Multifractal Detrended Cross-Correlation Analysis of Sunspot Numbers and River Flow Fluctuations

S. Hajian, M. Sadegh Movahed

Department of Physics, Shahid Beheshti University, G.C., Evin, Tehran 19839, Iran

We use the Detrended Cross-Correlation Analysis (DCCA) to investigate the influence of sun activity represented by sunspot numbers on one of the climate indicators, specifically rivers, represented by river flow fluctuation for Daugava, Holston, Nolichucky and French Broad rivers. The Multifractal Detrended Cross-Correlation Analysis (MF-DXA) shows that there exist some crossovers in the cross-correlation fluctuation function versus time scale of the river flow and sunspot series. One of these crossovers corresponds to the well-known cycle of solar activity demonstrating a universal property of the mentioned rivers. The scaling exponent given by DCCA for original series at intermediate time scale, \((12 - 24) \leq s \leq 130\) months, is \(\lambda = 1.17 \pm 0.04\) which is almost similar for all underlying rivers at 1σ-confidence interval showing the second universal behavior of river runoffs.

To remove the sinusoidal trends embedded in data sets, we apply the Singular Value Decomposition (SVD) method. Our results show that there exists a long-range cross-correlation between the sunspot numbers and the underlying streamflow records. The magnitude of the scaling exponent and the corresponding cross-correlation exponent are \(\lambda \in (0.76, 0.85)\) and \(\gamma_{\text{cross}} \in (0.30, 0.48)\), respectively. Different values for scaling and cross-correlation exponents may be related to local and external factors such as topography, drainage network morphology, human activity and so on. Multifractal cross-correlation analysis demonstrates that all underlying fluctuations have almost weak multifractal nature which is also a universal property for data series. In addition the empirical relation between scaling exponent derived by DCCA and Detrended Fluctuation Analysis (DFA), \(\lambda \approx (h_{\text{sun}} + h_{\text{river}})/2\) is confirmed.

I. INTRODUCTION

Recently, due to the developments in the area of complex systems as well as data measurements and data analysis, one can find many opportunities for examination and interpretation of climate change which exhibit irregular systems [1–8]. It is well shown that the climate system is enforced by the well-defined seasonal periodicity, however the existence of unpredictable perturbation and chaotic functioning lead to extreme climate events. Indeed the climate is a dynamical system affected by tremendous factors and variables, such as solar activity which is represented by Sunspot numbers in this research [4–10]. All factors that control the trajectory of such mentioned systems have enormously large phase space and evolve as non-stationary processes, consequently we should explore it with stochastic tools to achieve reliable results. Nowadays, it has been clarified that a remarkably wide variety of natural systems can be characterized by long-rangeSuch cross-correlations address scientists toward fractal geometry of the underlying dynamical systems and can hopefully help us to predict future events. Existence and determination of power-law cross-correlations would help to promote our understanding of the corresponding dynamics and their future evolutions [11–14]. Beside, many events which controls earths climate, water runoff records assigned by rivers and sun activity play a crucial and survival roles for human life. The runoff water fluctuations are excellent climate criteria because they integrate evapotranspiration (output) and precipitations (input) over large areas. It is well accepted that the prediction of water runoff is fundamental for different aspects of social and economical reasons, ranging from the prediction of floods and droughts to planning of agricultural conditions. As a result of the periodicity in precipitation, river flow has also strong seasonal periodicity [12]. It is worth to note that unlike other climate components, water runoff may be directly influenced by human activity, like agriculture, drainage network morphology and so on, consequently makes it hard to distinguish the artificial and natural effects on the river flow data. Finding some or at least a universal behavior for different streamflow fluctuations as well as quantifying the impact of sun activity on various temporal and spatial scales of water runoff fluctuations can improve the recent hydrological models [15].

To this end, the statistical and multifractal analysis of river flows as well as influence of sun activity due to the interior and exterior chemical and physical properties of sun should be an important issue in the geophysical and hydrological systems.

