OPEN

Detrended Partial-Cross-Correlation Analysis: A New Method for Analyzing Correlations in Complex System

Naiming Yuan[1,2,3], Zuntao Fu[3], Huan Zhang[2], Lin Piao[3], Elena Xoplaki[2] & Juerg Luterbacher[2]

[1]Chinese Academy of Meteorological Science, Beijing, 100081, China, [2]Department of Geography, Climatology, Climate Dynamics and Climate Change, Justus Liebig University Giessen, Senckenbergstrasse 1, 35390 Giessen, Germany, [3]Lab for Climate and Ocean-Atmosphere Studies, Dept. of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, 100871, China.

In this paper, a new method, detrended partial-cross-correlation analysis (DPCCA), is proposed. Based on detrended cross-correlation analysis (DCCA), this method is improved by including partial-correlation technique, which can be applied to quantify the relations of two non-stationary signals (with influences of other signals removed) on different time scales. We illustrate the advantages of this method by performing two numerical tests. Test I shows the advantages of DPCCA in handling non-stationary signals, while Test II reveals the "intrinsic" relations between two considered time series with potential influences of other unconsidered signals removed. To further show the utility of DPCCA in natural complex systems, we provide new evidence on the winter-time Pacific Decadal Oscillation (PDO) and the winter-time Nino3 Sea Surface Temperature Anomaly (Nino3-SSTA) affecting the Summer Rainfall over the middle-lower reaches of the Yangtze River (SRYR). By applying DPCCA, better significant correlations between SRYR and Nino3-SSTA on time scales of 6 ∼ 8 years are found over the period 1951 ∼ 2012, while significant correlations between SRYR and PDO on time scales of 35 years arise. With these physically explainable results, we have confidence that DPCCA is an useful method in addressing complex systems.

C omplex systems, such as in climatology, ecology, economy or finance, usually contain a large number of interactions[1–4]. By analyzing the cross-correlations between signals observed from the complexity, one aim is to better diagnose and understand the whole system. Duo to its simplicity, traditional cross-correlation analysis (TCA) has become the most widely used method in statistics. Especially in statistical climatology, TCA is used in various fields, including dynamical diagnosing and climate forecasting[5–7]. However, due to the effects of many nonlinear processes and external forcings, it should be noted that signals obtained in nature are usually characterized by multi-scale structures and non-stationarities[8,9]. Therefore, traditional cross-correlation analysis is not always appropriate and can provide erroneous results.

For example, suppose we are interested to diagnose the relations between the Summer (June, July and August) Rainfall over the middle-lower reaches of the Yangtze River (SRYR) and the previous winter-time (December, January and February) Nino3 Sea Surface Temperature Anomaly (Nino3-SSTA). Both records range from 1951 to 2012. SRYR is calculated according to the precipitation station data provided by the Chinese National Climate Center (http://ncc.cma.gov.cn), see Figure 1, while Nino3-SSTA is downloaded from the National Oceanic & Atmospheric Administration (NOAA, http://www.esrl.noaa.gov/psd/data/climateindices/). It is a well known fact that SRYR are teleconnected with the previous winter-time East Pacific SSTA[10,11], thus significant correlations are expected. However, by simply applying TCA to the two records, the calculated correlation coefficient is only 0 : 19, which is not statistically significant according to the student's t-test. To explain this low correlation coefficient, one reason could be that the connection between SRYR and Nino3-SSTA is nonstationary over time. Thus, it may be an unreasonable choice to analysis their relation over the whole length (1951–2012) since the climate regime is considered changed around the end of 1970s[12]. However, it may be also a "*time scale*" problem. Since Nino3-SSTA is an El Niño indicator with a typical period of 2 ∼ 7 years, its connections with SRYR may be only significant on these time scale. On other time scales, significant correlations may disappear, which further result in a low cross-correlation coefficient calculated over the whole length. Thus, the correlations between two sub-systems can be different on different time scales.
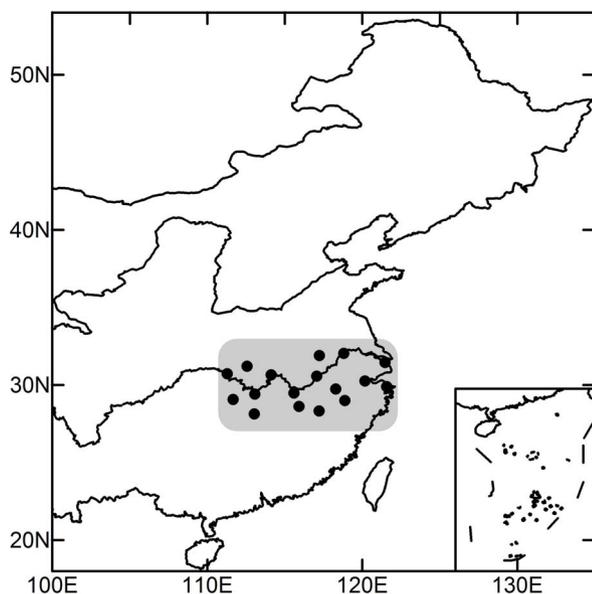
Figure 1 | Geographic locations of the stations around the Yangtze River. 17 stations over the middle-lower reaches of the Yangtze River are chosen for the calculation (average over the 17 stations) of SRYR. Their locations are shown as the solid circles. We generate the figure by using Surfer 8.0 (Golden Software, http://www.goldensoftware.com/products/surfer).

