A Skewed perspective of the Indian rainfall-ENSO Relationship

Justin Schulte[1*], Fredrick Policielli[2], and Benjamin Zaitchik[3]

1. Science Systems and Applications, Inc.

2. NASA Goddard Space Flight Center

3. John Hopkins University

***corresponding Author: Justin Schulte (justin.a.schulte@nasa.gov)**

Abstract

Wavelet coherence is a method that is commonly used in hydrology to extract scale-dependent, non-stationary relationships between time series. However, we show that the method cannot always determine why the time-domain correlation between two time series changes in time. We show that even for stationary coherence, the time-domain correlation between two time series weakens if at least one of the time series has changing skewness. To overcome this drawback, a nonlinear coherence method is proposed for quantifying the cross-correlation between nonlinear modes embedded in time series. It is shown that using nonlinear coherence and auto-bicoherence spectra together with auto-bicoherence spectra can provide additional insight into changing time-domain correlations. The new method is applied to the El Niño /Southern Oscillation (ENSO) and All-India rainfall (AIR), which is intricately linked to hydrological processes across the Indian sub-continent. The nonlinear coherence analysis showed that the skewness of AIR is weakly correlated with that of two ENSO time series after the 1970s, indicating that increases in ENSO skewness after the 1970s at least partially contributed to the weakening ENSO-AIR relationship in recent decades. The implication of this result is that the intensity of skewed El Niño events is likely to overestimate Indian drought severity, which was the case in the 1997 monsoon season, a time point when the nonlinear wavelet coherence between AIR and ENSO reached its lowest value in the 1871-2016 period. We determined that the association between the weakening ENSO-AIR relationship and ENSO nonlinearity could reflect the contribution of different nonlinear ENSO modes to ENSO diversity.

1. Introduction

The South Asian Monsoon, which is the dominant source of precipitation source for the Indian subcontinent, has been a target for seasonal prediction for well over a century (Blanford, 1884). Despite this long heritage of research, skillful prediction remains a challenge, driving extensive and ongoing research on statistically and dynamically based prediction methods (e.g., REFS). It is difficult to overstate the importance of the South Asian Monsoon to the well-being of citizens in India. Strong monsoon years have caused are associated with catastrophic flooding (Kale, 2012; Sanyal and Lu, 2005) and large landslides (Dortch et al., 2009), while weak monsoons have led to water shortages (Mishra et al., 2016) and crop losses (Parthasarathy et al., 1988; Prasanna, 2014) that, in historical times, were known to resulted in significant food shortages in historical times (Fagan, 2009). Thus, while the majority of monsoon forecast studies target prediction of rainfall totals, the hydrological and agricultural impacts of monsoon variability provide the most pressing motivation for the work.

Much of the research on South Asian Monsoon prediction has focused on the relationship between the El Niño/Southern Oscillation (ENSO; Walker and Bliss, 1932) and monsoon strength. During El Niño years, droughts are favored, while rainfall surpluses are favored during La Niña years (Shukla and Paolino, 1983; Kripalani and Kulkarni, 1997). However, there is no one-to-one relationship between ENSO and Indian rainfall. As a result, summer rainfall predictions based on ENSO have proven challenging. For example, the 1997/1998 El Niño event was extremely strong yet climatological Indian monsoon conditions were observed (Shen and Kimoto, 1999; Slingo and Annamalai, 2000). It is therefore important to understand why certain El Niño events are not accompanied by monsoon failures.

There are a few reasons for the challenges faced when predicting Indian rainfall using ENSO. The first reason is that the relationship between ENSO and India rainfall is non-stationary. As shown by Torrence and Webster (1999), the relationship between ENSO and India rainfall cycles between periods of high and low coherence. Kumar et al.

(1999) found that the relationship between Indian rainfall and ENSO weakened in the 1970s and hypothesized that a southward shift in Walker circulation anomalies associated with ENSO events and increased Eurasian spring and winter surface temperatures were was responsible for the weakening relationship. Other works suggest that the changing ENSO-Indian rainfall relationship was the result of the modulating influence of tropical Atlantic sea surface temperatures (SSTs) and the Atlantic Multi-decadal Oscillation modulating the relationship (Lu et al., 2006; Kucharski et al. 2007; Kucharksi et al., 2009; Chen et al., 2010). In contrast, Kumar et al. (2006) and Fan et al. (2017) argued that the occurrence of different ENSO flavors (Johnson, 2013) such as the Eastern Pacific and Central Pacific types could explain the changes in the ENSO-Indian rainfall relationship changes. Other investigators adopted another perspective to explain changes in the ENSO-Indian rainfall relationship and concluded that temporal undulations in the ENSO-Indian rainfall relationship are related to statistical under-sampling and stochastic fluctuations (Gershunov et al. 2001; van Oldenborgh and Burgers, 2005; Delsole and Shukla, 2006; Cash et al., 2017). In a recent analysis, Yun and Timmermann (2018) showed that changes in the ENSO-Indian rainfall relationship are consistent with a stochastically perturbed ENSO signal and argued that changes in the ENSO-Indian monsoon relationship may not be related to external climate forcing mechanisms.

The second reason for the ENSO-related prediction challenges is that ENSO itself is a non-stationary phenomenon. Using wavelet analysis, Kestin et al. (1998) found that the interannual variability of ENSO from 1930 to 1960 was dominated by a 4- to 7- year periodicity, whereas for the time period from 1960 to 1990, the interannual variability was also dominated by a 2- to 5- year periodicity. A wavelet power spectral analysis conducted by Torrence and Webster (1999) and Schulte (2016a) showed that ENSO signal energy in the 2- to 7-year period band undulates throughout the historical period and increases after the 1960s (Schulte 2016a).These changes in spectral characteristic are relevant to Indian Monsoon prediction because differing spectral characteristics among predictors (e.g. ENSO) and predictands (e.g. Indian rainfall) can negatively impact the predictive skill of statistical models (Jiang et al., 2020). Using a wavelet-based variance transformation method, Jiang et al. (2020) demonstrated that accounting for differences in spectral characteristics can improve prognostic skill. That study suggests that Indian Monsoon prediction could be improved using wavelet-based methods instead of time-domain correlation and regression methods.

The nonlinear characteristics (e.g. skewness) of ENSO are also non-stationary and undergo interdecadal changes (Wu and Hsieh, 2003). Numerous studies have reported an ENSO regime shift in the 1970s in which ENSO began to evolve more nonlinearly than in previous decades (An, 2004; An and Jin 2004; An, 2009). It is a curious fact that the ENSO regime shift of the 1970s coincided with the weakening ENSO-Indian rainfall relationship as documented by Kumar et al. (1999). This observation begs the question as to whether nonlinear ENSO regime changes are related to changes in the ENSO-Indian rainfall relationship.

Various mechanisms have been proposed for explaining the cause of ENSO skewness. Kang and Kug (2002) suggested that the asymmetry between the magnitude of El Niño and La Niña events is related to the relative westward displacement of zonal wind stress anomalies during La Niña events compared to El Niño events. Jin et al., (2003) and An and Jin (2004) found that ENSO asymmetry is related to nonlinear dynamical heating (NDH), where the magnitude of NDH is related to the propagation characteristics of ENSO. As shown by An and Jin (2004), NDH during strong El Niño events like the 1982/1983 and 1997/1998 events tends to be stronger than that of weak El Niño events because SST anomalies tend to propagate eastward. Since the late 1970s there has been a propensity for eastward propagation characteristics of ENSO (Santoso et al., 2013), contrasting with the time period before the 1970s that consisted of the relatively weak El Niño events of 1957/1958 and 1972/1973 (An and Jin, 2004; An, 2009). More recently, Su et al. (2010) showed that vertical temperature advection may have an opposing effect on ENSO asymmetry and that the asymmetry in the extreme eastern equatorial Pacific is related to meridional ocean temperature advection.

Previous investigators have used different metrics to quantify ENSO asymmetry. To measure the nonlinear character of ENSO, An and Jin (2004) used time-domain metrics such as skewness and maximum potential intensity (MPI) to quantify the skewness of SST anomalies and the skewness of individual ENSO events, respectively. An (2004) applied a principal component analysis (PCA) to a 21- year moving window of tropical Pacific SST skewness and found that the first PCA mode is characterized by positive skewness across the eastern equatorial Pacific and negative skewness across the central equatorial Pacific. This pattern means that interdecadal changes in the

nonlinearity of ENSO is associated with positively skewed SST anomalies across the eastern equatorial Pacific, implying that El Niño events are stronger than La Niña events. While the methods implemented in the aforementioned studies provided important insights, they cannot reveal the frequency modes of ENSO that are contributing to the skewness.

Recognizing the limitations of time-domain approaches, Timmermann (2003) conducted a bi-spectral analysis of the Niño 3 anomaly time series, where a peak (f_1,f_2) in the bi-spectrum means there is statistical phase dependence among oscillators with frequencies f_1, f_2, and $f_1 + f_2$. That bi-spectral analysis revealed statistically significant bi-spectral power at several frequency pairs, including (0.038, 0.038), (0.028, 0.028), (0.0225, 0.0225), (0.0045, 0.032), and (0.0045, 0.045) [month-^{1}]. The peaks (0.0045, 0.032), and (0.0045, 0.045) [months^{-1}] were identified with the nonlinear interactions among 18-year and 2-year variability. Although the analysis provided new insights, the Fourier-based analysis could not reveal how the nonlinear nature of ENSO changed with time, an important property to capture given how the nonlinear characteristics of ENSO are non-stationary (Santoso et al., 2013). Much like the cross-wavelet power (Maraun and Kurths, 2004) and time-domain covariance, bi-spectral power is not a bounded quantity and so high bi-spectral power does not always mean strong phase dependence.

In this study, the deficiencies associated with the above-mentioned techniques are addressed using higher-order wavelet analysis, which allows for the quantification of frequency-dependent and non-stationary nonlinearities in time series (Van Millagan, 2004, Elsayad, 2006; Schulte, 2016b). More specifically, the objectives of the paper are the following: 1) quantify the nonlinearity of ENSO using higher-order wavelet analysis together with recently developed statistical tests; (2) Determine if different nonlinear modes of ENSO are associated with distinct SST patterns; and (3) develop nonlinear wavelet coherence methods to test the hypothesis that the breakdown of the ENSO-Indian rainfall relationship in recent decades is related to the shift of ENSO from a linear regime to a nonlinear one. The paper is organized as follows: In Section 2, data used are described. Section 3 includes the description of the implemented methodologies. Results are presented in Section 4 and concluding remarks are provided in Section 5.

2. Data

The variability of India rainfall from 1871-2016 was analyzed using the All-India rainfall (AIR; Parthasarathy et al. 1994) time series, which was created by averaging representative rain gauges at various locations across India. The full monsoon season (June-September) and the late monsoon (August-September) season were used to identify possible within-season variations in the ENSO-All-India relationships. To remove the influence of the annual cycle, AIR time series was converted into anomaly time series by subtracting the 1871-2016 long-term mean for each month from the individual monthly values. The AIR anomaly (AIR, hereafter) time series were subsequently standardized by dividing it by its 1871-2016 standard deviations. Because wavelet analysis focuses on specific frequency components that are not impacted by long-term time-domain trends, no detrending of the data was performed.

The monthly data for the Niño 1+2, Niño 3, Niño 3.4, and Niño 4 indices (available at: https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino34.long.data) from 1871 to 2016 were used to understand how the nonlinear characteristics of SSTs varied from one ENSO region to another. The Niño 1+2 index is the average SST in the region with latitudinal boundaries 0° and 10°S and longitudinal boundaries 90°W and 80°W and the Niño 3 index is the average SST in the region with latitudinal boundaries 5°N and 5°S and longitudinal boundaries 150°W and 90°W. Variations in SSTs further west were described using the Niño 3.4 and Niño 4 indices, where the Niño 3.4 index is defined as the average SST in the region bounded by 5°N and 5°S and 170°W and 120°W and the Niño 4 index is defined as average SSTs in the region bounded by 5°N and 5°S and 160°E and 150°W. The seasonal cycle was removed from these time series in the same way as it was removed from the All-India rainfall time series.

The monthly SST data from 1871-2016 were based on the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST1; Rayner et al., 2003) The data at each grid point were converted to monthly anomalies in the same way as they were computed for the ENSO and AIR All-India time series.

3. Methods

3.1 Wavelet Analysis

To better diagnose changes in time series statistics associated with AIR and ENSO, we adopted a wavelet analysis. For a time series, X, comprising data points $x_1, x_2, \ldots, x_N$, the continuous wavelet transform is given by

$$W_n(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0 \left[(n' - n) \frac{\delta t}{s} \right] \quad (1)$$

where s is wavelet scale, ψ_0 is an analyzing wavelet, δt is a time step (1 month in this study), and n is time. The sample wavelet power spectrum $|W_n(s)|^2$ measured the energy content of a signal at time n and scale s. The commonly used Morlet wavelet with angular frequency ω = 6 was used throughout this paper because it balances time and frequency localization and because it is commonly used in hydrological and climate studies (Schaefli et al., 2007; Zhang et al., 2007; Holman et al., 2001; Carey et al., 2013). The readers are referred to Torrence and Compo (1998) and Grinsted et al. (2004) for more details about wavelet analysis.

Linear wavelet coherence (Table 1) was used to quantify the linear relationship between two time series as a function of frequency and time. Linear wavelet coherence between two time series X and Y is given by

$$R_n^2(s) = \frac{\left| S s^{-1} W_n^{XY}(s) \right|^2}{S\left(s^{-1} |W_n^X(s)|^2 \right) S\left(s^{-1} |W_n^Y(s)|^2 \right)}, \quad (2)$$

where S is a smoothing operator (Grinsted et al., 2004) and $W_n^{XY}(s)$ is the cross-wavelet power spectrum. Two time series are perfectly coherent ($R_n^2(s)$ = 1) at s if $\phi_n^X(s) - \phi_n^Y(s) = c$ over a sufficiently long time interval, where c is a constant, $\phi_n^X(s)$ is the phase associated with X, and $\phi_n^Y(s)$ is the phase associated with Y.

In the context of the Indian monsoon, strong coherence between rainfall and a climate pattern (e.g. ENSO) at a scale s indicates shared temporal characteristics between a climate pattern and rainfall. Because theory supports a causal link between ENSO and monsoon variability through changes in the Walker Circulation (Ropelewski and Halpert, 1987), strong coherence means that when ENSO is in a warm (cool) phase at the scale s, negative (positive) rainfall anomalies are preferred. Thus, a periodic climate forcing could create periodicities in an otherwise noisy rainfall time series.

3.2 Higher-order Wavelet Analysis

Although the wavelet power spectrum is useful for quantifying the signal energy at a scale s and time n, it cannot determine if there is a nonlinear relationship among different frequency components. In fact, the power spectrum can only fully describe time series in frequency space in the case of linear systems in which the output is proportional to the input (King, 1998). As ENSO is nonlinear, we adopted higher-order wavelet methods to address the deficiencies of traditional wavelet methods.