The streamflow of rivers and sun activity have been studied from various point of views such as: the probability distribution [16,17], correlation and fractal behaviors [18–23], connection between volatility and nonlinearity of fluctuations [11,12,24,25], scale invariance for distribution function [26]. In addition, sun activity have been investigated by some methods in chaos theory [27] and also
multifractal analysis \[28,29\], wavelet analysis \[30\], cross-correlation functions between monthly mean sunspot areas and sunspot numbers \[31,32\], the relation between sunspot numbers fluctuation and number of flares, their evolution step \[33,34\], principal components and neural network methods to predict sunspots \[35,36\], sunspot areas time series and solar irradiance reconstructions \[37\], magnetic and dynamic properties of sunspots at the photospheric level \[38\] and the hydraulic-geometric similarity of river \[39–41\].

More recently, Pablo J. D. Mauas et. al., have investigated the solar forcing on climate, using the quantification of cross-correlation between the yearly sunspot numbers, irradiance reconstruction and streamflow of Paraná river \[9,10\]. On the other hand, the mechanisms for solar influence on the earth’s climate has been clarified in detailed from various point of views in \[42\]. Q. Zhang et. al., have investigated the universal behavior of streamflow records of the Pearl river \[15\].

After innovation of Hurst to propose the self-similar processes and its criteria, namely “Hurst exponent” \[18–23\], long-range correlated fluctuation behavior has also been reported for vast category of sciences, specifically the geophysical records (for more discussion see \[18–23,43–46\]). In the last decade, the modification prescription which is required for a full characterization of many data sets such as the runoff records, the various moments of the so-called fluctuation functions, have been introduced \[18–23\]. The effect of non-stationarity on the detrended fluctuation analysis has been investigated in \[47,48\].

Here we take a new approach and rely on the state-of-the-art algorithm to investigate the contribution of sinusoidal trends embedded in the data set as well as non-stationarity properties of the underlying series. We implement robust methods to explore the multifractal nature of cross-correlation between two important climate variables, the monthly streamflow of some rivers and sun activity represented by sunspot numbers (see Figure (1)), by using the novel approach in the fractal analysis, Detrended Cross-Correlation Analysis (DCCA) and its multifractal modification, the Multifractal detrended Cross-Correlation Analysis (MF-DXA) \[49,50\]. We restrict this article to use the sunspot numbers as the solar activity indicator, since there are many large and reliable data sets which can be considered as solar influence on the climate. Due to the presence of the sinusoidal trends in both sunspot numbers and the runoff river fluctuations and based on previous researches, one cannot expect to find a unique scaling behavior for fluctuation functions in all time scales (see section III), consequently, we have been motivated to use the well-known method, namely Singular Value Decomposition (SVD) method to exclude dominant trends in data set (see section II for more details). So after, clean data set will be used in the DCCA and MF-DXA methods.

This paper is organized as follows: in section II, we describe the methods which are used to determine the cross-correlation of two non-stationary time series, the Detrended Cross-Correlation Analysis (DCCA), and investigate the corresponding multifractal properties by using the Multifractal Detrended Cross-Correlation Analysis (MF-DXA). Section II will be continued by introducing a method to eliminate trends from the original data set, the Singular Value Decomposition (SVD), and describing data used in this paper. In section III, the multifractal cross-correlation of the underlying data sets will be examined. Section IV, will be devoted to the results and summary.
II. ANALYSIS TECHNIQUES AND DATA DESCRIPTION

Time series measured in the nature are usually affected by non-stationarities such as trends and artificial noises which must be well distinguished from the intrinsic fluctuations of the series. In many cases also, intrinsic fluctuations behave as non-stationary processes. Consequently, common methods in data analysis will be encountered with spurious or at least unreliable results. One of the most famous and well-known approach used in many studies is Multifractal Detrended Fluctuations Analysis (MF-DFA) [51,52]. This method has been applied to various areas, such as economical time series [53–55,56,57], river flow [13] and sunspot fluctuations [57,58], cosmic microwave background radiations [59], music [60,61], plasma fluctuations [62].