In traditional statistics, one can apply filter methods (including low-pass filter, high-pass filter, and band-pass filter) to discuss correlations of two considered time series on different time scales[13–15]. However, the low (high) pass frequency, or the band-width are usually chosen subjectively, which make these simple filter methods not appropriate in performing cross-correlation research over different time scales. Another method, cross-spectral analysis (CSA)[16,17], may also be useful in discussing connections of two time series on different time scales, but it requires the analyzed data to be stationary with no external trends, which of course are rare in nature. Recently, a new method based on detrended covariance, detrended cross-correlation analysis (DCCA), has been proposed and widely used[18]. DCCA is a modification of the standard covariance analysis, but can be used in the research of non-stationary time series[19]. DCCA is also a generalization of detrended fluctuation analysis (DFA)[20,21], but can be used to investigate the power-law cross-correlations between *two* simultaneously recorded time series. By further calculating the DCCA cross-correlation coefficient ρ_{DCCA} according to the procedure proposed by[22],

$$\rho_{DCCA} = \frac{F_{DCCA}^2}{F_{DFA}(x_i^1)F_{DFA}(x_i^2)}, \qquad (1)$$

where F_{DCCA} is the fluctuation function obtained from DCCA[18], F_{DFA} is the fluctuation function obtained from DFA[20], and $\{x_i^1\}$, $\{x_i^2\}$ are the two considered time series, one can quantify the level of cross-correlations on different time scales. Therefore, during the past few years, signals from various fields such as economics[23], seismic studies[24], traffic flows[25], as well as geophysical systems[26], have been analyzed by using DCCA and its multifractal version MFDCCA[27]. In this study, we will mainly focus on the DCCA cross-correlation coefficient ρ_{DCCA} derived from DCCA.

In Figure 2 (top panel), we analyze the relations between SRYR and lead Nino3-SSTA by calculating DCCA cross-correlation coefficient ρ_{DCCA}. Apparently, SRYR is correlated with Nino3-SSTA on time scale of 5 ~ 7 years with cross-correlation coefficient larger than 0.3. While on other time scales, the cross-correlations drop to a very low level (around 0.1). This result is in

line with our discussion above, and indicates that studying correlations on different time scales is very important for better understanding the whole complex system.

However, it should be further noted, that signals observed from a complex system are normally linked via interwoven heterogeneous ties. Quantifing cross-correlations between only *two* signals may be not sufficient and can provide erroneous results. Especially in the case, when the two signals are both correlated with other signals simultaneously. Such as the relations among SRYR, Nino3-SSTA, and the Pacific Decadal Oscillation (PDO). PDO is a pattern of warm or cold anomalous surface waters in north Pacific (north of 20°N), with inter-decadal time scale of about 30 years[28]. It is a well known fact that both the winter-time Nino3-SSTA and the winter-time PDO index can be considered as important precursor factors of the following summer rainfall over Yangtze River[29,30]. However, since PDO and El Niño are also coupled with each other[31,32], simple analysis based on either PDO index or Nino3-SSTA may provide biased information. As shown in Figure 2 (bottom panel), we calculated DCCA cross-correlation coefficient ρ_{DCCA} to study the relations between SRYR and PDO index. The PDO index is downloaded from the National Oceanic & Atmospheric Administration (NOAA, http://www.esrl.noaa.gov/psd/data/climateindices/), with only winter-time data selected. It is obvious that the results from PDO and that from Nino3-SSTA have similar pattern, especially on small time scale of 4 ~ 8 years (El Niño typical scale) and large time scale of 30 ~ 45 years (PDO typical scale), which indicates strong coupling between PDO and Nino3-SSTA. When making diagnostic analysis or prediction, one usually prefers to take as many related factors as possible into account to improve the accuracy. Here we argue that, first we need to state to what extent and on which time scales each related factor is independently affecting the system of interest, then further consider how these factors are connected to each other. One way to address this is by applying detrended partial-cross-correlation analysis (DPCCA). DPCCA is based on DCCA, thus can provide information on different time scales. Compared to DCCA cross-correlation coefficient ρ_{DCCA}, ρ_{DPCCA} calculated from DPCCA is further upgraded by combining partial-correlation technique, therefore it is expected to be useful in quantifying correlations of multi-signals (not only two signals) in a complex system.

In this report, we will first illustrate the advantages of DPCCA by conducting two numerical tests. Test I shows the advantages of DPCCA in handling non-stationary signals, while Test II illustrate the advantage of DPCCA in revealing "intrinsic" relations between two time series of interest, with potential influences of other unconsidered signals removed. Furthermore, the utility of DPCCA is confirmed by revisiting the climatic example mentioned above. Results and discussions are shown in the next sections. In the last part of this report, we will show explicitly how the DPCCA is designed.

Results

Advantages of DPCCA. Since DPCCA is based on the DCCA method but improved by combining the partial cross-correlation analysis (PCCA), it is expected to have the advantages of both methods. Therefore, we will perform two tests to verify the utility of DPCCA as indicated below.

Test I: According to[18], DCCA is designed to investigate cross-correlations between two time series with nonstationarity. When nonstationarity such as local trends or periodic background exist, without detrending, there will be crossovers in the fluctuation function F_{DCCA} as a function of time scale[33,34], and the DCCA cross-correlation coefficient ρ_{DCCA} calculated from Eq.(1) will be spuriously high[35]. Fortunately, by choosing an appropriate detrending order, DCCA is able to remove the effects of nonstationarity, and further provide us reliable information on the cross-correlation[36]. Similarly, DPCCA should also have this advantage. Suppose we have three independent and identically distributed (i.i.d) Gaussian vari-
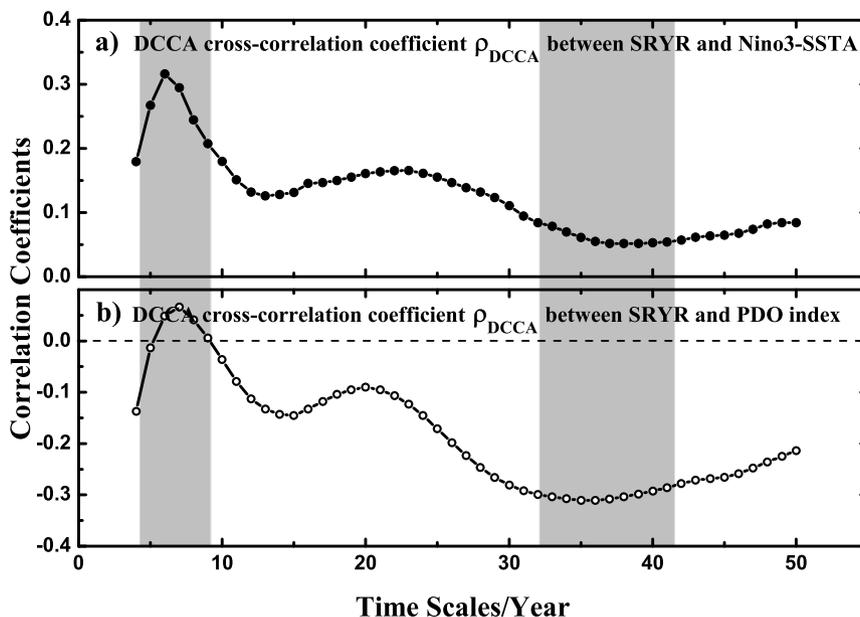
Figure 2 | DCCA cross-correlation coefficient ρ_{DCCA} between SRYR and Nino3-SSTA (a), as well as SRYR and PDO index (b). The grey region covers the time scale of 5 ~ 8 years (El Niño typical scale) and the time scale of 33 ~ 42 years (PDO typical scale). The two curves have similar pattern.