The type of nonlinearities considered in this study were quadratic nonlinearities in which the scales s_1, s_2, and s_3 satisfied the sum rule

$$\frac{1}{s_3} = \frac{1}{s_1} + \frac{1}{s_2} \quad (3)$$

and the wavelet phases satisfied

$$\phi_n(s_3) = \phi_n(s_1) + \phi_n(s_2). \quad (4)$$

These types of nonlinearities arise, for example, when a sinusoid is squared, in which case a harmonic is produced. More generally, quadratic nonlinearities induce time series skewness, which was computed in this study using

$$skewness = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^3}{S}, \quad (5)$$

where S in the standard deviation of the time series and $\bar{x}$ is the mean of the time series. Positive (negative) skewness meant that the right (left) tail of the time series distribution is longer than the left (right). In other words, positive

(negative) skewness meant that there was a tendency for positive (negative) time series events (i.e. anomalies) to be more intense than negative (positive) ones.

In this paper, quadratic nonlinearities giving rise to time series skewness were quantified using local and global wavelet-based auto-bicoherence methods (Schulte, 2016b). Global auto-bicoherence (Table 1) was computed using the equation

$$bi^X_{global}(s_1, s_2) = \frac{\left|B^X_{global}(s_1 s_2)\right|^2}{\left(\sum_{n=1}^{N}\left|W_n^X(s_1)W_n^X(s_2)\right|^2\right)\left(\sum_{n=1}^{N}\left|W_n^X(s_3)\right|^2\right)}, \quad (6)$$

where

$$B^X_{global}(s_1, s_2) = \sum_{n=1}^{N} \widehat{W}_n^X(s_3) W_n^X(s_1) W_n^X(s_2) \quad (7)$$

is the global bi-spectrum and the hat denotes the complex conjugate. Identical to wavelet coherence, auto-bicoherence is bounded by 0 and 1, a value of 1 indicating the strongest possible phase coupling among the phases $\phi_n(s_3)$, $\phi_n(s_2)$, and $\phi_n(s_1)$ such that sum rules Eq. (3-4) is satisfied. A peak in the auto-coherence spectrum at (s_1, s_2) indicated that there was quadratic phase coupling among oscillatory modes with scales s_1, s_2, and s_3 so the oscillatory modes was contributing to time series skewness. It is important to note that the auto-bicoherence method cannot detect other types of nonlinearities such as cubic nonlinearities whose detection would require trisepctra (Collis et al., 1998).

To determine if the strength of the quadratic phase coupling was a function of time changed temporally, the local auto-bicoherence spectrum (Schulte, 2016b) given by

$$bi^X_n(s_1, s_1) = \frac{\left|s s_1^{-1} {B^X}_n(s_1, s_1)\right|^2}{s\left(s_1^{-1}\left|W_n^X(s_1)W_n^X(s_1)\right|^2\right) s\left(s_1^{-1}\left|W_n^X(\frac{s_1}{2})\right|^2\right)} \quad (8)$$

was computed, where $B_n^X(s_1, s_1)$ is the local bi-spectrum given as

$$B_n^X(s_1, s_1) = \widehat{W}_n^X(s_3) W_n^X(s_1) W_n^X(s_1). \quad (9)$$

and $s_3 = s_1/2$. In this special case, the local auto-bicoherence spectrum revealed the time-evolution of auto-bicoherence estimates located along the diagonal slices of the global auto-bicoherence spectra. Local bi-phase

$$\psi_n(s_1, s_1) = \phi_n(s_1) + \phi_n(s_1) - \phi_n(s_3) \quad (10)$$

was used to measure the skewness and asymmetries of waveforms. A bi-phase of 0◦ meant that the relationship among the scale components produced positive skewness with respect to a horizonal axis so that positive deviations from the mean are larger than negative deviations from the mean. On the other hand, a bi-phase of 180◦ indicated negative skewness with respect to the mean. Bi-phases near -90◦ or 90◦ indicated the presence of asymmetric cycle geometry (King, 1998; Maccarone, 2014; Schulte, 2016b) such that a time series rose (fell) more quickly than it fell (rose) (King, 1998; Maccarone, 2014; Schulte, 2016b).

To be consistent with the wavelet power and coherence analyses, results for the higher-order wavelet analysis were casted in terms of Fourier period rather than wavelet scale. The Fourier period corresponding to s_i was denoted by p_i, where the Fourier period is obtained by multiplying s_i by 1.03 for the Morlet wavelet (Torrence and Compo, 1998). Thus, the local diagonal slice of the auto-bicoherence spectra were plotted using the Fourier period p_1 corresponding to s_1 as the vertical axis and time as the horizonal axis.

3.3 Statistical Hypothesis Testing

The statistical significance of all wavelet spectra was evaluated using the cumulative area-wise test (Schulte, 2016a; Schulte, 2019) to account for the simultaneous testing of multiple hypotheses (Maraun and Kurths, 2004; Maraun et al., 2014). This test evaluated the statistical significance of points in the wavelet domain based on the area of contiguous regions of point-wise significance (i.e. patches) to which they belong so that larger area implies greater statistical significance. Given that patch area can change as the point-wise significance changes, the cumulative area-

wise test was used to evaluate significance based on patch area the area of a patch averaged across a set of point-wise significance levels (Schulte, 2019). The test was applied at the 5% cumulative area-wise significance level using point-wise significance levels ranging from 0.02 to 0.18 because this choice of point-wise significance levels was shown to result in the cumulative area-wise test outperforming the point-wise test in terms of true positive detection for moderate to high signal-to-noise ratios even though the cumulative area-wise test is more stringent. The test was performed using the Advanced Biwavelet Wavelet R software Package (available at: http://justinschulte.com/wavelets/advbiwavelet.html). Technical details of the testing procedure can be found in Schulte (2019) and in Appendix A.

To assess the statistical significance of the global auto-bicoherence estimates, a modified version of the cumulative area-wise test was applied. In the modified version of the cumulative area-wise test, the normalized area of patches was computed by dividing patch area by the product $\widehat{s_1}\widehat{s_2}$, where $\widehat{s_1}$ is the mean first-coordinate of the patch and $\widehat{s_2}$ is the mean second coordinate. The reason for this modified normalized area is that dividing area by say, $\widehat{s_1}^2$, retained the correlation between normalized area and s_2. The test was applied using the same point-wise significance levels that were used to assess the statistical significance of wavelet power and coherence.

3.4 Higher-order Coherence

Although wavelet coherence spectra can provide information regarding how the relationship between two climate variables changes at a scale *s*, it cannot completely explain why the time-domain correlation between the climate variables temporally fluctuates. The reason is that linear wavelet coherence only examines how well the variance of one time series corresponds to the variance of another at a scale *s* (Table 1) because linear coherence is determined by the wavelet power spectra of the time series. However, for two time series to be perfectly correlated in the time domain, higher skewness of one climate variable must also correspond to higher skewness of the other climate variable.

Recognizing that skewness is important for better understanding time-domain correlation changes, the quantity

$$Bi_n^2(s_1,s_2) = \frac{\left|Ss_{smooth}^{-1}B_n^{XYY}(s_1,s_2)\right|^2}{S\left(s_{smooth}^{-1}\left|B_n^X(s_1,s_2)\right|^2\right)S\left(s_{smooth}^{-1}\left|B_n^Y(s_1,s_2)\right|^2\right)}, \quad (11)$$

called third-order coherence (nonlinear coherence, hereafter) was used to determine if changes in the skewness of *X* are associated with changes in the skewness of *Y* (see Appendix B for a more general definition). In Eq, (11), s_{mooth} is one of the three scales, and $B_n^{XY}(s_1,s_2)$ is the third-order cross-wavelet power spectrum, which is the product of the bi-spectrum of *X* and the conjugate of the bi-spectrum of *Y*, the higher-order analog of the cross-wavelet power spectrum. The word cross-bispectrum was not used to avoid confusion with cross-bicoherence analysis (Van Milligen,1995). Like wavelet coherence, the nonlinear coherence is bounded by 0 and 1, a value of 1 indicating that the bi-spectra of *X* and *Y* at (s_1,s_2) are perfectly and linearly correlated. The statistical significance of nonlinear coherence was assessed using Monte Carlo methods and the cumulative area-wise test in the same way as it was used to assess the statistical significance of linear wavelet coherence.

Higher-order wavelet analysis can also be interpreted in terms of Another way to interpret higher-order wavelet coherence is using linear and nonlinear modes. A linear mode $\gamma_{s_i}^X$ is the signal component of *X* at the scale s_i obtained by setting all wavelet coefficients to zero except those at s_i and taking the inverse wavelet transform of the result. Because linear modes are only composed of a single frequency component, the local cross-correlation (coherence) between $\gamma_{s_i}^X$ and $\gamma_{s_i}^Y$ is only impacted by the variances of *X* and *Y* at s_i. On the other hand, nonlinear coherence measures the local cross-correlation between the skewness of $\gamma_{s_1}^X+\gamma_{s_2}^X+\gamma_{s_3}^X$ and $\gamma_{s_1}^Y+\gamma_{s_2}^Y+\gamma_{s_3}^Y$ or between $\gamma_{s_1}^X+\gamma_{s_1/2}^X$ and $\gamma_{s_1}^Y+\gamma_{s_1/2}^Y$ in the case that $s_1=s_2$.

To better understand nonlinear coherence, we supposed that

$$\phi_n^X(s_1)-\phi_n^Y(s_1)=c_1 \quad (12)$$

$$\phi_n^X(s_2) - \phi_n^Y(s_2) = c_2 \quad (13)$$

$$\phi_n^X(s_3) - \phi_n^Y(s_3) = c_3 \quad (14)$$

for constants c_1, c_2, and c_3. Adding Eqs. (12) and (13) and subtracting Eq. (14) from the result produced the equality

$$\phi_n^X(s_1) + \phi_n^X(s_2) - \phi_n^X(s_3) - (\phi_n^Y(s_1) + \phi_n^Y(s_2) - \phi_n^Y(s_3)) =$$

$$\psi_n^X(s_1, s_2) - \psi_n^Y(s_1, s_2) = \psi_n^{bi}(s_1, s_2) = K\ , \quad (15)$$

for some constant $K = c_1 + c_2 - c_3$. Thus, if *X* was found to be perfectly nonlinear coherent with *Y*, then *X* and *Y* must be were perfectly coherent at the three scales participating in the quadratic phase coupling. Even if the coherence was perfect at two scales, the relative bi-phase $\psi_n^{bi}(s_1, s_2)$ will fluctuated randomly if the relative phase difference at the remaining scale fluctuated randomly so that the nonlinear coherence was will be low. Thus, if nonlinear coherence was high, then there was some non-random relationship between *X* and *Y* at all three scales even if high linear coherence was not identified at one or more scales. This theoretical idea indicates that nonlinear coherence can uncover relationships that linear coherence cannot (see Figure S1 in supplementary material).

The relative bi-phase difference $\psi_n^{bi}(s_1, s_2)$ is the higher-order analog of the relative phase difference between two time series. It measures how much the cycle geometry of one time series lags that of another. A lagged bi-phase of 180◦ means that the skewness or asymmetry of the forcing time series is opposite to that of the response. For example, if the forcing has positive skewness, then the response will have negative skewness. If the relative bi-phase is 0◦, then negative (positive) skewness of the forcing produces negative (positive) skewness of the response, contributing to the positive time-domain correlation between the time series. Scales and time points for which nonlinear coherence is high are where the relative bi-phase is stable.

In this paper, we focused on nonlinear coherence computed along the diagonal slices ($p_1 = p_2$) of the time series bi-spectra. The nonlinear coherence spectra was then plotted using p_1 as the vertical axis and time as the horizonal axis. High nonlinear coherence at p_1 and *n* meant that the skewness or asymmetry between $\gamma_{p_1}^X + \gamma_{p_1/2}^X$ and $\gamma_{p_1}^Y + \gamma_{p_1/2}^Y$ were locally cross-correlated.

To demonstrate the concept of nonlinear coherence, we considered a simple example in which the nonlinear climate forcing time series was given by

$$F(t) = \cos(\frac{2\pi}{p_1} t + \varphi) + \gamma\,(t) \cos(\frac{2\pi}{p_3} t + 2\varphi) + W_F(t) \quad (16)$$

and the response to the forcing was given as

$$R(t) = \cos(\frac{2\pi}{p_1} t + \varphi) + W_R(t), \quad (17)$$

In Eq. (16) and Eq. (17), $\gamma\,(t)$ is a time-varying nonlinear coefficient, $W_F(t)$ is Gaussian white noise associated with the forcing, $W_R(t)$ is Gaussian white noise associated with the response, $\varphi = 0$ is phase, and $p_1 = 2p_3 = 32$. The nonlinear coefficient was assumed to be a linear function of time, i.e.,

$$\gamma(t) = t/500. \quad (18)$$

The effect of the coefficient was to linearly increase the variance of $F(t)$ at $p_3 = 16$ and increase the strength of the quadratic phase coupling between the modes with periods $p_3 = p_1/2 = 16$ and $p_1 = 32$.

As shown in Figure 1a, $F(t)$ (black curve) and $R(t)$ (thick green curve) evolve coherently from $t = 0$ to $t = 200$. After $t = 200$, $F(t)$ begins to noticeably exceed $R(t)$ at certain time points (e.g. $t = 430$) while the relationship between them at other points is reversed (e.g. $t = 450$) in the sense that a positive forcing produces a negative response. As a result, the correlation between $F(t)$ and $R(t)$ weakens (Figure 1b). An inspection of the wavelet coherence spectrum (Figure 2a) reveals that the coherence at $p_1 = 32$ is strong and stable so that changes in the relationship strength at that time scale is not the cause of the weakening time-domain correlation. The coherence at all other periods

is also stationary by construction so that it is not the changing relationship strength at any time scale that is causing the time-domain correlation weakening. However, the variance of $F(t)$ at $p_3 = 16$ increases with time (not shown) and the coherence between $F(t)$ and $R(t)$ is also weak at that time scale, implying that larger fluctuations in $F(t)$ at $p_3 = 16$ are not accompanied by larger fluctuations in $R(t)$. Thus, variance increase of $F(t)$ is one reason for the weakening time-domain correlation, though the linear coherence and wavelet power methods cannot explain why the skewness of $F(t)$ increases without a corresponding increase in the skewness of $R(t)$ (Figure 1c).