For many reasons, we are interested in studying the mutual influence of two series in the presence of non-stationarities. Obviously, traditional methods for this investigation become inaccurate procedures. Recently Jun et. al. have proposed an approach for analyzing correlation properties of a series by decomposing the original signal into its positive and negative fluctuation components [63]. Based on the previous study, Podobnik et. al. have modified the mentioned correlation method and improved it to explore the cross-correlation between two non-stationary fluctuations, named Detrended Cross-Correlation Analysis (DCCA) [49] and its generalized, the Multifractal Detrended Cross-Correlation Analysis (MF-DXA) which also examine higher orders detrended covariance [50].

As mentioned before, trends in data set may influence the accuracy of results. For reliable detection of the cross-correlations, it is essential to distinguish trends from the intrinsic fluctuations in data. Generally, trends embedded in measurements are of two types: Polynomial and Sinusoidal trends. Although the MF-DFA and MF-DXA methods eliminate the polynomial trends, the sinusoidal trends remain [47,48]. There are several robust methods to eliminate the sinusoidal one such as Fourier Detrended Fluctuations Analysis (F-DFA) [62,64], which is actually a high-pass filter, and Singular Value Decomposition (SVD) [65,66]. One of the most disadvantage of the F-DFA method is the reduction of the size of underlying data set. To apply the MF-DXA (see the following subsection), we have to synchronize two underlying data set which may not be done using the F-DFA method. On the other hand, the SVD method promises to remain the length as well as synchronization [65,66]. Here in order to eliminate the effect of sinusoidal trends, we apply the Singular Value Decomposition (SVD) [65,66]. After trends elimination, we use the MF-DXA to analyze the cleaned data sets.

A. DCCA and MF-DXA

One of the newly methods in analyzing two non-stationary time series is Detrended Cross-Correlation Analysis (DCCA) [49,63]. This method is a generalization of the DFA method in which only one time series was analyzed. Recently a generalized version of the DCCA method which is so-called MF-DXA, has been introduced [50]. Just same as the MF-DFA method, MF-DXA consists of the 4 steps (see [49,50,67] for more details):

(I): Computing the profiles of the underlying data series, \(x_k\) and \(y_k\), as

\[
X(i) \equiv \sum_{k=1}^{i} [x_k - \langle x \rangle] \quad i = 1, \ldots, N
\]

\[
Y(i) \equiv \sum_{k=1}^{i} [y_k - \langle y \rangle] \quad i = 1, \ldots, N
\]

the subtraction of mean is not compulsory. Since we are going to compare two different time series, we construct data sets with zero mean and unit variance using initial ones.

(II): Dividing each of the profile into \(N_s = \text{int}(N/s)\) non-overlapping segments of equal lengths \(s\), and then computing the fluctuation function for each segments. In order to take the whole series into account when the size of the data sets is not a multiple of considered time scale, \(s\), we do the same procedure from the opposite end, consequently one finds \(2N_s\) segments.

\[
F(s, m) = \frac{1}{s} \sum_{i=1}^{s} \{Y[(m-1)s+i] - y_m(i)\} \times \{X[(m-1)s+i] - x_m(i)\}
\]

for \(m = 1, \ldots, N_s\) and:

\[
F(s, m) = \frac{1}{s} \sum_{i=1}^{s} \{Y[N-(m-1)s+i] - y_m(i)\} \times \{X[N-(m-1)s+i] - x_m(i)\}
\]

for \(m = N_s+1, \ldots, 2N_s\), where \(x_m(i)\) and \(y_m(i)\) are a fitting polynomial in segment \(m\)th. Usually, a linear function is selected for fitting the function. If there is no trend in the data, a zeroth-order fitting function might be enough [68].

(III): Averaging the local fluctuation function over all the part, given by:

\[
F_q(s) = \left\{ \frac{1}{N_s} \sum_{m=1}^{N_s} [F(s, m)]^{q/2} \right\}^{1/q}
\]

Generally, \(q\) can take any real value, except zero. For \(q = 0\), equation (4) becomes:
\[ F_0(s) = \exp \left( \frac{1}{2N_s} \sum_{m=1}^{N_s} \ln F(s, m) \right) \] (5)

For \( q = 2 \), the standard DCCA is retrieved.