ables: $\{x_i^A\}$, $\{x_i^B\}$ and $\{x_i^C\}$ (with length of 10,000). They are not related to each other. However, if we generate another two time series: $\{x_i^{A'}\}$ and $\{x_i^{B'}\}$, which are combinations of time series $\{x_i^A\}$ with $\{x_i^C\}$ (as $\{x_i^{A'}\} = \{x_i^A\} + \{x_i^C\}$), and $\{x_i^B\}$ with $\{x_i^C\}$ (as $\{x_i^{B'}\} = \{x_i^B\} + \{x_i^C\}$), respectively, the two new generated time series will be correlated, and are both related to $\{x_i^C\}$. By applying the partial cross-correlation analysis to the three time series $\{x_i^{A'}\}$, $\{x_i^{B'}\}$ and $\{x_i^C\}$, one can remove the influence of $\{x_i^C\}$ on $\{x_i^A\}$ and $\{x_i^B\}$, and further discover the "intrinsic" correlations between $\{x_i^A\}$ and $\{x_i^B\}$ (which should be zero). However, if the three variables are all nonstationary with nonlinear trends, traditional PCCA may fail in detecting the "intrinsic" correlation information. As shown in Figure 3a, 3b, $\{x_i^{A'}\}$ and $\{x_i^{B'}\}$ are shown with quadratic trends. The cross-correlation coefficient between $\{x_i^{A'}\}$ and $\{x_i^{B'}\}$ is 0.71 (Figure 3c, the blue line). After removing the influence of $\{x_i^C\}$, the PCCA coefficient between the two variables drops to 0.59 (the red line), which is still significantly higher than zero. If the influence of quadratic trends is removed by DCCA, the coefficient ρ_{DCCA} further decreases to 0.51 (the yellow line), but still not the expected result. In this case, however, if we apply DPCCA, "intrinsic" relation between $\{x_i^A\}$ and $\{x_i^B\}$ is finally obtained (the black curve in Figure 3c). In fact, not only for the case when quadratic trends exist, for cases with cubic trends, or even sinusoidal trends, DPCCA still shows reliable and accurate results, as shown in Figure 3d. We show the cases with "No Trend", "Linear Trend", "Quadratic Trend", "Cubic Trend", as well as "Sinusoidal Trend". By applying DPCCA with appropriate detrend order (see the discussions in[36], and also in the "Methods" section. One can remove the non-stationary effects by substracting local trends with appropriate polynomial order. Normally, DPCCAn means the polynomial order of n), expected results still arise (Black line). While other methods failed, such as PCCA (the red line) and DCCA cross-correlation coefficient ρ_{DCCA} (the yellow line). Therefore, from this test we confirm that DPCCA has the advantages of DCCA.

Test II: Another advantage of DPCCA should originated from the partial cross-correlation analysis. Compared with the DCCA cross-correlation coefficient ρ_{DCCA}, DPCCA can be used to investigate the correlations of multi-signals in a complex system, and find the

"intrinsic" relations between two considered signals. Suppose we have two independent and identically distributed (i.i.d) Gaussian variables: $\{x_i^A\}$ and $\{x_i^B\}$ (with length of 10,000). By adding sinusoidal signals $\{x_i^{S_1}\}$ and $\{x_i^{S_2}\}$ (as shown in Figure 4a–b), two new time series $\{x_i^{A'}\}$ and $\{x_i^{B'}\}$ can be generated, as $\{x_i^{A'}\} = \{x_i^A\} + \{x_i^{S_1}\} + \{x_i^{S_2}\}$, and $\{x_i^{B'}\} = \{x_i^B\} + \{x_i^{S_2}\}$. The two sinusoidal signals have different frequencies, as shown in Figure 4, $\{x_i^{S_1}\}$'s cycle is 1000 (days), while $\{x_i^{S_2}\}$'s cycle is 100 (days). Therefore, in the newly generated time series, $\{x_i^{S_1}\}$ can be considered as a background field, and the 100 (days) periodic signal (corresponding to $\{x_i^{S_2}\}$) can be detected by DCCA cross-correlation coefficient ρ_{DCCA}, as shown in Figure 4e (the red line). However, if another sinusoidal signal $\{x_i^{S_3}\}$ (as shown in Figure 4c, which has the same frequency with $\{x_i^{S_2}\}$, but with different phases), is also added to $\{x_i^{A'}\}$, by simply calculating DCCA cross-correlation coefficient ρ_{DCCA}, one may underestimate the 100 (days) periodic signal due to the "offset" effect between $\{x_i^{S_2}\}$ and $\{x_i^{S_3}\}$ (Figure 4d and Figure 4e, the open circles). Therefore, in this case, ρ_{DPCCA} should be a more appropriate choice for our analysis and diagnose. See Figure 4f, by applying DPCCA to the three time series $\{x_i^{A'}\}$, $\{x_i^{B'}\}$, and $\{x_i^{S_3}\}$, we can remove the influence of $\{x_i^{S_3}\}$ on $\{x_i^{A'}\}$ successfully, and reveal the 100 (days) periodic signal accurately. Therefore, from this test we confirm that DPCCA inherits the advantages of the partial-correlation technique.

