ENSO on streamflows in the Amazon region of Colombia does not show a clear-cut pattern of positive or negative anomalies [*Arango-Ruda and Poveda*(2018)]. Hence, ENSO's influence on rainfall anomalies in Colombia depends on the phase of ENSO (La Niña or El Niño) and the region of the country.

Second, we quantify the linear association between SST anomalies in the tropical Pacific and rainfall anomalies over Colombia during ENSO using Pearson correlations, r. We calculate r using the complete monthly time series to capture the big picture. Fig. **4-5** suggests that positive (negative) rainfall anomalies over the Pacific, the Andean and the Caribbean regions are significantly correlated with negative (positive) SST anomalies in the tropical Pacific Ocean. Furthermore, the spatial patterns of statistical significance of r are in agreement with the spatial patterns of positive (negative) rainfall anomalies in Colombia during ENSO (see Fig. **C-11**). However, it is worth noting that correlation patterns between SST3.4 anomalies and rainfall anomalies could differ during the phases of ENSO, and that Pearson correlations only quantify the linear association. This points out the need to explore non-linear techniques to quantify the dependence between hydrological anomalies in Colombia and the ENSO signal considering La Niña and El Niño phases.

4.5.2. Moisture Advection Anomalies by LLJs during ENSO

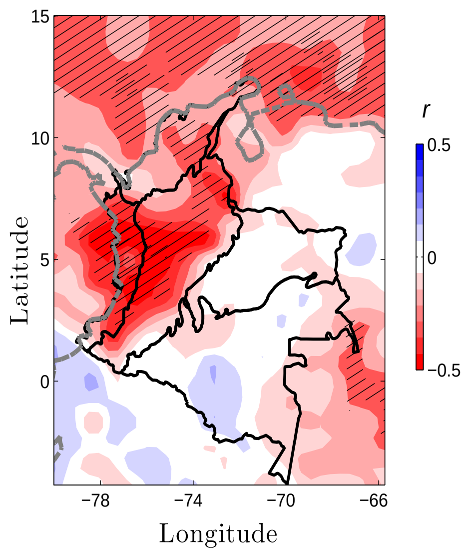
Low-Level Jets (LLJs) are important rainfall-formation mechanisms and their non-linear interdependence with hydroclimatology requires an adequate characterization of moisture advection anomalies [*Bonner* (1968), *Stensrud* (1996), *Gimeno et al.* (2016), *Hoyos et al.*(2017)]. To that end, we characterize the annual cycle of moisture advection anomalies by the Caribbean LLJ, the CHOCO LLJ, and the Orinoco LLJ as well as by the Cross-Equatorial Flow (CEF) during the phases of ENSO. To this end, we use data from the NCEP/NCAR Reanalysis Project [*Kalnay et al.*(1996)] and calculate moisture advection as the product



Figures 4-4.: Seasonal average rainfall anomalies (HyAns) in Colombia during ENSO. Data set No. 1. Rainfall anomalies were calculated using F-Filtering of the annual cycle by moving averages, which is previously mentioned as the HyAns–FAC method. (gray line) northern South American coast line. (black line) Borders of the Colombia's natural regions.

between the mean zonal or meridional winds and the mean specific humidity at 925 hPa averaged over the geographic areas described in section 4.3.3. We estimate moisture advection anomalies in such a way that positive or negative anomalies represent increase or decrease of moisture advection, respectively. Furthermore, after an exhaustive verification of NCEP/NCAR 925 hPa data in the period 1970-2015, the moisture advection by the Caribbean and the Orinoco LLJs have a predominant easterly direction whereas the moisture advection by the CHOCO LLJ exhibits a clear-cut westerly direction during 80% of time. Likewise, the CEF exhibits dominant northerly (southerly) moisture advection from October to March (April to September).

Fig. **4-6**a shows that the Caribbean LLJ exhibits positive (negative) zonal moisture advection anomalies during El Niño (La Niña) for the period from May to November, with a peak in July for both positive and negative cases, which coincides with the variation of Caribbean LLJ winds during June-July-August (JJA) and December-January-February