(IV): The final step is calculating the slope of the log-log plot of \( F_q(s) \) versus \( s \) which directly determines the scaling exponent \( \lambda(q) \), as:

\[ F_q(s) \sim s^{\lambda(q)} \] (6)

If both underlying series are equal then \( \lambda(q) \) is nothing else but so-called generalized Hurst exponent, \( h(q) \). In the absence of sinusoidal trends embedded in data sets, if one finds no scaling behavior for the fluctuation function in equation (6) or at least, there does not exist any unique exponent for all scaling ranges then there exists either short-range cross-correlation or not at all any cross-correlation. For a series of size \( N \), the minimum number of windows will be \( N_s = 2 \) corresponding to the maximum value of \( s = \text{int}(N/2) \). In addition, it has been demonstrated that to find the most correct value of scaling exponent by using DDA and DCCA methods, we should set \( s \leq (N/2) \), namely \( N_s \geq 2 \) [52].

To determine the slope of curve in the log-log plot of fluctuation function versus scale (equation (6)), we use Bayesian statistics [69]. We introduce measurements and model parameters as \( \{X\} : \{F_q(s)\} \) and \( \{\Theta\} : \{\lambda(q)\} \), respectively. Based on the Bayesian theorem, the conditional probability of the model parameters given data set (observation) is so-called posterior probability and is given by:

\[ P(\lambda(q)|X) = \frac{\mathcal{L}(X|\lambda(q))P(\lambda(q))}{\int \mathcal{L}(X|\lambda(q))d\lambda(q)} \] (7)

where the first term in the nominator of the right hand side is Likelihood and the second term contains all initial constraints concerning model parameters, so-called prior distribution. This term expresses the degree of belief about the model. In the absence of every prior constraints, the posterior function, \( P(\lambda(q)|X) \) is proportional to the Likelihood function. If there is no correlation between various measurements, consequently according to the central limit theorem, Likelihood function is given by a product of Gaussian functions as follows:

\[ \mathcal{L}(X|\lambda(q)) \sim \exp \left( -\frac{\chi^2(\lambda(q))}{2} \right) \] (8)

where:

\[ \chi^2(\lambda(q)) = \int ds \frac{[F_{\text{obs.}}(s) - F_{\text{The.}}(s; \lambda(q))]^2}{\sigma_{\text{obs.}}^2(s)} \] (9)

Here \( F_{\text{obs.}}(s) \) and \( F_{\text{The.}}(s; \lambda) \) are fluctuation function computed directly from the data set by using DFA or DCCA and determined by equation (6), respectively. Also, \( \sigma_{\text{obs.}}(s) \) is the mean standard deviation, associated to \( F_{\text{obs.}}(s) \). Apparently, this Likelihood function to be maximum when for a value of the scaling exponent, \( \lambda(q) \), \( \chi^2 \) reaches to its global minimum. The value of error-bar at 1σ confidence interval of \( \lambda(q) \) is determined by the likelihood function based on the following condition:

\[ 68.3\% = \int_{-\sigma}^{+\sigma} \mathcal{L}(X|\lambda(q))d\lambda(q) \] (10)

Finally we report the best value of scaling exponent at 1σ confidence interval according to \( \lambda^{+\sigma} \) for each moment, \( q \)’s.