Application of DPCCA to natural complex system. Considering natural signals are normally recorded from complex systems, they are usually characterized by non-stationary, and are always correlated with other multi-signals. Therefore, the DPCCA method proposed in this report could be widely used in various fields. In the following, we will further illustrate the utility of DPCCA by revisiting the climatic example we have mentioned in the introduction.

We study how the winter-time Pacific Decadal Oscillation (PDO) and winter-time Nino3 Sea Surface Temperature Anomaly (Nino3-SSTA) affect the Summer Rainfall over the middle-lower reaches of the Yangtze River (SRYR) over the past 60 years. It has been well recognized that the summer rainfall over China is influenced by two main modes of Pacific SST variation: PDO and El Niño (with Nino3-
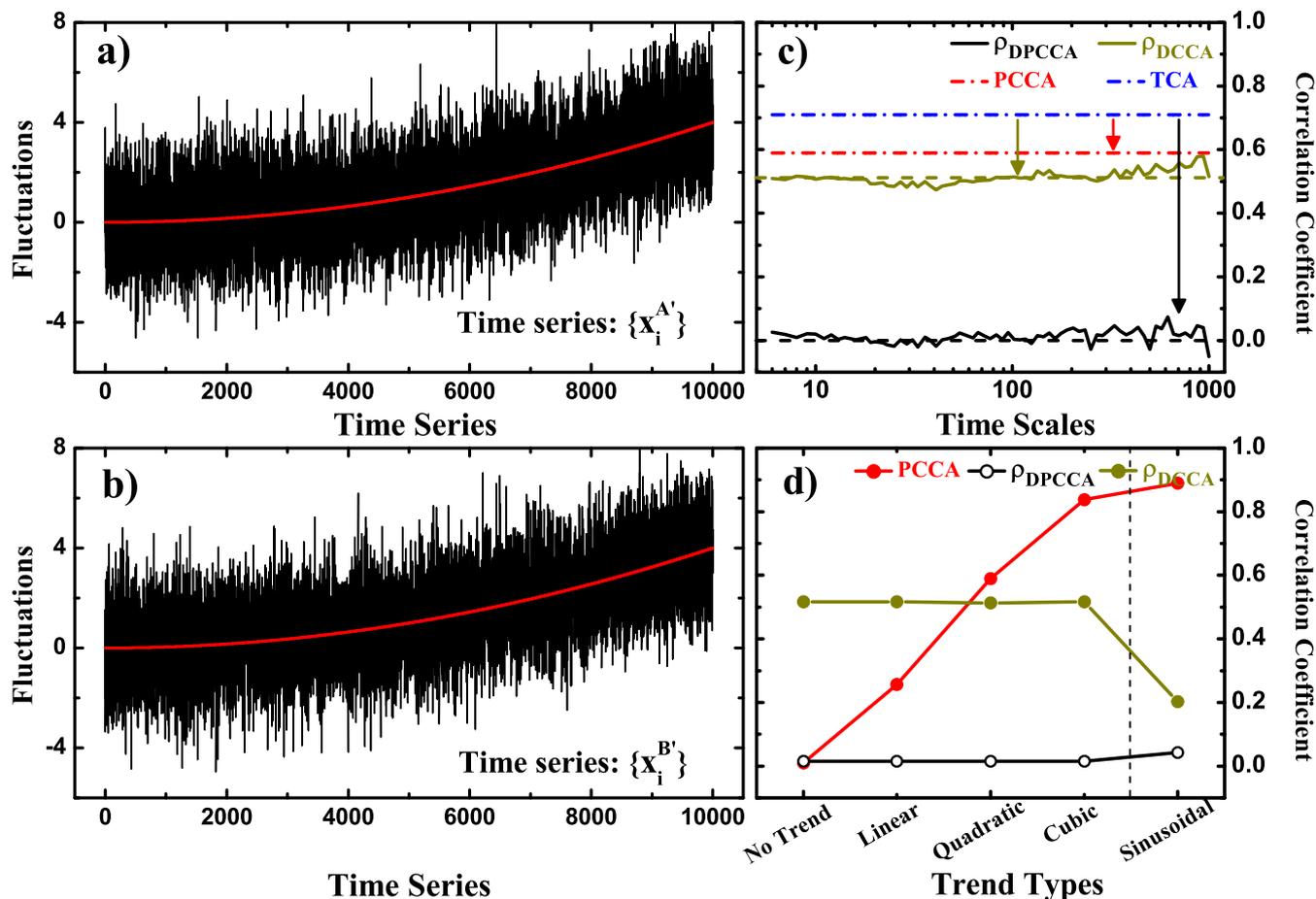
Figure 3 | Time series and related results in Test I. (a, b) show the two time series $\{x_i^{A'}\}$ and $\{x_i^{B'}\}$ in test I, but with quadratic trends. The results of traditional cross-correlation analysis (TCA, blue), partial cross-correlation analysis (PCCA, red), detrended cross-correlation analysis (DCCA cross-correlation coefficient ρ_{DCCA}, yellow), as well as the detrended partial-cross-correlation analysis (DPCCA, black) are shown in (c). In (d), more cases with higher-order trends are shown. For each "Trend Type", the detrend order of DPCCA (DCCA) are different, as DPCCA1 (DCCA1), DPCCA2 (DCCA2), DPCCA3 (DCCA3), DPCCA4 (DCCA4), and DPCCA16 (DCCA16), respectively. The coefficient of each "Trend Type" is actually the mean correlation coefficient averaged over all the time scales in (c).