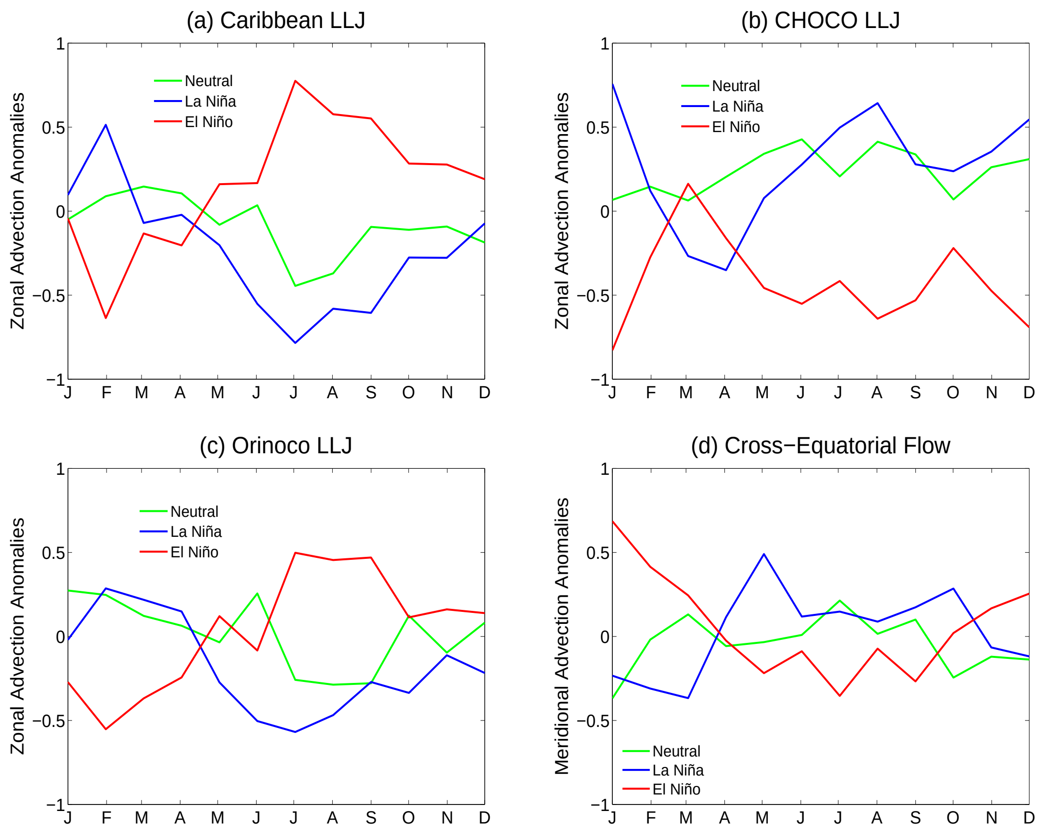


Figures 4-5.: Pearson correlation, r, between rainfall anomalies in Colombia and SST3.4 anomalies. (dashed gray line) nothern South American coast line. (black line) Borders denote Colombia's natural regions. The hatched areas indicate that r is significant at 5%. The significance levels are calculated considering the Effective Degrees of Freedom due to autocorrelation in timeseries [Afyouni et al. (2019)]

(DJF) [*Wang*(2007),*Martin and Schumacher*(2011)]. In particular, this result coincides with the observation that during La Niña, the easterly Caribbean LLJ winds are weak (strong) in JJA (DJF), whereas during El Niño, the easterly Caribbean LLJ winds are strong (weak) in JJA (DJF) [*Wang*(2007)]. Furthermore, Fig. **4-6**b shows that zonal moisture advection anomalies by the CHOCO LLJ are positive (negative) during La Niña (El Niño) from July to December. These results are consistent with those reported by [*Poveda and Mesa*(2000)]. Moreover, Fig. **4-6**c shows that moisture advection anomalies by the Orinoco LLJ are positive (negative) during El Niño (La Niña) in JJA. Finally, Fig. **4-6**d shows that the CEF exhibits positive (negative) moisture advection anomalies during La Niña (El Niño) from April to October for the southerly winds, whereas the CEF exhibits negative (positive) during La Niña (El Niño) from November to March for the northerly winds.
The annual cycles of moisture advection anomalies, as well as the annual cycle of hydrological anomalies during ENSO, are important information sources for the interpretation of results of non-linear metrics such as GS. In the following sections, we will revisit these findings in order to interpret the GS results from a physical point of view.