An inspection of the local auto-bicoherence spectrum of $F(t)$ (Figure 2b) reveals that the auto-bicoherence at $p_1 = 32$ is increasing with time, indicating that the phase coupling between modes with periods $p_3 = 16$ and $p_1 = 32$ is strengthening with time. The bi-phase of 0◦, as indicated by arrows pointing to the right, confirms that the quadratic phase coupling is contributing to the positive skewness seen in Figure 1a to an increasing degree. Furthermore, the nonlinear coherence between $R(t)$ and $F(t)$ is weak and mostly statistically insignificant at $p_3 = 32$ (Figure 2c), implying the skewness of $\gamma_{16}^F + \gamma_{32}^F$ is uncorrelated with the skewness of $\gamma_{16}^R + \gamma_{32}^R$, where $\gamma_{16}^F + \gamma_{32}^F$ is the sum of the cosines in Eq. (16) and the components of $W_F(t)$ at $p_3 = 16$ and $p_1 = 32$. The nonlinear mode $\gamma_{16}^R + \gamma_{32}^R$ is the sum of the cosine in Eq. (17) and the components of $W_R(t)$ at $p_3 = 16$ and $p_1 = 32$. Thus, the skewness of $R(t)$ in the time-domain is practically uncorrelated with the skewness of $F(t)$ because the skewness of $F(t)$ is solely related to the phase coupling between the modes with periods $p_3 = 16$ and $p_1 = 32$ Thus, the increase in skewness of $F(t)$ also contributes to the weakening time-domain correlation.

The lack of nonlinear coherence at time scales for which $F(t)$ is nonlinear has implications for empirical prediction. At time points when $F(t)$ is positively skewed, $R(t)$ is overestimated because $R(t)$ is not inheriting the skewness of $F(t)$. That is, if one created a linear regression model based on the relationship between $F(t)$ and $R(t)$ from $t = 0$ to $t = 200$ one would find that a forcing value of, say, 1 would produce a response close to 1. If the same model was used to predict $R(t)$ at, say, $t = 430$ one would predict that the forcing with value around 2 should result in a response near 2. However, because the relatively large value $F(430)$ results from skewness and $R(t)$ is uncorrelated with its skewness, the response is only as strong as the part of $F(t)$ not resulting from the quadratic phase coupling. The more nonlinear $F(t)$ becomes, the more $F(t)$ will overestimate $R(t)$ when $F(t)$ is positively skewed. Similarly, the positive forcing produces a negative response at $t = 450$ because of skewness and not simply a change in variance. Nonlinear coherence allows for the quantification and identification of these time-domain aberrations.

The weakening relationship shown in Figure 1b could lead a researcher studying a hydrological process to believe that another direct forcing mechanism must be influencing the hydrological process. This belief could lead to the applications of partial wavelet coherence (Ng, and Chan, 2012) and partial correlation analyses to identify another influential forcing mechanism. However, in this case, there are no other direct forcing mechanisms; the weakening time-domain relationship is solely related to how $F(t)$ transitioned from a linear regime to a nonlinear regime. This theoretical result suggests that hydrological studies using wavelet coherence should also consider the nonlinearity of the times series.

4. Results

4.1 The ENSO and Indian Monsoon time series and their time-domain relationship

The contrasting Niño 1+2 and Niño 4 indices are shown in Figures 3. For the Niño 1+2 time series (Figure 3a), a few recent notably intense warm events are located around 1982/1983, 1997/1998, and 2015/2016, coinciding with the strongest El Niño events in recent decades (McPhaden, 1999, Hu and Fedorov, 2017; Santoso et al, 2017). A few notably intense Niño 1+2 events are also seen in the 1800s, indicating that intense ENSO events are not unique to recent decades. An inspection of Figure 3a also reveals that the recent intense warm Niño 1+2 events are also skewed in the sense that they are stronger than the surrounding cool Niño 1+2 events. Unlike the 1982/1983 and 1997/1998 Niño 1+2 events, the 1982/1983 and 1997/1998 warm events for the Niño 4 time series are unremarkable. (Figure 3b). Furthermore in contrast, there is a tendency for cool Niño 4 events are preferentially stronger to be stronger than warm events the warm events after the 1960s, suggesting an intensification of negative skewness. has intensified.

The 20-year sliding skewness time series of the Niño 1+2 index (Figure 4) reveals enhanced skewness during the early 1800's, near zero skewness around the 1930's and early 1940's, and especially enhanced skewness after the 1970s associated with an upward trend in skewness beginning around the 1940s. In contrast to the Niño 1+2 index, the skewness of the Niño 4 index becomes more negative after the 1960s, and the magnitude of the skewness is generally smaller than that of the Niño 1+2 time series. This finding suggests that the transition of the Niño 1+2 time series to a nonlinear regime was is more pronounced than the transition associated with the Niño 4 time series.

Interestingly, a 20-year sliding skewness analysis of AIR (Figure 4) reveals that the skewness of June-September AIR remains close to zero until the 1990s despite the upward trend in Niño 1+2 skewness beginning in the 1940s (Figure 4a). Similarly, the skewness of August-September AIR does not increase to the extent that Niño 1+2 skewness does (Figure 4b). On the other hand, the skewness of August-September AIR could be negatively correlated with the skewness of the Niño 4 index after the 1960s, consistent with how August-September AIR and the Niño 4 index are negatively correlated. The skewness of June-September AIR becomes more negative in the 1990s and 2000s, but it is unclear if that negative skewness is related to ENSO, noise, or another climate pattern because the skewness of the Niño 1+2 and Niño 4 indices do not change as abruptly. Negative June-September AIR skewness is accompanied by enhanced positive skewness of the Niño 1+2 indices index prior to the 1940s, which is consistent with how June-September AIR is negatively correlated with the Niño 1+2 index time series during that time period (Figures 5a).

The differences in skewness shown in Figure 4 suggests that the correlation between the ENSO time series and AIR degrades after the 1970s, which is confirmed by the 20-year sliding correlation between June-September AIR and ENSO time series (Figure 5a). The relationship with the Niño 1+2 generally weakens from the 1800's to the 2000s. In contrast, the June-September AIR-Niño 4 index relationship appears to have no long-term trend, resulting in the Niño 4 index becoming more strongly correlated with AIR than the Niño 1+2 index after the 1970's. The relationship between AIR and time series for the Niño 3 and Niño 3.4 indices are also relatively weak after the 1970s (not shown).

The stronger AIR-Niño 4 index relationship compared to the AIR-Niño 1+2 index relationship after the 1970s is more evident in the August-September analysis (Figure 5b). An abrupt weakening of the August-September AIR-Niño 1+2 index relationship occurs around the 1970's, with the relationship reversing around the 1990s. A comparison of Figures 4b and Figures 5b reveals that the weakening and reversal of the relationship occurs during the time period when the Niño 1+2 index is especially skewed. The difference in the magnitudes of Niño 1+2 and Niño 4 skewness after the 1970s could explain why the August-September AIR-Niño 1+2 index relationship weakens more abruptly than the AIR-Niño 4 index relationship. Thus, a further investigation is needed to better understand the temporal changes in ENSO statistics and their impact on the ENSO-AIR relationship.

4.2. Wavelet Power Analysis and Coherence

The wavelet power spectra associated with the Niño 1+2 and Niño 4 time series (Figure 6) reveal enhanced variance in the 16- to 64-month band after 1965 for all the time series. For the Niño 3 and Niño 4 time series, there is also enhanced variance in the 16- to 64-month period band from 1875 to 1895. The appearance of holes in contoured regions suggests that there are oscillatory modes with nearby frequencies (Schulte, et al., 2015), though the wavelet power spectra cannot determine if there is quadratic phase coupling between the oscillatory modes.

The linear wavelet coherence spectrum shown in Figure 7, indicates that the AIR relationship with the Niño 1+2 and Niño 4 indices in the 16- to 64-month period band breaks down after 1995, which is consistent with the findings from the sliding correlation analysis shown in Figure 5. The relationship between AIR and these ENSO indices also weakens around 1925, but this weakening does not appear in the sliding correlation analysis. Note that the lack of linear coherence after 1995 coincides with the enhanced ENSO variance (Figure 6), implying that higher ENSO variance need not be associated with higher AIR variance at those time scales so that changes in ENSO variance could be contributing to the weakening ENSO-AIR time-domain correlation. However, ENSO skewness is also enhanced during this time period (Figure 4) so that the weakening relationships may not be simply related to ENSO variance. Thus, a further analysis is needed to extract information unrevealed by the linear wavelet power and coherence methods.

4.3 Local auto-bicoherence of ENSO

Figure 8 shows that the local auto-bicoherence spectra of all ENSO time series contain statistically significant local auto-bicoherence, but the spectrum of the Niño 4 index is only associated with a few statistically significant regions such as the one around 2015 at a period of 32 months. For the Niño 3 and Niño 3.4 time series, two features of interest are seen in the time period extending from 1973 to 2016 in the 16- to 64-month period band. The first feature is the time-elongated region of statistical significance extending from 1973 to 2016 around a period of 61 months. This result implies that after 1973 the nonlinear phase coupling between modes with periods of approximately 30.5 and 61 intensifies., consistent with the ENSO regime shift (Santoso et al., 2013). For the Niño 1+2 index, there is an intensification auto-bicoherence spectrum after the 1970s, which is consistent with the ENSO regime shift (Santoso et al., 2013). Although the exact periods associated with the phase-coupled oscillatory modes are more difficult to discern, Nevertheless, a comparison of Figures 4 and 8 reveals that enhanced skewness coincides with stronger auto-bicoherence in the 32- to 64-month period, suggesting that the skewness partially arises from the stronger quadratic phase coupling among oscillatory modes with periods ranging from 32 to 64 months. The correspondence between auto-bicoherence and time-domain skewness also holds for the Niño 3 and Niño 3.4 time series (not shown).

The second feature of interest in the Niño 3 and Niño 3.4 auto-bicoherence spectra is the one that emerges around 1995 at a period of 31 months, indicating that the onset of this quadratic phase coupling occurred well after the 1970s regime shift just before the 1997/1998 El Niño event. Thus, the nonlinear character of, say, the 1982/1983 El Niño is different from that of both the 1997/1998 and 2015/2016 El Niño events because of the additional quadratic phase coupling between the 15.5- and 31-month modes.

To confirm that the nonlinear phase coupling identified in Figure 8 is associated with skewed waveforms, we inspected the corresponding local bi-phase spectra (not shown). It was found that the bi-phase in the 42- to 64-month period band is generally 0° so that the nonlinear phase coupling in that period band contributes to the positive skewness of the 1982/1983, 1997/1998, and 2015/2016 events.

4.4 Nonlinear Coherence between All-India Rainfall and ENSO

The results shown in Figure 9 indicate that the nonlinear wavelet coherence between AIR and the time series for the all four ENSO indices Niño 1+2 and Niño 4 indices is statistically significant in the 32- to 64-month period band mainly prior to the 1980s. The nonlinear coherence in this period band appears to peak around the 1972/1973 El Niño event, indicating that an increase in positive skewness of ENSO should tend to coincide with enhanced negative skewness of AIR around this time. However, much of the statistically nonlinear coherence is located during the time period when ENSO is more linear than it has been in recent decades (Figure 8) so that the effects of nonlinearities are small regardless of the nonlinear wavelet coherence. In contrast, the auto-bicoherence of the Niño 1+2 time series in the 32- to 64-month period band is statistically significant and high after the 1970s (Figure 8) so that the lack of nonlinear coherence after 1980s shown in Figure 9 is expected to impact the time-domain correlation more strongly, much like the theoretical situation shown in Figures 1 and 2. Our results are consistent with this theoretical idea because the AIR-Niño 1+2 relationship weakens more than the AIR-Niño 4 relationship after the 1970s (Figure 5), which is expected because the Nino 1+2 index is more nonlinear than the Nino 4 index during this time period. However, unlike the theoretical example shown in Figure 2, the linear coherence between the ENSO time series and AIR also weakens around the 1990s (Figure 7) so that the weakening relationship could be the result of a combination of factors that includes ENSO nonlinearity.

The 20-year sliding mean of the ENSO auto-bicoherence, coherence, and nonlinear coherence averaged in the 32 to 64-month period band further highlights the impact of ENSO nonlinearity. As shown in Figure 10a, the sliding mean nonlinear coherence between the Niño 1+2 index and AIR fluctuates less than linear coherence and reaches a clear global maximum around the 1970s before rapidly declining to a global minimum around the late 1990s when the Niño 1+2 index is very nonlinear. As shown in Figure 10a, the Niño 1+2 auto-bicoherence peaks around the same time that the August-September AIR-Niño 1+2 index correlation is positive. In fact, the correlation between the sliding September-August AIR-Niño 1+2- correlation time series and the sliding Niño 1+2 auto-bicoherence time series is 0.81, much higher than the correlation with the linear coherence (r = -0.11) and nonlinear coherence (r = -0.34) time series. These results support the idea that the Niño 1+2 regime shift impacted the weakening time-domain

correlation. On the other hand, the correlation between Niño 4 auto-bicoherence and the August-September AIR-Niño 4 correlation time series is weak so that changes in the nonlinearity of the Niño 4 index unlikely contributes substantially to changes in the AIR-Niño 4 relationship. Nevertheless, this result agrees with theory that suggests that nonlinearity is only an important contributor when the timeseries is highly nonlinear, which is not the case for the Niño 4 index because of the low auto-bicoherence (Figures 8b and 10b). Because the nonlinear coherence between AIR and indices for the Niño 1+2 and Niño 4 is weak (Figures 9 and 10), the more pronounced change in the August-September AIR-Nino 1+2 correlation reflects the more intense increase in Niño 1+2 nonlinearity compared to that of the Niño 4 index in recent decades.

4.5. A possible explanation for the ENSO Nonlinearity Impacts

To better understand the association between ENSO nonlinearity and the AIR-ENSO relationship, the global auto-bicoherence spectra associated with the ENSO time series were first computed (Figure 11). Then, the auto-bicoherence of SSTs associated with a few select peaks (p_1, p_2) in Figure 8 were computed at each grid point in the domain bounded by 20°N and 20°S and by 146°E and 80°W. The peaks were selected based on the auto-bicoherence spectra of the Niño 3.4 and Niño 1+2 indices. To select the peaks, local maxima in auto-bicoherence within the statistically significance regions shown in Figure 11 were identified, where points associated with local maxima were chosen because they were associated with the clearest patterns.

The spatial structure of global auto-bicoherence corresponding to the peaks in the Niño 3.4 auto-bicoherence spectrum are shown in Figure 12. The auto-bicoherence associated with the pair (31, 31) is greatest across the central equatorial Pacific, with the overall spatial pattern being reminiscent of a central Pacific El Niño (Lee and McPhaden, 2010). This result suggests that the phase coupling between the 31-month mode and the 15.5-month mode could be related to the occurrence of central Pacific El Niño events (Section 5). In contrast, the auto-bicoherence pattern associated with the pair (56, 56) is more uniform, with auto-bicoherence slightly greater across the extreme eastern equatorial Pacific than the central equatorial Pacific. This pattern is reminiscent of an eastern Pacific El Niño. Like the pattern corresponding to the pair (31, 31), the auto-bicoherence for the pair (105, 57) tends to be greater across the central equatorial Pacific. Our findings suggest that different nonlinear modes contribute to different ENSO flavors. Although An and Jin (2004) and Burgers and Stephenson (1999) showed that skewness is greatest across the eastern equatorial Pacific, we determined that such a time-domain approach is unable to capture frequency-dependent patterns in nonlinearity.