SSTA as an indicator[37]). Their winter signals are both considered as important precursor factors of the summer rainfall over China[29,30]. However, since PDO and El Niño are also coupled with each other[31,32] (Figure 2), simple predictions based on either PDO or Nino3-SSTA are not entirely reliable. Therefore, we need to reveal the "intrinsic" relations between SRYR and PDO, as well as the "intrinsic" relations between SRYR and Nino3-SSTA. Figure 5 shows the results, where significant differences between the output of DPCCA and DCCA are presented. For the relations between SRYR and Nino3-SSTA, after removing the influence of PDO, much higher (positive) cross-correlation coefficients ρ_{DPCCA} over all the time scales are found. Especially on time scales of 5 ~ 8 years (the gray area), more significant cross-correlations between SRYR and Nino3-SSTA are found (exceeding the 95% confidence level), which corresponds to the typical period of El Niño. As for the relations between SRYR and PDO, after removing the influence of Nino3-SSTA, much lower (negative) cross-correlation coefficients ρ_{DPCCA} over all time scales are obtained. If we calculate DCCA cross-correlation coefficient ρ_{DCCA} only, positive correlations between SRYR and PDO on time scales of 6 ~ 8 years are found, however not significant. After removing the effect of Nino3-SSTA, the positive correlations disappear. Interestingly, on time scales of about 35 years (the grey area), significant (negative) correlations between SRYR and PDO arise (exceeding the 95% confidence level), which corresponds to the typical period of PDO. However, masked by the El Niño, this

signal cannot be revealed from ρ_{DCCA}. From these results it becomes obvious that El Niño has important impacts on SRYR during its typical period (5 ~ 8 years), while at the multidecadal scale, the SRYR may be modulated by the PDO. This finding is in line with previous studies. In fact, it has been well accepted that during the period of El Niño, a persistent anomalous anticyclone over the Western North Pacific (WNP) can bring a large amount of water vapor to East Asia, which leads to an increase of precipitation over the Yangtze River[38–41]. However, modulated by the locations and strengths of WNP monsoon trough and the WNP subtropical high (WNPSH)[41,42], which maybe related to the variations of PDO, the effects of El Niño on East Asia can also vary on multidecadal scale. Such as the time before the late 1970s, positive (negative) winter-time Nino3-SSTA usually corresponds to less (more) rainfall over the Yangtze River. Due to a westward expansion of the WNPSH after late 1970s, summer precipitation increased over the Yangtze River[43,44]. Therefore, for better understanding the Summer Rainfall over the middle-lower reaches of the Yangtze River, different mechanisms on different time scales should be considered carefully. From this example, cross-correlation coefficient ρ_{DPCCA} obtained from DPCCA shows better performance than ρ_{DCCA} from DCCA.

Discussion
In this report, we proposed a new method, Detrended Partial-Cross-Correlation Analysis (DPCCA), which can be used to diagnose
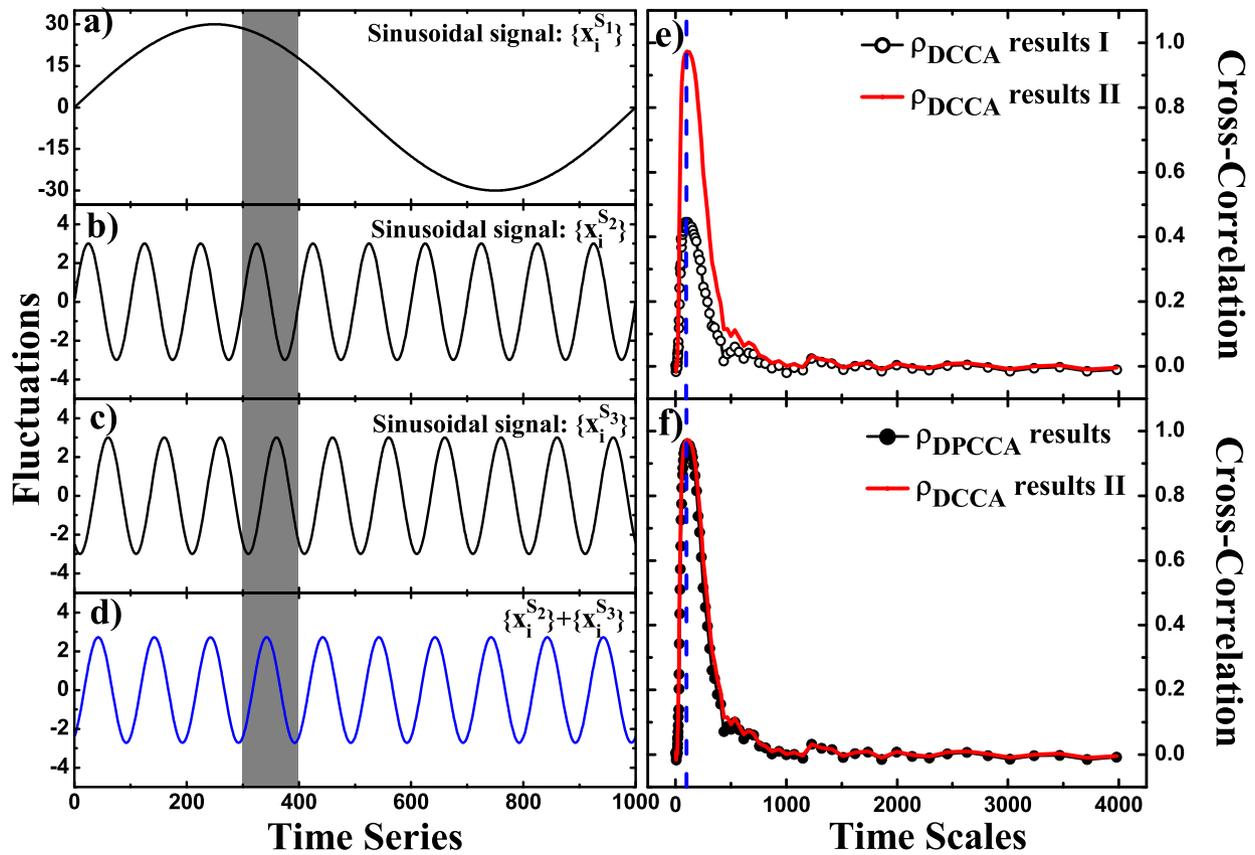
Figure 4 | Time series and related results in Test II. (a–c) show fractions of the three sinusoidal signals in test II: $\{x_i^{S_1}\}$, $\{x_i^{S_2}\}$, and $\{x_i^{S_3}\}$. $\{x_i^{S_1}\}$ acts as a background field, with low-varying frequency and larger amplitude. $\{x_i^{S_2}\}$, and $\{x_i^{S_3}\}$ have different phases (see the gray part), and their combination is shown in (d). The red curve in (e) is the DCCA cross-correlation coefficient ρ_{DCCA} between $\{x_i^{A'}\} = \{x_i^A\} + \{x_i^{S_1}\} + \{x_i^{S_2}\}$ and $\{x_i^{B'}\} = \{x_i^B\} + \{x_i^{S_2}\}$ (denoted as ρ_{DCCA} results II), and the blue dashed line shows the time scale of 100 (days). If the signal $\{x_i^{S_2}\}$ in $\{x_i^{A'}\}$ is offset by $\{x_i^{S_3}\}$, ρ_{DCCA} fails in providing accurate results, as shown in (e), the black open-circle curve (denoted as ρ_{DCCA} results I). But DPCCA succeeds, as shown in (f), the black solid-circle curve.