B. Singular Value Decomposition (SVD)

Determining trends and construction proper detrending operations are important step toward robust analysis, specially in climatic data analysis. As given by Z. Wu and his collaborators [70], there is no unique definition of trend and any proper algorithm for extracting it from underlying stationary as well as non-stationary data sets. In another aspect, the trend in a real world data series, non-stationary one, is an intrinsic function imposed by the nature on data set. To identify the trend on a data set, we can investigate the series in whole domain or on some specific span of domains. For linear and stationary data sets choosing the length of data set as domain of trend may be suit but for a real world data set which is non-stationary and nonlinear, we need more precise definition of trend. As of the importance of investigating trends and probably removing them from series, one can point out to two following aims:

I) In some cases, there exists one or more crossover (time) scales, \( s_x \), in the log-log plot of \( F_q(s) \) versus \( s \) (equation (6)), segregating regimes with different scaling exponents. These patterns demonstrates that underlying fluctuation has different correlation behavior in various values of scales [47,48,65,66,71].

II) In many cases, crossovers are produced by the embedded sinusoidal trends, e.g. seasonal trends in the climate time series. Subsequently, to find scaling exponent of the intrinsic fluctuations, we should remove sinusoidal trends by using most robust detrending method so after, produced clean data set is used as an input for DCCA and MF-DXA methods [47,48,62].

In order to remove trends corresponding to the low frequency periodic behavior, we use Singular Value Decomposition method [72] instead of transformation a recorded data to the Fourier space using the method proposed in [73] (see also [47,48,71]). Using the SVD method not only we can track the influence of sinusoidal trends on the results, but also the synchronization of two underlying data
sets will be guaranteed [65,66]. In addition, we can determine over which scale, noises or trends have dominant contribution also the value of so-called crossover in the fluctuation function, in terms of the scale is computed by using DCCA method. [64–66,74,75]. After removing the dominant periodic functions, such as sinusoidal trends, we obtain the fluctuation exponent by direct application of the MF-DXA.

The SVD method consists of the following steps:

(I) Consider a data set which is superimposed with periodic trends \( \{x_i\}; i = 1, ..., N \). Embed \( x_i \) with parameters \((d, \tau)\) where \( d \) and \( \tau \) are the embedding dimension and the time delay, respectively. The embedded data can be represented as a matrix \( \Gamma \) given by:

\[
\Gamma = \begin{pmatrix}
x_1 & x_{1+\tau} & \ldots & x_{1+N-(d-1)\tau-1} \\
\vdots & \vdots & \ddots & \vdots \\
x_i & x_{i+\tau} & \ldots & x_{i+N-(d-1)\tau-1} \\
\vdots & \vdots & \ddots & \vdots \\
x_d & x_{d+\tau} & \ldots & x_{d+N-(d-1)\tau-1}
\end{pmatrix}
\]

as one can see, \( 1 \leq i \leq d \). For a periodic embedded trend, we set \( p \) as the number of dominant frequency components of the form \( e^{i(\omega_k t)} \), \( k = 1, \ldots, p \) in their power spectrum. For a fixed value of sample size, \( N \), the maximum value of \( d \) in the so-called trajectory matrix, \( \Gamma \), equals to \( d \leq N - (d - 1)\tau + 1 \) [66,74,76].

(II): Matrix \( \Gamma \) will be decomposed to the two left and right orthogonal matrices as follows:

\[
\Gamma = USV^\dagger
\]

where \( U_{d \times d} \) and \( V_{(N-(d-1)\tau) \times (N-(d-1)\tau)} \) are the left and right orthogonal matrices, respectively. The diagonal elements of \( S_{d \times (N-(d-1)\tau)} \) are the desired singular values, also known as eigenvalues. The Singular Value Decomposition of \( \Gamma \) is related to the eigendecomposition of the symmetric matrices \( \Gamma^\dagger \Gamma \) and \( \Gamma \Gamma^\dagger \), as \( \Gamma^\dagger \Gamma v_i = \lambda_i^2 v_i \) and \( \Gamma \Gamma^\dagger u_i = \lambda_i^2 u_i \). The nonzero eigenvalues of \( \Gamma^\dagger \Gamma \) are the same as that of \( \Gamma \Gamma^\dagger \); and determine the rank of \( \Gamma \). The eigenvalues are ordered such that \( \lambda_1 > \lambda_{i+1} \geq 0 \). The diagonal elements of \( S \) will be constructed by the ordered eigenvalues (the others will be set to zero). The rank of \( \Gamma \) is equal to the number of nonzero eigenvalues. The columns of \( U \) and \( V \) are constructed by the \( u_i \) and \( v_i \), namely the eigenvectors mentioned before, corresponding to the ordered eigenvalues.