The spatial auto-bicoherence plots associated with the peaks in the Niño 1+2 auto-bicoherence spectrum are shown in Figure 13. The auto-bicoherence associated with the pairs (148, 53) and (148, 105) is strong across the eastern equatorial Pacific but weak across the central equatorial Pacific, suggesting that the quadratic phase coupling between the 148- and 105-month modes and between the 148- and 53-month modes are associated with the skewness of eastern equatorial Pacific SSTs. The pattern associated with the pair (62, 44) is reminiscent of an eastern Pacific El Niño and the auto-bicoherence associated with the pair (88, 88) is relatively weak across the entire equatorial Pacific. A comparison of Figures 12 and 13 shows that there is a tendency for auto-bicoherence to be greater across the eastern equatorial Pacific than the central equatorial Pacific, which is consistent with how SSTs across the eastern equatorial Pacific are most skewed (Burgers and Stephenson, 1999; An and Jin, 2004).

5. Discussion/Conclusion

The nonlinear nature of ENSO was examined using higher-order wavelet methods. The auto-bicoherence spectra of the four ENSO Niño 1+2 and Niño 4 indices time series revealed that ENSO skewness arose from the quadratic phase coupling of modes with various periods. For the Niño 1+2 index, the quadratic phase coupling after the 1970s was especially strong, which is consistent with how ENSO underwent a regime shift around the 1970s (Santoso et al., 2013) marked by an increase in ENSO skewness. Although the Niño 3.4 time series was not considered in detail in this study, an auto-bicoherence analysis of the time series (not shown) revealed phase coupling between modes with periods of 31 and 15.5 months in addition to coupling between modes with periods of 61 and 30.5 months. The phase coupling between the 31 and 15.5 modes was found to be especially strong after 1995, whereas the quadratic phase coupling between the 61- and 30.5-month modes was found to intensify after the 1970s. These additional results suggest that nonlinear modes vary in intensity and time of occurrence.

The evolution of SSTs across the Niño 4, Niño 3.4, Niño 3, and Niño 1+2 regions was found to be nonlinear, but the degree to which the time series are nonlinear are different (Figure 11). Overall, the Niño 1+2 time series was found to be the most nonlinear, while the Niño 4 index was found to be the most linear. The spatial patterns associated with the nonlinearities depend on the frequency components contributing to the nonlinearities. For example, the quadratic phase coupling between the modes with periods of 31 and 15.5 months was found to be strongest in the central equatorial Pacific and weakest across the eastern equatorial Pacific. This finding suggest that the more frequent occurrence of central Pacific El Niño events in recent decades (Lee and Mcphaden, 2010) could be linked to the strengthening of this quadratic phase coupling, which could explain the relationship between ENSO nonlinearity and changes in the ENSO-AIR because central Pacific El Niño events have been shown to be more effective at creating drought-inducing subsidence over India (Kumar et al., 2006).

The results from the present and previous studies (Fan et al. 2017) supports the idea that changes in the ENSO-AIR relationship are related to ENSO flavors because ENSO nonlinearity appears to be related to ENSO flavors (Figures 12 and 13), opposing the findings of other work showing that the changes are related to sampling variability or to noise. According to Yun and Timmermann (2018), the changes in the time-domain correlation between AIR and ENSO is consistent with the assumption that AIR is the sum of the ENSO signal and Gaussian white noise (i.e., AIR = ENSO + white noise). However, for this hypothesis to hold, the difference AIR – ENSO must be Gaussian white noise. As shown in this study, the nonlinear wavelet coherence between ENSO metrics and AIR is weak so that ENSO – AIR contains periodicities (Figure S2), which means that AIR is not simply a stochastically perturbed ENSO signal, as noise does not contain periodicities. The retention of non-Gaussian noise features was certainly the case for $R(t)$ – $F(t)$ in the example in Section 3.5 because the difference retains the cosine function with a period of 16.

The fact that nonlinear coherence between rainfall and ENSO is determined by linear coherence between ENSO and rainfall at two or three frequencies means that the changing time-domain correlation could be more fully understood by determining why linear coherence changes at the frequencies that contribute to ENSO skewness. Such an analysis could provide a more mechanistic perspective than the theoretical perspective adopted in this study. A preliminary analysis showed that enhanced linear coherence between the North Atlantic Oscillation index and AIR after 1995 in the 16- to 64-month period band associated with ENSO nonlinearity. This result suggests that conditions across the North Atlantic (Kakade, 2000, Bhatla, 2016) could influence the nonlinear coherence between ENSO and AIR and thus the corresponding time-domain correlation.

The tools used and developed in this study may have important applications in understanding how forecasting systems replicate Indian rainfall and its associated teleconnections. These methods, for example, could determine if forecasting systems can reproduce nonlinear characteristics of climate time series. As such, a R software package has been developed to implement these methods (available at: http://justinschulte.com/wavelets/advbiwavelet.html; last access: 22 November 2019). These methods could provide new directions for improving current forecasting systems and ultimately predictions of Indian rainfall.

Appendix A

The first step (STEP 1) in assessing the cumulative-area significance of a point was the calculation of the $N = 12$ sets

$$P_{pw}^{i} = \{(b,a): \rho_{pw}(b,a) < \alpha_i\}, \quad \text{(A1)}$$

where each set is the subset of the wavelet domain consisting of points whose wavelet quantities are point-wise statistically significant at the α_i significance level (i.e. the point-wise test p-value, ρ_{pw}, is less than α_i). In this paper, $\alpha_1 = 0.02$, $\alpha_{12} = 0.18$, and $\alpha_{i+1} - \alpha_i = 0.02$. In the second step (STEP 2), a geometric pathway about x was computed, where a geometric pathway was a nested sequence

$$P_1^x \subseteq P_2^x \subseteq \cdots \subseteq P_N^x \quad \text{(A2)}$$

such that the

$$P_i^x = \{(b,a): (b,a) \in P_{pw}^{i}, (b,a) \sim x\} \quad \text{(A3)}$$

were path-components of P_{pw}^{i} containing x. The equivalence relation $\sim$ on P_{pw}^{i} made two points in P_{pw}^{i} equivalent because they could be connected by a continuous path in P_{pw}^{i}. The third step (STEP 3) involved the calculation of the normalized area corresponding to P_i^x. The normalized area, A_i^x, was defined as patch area divided by the square of mean scale coordinate of the patch, where A_i^x was assumed to be 0 if $P_i^x = \phi$ or $P_i^x = \{x\}$. The critical area A_i^{crit} was obtained by computing the $(1 - \alpha_c)$th percentile of the null distribution of normalized areas corresponding to the significance level α_i, where α_c is the significance level of the cumulative area-wise test. The null distributions were constructed by generating 1000 patches at the α_i significance level under the null hypothesis of red noise. The final step (Step 4) was to compute

$$r^x = \frac{1}{N}\sum_{j=1}^{N} \lambda_j^x, \quad \text{(A4)}$$

where $\lambda_i^x = 2$ if $A_i^x / A_i^{crit} > 1$ and $\lambda_i^x = 0$ if $A_i^x / A_i^{crit} <= 1$. The wavelet quantity at the point x was deemed statistically significant at the α_c cumulative area-wise level if $r^x > 1$.

Appendix B

For $p > 1$, the $(p+1)$-th order poly spectrum of a time series X is given by

$$B_n^X(s_1, s_2, \dots, s_p) = \widehat{W}_n^X(s_{p+1})\left(\prod_{k=1}^{p} W_n^X(s_k)\right) \quad \text{(B1)}$$

where

$$\frac{1}{s_{p+1}} = \sum_{k=1}^{p} \frac{1}{s_k} \quad \text{(B2)}$$

The third-order poly spectrum is the bi-spectrum, and the fourth-order poly spectrum is the tri-spectrum (Collis et al., 1998), which identifies the frequency components contributing to kurtosis. The $(p+1)$-th order coherence between two time series is given as

$$R_n^2(s) = \frac{\left|S s_{smoth}^{-1} B_n^{XY}(s_1, s_2, \dots, s_p)\right|^2}{S\left(s_{smooth}^{-1}\left|B_n^X(s_1, s_2, \dots, s_p)\right|^2\right) S\left(s_{smooth}^{-1}\left|B_n^Y(s_1, s_2, \dots, s_p)\right|^2\right)}, \quad \text{(B3)}$$

where $B_n^{XY}(s_1, s_2, \dots, s_p)$ is the $(p+1)$-th-order cross-spectrum given by

$$B_n^{XY}(s_1, s_2, \dots, s_p) = B_n^X(s_1, s_2, \dots, s_p)\hat{B}_n^Y(s_1, s_2, \dots, s_p). \quad \text{(B4)}$$

When $p = 2$, Eq. (B3), measures the local cross-correlation between skewness, and when $p = 3$ the equation measures the local cross-correlation between kurtosis.

Data/code availability

Data for Indian rainfall can be accessed through https://www.tropmet.res.in/DataArchival-51-Page (Indian Institute of Tropical Meteorology, 2019). The monthly ENSO indices are available at https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/ (NOAA/OAR/ESRL PSD, 2019). An R software package used to implement the new methods can be found on the author's webpage. (available at: http://justinschulte.com/wavelets/advbiwavelet.html; last access: 22 November 2019).

References

An, S.-I., and Jin, F.-F: Nonlinearity and asymmetry of ENSO., J. Climate, 17, 2399–2412, 2004.

An, S.-I., Ham, Y. G., Kug, J. S., Jin, F.-F., and Kang, I.: El Nino-La Nina asymmetry in the coupled model intercomparison project simulations, J. Climate, 18, 2617-2627, 2005.

An, S-I.: Interdecadal changes in the El Niño-La Niña symmetry, Geophys Res. Lett., 31, L23210, doi:10.1029/2004GL021699, 2004.

An, S.-I.: A review of interdecadal changes in the nonlinearity of the El Nino–Southern Oscillation, Theor. Appl. Climatol., 97, 29–40, 2009.

Ashok, K., Guan, Z., and Yamagata, T.: Impact on the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO, Geophys Res. Lett., 28, 4499–4502, 2001.

Ashok, K, Guan, Z, Saji, N. H., and Yamagata, T.: Individual and combined influences of ENSO and the Indian Ocean dipole on the Indian summer monsoon, J. Climate, 17, 3141–3155, 2004.

Bhatla, R., Singh, A. K., Mandal, B., Ghosh, S., Pandey, S. N., and Abhijit, S.: Influence of North Atlantic Oscillation on Indian Summer Monsoon Rainfall in Relation to Quasi-Binneal Oscillation, Pure and Applied Geophysics, 173, 2959-2970, 2016.

Blanford, H. F.: On the connexion of the Himalaya snowfall with dry wind and seasons of drought in India, Proc R Soc Lond., 37, 3–22, 1884.

Burgers, G., and Stephenson, D. B: The 'Normality' of ENSO, Geophys Res. Lett., 26, 1027–1030, 1999.

Carey, S. K., Tetzlaff, D., Buttle, J., Laudon, H., McDonnell, J,. McGuire, K., Seibert, J., Soulsby, C., Shanley, J. : Use of color maps and wavelet coherence to discern seasonal and inter annual climate influences on streamflow variability in northern catchments. Water Resources Research, 49, 6194–6207, 2013.

Cash, B. A., Barimalala, R., Kinter, J. L., Altshuler, E. L., Fennessy, M. J., Manganello, J.V., Molteni, F., Towers, P., Vitart, F.: Sampling variability and the changing ENSO–monsoon relationship, Clim. Dyn., 48, 4071–4079, 2017.

Chen, W., Dong, B., and Lu, R.: Impact of the Atlantic Ocean on the multidecadal fluctuation of El Niño-Southern Oscillation-South Asian monsoon relationship in a Coupled General Circulation Model, Journal of Geophysical Research, 115, D17109, https://doi.org/10.1029/2009JD013596, 2010.

Collis, W. B., White, P. R., and Hammond, J. K.: Higher-order Spectra: The Bispectrum and Trispectrum, Mech. Syst. Signal Pr., 12, 375–394, 1998.

Dortch, J. M., Owen, L. A., Haneberg, W. C., Caffee, M. W., Dietsch, C., and Kamp, U.: Nature and timing of large landslides in the Himalaya and Transhimalaya of northern India, Quaternary Science Reviews, 28, 1037-1054, 2009.

Duan, W., Huang, C., Xu, H.: Nonlinearity modulating intensities and spatial structures of central Pacific and eastern Pacific El Niño events, Adv Atmos Sci., 34, 737–756, 2017.

DelSole, T., and Shukla, J.: Climate models produce skillful predictions of Indian summer monsoon rainfall, Geophys Res. Lett, 39, L09703, https://doi.org/10.1029/2012GL051279, 2012.

Elsayed, M. A. K.: Wavelet Bicoherence Analysis of Wind–wave Interaction, Ocean Eng., 33, 458–470, 2006.

Fagan, B.: Floods, famines, and emperors: El Niño and the fate of civilizations, Basic Books, 2009.

Fan, F., Dong, X., Fang, X., Xue, F., Zheng, F., and Zhu, J.: Revisiting the relationship between the south Asian summer monsoon drought and El Niño warming pattern, Atmospheric Sci. Lett., 18, 175–182, 2017.

Gershunov, A., Schneider, N., and Barnett, T.: Low-frequency modulation of the ENSO-Indian monsoon rainfall relationship: Signal or noise?, J. Climate, 14, 2486–2492, 2001.

Grinsted, A., Moore, J. C., and Jevrejeva, S.: Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series, Nonlinear Proc. Geophys., 11, 561–566, 2004.

Holman, I. P., Rivas-Casado, M., Bloomfield, J.P., Gurdak, J. J.: Identifying nonstationary groundwater level response to North Atlantic ocean–atmosphere teleconnection patterns using wavelet coherence. Hydrogeol. J. http:// dx.doi.org/10.1007/s10040-011-0755-9, 2011.

Hu, S. and Fedorov, A. V.: The extreme El Nino of 2015–2016˜and the end of global warming hiatus, Geophys. Res. Lett., 4415 3816–24, 2017

Jiang, Z., Sharma, A., & Johnson, F. (2020). Refining Predictor Spectral Representation Using Wavelet Theory for Improved Natural System Modeling. Water Resources Research, 56(3), e2019WR026962. doi:10.1029/2019wr026962

Jin, F.-F., An, S-I., Timmermann, A., and Zhao, J.: Strong El Nino events and nonlinear dynamical heating, Geophys. Res. Lett., 30, 1120, doi:10.1029/2002GL016356, 2003.