"intrinsic" relations of two nonstationary signals (with influences of other signals removed) on different time scales. This method is based on the Detrended Cross-Correlation Analysis (DCCA), but improved by including the Partial-Cross-Correlation Analysis (PCCA), which therefore has the advantages of both DCCA and PCCA. To illustrate the advantages, we made two simple tests in our study. Test I proved DPCCA indeed can provide robust results even when nonlinear trends are mixed in the data we are analyzing, and further show relations between the two considered data on different time scales. While Test II illustrated the ability of DPCCA in investigating correlations when multi-signals are linked via interwoven ties, as shown in Figure 4. In general, DPCCA has better performance in dealing with correlations in complex system. However, when applying it, there are two points that need to be considered.

i) Significance testing. With DPCCA, one can obtain cross-correlations on different time scales. However, to determine whether the calculated correlations are statistically significant, one can not simply apply the student's t-test due to the changing degree of freedom. Normally, Monte-Carlo tests have to be applied to decide whether the obtained cross-correlations are significant on a given time scale[45] (Figure 5, the blue line).

ii) Background assumptions. When applying DPCCA, one has to pay attention to the background assumptions of partial cross-correlation analysis. That is, the considered multi-signals should have linear relationships with each other. This is the main deficiency of PCCA. However, by using DCCA, we believe this

deficiency can be reduced to some extent, since only the relationships on different time scales are discussed, but not on the whole length. Nevertheless, we would like to stress that more advanced analytical methods are needed, especially the methods based on nonlinear frameworks.

We applied DPCCA to a climatic example, that deals with the winter-time Pacific Decadal Oscillation (PDO) and the winter-time Nino3 Sea Surface Temperature Anomaly (Nino3-SSTA) affecting the Summer Rainfall around Yangtze River (SRYR) over past decades. Since PDO has longer variation period (≈ 30 years), it can be considered as a variable background. With PDO controlled, the relations between Nino3-SSTA and SRYR seems to be more apparent on the time scales of $5 \sim 8$ years. Similarly, with Nino3-SSTA controlled, significant relations between PDO and SRYR emerges on time scales of about 35 years. Although, due to possible influences of nonlinear effects, our results may still be problematic, considering traditional Cross-Correlation Analysis is still the main analytical method in various fields, our study still improves our ability in analyzing cross-correlations among multi-variables on different time scales. From the two numerical tests and the climatic example, we can summarize the advantages of DPCCA. i) DPCCA can be used to reveal the "intrinsic" relations between two considered variables, by removing the possible influences of other unconsidered signals, ii) DPCCA is appropriate in the research of non-stationary variables, and iii) DPCCA can show the correlation levels on different time scales. Based on these advantages, we are convinced that this method will have extensive application prospects.
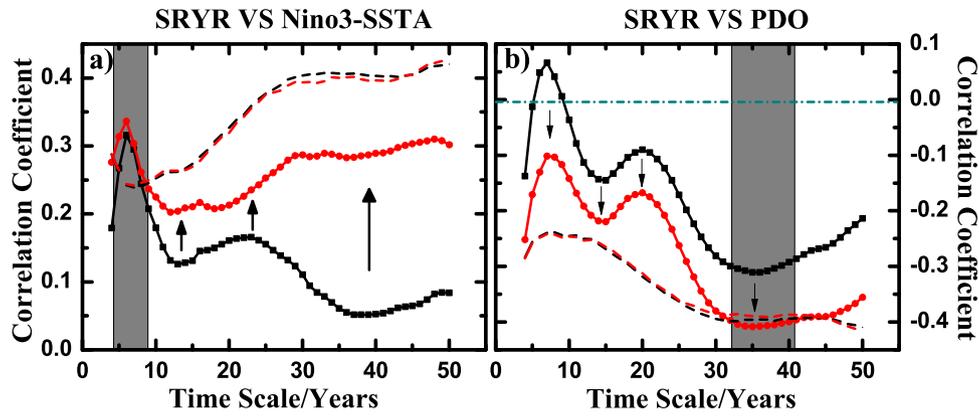
Figure 5 | Application of DPCCA to the climatic example. DCCA (black) and DPCCA (red) outputs between (a) SRYR and Nino3-SSTA, (b) SRYR and PDO. The black dashed line represents the 95% significance criterions of ρ_{DCCA} on different time scales, while the red dashed line represents the criterions of ρ_{DPCCA}. They are both obtained from Monte-Carlo Simulations according to[45]. We shuffled the considered time series and repeated the DCCA/DPCCA calculations for 10,000 times. The top 2.5% largest values (on different time scales) give the dashed lines in (a), while the top 2.5% smallest values give the dashed lines in (b). The grey area covers the typical time scales where (a) the Nino3-SSTA has significant positive correlation with SRYR on time scales of $5 \sim 8$ years, and (b) the PDO has significant negative correlation with SRYR on time scales of about 35 years.

Methods

In this section, we will show the details on how the method, DPCCA, is designed.