By applying the SVD to the matrix \( \Gamma \) (i.e., \( \Gamma = USV^\dagger \)) we will get \( U \), \( S \) and \( V \). We consider the number of frequency components in the periodic trend to be \( p \). Set the dominant \( 2p + 1 \) eigenvalues in the matrix \( S \) to zero and hereafter named as \( S^* \). The filtered matrix \( \Gamma'^*_{d \times (N-(d-1)\tau)} \) determined by \( \Gamma'^* = US^*V^\dagger \) with elements \( \Gamma'^*_{ij} \). This in turn is mapped back on to a one-dimensional or filtered data given by:

\[
x^*_{i+j-1} = \Gamma'^*_{ij}
\]

where \( 1 \leq i \leq d \) and \( 1 \leq j \leq N - (d - 1) \tau \). The \( p \) dominant eigen-values and associating eigendecomposed vectors, represent trends subspace subsequently, the remaining \((d - p)\) eigenvalues and corresponding eigenvectors demonstrate intrinsic fluctuations subspace. In order to determine the value of \( p \) for a typical series such
as monthly sunspot data set, at first, we compute the power spectrum of monthly sunspot numbers. As shown in Figure (2), there are at least two dominant sinusoidal trends embedded in the monthly sunspot numbers. The first one corresponds to the well-known sun activity and second period belongs to the so-called Schwabe cycle interval. In order to eliminate mentioned trends, we set \( p = 2 \) in the SVD algorithm and compute the power spectrum of the cleaned data again. Our expectation is that there must be no deviation from scaling function as \( \nu^{-\beta} \) with \( \beta = 2h(q = 2) - 1 = 1.24 \pm 0.02 \) as one can see in the lower plot in Figure (2). It is worth to note that by increasing \( p \) from its optimum value in the SVD method, probably some intrinsic statistical properties of underlying data set will be destroyed. After which we use the detrended sunspot series as an input data for common DFA method (see Figure (3)). As shown in Figure (3), all of the crossover time scales which are produced due to the competition between sinusoidal trends embedded in sunspot series and intrinsic fluctuations, were diminished and intrinsic fluctuations will be retrieved. Our result is in agreement with the previous result regarding to scaling behavior of sunspot based on MF-DFA method accompanying by Fourier-Detrended Analysis, reported in [57]. Figure (3) confirms that, SVD method could remove sinusoidal trends.

In the log-log plot of fluctuation function versus time scale given by DFA method, also one can find three crossovers (see Figure (3)). To determine their value, we define error function as:

\[
\Delta(s) = \sqrt{|F_{\text{obs}}(s) - F_{\text{Linear}}(s)|^2}
\]

(14)

for each \( q \), where \( F_{\text{obs}}(s) \) and \( F_{\text{Linear}}(s) \) are the fluctuation functions for the original data and the filtered data produced by the SVD method, respectively. In Figure (4) we plot \( \Delta \) as a function of \( s \) for the sunspot numbers fluctuations. The first crossovers occurs at \( s_{1x} \sim [12 - 18] \) months corresponding to the annual period. The second crossovers is equal to \( s_{2x} \sim [130 - 170] \) months which is related to the well-known solar activity period. We cannot determine the value of third crossover with the acceptable uncertainty by using DFA method due to small size of current sunspot series.