Johnson, N.C.: How Many ENSO Flavors Can We Distinguish?, J. Climate, 26, 4816–4827, 2013.

Johnson N. C., and Kosaka, Y.: The role of eastern equatorial Pacific convection on the diversity of boreal winter El Niño teleconnection patterns, Clim. Dyn., 47, 3737–3765, 2016.

Johnson, S. J., Turner, A., Woolnough, S., Martin, G., and MacLachlan, C.: An assessment of Indian monsoon seasonal forecasts and mechanisms underlying monsoon interannual variability in the Met Office GloSea5-GC2 system. Climate Dynamics, 48(5-6), 1447-1465, 2017.

Kakade, S. B., and Dugam, S. S.: The simultaneous effect of NAO and SO on the monsoon activity over India, Geophys Res Lett., 27, 3501–3504, 2000.

Kale, V. : On the link between extreme floods and excess monsoon epochs in South Asia, Climate dynamics, 39, 1107-1122, 2012.

Kalnay, E., and Dalcher, A.: Forecasting forecast skill., Mon. Wea. Rev.,115, 349–356, 1987.

Kang, I.-S., and Kug, J-S.: El Nino and La Ni ~ na sea surface ~ temperature anomalies: Asymmetry characteristics associated with their wind stress anomalies, J. Geophys. Res., 107, 4372, doi:10.1029/2001JD000393, 2002.

Kestin, T. A., Karoly, D. J., Yano, J.-I., and Rayner, N. A.: Time– frequency variability of ENSO and stochastic simulations, J. Climate, 11, 2258–2272, 1998.

King, T.: Quantifying Nonlinearity and Geometry in Time Series of Climate, Quaternary Sci. Rev., 15, 247–266, 1996.

Kripalani, R. H., and Kulkarni, A.: Climatic impact of El Nino/La Nina on the Indian monsoon: A new perspective, Weather, 52, 39-46, 1997.

Kucharski, F., Bracco, A., Yoo, J. H., and Molteni, F.: Low-frequency variability of the Indian monsoon–ENSO relationship and the tropical Atlantic: The "weakening" of the 1980s and 1990s, J. Climate., 20, 4255–4266, 2007.

Kucharski, F., Bracco, A., Yoo, J. H., Tompkins, A. M., Feudale, L., Ruti, P., and Dell'Aquila, A.: A gill-Matsuno-type mechanism explains the tropical Atlantic influence on African and Indian monsoon rainfall, Quarterly Journal of the Royal Meteorological Society, 135, 569–579, 2009.

Kumar, K. K., Soman, M. K., and Kumar, K. R.: Seasonal forecasting of Indian summer monsoon rainfall: a review, Weather, 50, 449-467, 1995

Krishnamurthy, V., and Goswami, B. N.: Indian monsoon–ENSO relationship on interdecadal timescale, J. Climate, 13, 579–595, 2000.

Kripalani, R. H., and Kulkarni, A.: Climatic impact of El Nino/ La Nina on the Indian monsoon: A new perspective, Weather, 52, 39–46, 1997.

Kumar, K. K., Rajagopalan, B., and Cane, M. A.: On the weakening relationship between the Indian monsoon and ENSO, Science, 284, 2156–2159, 1999.

Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G., Cane, M. A.: Unraveling the Mystery of Indian Monsoon Failure During El Niño, Science, 314, 115-119, 2006.

Lee, T., and McPhaden, M. J.: Increasing intensity of El Niño in the central-equatorial Pacific, Geophys. Res. Lett., 37, L14603, doi: 10.1029/2010GL044007, 2010.

Lu, R., Dong, B., and Ding, H.: Impact of the Atlantic Multidecadal Oscillation on the Asian summer monsoon, Geophys. Res. Lett., 33, L24701, https://doi.org/10.1029/2006GL027655, 2006.

McPhaden, M. J.: Genesis and evolution of the 1997–98 El Nino, Science, 283, 950–954, 1999.

Maccarone, T. J.: The Biphase Explained: Understanding the Asymmetries in Coupled Fourier Components of Astronomical Timeseries, Mon. Not. R. Astron. Soc., 435, 3547, doi:10.1093/mnras/stu1824, 2013.

Maraun, D.,and Kurths, J.: Cross wavelet analysis: significance testing and pitfalls. Nonlinear Process. Geophys., 11, 505–514., 2004.

Maraun, D., Kurths, J., and Holschneider, M.: Nonstationary Gaussian Processes in the Wavelet Domain: Synthesis, Estimation, and Significance Testing, Phys. Rev. E, 75, 016707, doi:10.1103/PhysRevE.75.016707, 2007.

Mishra, V., Aadhar, S., Asoka, A., Pai, S., & Kumar, R.: On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain, Geophysical Research Letters, 43, 12-102, 2016.

Munot, A. A., and Kumar, K. K.: Long range prediction of Indian summer monsoon rainfall. Journal of earth system science, 116, 73-79, 2007.

Ng, E. K. W. and Chan, J. C. L.: Geophysical applications of partial wavelet coherence and multiple waveletcoherence, J. Atmos. Ocean. Tech., 29, 1845–1853, 2012.

Parthasarathy, B., Munot, A. A., and Kothawale, D. R.: Regression model for estimation of Indian foodgrain production from summer monsoon rainfall, Agricultural and Forest Meteorology, 42, 167-182, 1988.

Parthasarathy, B., Munot, A. A., and Kothawale, D. R.: All-India monthly and seasonal rainfall series: 1871–1993, Theoretical and Applied Climatology, 49, 217–224. 1994.

Pokhrel, S., Saha, S. K., Dhakate, A., Rahman, H., Chaudhari, H. S., Salunke, K.,Hazra, A., Sujith, K., Sikka, D. R.: Seasonal prediction of Indian summer monsoon rainfall in NCEP CFSv2: forecast and predictability error, Climate dynamics, 46, 2305-2326, 2016.

Prasanna, V. (2014). Impact of monsoon rainfall on the total foodgrain yield over India. *Journal of earth system science*, *123*(5), 1129-1145, 2014.

Rajeevan, M., Pai, D. S., Kumar, R. A., and Lal, B.: New statistical models for long-range forecasting of southwest monsoon rainfall over India, Climate Dynamics, 28, 813-828, 2007.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, 4407, doi: 10.1029/2002JD002670,2003.

Ropelewski, C. F, Halpert ,M. S. : Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation., Monthly Weather Review, 115, 1606–1626, 1987.

Roy, I. and Tedeschi, R. G.: Influence of ENSO on regional indian summer monsoon precipitation—local atmospheric influences or remote influence from pacific, Atmosphere, 7, 25, 2016.

Santoso, A., McGregor, S., Jin, F.-F., Cai, W., England, M. H., An, S.-I., McPhaden, M. J., and Guilyardi, E.: Late-twentieth-century emergence of the El Niño propagation asymmetry and future projections., Nature, 504, 126–130, 2013.

Santoso, A., McPhaden, M. J., Cai, W.: The defining characteristics of ENSO extremes and the strong 2015/2016 El Niño, Rev. Geophys, 55, 1079–1129, 2017.

Sanyal, J., and Lu, X. X.: Remote sensing and GIS-based flood vulnerability assessment of human settlements: a case study of Gangetic West Bengal, India. Hydrological Processes: An International Journal, 19, 3699-3716, 2005.

Schaefli, B., Maraun, D., and Holschneider, M.: What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology, Adv. Water Resour., 30, 2511–2525, 2007.

Schulte, J. A.: Statistical hypothesis testing in wavelet analysis: theoretical developments and applications to Indian rainfall, Nonlin. Processes Geophys., 26, 91-108, https://doi.org/10.5194/npg-26-91-2019, 2019.

Schulte, J. A.: Cumulative areawise testing in wavelet analysis and its application to geophysical time-series, Nonlin. Processes Geophys., 23, 45-57, 2016.

Schulte, J. A.: Wavelet analysis for non-stationary, nonlinear time series, Nonlin. Processes Geophys., 23, 257-267, https://doi.org/10.5194/npg-23-257-2016, 2016.

Schulte, J. A., Duffy, C., and Najjar, R. G.: Geometric and Topological Approaches to Significance Testing in Wavelet Analysis, Nonlin. Processes Geophys., 22, 139-156, 2015.

Shen, X., and Kimoto, M.: Influence of El Niño on the 1997 Indian summer monsoon, J. Meteor. Soc. Japan, 77, 1023–1037, 1999.

Shukla, J., and Paolino, D. A.: The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India. Monthly Weather Review, 111, 1830-1837, 1983.

Slingo, J. M., and Annamalai, H.: 1997: The El Niño of the century and the response of the Indian summer monsoon, Mon. Wea. Rev., 128, 1778–1797, 2000.

Timmermann, A.: Decadal ENSO amplitude modulations: A nonlinear mechanism, Global Planet. Change, 37, 135–156, 2003.

Torrence, C., and Webster, P. J.: Interdecadal changes in the ENSO-monsoon system, J. Climate, 12, 2679–2690. 1999.

Van Milligen, B. P., Saìnchez, E., Estrada, T., Hidalgo, C., Brañas,B., Carreras, B., and Garcìa, L.: Wavelet Bicoherence: A New Turbulence Analysis Tool, Phys. Plasmas, 2, 3017–3032, 1995.

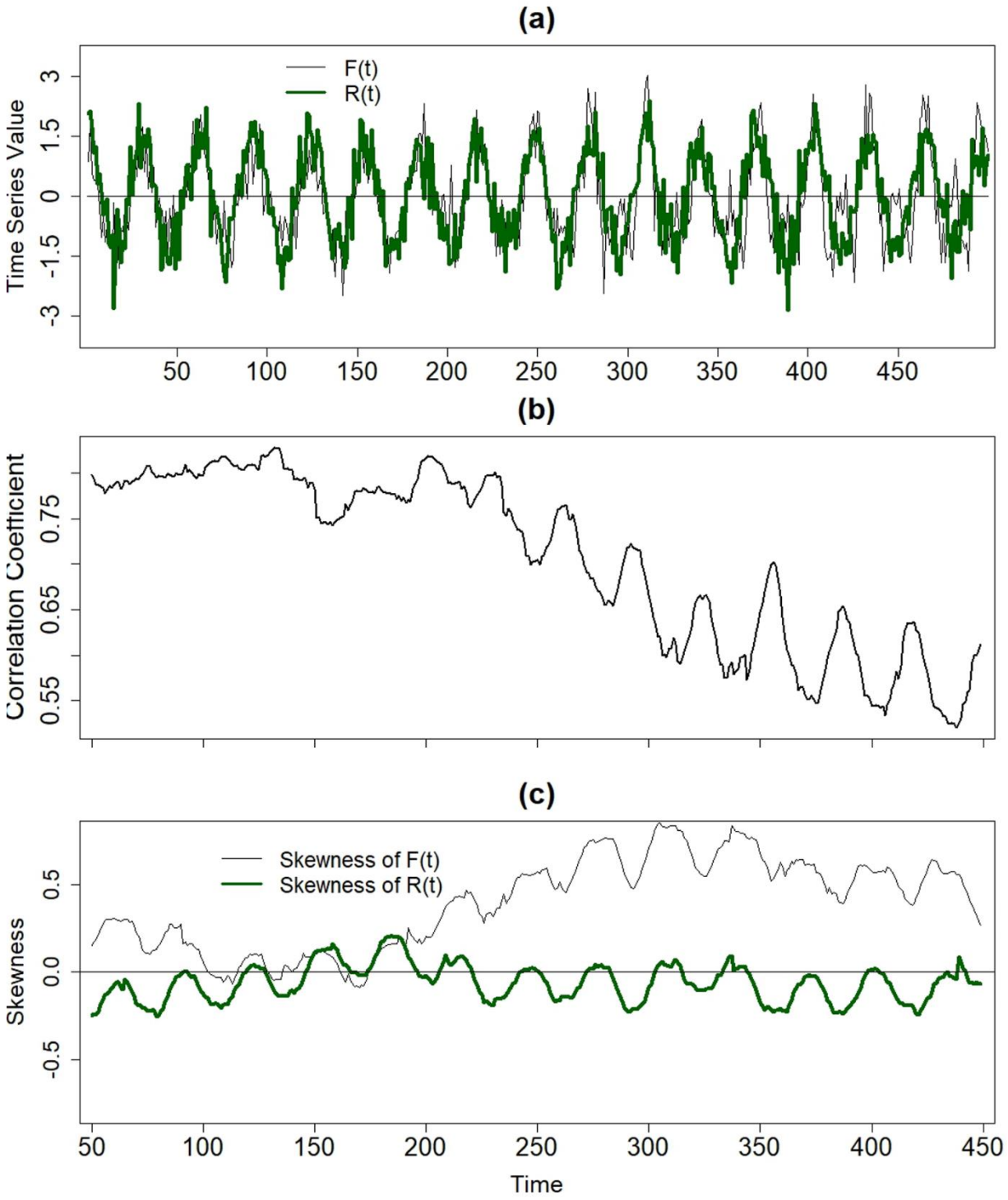
Walker, G. T., and Bliss, E. W.: World weather V, Mem., R. Meteorol. Soc., 4, 53–84, 1932.

Wu, A., and Hsieh, W. W.: Nonlinear interdecadal changes of the El Niño-Southern Oscillation, Climate Dyn., 21, 719–730, 2003.

Yun, K. S and Timmermann, A.: Decadal monsoon-ENSO relationships reexamined, Geophys Res Lett., 45, 2014–2021, 2018.

Zhang, Q., Xu, C., Jiang, T., Wu, Y.: Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China, Journal of Hydrol, 333, 265–274. doi:10.1016/ j.jhydrol.2006.08.010, 2007.

774



775

776 Figure 1. (a) An idealized nonlinear forcing time series together with an idealized response $R(t)$. The 120-point
777 sliding correlation between $F(t)$ and $R(t)$. (c) The 120-point sliding skewness of $F(t)$ and $R(t)$.