Suppose we have m time series $\{x_i^1\}$, $\{x_i^2\}$, $\{x_i^3\}$, $\cdots$, $\{x_i^m\}$, where $i = 1,2,3,\cdots,N$. Each time series can be considered as a random walk, and we can define the so called profile as:

$$P_k^j \equiv \sum_{i=1}^{k} x_i^j, \qquad (2)$$

where $j = 1,2,3,\cdots,m$, $k = 1,2,3,\cdots,N$. Similar to the procedures in DCCA, one first divide the entire profile into $N - s$ overlapping boxes. Each box i contains $s + 1$ values, starts at i and ends at $i + s$. In each box i, we can determine the "local trend" $\widetilde{P_{k,i}^j}$ ($i \leq k \leq i + s$) by using a polynomial fit, and further define the "detrended walk" as the difference between the original profile and the local trend, as:

$$Y_{(i-1)(s+1)+k-i+1}^j = P_k^j - \widetilde{P_{k,i}^j}, \qquad (3)$$

In this way, we can get one detrended residual series Y_l^j, $l = 1,2,3,\cdots,(N-s)(s+1)$, for each time series $\{x_i^j\}$. By calculating the covariance between any two residuals,

$$F_{j_1,j_2}^2(s) \equiv \frac{\sum_{l=1}^{(N-s)(s+1)} Y_l^{j_1} Y_l^{j_2}}{(N-s)(s-1)}, \qquad (4)$$

where $j_1,j_2 = 1,2,3,\cdots,m$, we can obtain a covariance matrix,

$$\mathbf{F^2(s)} = \begin{pmatrix} F_{1,1}^2(s) & F_{1,2}^2(s) & \dots & F_{1,m}^2(s) \\ F_{2,1}^2(s) & F_{2,2}^2(s) & \dots & F_{2,m}^2(s) \\ \vdots & \vdots & & \vdots \\ F_{m,1}^2(s) & F_{m,2}^2(s) & \dots & F_{m,m}^2(s) \end{pmatrix}. \qquad (5)$$

Obviously, according to[22], the cross-correlation levels between any two time series, $\{x_i^{j_1}\}$ and $\{x_i^{j_2}\}$, can be estimated as,

$$\rho_{j_1,j_2}(s) \equiv \frac{F_{j_1,j_2}^2(s)}{F_{j_1,j_1}(s) \cdot F_{j_2,j_2}(s)}, \qquad (6)$$

and a coefficients matrix can further be obtained as,

$$\boldsymbol{\rho(s)} = \begin{pmatrix} \rho_{1,1}(s) & \rho_{1,2}(s) & \dots & \rho_{1,m}(s) \\ \rho_{2,1}(s) & \rho_{2,2}(s) & \dots & \rho_{2,m}(s) \\ \vdots & \vdots & & \vdots \\ \rho_{m,1}(s) & \rho_{m,2}(s) & \dots & \rho_{m,m}(s) \end{pmatrix}. \qquad (7)$$

Where $\rho_{j_1,j_2}(s)$ ranges from -1 to $+1$, and represents the level of cross-correlation on time scales of s. However, it should be noted that it only shows the relations between time series $\{x_i^{j_1}\}$ and $\{x_i^{j_2}\}$. This may provide spurious correlation information if the two time series are both correlated with other signals. Therefore, to exclude the

possible influence of other time series, we need to combine the partial-correlation technique with the calculations above.

To apply the partial-correlation technique, we first need to calculate the inverse matrix of $\boldsymbol{\rho}(s)$,

$$C(s) = \boldsymbol{\rho}^{-1}(s) = \begin{pmatrix} C_{1,1}(s) & C_{1,2}(s) & \dots & C_{1,m}(s) \\ C_{2,1}(s) & C_{2,2}(s) & \dots & C_{2,m}(s) \\ \vdots & \vdots & & \vdots \\ C_{m,1}(s) & C_{m,2}(s) & \dots & C_{m,m}(s) \end{pmatrix}, \qquad (8)$$

for any two time series $\{x_i^{j_1}\}$ and $\{x_i^{j_2}\}$, the partial-cross-correlation level can thus be determined as,

$$\rho_{DPCCA}(j_1,j_2;s) = \frac{-C_{j_1,j_2}(s)}{\sqrt{C_{j_1,j_1}(s) \cdot C_{j_2,j_2}(s)}}. \qquad (9)$$

where the coefficients $\rho_{DPCCA}(j_1, j_2; s)$ can be used to characterize the "intrinsic" relations between the two time series on time scales of s. It is worth to note that we use the word "intrinsic" here, is to indicate a condition when the influences of other time series have been removed, or assume a situation that other time series remain unchanged. By changing s, similar to the DCCA cross-correlation coefficient ρ_{DCCA}, we can further estimate the partial cross-correlation levels on different time scales.

1. Holling, C. S. Understanding the complexity of economic, ecological, and social systems. *Ecosystems* **4**, 390–405 (2001).
2. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
3. José, A. R. *et al.* Nonlinearities, feedbacks and critical thresholds within the Earth's climate system. *Clim. Change* **65**, 11–38 (2004).
4. Shao, Y., Gu, G., Jiang, Z., Zhou, W. & Sornette, D. Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series. *Sci. Rep*, **2**, 835 (2013).
5. Ward, M. N. Diagnosis and short-lead time prediction of summer rainfall in tropical north Africa at interannual and multidecadal timescales. *J. Climate* **11**, 3167–3191 (1998).
6. Czaja, A., van der Vaart, P. & Marshall, J. A diagnostic study of the role of remote forcing in tropical Atlantic variability. *J. Climate* **15**, 3280–3290 (2002).
7. Ludescher, J. *et al.* Improved El Niño forecasting by cooperativity detection. *PNAS* **110**, 11742–11745 (2013).
8. Barnett, T. P. The interaction of multiple time scales in the tropical climate system. *J. Climate* **4**, 269–285 (1991).
9. Huang, N. E. *et al.* The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **454**, 903–995 (1998).
10. Huang, R. & Wu, Y. The influence of ENSO on the summer climate change in China and its mechanism. *Adv. Atmos. Sci.* **6**, 21–32 (1989).
11. Zhang, R., Sumi, A. & Kimoto, M. A diagnostic study of the impact of El Niño on the precipitation in China. *Adv. Atmos. Sci.* **16**, 229–241 (1999).
12. Gong, D.-Y. & Ho, C.-H. Shift in the summer rainfall over the Yangtz River valley in the late 1970s. *Geophys. Res. Lett.* **29**, 1436 (2002).