4.5.3. Time Delay and Embedding Dimension

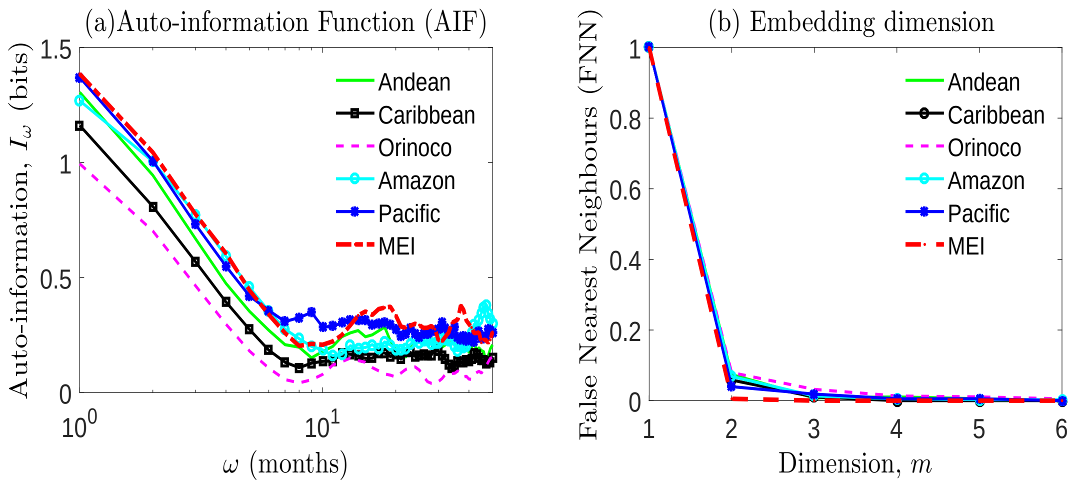
Detection of GS using the recurrence quantification analysis (RQA) requires to adequately



Figures 4-6.: Annual cycle of standardized moisture advection anomalies during ENSO using data from the NCEP/NCAR Reanalysis Project for the period 1970-2015. (a) Zonal moisture advection by the Caribbean LLJ. (b) Zonal moisture advection anomalies by the CHOCO LLJ. (c) Zonal moisture advection anomalies by the Orinoco LLJ. (d) Meridional moisture advection by the Cross-Equatorial Flow.

select the time delay, ω, and the embedding dimension, m. The time delay, ω, for each time series is calculated as the time that the *auto information function* (AIF) takes to decay to $1/e$ of its value at zero lag (see section 4.4.1). Figure **4-7** (panel a) shows that all the time series exhibit an exponential decay until a time delay $\omega \approx 8$ months. The embedding dimension, m, is calculated using the *false nearest neighbors* (FNN). Figure **4-7** (panel b) shows that according to FNN an appropriate embedding dimension for all time series is $m = 2$.

Hereafter, all analyses are carried out using the fixed amount of nearest neighbors (FAN) whereas the significance tests are based on 100 twin surrogates. On the other hand, in this work, the threshold of proximity between the phase space vectors was selected as $\epsilon = 10\%$ according to Refs. [*Webber*(2014), *Mindlin and Gilmore*(1992), *Zbilut and Webber*(1992)].



Figures 4-7.: Estimation of the time delay and the embedding dimension for monthly precipitation averaged over the natural regions of Colombia (from data set No. 1) and MEI index. (a) Time delay, ω, using the auto-information function (AIF). (b) Embedding dimension, m, using *false nearest neighbors* (FNN).

4.5.4. Sensitivity of the GS-metrics regarding diverse HyAns methods

We use a time series of rainfall anomalies for the Andean region (monthly precipitation spatially averaged over the Andean region) and quantify the GS-index using the four HyAns methods: HyAns–TAC, HyAns–FAC, HyAns–SSAC, and HyAns–MAC (see Chapter 2 and section 4.4.4). To that end, we select sample data classified according to El Niño and La Niña months, as per the MEI definition (see section 4.3.1). Fig. **4-8** shows the GS-index for the mentioned HyAns methods. In general, our findings reveal that the HyAns–TAC method exhibits the lowest GS values in relation to the other methods, during both La Niña and El Niño. This behavior can be explained by the presence of noise, intra-annual and significant-residual seasonal frequencies in anomalies, when those are estimated using HyAns–TAC (see Chapter 2). Hence, we consider that the HyAns–TAC method is not reliable for the RQA and purposes of this study.

In addition, Fig. **4-8** shows the GS-index between the ENSO signal and rainfall anomalies using diverse HyAns methods. In general, the GS-index exhibits higher values during La Niña than during El Niño, for the HyAns–FAC, HyAns–SSAC and HyAns–MAC methods. In this sense, [*Poveda and Mesa*(1996),*Poveda and Mesa*(1997)] showed that the ratio between hydrological anomalies and the long-term mean is higher during La Niña than during El Niño. In agreement with this finding, we showed in Chapter 3 that the strength of phase synchronization index between rainfall anomalies and the ENSO signal also is stronger during La Niña than during El Niño.

In terms of the time-lag between the ENSO signal and rainfall anomalies over the Andean region of Colombia, Fig. **4-8** shows that during La Niña, the HyAns–FAC, HyAns–SSAC,