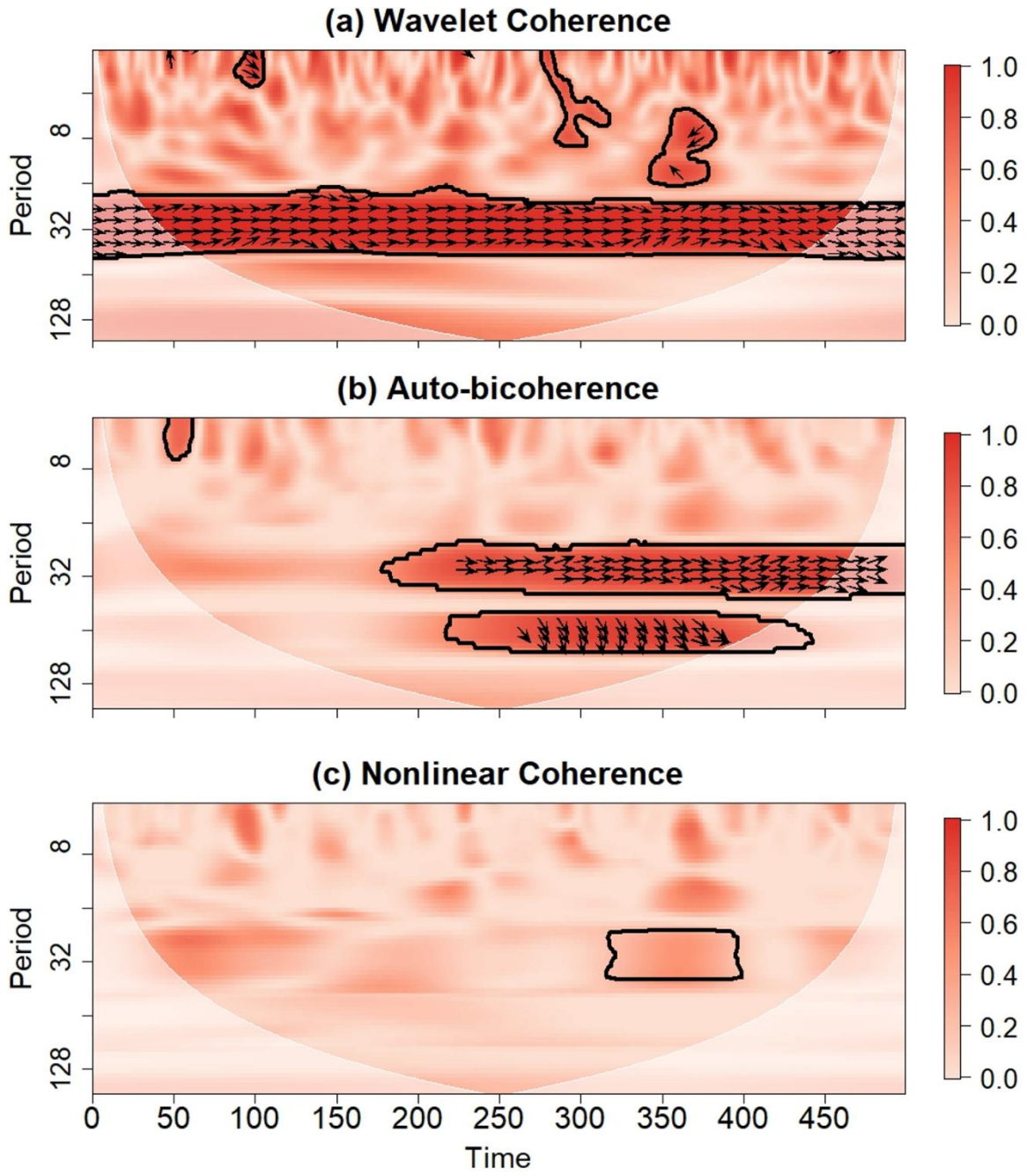


Figure 2. (a) Wavelet coherence between the time series of $F(t)$ and $R(t)$ shown in Figure 1. Arrows indicate the relative phase difference, where arrows pointing to the right mean that the time series are in phase. (b) The local diagonal slice of the auto-bicoherence spectrum of $F(t)$. Arrows represent the bi-phase, where arrows pointing to the right mean that the quadratic phase coupling between the mode with period indicated on the vertical axis and its harmonic contributes to positive skewness. (c) Nonlinear coherence between $F(t)$ and $R(t)$. Contours in all panels enclose regions of 5% cumulative area-wise significance. Light-shaded region represents the cone of influence where edge effects may be important.

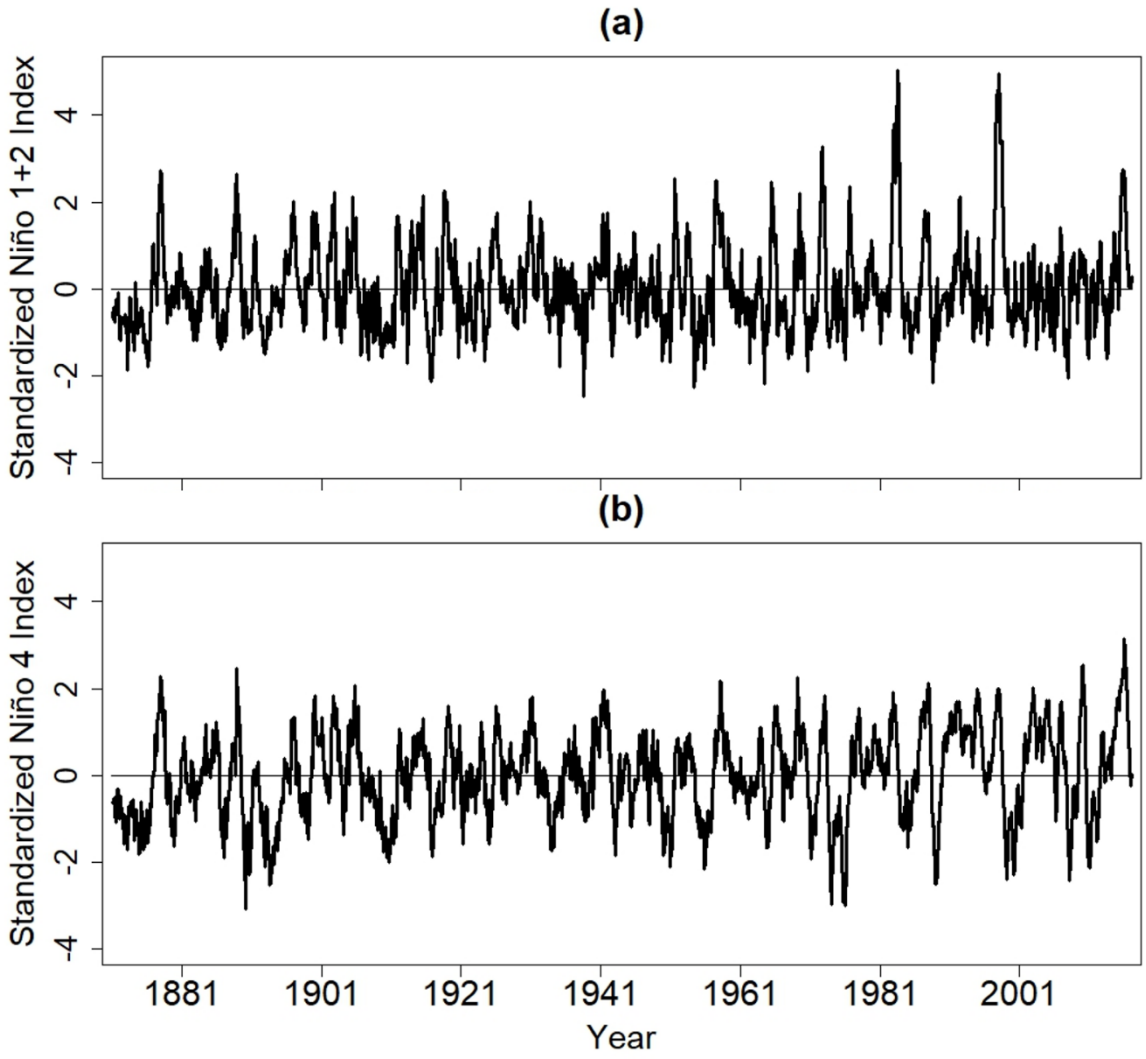
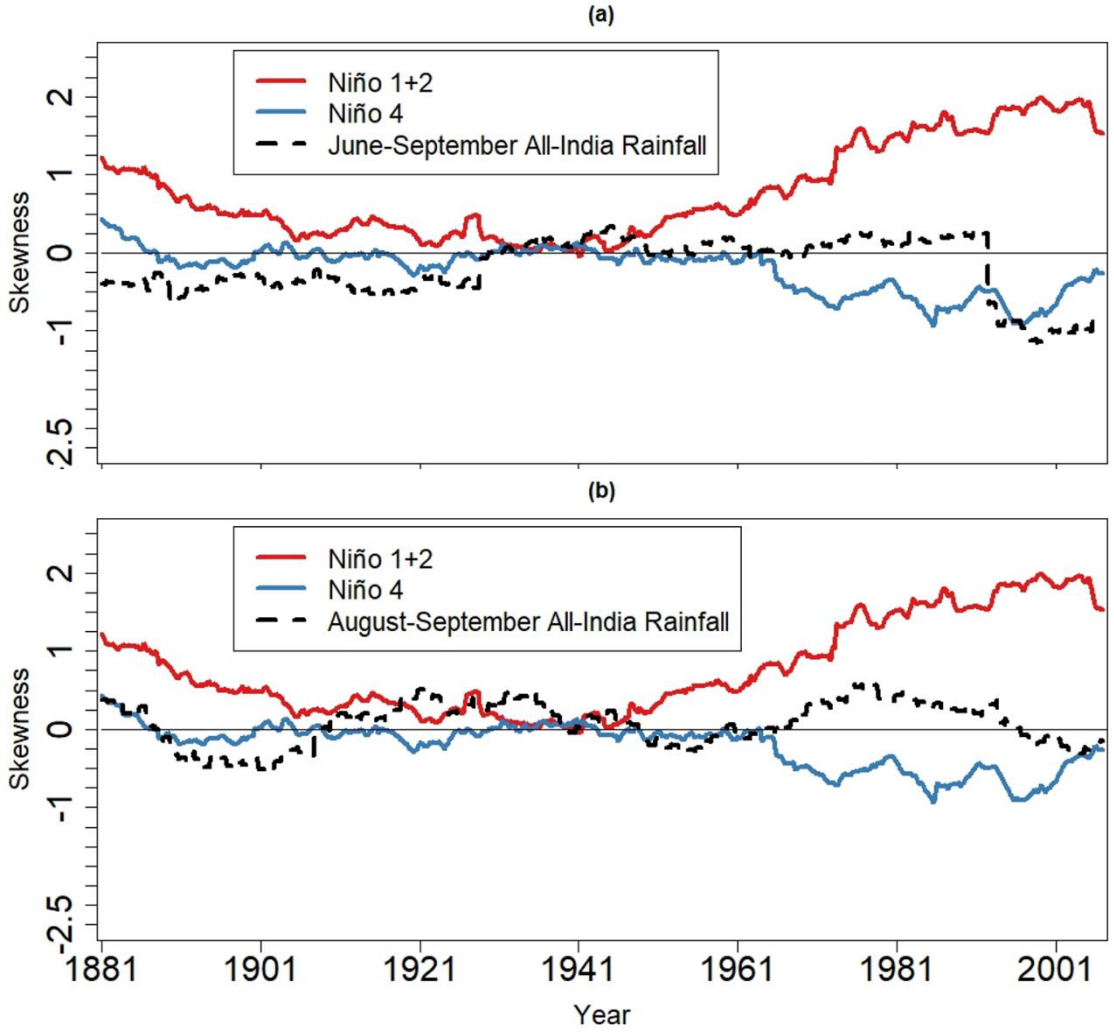


Figure 3. The time series of the (a) Niño 1+2 and (b) Niño 4 indices.

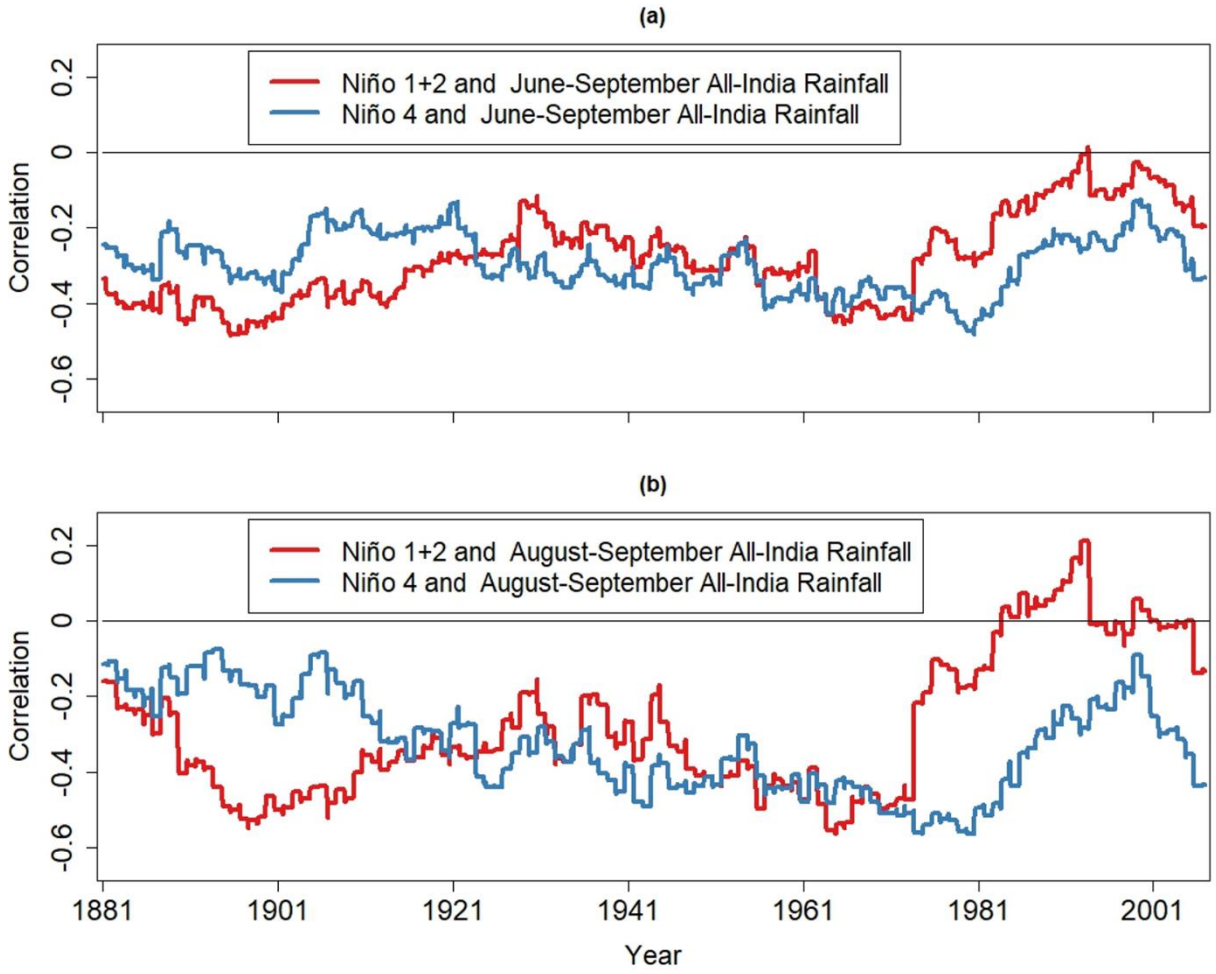


788

789 Figure 4. 20-year sliding skewness of (a) June-September and (b) August-September AIR and full time series for the
790 Niño 1+2 and Niño 4 indices.