C. Data Description

We use the up-to-date monthly sunspot numbers \((S_N)\) data series released by National Oceanic and Atmospheric Administration (NOAA) [77] and the Sunspot Index Data Center (SIDC) [78]. The monthly flow fluctuations of four famous rivers, namely Daugava at Latvia, Holston near Damascus, Nolichucky at Embreeville and French Broad at Asheville were collected from National Water Information System [79]. The original Daugava river data source is Latvian Environmental Geological and Meteorological Agency database [80]. The runoff dimension of the underlying rivers is \( m^3/s \) (see Figure (1)). The river flow fluctuations have been measured independently in corresponding area. All mentioned rivers are mixed feeding from rain, snowmelt water and groundwater. Table (I) reports some characteristics of river flow fluctuations used in this study. These data sets are almost long in length and most available series. In Figure (5) we plot underlying streamflows and sunspot numbers detrended data sets. In this figure, we use SVD method explained in the previous subsection for detrending data sets, also the mean and variance of series equate to zero and unity, respectively. According to Figure (5), one may find a remarkable correlation for streamflows and sunspot numbers. To quantify this correlation we use the Pearson’s correlation coefficients for \( X \) and \( Y \) defined by:

\[
r = \frac{(X - \langle X \rangle)(Y - \langle Y \rangle)}{\sqrt{(X - \langle X \rangle)^2(Y - \langle Y \rangle)^2}}
\]

The value of \( r \) for detrended river flows and sunspot numbers correspond to \( r_{\text{Daugava}} = +0.44 \), \( r_{\text{Holston}} = +0.94 \), \( r_{\text{Nolichucky}} = +0.88 \) and \( r_{\text{French-Broad}} = +0.94 \).

III. MF-DXA OF SUNSPOTS AND RIVER FLOW FLUCTUATIONS

To examine the multifractal properties and cross-correlation of sunspot numbers and river flow fluctuations, we apply the DCCA and MF-DXA methods. As mentioned in the previous section, we have to synchronize two time series of interests. The length of the river flow fluctuations may vary from 600 to 1300 months for the studied rivers, Daugava, Nolichucky, Holston and French Broad. On the contrary, there is more information avail-
FIG. 5. The detrended fluctuations of rivers (dashed line) and the corresponding sunspot (solid line) time series. Here we use a SVD method to trend series. The mean and variance of each series became zero and unity, respectively. The Pearson’s correlation coefficients are $r_{\text{Daugava}} = +0.44$, $r_{\text{Holston}} = +0.94$, $r_{\text{Nolichucky}} = +0.88$ and $r_{\text{French Broad}} = +0.94$.

able for sunspot databases, ranging from daily to annually data sets. To find reliable results given by DCCA, we synchronized the sunspot monthly series to all four river’s data sets.

Before applying any detrending program we apply DFA, DCCA and MF-DXA method. Log-log plot of $F(s)$ versus $s$ for original streamflow and sunspot data sets by using DFA method (Figure (6)) confirms that all underlying runoff rivers behave as the anti-persistent long-range correlated series for time scale $s < 12$ months, the value of Hurst exponent at this scale for Daugava, Holston, Nolichucky, French Broad and sunspot correspond to $H_{\text{Daugava}} = h(2) - 1 = 0.21 \pm 0.02$, $H_{\text{Holston}} = h(2) - 1 = 0.22 \pm 0.02$, $H_{\text{Nolichucky}} = h(2) - 1 = 0.21 \pm 0.02$, $H_{\text{French Broad}} = h(2) - 1 = 0.18 \pm 0.03$ and $H_{\text{Sunspot}} = h(2) - 1 = 0.11 \pm 0.03$, respectively.

Figure (7) shows the fluctuation function given by multifractal DCCA method for each river flow fluctuations versus sunspot numbers without applying any detrending procedure. Some crossover time scales in fluctuation functions are detected in Figure (7). One of those crossovers is known as cycle of solar activity. In addition the scaling behavior of fluctuation function at intermediate time scales, namely $[12 - 24] \leq s \leq 130$ months, is similar at $1\sigma$ confidence interval for $q = 2$ for all rivers used in this paper (see Figure (8)). The scaling exponent at this regime is $\lambda(q = 2) = 1.17 \pm 0.04$. As shown in Ref. [48], trends become dominant at intermediate regimen while at small and very large time scales, the intrinsic fluctuations to be dominated, therefor, we conclude that at $[12 - 24] \leq s \leq 130$ months the cross-correlation of water runoff and corresponding sunspot numbers determined by DCCA method for $q = 2$, is characterized as the universal behavior. For $s < [12 - 24]$ and $s > 130$ months the slopes of fluctuation functions are different for various rivers and implies that at mentioned time scales the local effects such as morphology, human activities, various drainage areas become dominant [15,81]. Moreover, cycle of solar activity represented by sunspot is one of most robust effect which affects on streamflow fluctuations.