13. Baldwin, M. P. & Dunkerton, T. J. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *J. Geophys. Res.* **104**, 30937–30946 (1999).

14. Rittenour, T. M., Brigham-Grette, J. & Mann, M. E. El Niño like climate teleconnections in new England during the late Pleistocene. *Science* **288**, 1039–1042 (2000).

15. Mann, M. E. On smoothing potentially non-stationary climate time series. *Geophys. Res. Lett.* **31**, L07214 (2004).

16. Urban, F. E., Cole, J. E. & Overpeck, J. T. Influence of mean climate change on climate variability from a 155-year tropical Pacific coral record. *Nature* **407**, 989–993 (2000).

17. Wang, Y. *et al.* The Holocene Asian monsoon: links to solar changes and north Atlantic climate. *Science* **308**, 854–857 (2005).

18. Podobnik, B. & Stanley, H. E. Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Phys. Rev. Lett.* **100**, 084102 (2008).

19. Podobnik, B., Horvatic, D., Petersen, A. M. & Stanley, H. E. Cross-correlations between volume change and price change. *PNAS* **106**, 22079–22084 (2009).

20. Peng, C. K. *et al.* Mosaic organization of DNA nucleotides. *Phys. Rev. E* **49**, 1685–1689 (1994).

21. Koscielny-Bunde, E. *et al.* Indication of a universal persistence law governing atmospheric variability. *Phys. Rev. Lett.* **81**, 729–732 (1998).

22. Zebende, G. F. DCCA cross-correlation coefficients: Quantifying level of cross-correlation. *Physica A* **390**, 614–618 (2011).

23. He, L. & Chen, S. Multifractal detrended cross-correlation analysis of agricultural futures markets. *Chaos Soliton Fract.* **44**, 355–361 (2011).

24. Shadkhoo, S. & Jafari, G. R. Multifractal detrended cross-correlation analysis of temporal and spatial seismic data. *Eur. Phys. J. B* **72**, 679–683 (2009).

25. Xu, N., Shang, P. & Kamae, S. Modeling traffic flow correlation using DFA and DCCA. *Nonlinear Dyn.* **61**, 207–216 (2010).

26. Vassoler, R. T. & Zebende, G. F. DCCA cross-correlation coefficient apply in time series of air temperature and air relative humidity. *Physica A* **391**, 2438–2443 (2012).

27. Zhou, W. Multifractal detrended cross-correlation analysis for two nonstationary signals. *Phys. Rev. E* **77**, 066211 (2008).

28. Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M. & Francis, R. C. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.* **78**, 1069–1079 (1997).

29. Lau, K.-M & Weng, H. Coherent modes of global SST and summer rainfall over China: An assessment of the regional impacts of the 1997–98 El Niño. *J. Climate* **14**, 1294–1308 (2001).

30. Chan, J. C. L. & Zhou, W. PDO, ENSO and the early summer monsoon rainfall over south China. *Geophys. Res. Lett.* **32**, L08810 (2005).

31. Zhang, Y., Wallace, J. M. & Battisti, D. S. ENSO-like interdecadal variability: 1900–93. *J. Climate* **10**, 1004–1020 (1997).

32. Gershunov, A. & Barnett, T. P. Interdecadal modulation of ENSO teleconnections. *Bull. Amer. Meteor. Soc.* **79**, 2715–2725 (1998).

33. Hu, K., Ivanov, P. Ch., Chen, Z., Carpena, P. & Stanley, H. E. Effect of trends on detrended fluctuation analysis. *Phys. Rev. E* **64**, 011114 (2001).

34. Gao, J. *et al.* Assessment of long-range correlation in time series: How to avoid pitfalls. *Phys. Rev. E* **73**, 016117 (2006).

35. Yuan, N. & Fu, Z. Different spatial cross-correlation patterns of temperature records over China: A DCCA study on different time scales. *Physica A* **400**, 71–79 (2014).

36. Horvatic, D., Stanley, H. E. & Podobnik, B. Detrended cross-correlation analysis for non-stationary time series with periodic trends. *Europhys. Lett.* **94**, 18007 (2011).

37. Trenberth, K. E. The definition of El Niño. *Bull. Amer. Meteor. Soc.* **78**, 2771–2777 (1997).

38. Wang, H. The interannual variability of east Asian monsoon and its relationship with SST in a coupled atmosphere-ocean-land climate model. *Adv. Atmos. Sci.* **17**, 31–47 (2000).

39. Li, S., Lu, J., Huang, G. & Hu, K. Tropical indian ocean basin warming and east Asian summer monsoon: A multiple AGCM study. *J. Climate* **21**, 6080–6088 (2008).

40. Xie, S. *et al.* Indian ocean capacitor effect on Indo-Western pacific climate during the summer following El Niño. *J. Climate* **22**, 730–747 (2009).

41. Wu, B., Li, T. & Zhou, T. Relative contributions of the Indian Ocean and local SST anomalies to the maintance of the western North Pacific anomalous anticyclone during El Niño decaying summer. *J. Climate* **23**, 2974–2986 (2010).

42. Wang, B., Liu, J., Yang, J., Zhou, T. & Wu, Z. Distince principal modes of early and late summer rainfall anomalies in east Asia. *J. Climate* **22**, 3864–3875 (2009).

43. Chang, C.-P., Zhang, Y. & Li, T. Interannual and interdecadal variations of the east Asian summer monsoon and tropical Pacific SSTs. Part II: Meridional structure of the monsoon. *J. Climate* **13**, 4326–4340 (2000).

44. Zhou, T. *et al.* Why the western Pacific subtropical high has extended westward since the late 1970s. *J. Climate* **22**, 2199–2215 (2009).

45. Podobnik, B., Jiang, Z., Zhou, W. & Stanley, H. E. Statistical tests for power-law cross-correlated processes. *Phys. Rev. E* **84**, 066118 (2011).

Acknowledgments

Author contributions

N.Y. designed the study, N.Y. and Z.F. performed the study. N.Y., H.Z., L.P., J.L. and E.X. wrote the main manuscript. All the authors reviewed the manuscript.

Additional information