791

792



793

794 Figure 5. 20-year sliding correlation between anomalies for June-September AIR and the time series for the June-
795 September Niño 1+2 and Niño 4 indices. (b) Same as (a) but for the August-September season.

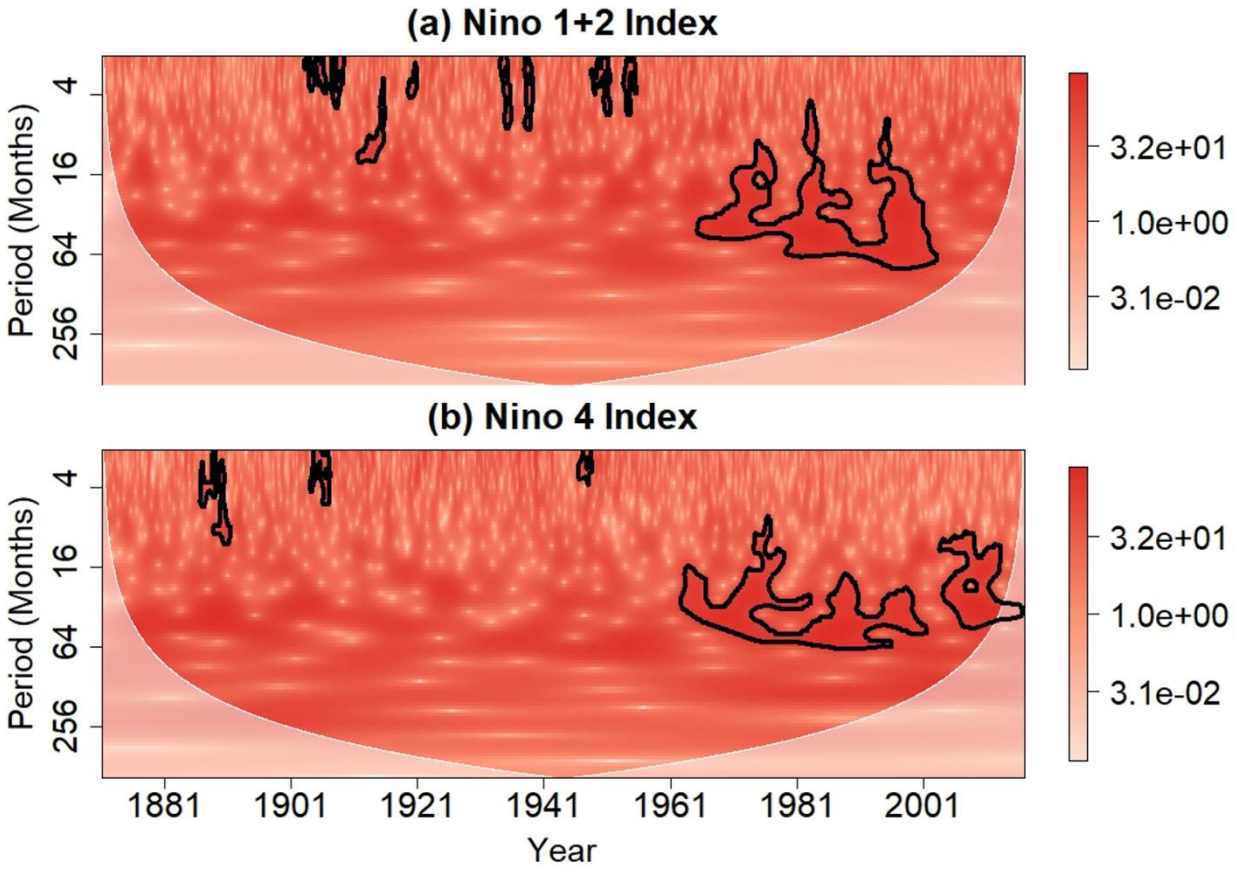


Figure 6. Wavelet Power spectrum of the (a) Niño 1+2 and (d) Niño 4 indices. Contours enclose regions of 5% cumulative area-wise significance. Light-shaded region represents the cone of influence, which is the region where edge effects are non-negligible.

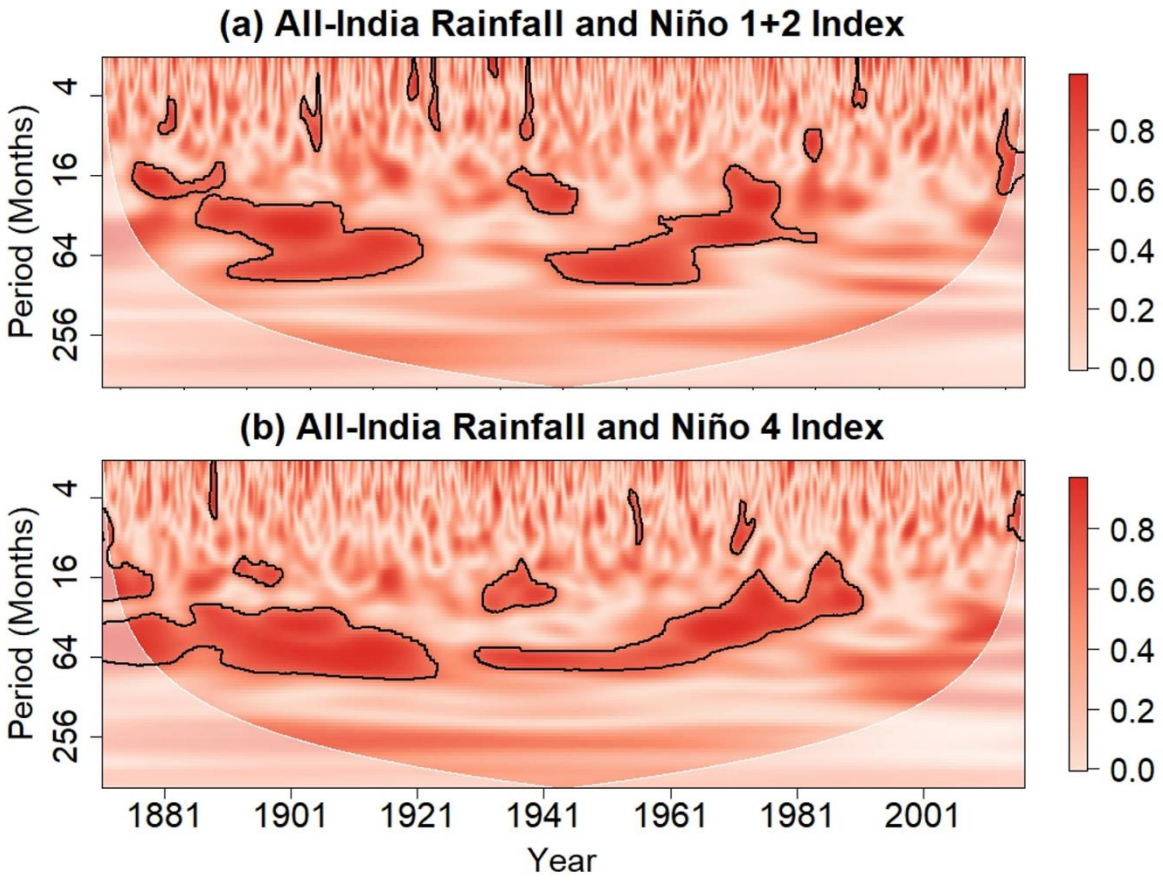
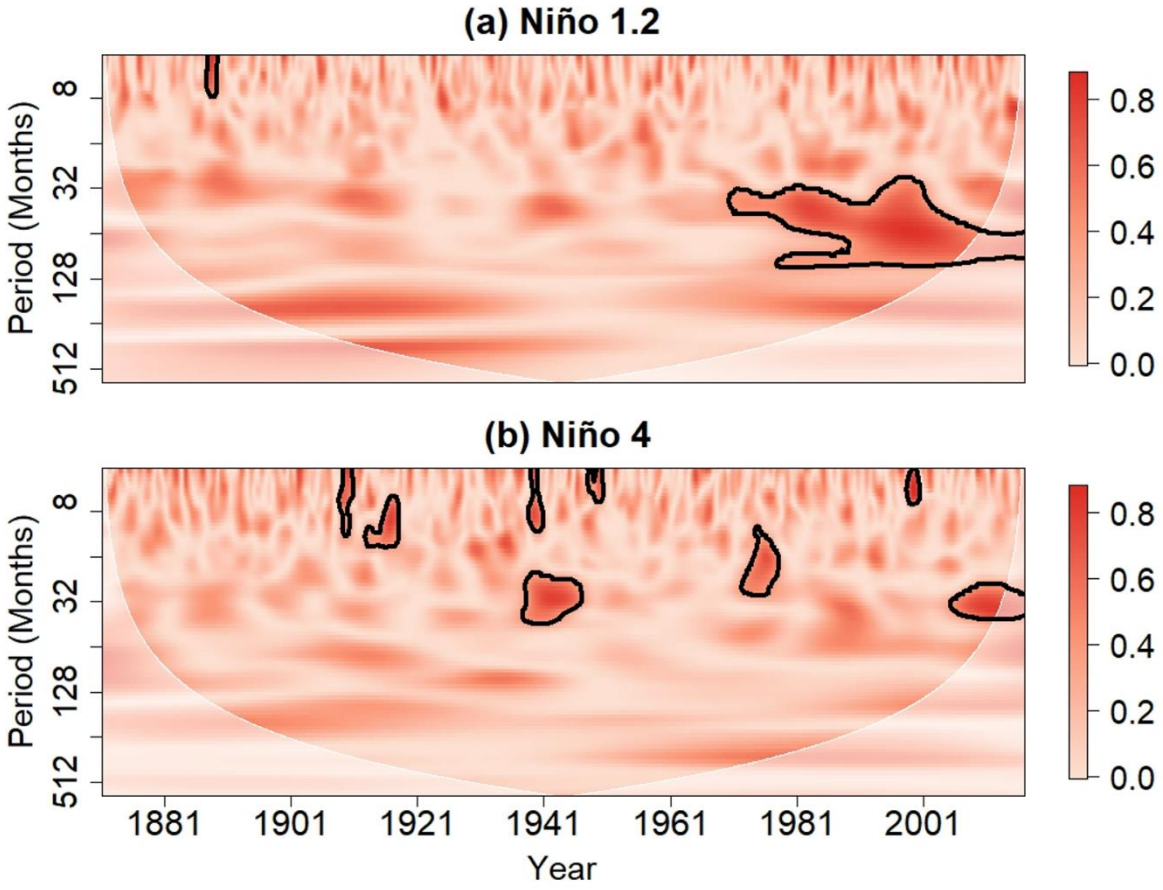


Figure 7. Wavelet coherence spectrum between AIR and time series for the (a) Niño 1+2 and (b) Niño 4 indices. Contours enclose regions of 5% cumulative area-wise significance. Light-shaded region represents the cone of influence, which is the region where edge effects are non-negligible.

807



808
809 Figure 8. Local auto-bicoherence spectra of the (a) Niño 1+2 and (b) Niño 4 indices. Contours enclose regions of 5%
810 cumulative area-wise significance and the light shading represents the cone of influence.

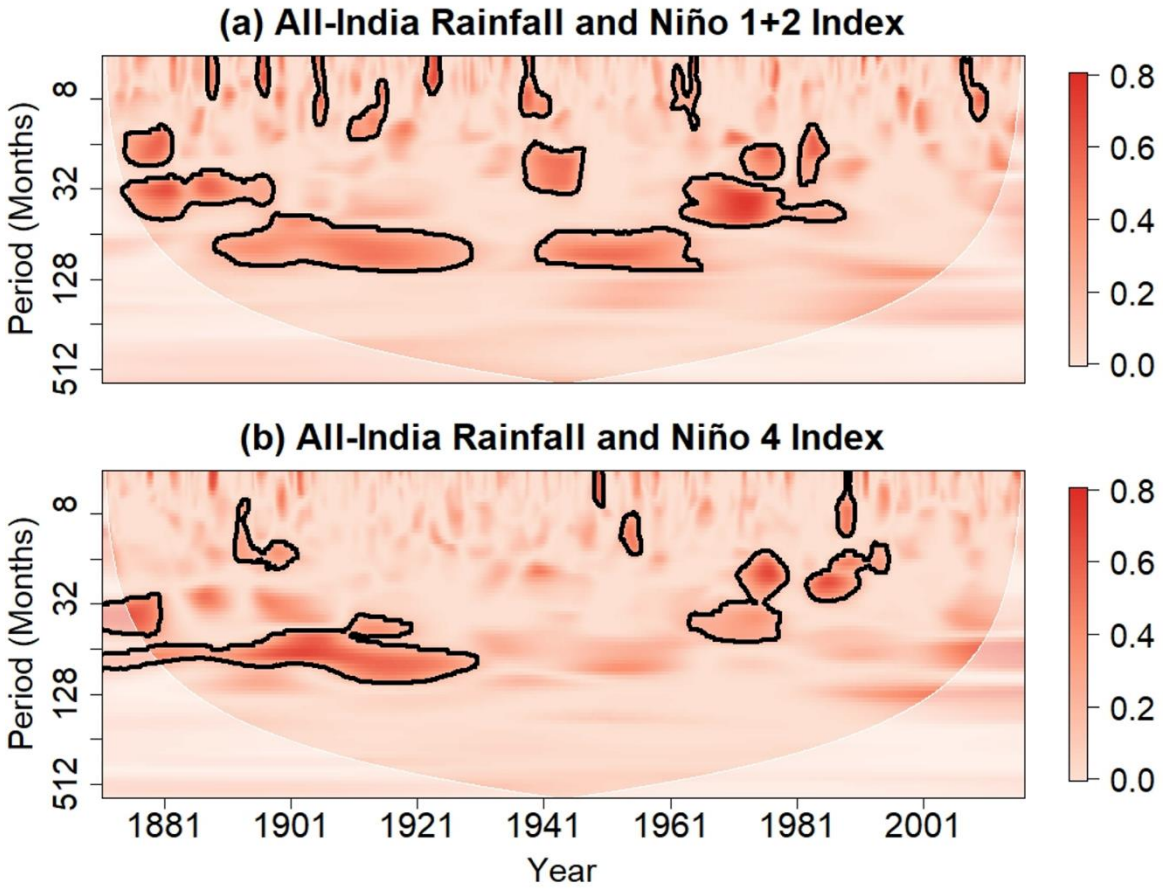


Figure 9. Nonlinear wavelet coherence between the full AIR time series and full times series for the (a) Niño 1+2 and (b) Niño 4 indices. Contours enclose regions of 5% cumulative area-wise significance and light shading represents the cone of influence.

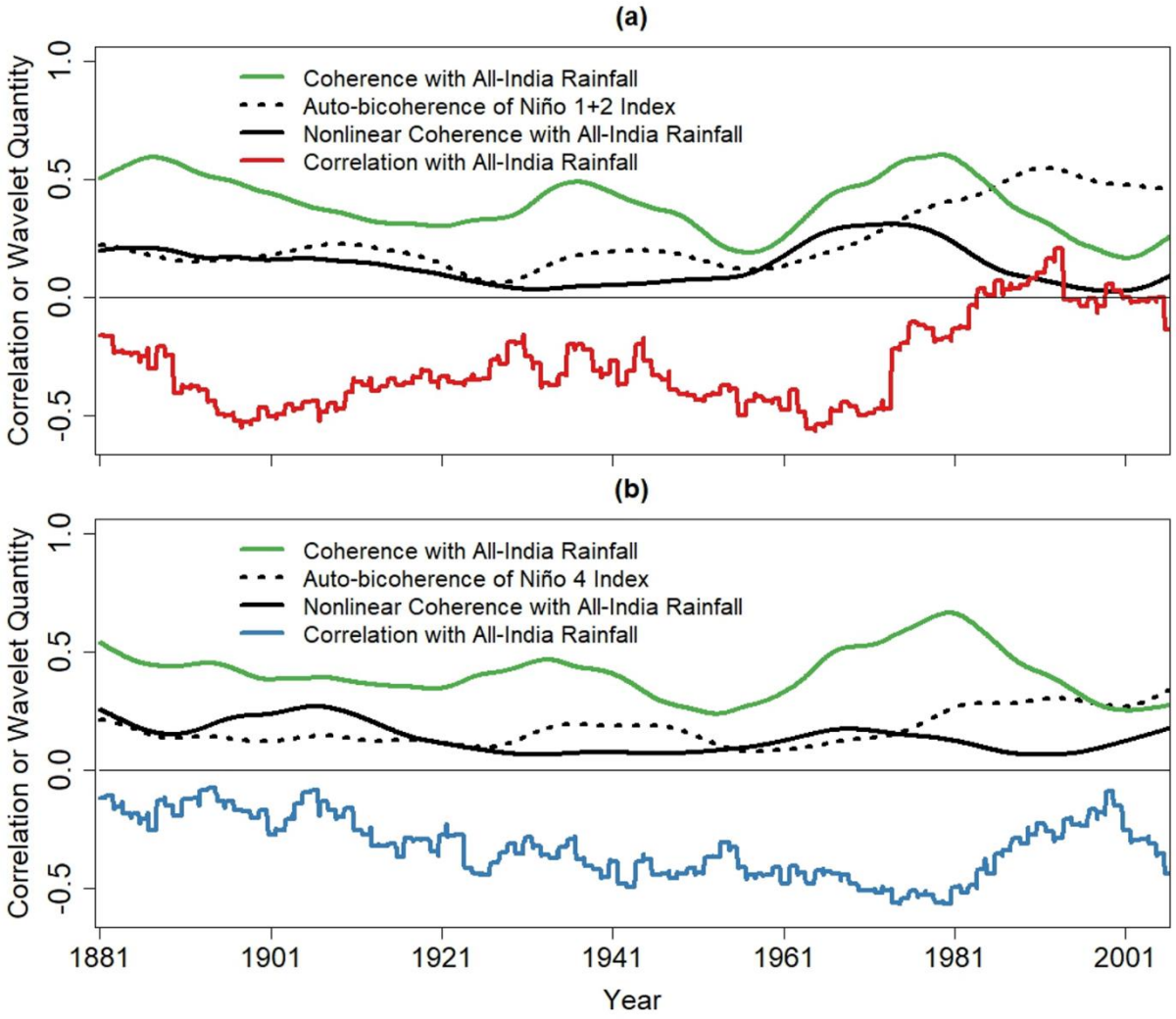


Figure 10. (a) 20-year sliding mean time series of the linear wavelet coherence between AIR and the Niño 1+ 2 index, the auto-bicoherence of the Niño 1+2 index, and the nonlinear coherence between the Niño 1+2 index and AIR after they have been averaged in the period band of 16 to 64 months. Red curve is the 20-year sliding correlation between the August-September Niño 1+2 index and AIR. (b) The same as (a) but for the Niño 4 index. Blue curve is the 20-year sliding correlation between the August-September Niño 4 Index and AIR.

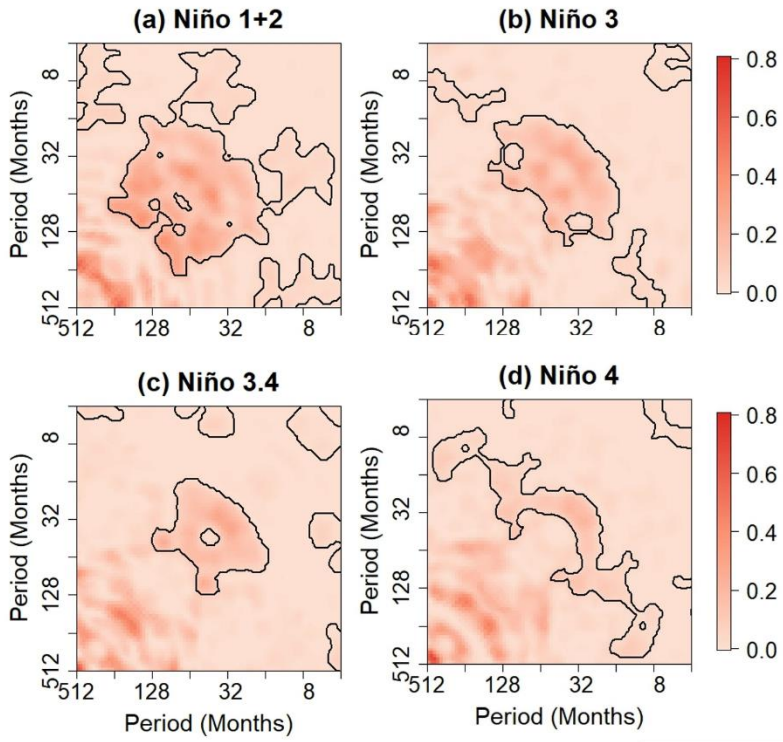


Figure 11. Global auto-bicoherence spectra of the (a) Niño 1+2, (b) Niño 3, (c) Niño 3.4, and (d) Niño 4 indices. Contours enclose regions of 5% cumulative area-wise significance.

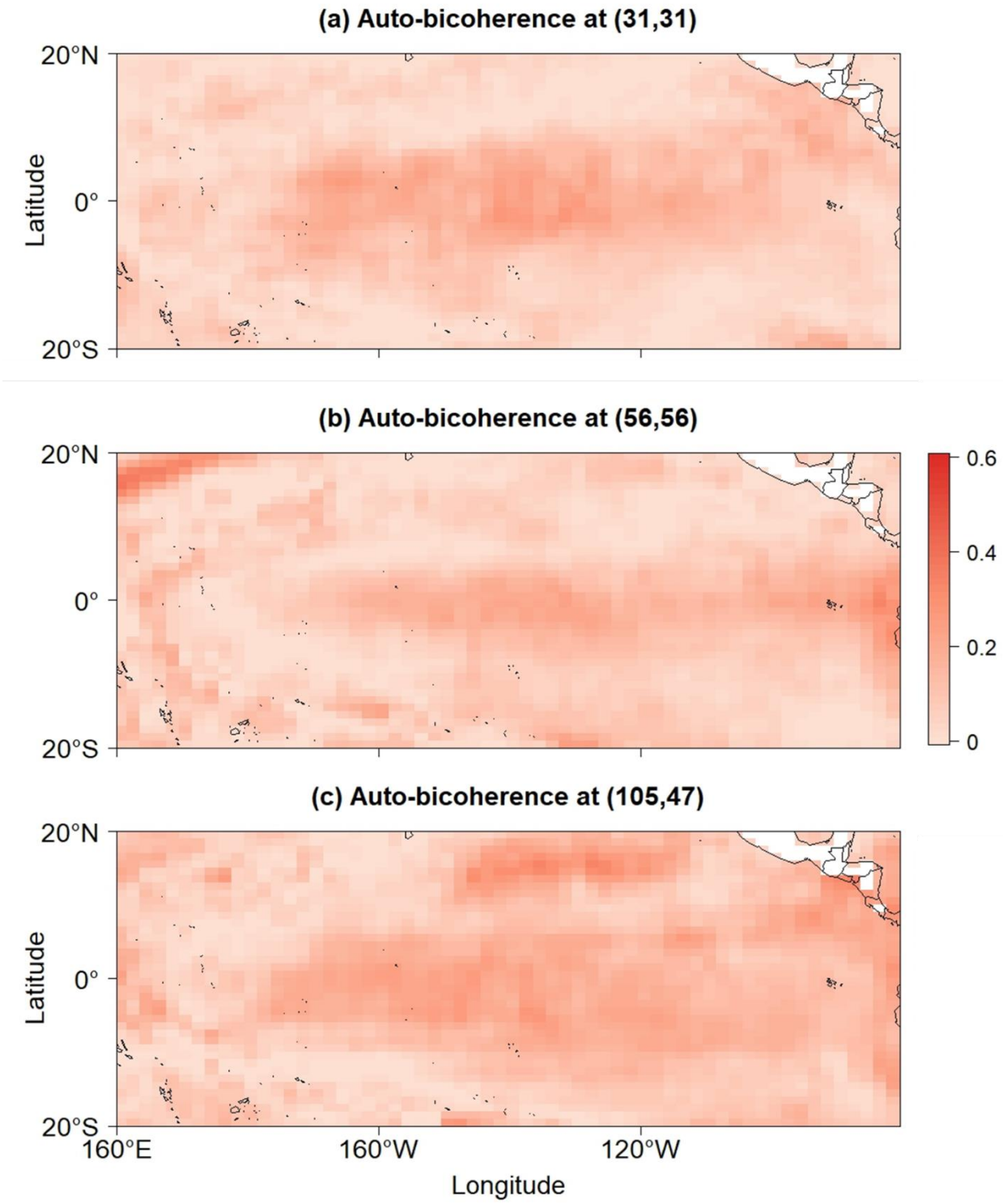
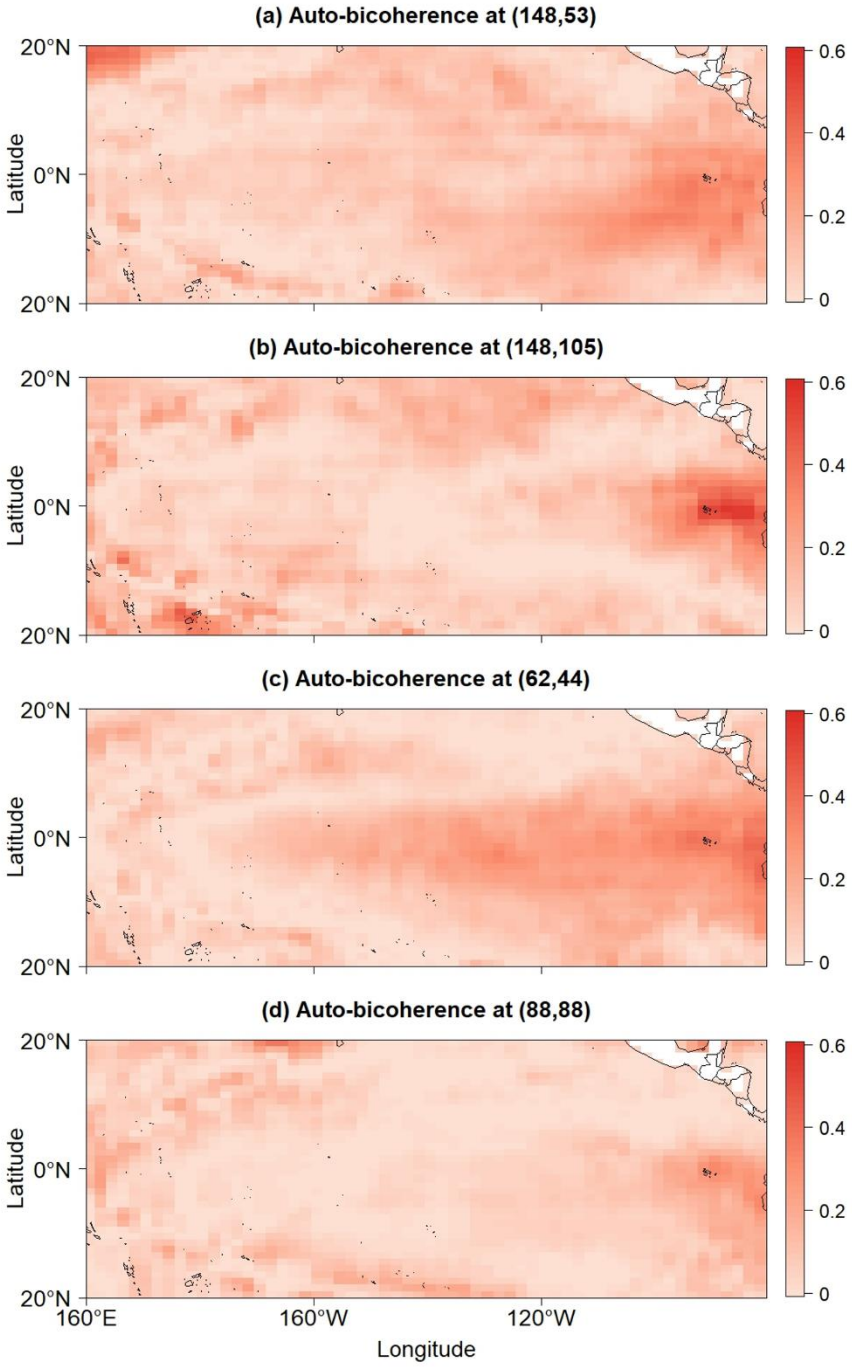


Figure 12. Global auto-bicoherence corresponding to the pairs (a) (31, 31), (b) (56, 56), and (c) (105, 47) [months].



832

833 Figure 13. Global auto-bicoherence corresponding to the pairs (a) (158, 43), (b) (148, 105), (c) (62, 44), and (d)
834 (88,88) [months].

835

836 Table 1. Wavelet quantities and the relationships they measure.

Wavelet Quantity	Quantified Relationship
Linear Coherence	Cross-correlation between the variance of two time series at a Fourier period
Global Auto-bicoherence	Time-averaged quadratic phase coupling among two or three linear modes
Local Auto-bicoherence	Quadratic phase coupling among two or three linear modes at a time point
Nonlinear Coherence	Cross-correlation between the skewness of nonlinear modes

837

838

839

840