Figure (9) shows the power spectrum of original and detrended mentioned rivers data sets and synchronized sunspot numbers series. The spectral density of French Broad and Nolichucky rivers behave in similar way
around the well-known solar activity. This outcome is also confirmed in Figure (7), since there is a sharp crossover in the fluctuation function versus sunspot numbers as a function of time scale. Here $s_{\text{max}} = \text{size of series}/2$. We shifted the vertical scale to make more sense.

Now, to explore the existence of cross-correlation in data set without sinusoidal trends from the DCCA and MF-DXA method we detrend each of the above synchronized data set by using the SVD method. Applying the MF-DXA method we detrend each of the above synchronized sunspot data set without sinusoidal trends from the DCCA and MF-DXA method for various values of $q$ for four rivers’ flow mentioned in the manuscript versus sunspot numbers as a function of time scale. Here $s_{\text{max}} = \text{size of series}/2$. We shifted the vertical scale to make sense.

![FIG. 7. Log-log plot of fluctuation functions given by the MF-DXA ($F_{\text{DCCA}}(s)$) method for various values of $q$ for four rivers’ flow mentioned in the manuscript versus sunspot numbers as a function of time scale. Here $s_{\text{max}} = \text{size of series}/2$. We shifted the vertical scale to make more sense.](image)

![FIG. 8. Log-log plot of fluctuation functions given by DCCA($F_{\text{DCCA}}(s)$) method for four streamflows fluctuations versus sunspot numbers as a function $s$ at intermediate time scale, $12 - 24 \leq s \leq 130$ months. For clearly, we shifted the vertical scale. Solid lines correspond to the power fit of fluctuation function and slopes have been computed by likelihood analysis.](image)

respectively. Filled triangle symbol indicates generalized scaling exponent, $\lambda(q)$. Weakly $q$-dependency of scaling exponent implies the rivers and sunspot numbers have almost multifractal behavior. For all studied rivers, the segments with fluctuations in the MF-DXA method near the mean value are larger than that of far from the mean as shown in Figure (12), namely the value of $\lambda(q)$ for $q < 0$ is almost larger than that of for $q > 0$. This shows the statistics of small fluctuations in water runoff are bigger than large fluctuations, rare events. According to Figure (12), the value of $\lambda(q)$ for $q > 0$ is almost constant, this indicates that the rare events in the cross-correlation function behave in similar way. On the other hand, the scaling behavior of fluctuation function, $F_q(s)$ in the MF-DXA method confirms the various type of fluctuations in different time scales have the same fractal features. This may be useful to extend the fractal properties of cross-correlation function which are derived in the short time scales to larger one in the absence of large time series. In addition, according to Figure (12), for $q = 2$, we obtain the empirical approximation, $\lambda(q = 2) \approx [h_{\text{sun}}(q = 2) + h_{\text{river}}(q = 2)]/2$.

According to auto-correlation function given by:

$$C(\tau) = \langle [x(i + \tau) - \langle x \rangle][x(i) - \langle x \rangle] \rangle \sim \tau^{-\gamma}$$

we can introduce the cross-correlation function for so-called long-range cross-correlation behavior as:

$$C_\times(\tau) = \langle [x(i + \tau) - \langle x \rangle][y(i) - \langle y \rangle] \rangle \sim \tau^{-\gamma_\times}$$
where $\gamma$ and $\gamma_x$ are the auto-correlation and cross-correlation exponents, respectively. Very often, direct calculation of these exponents are not recommended due to the non-stationarities and trends superimposed on the collected data. One of the reliable and proper statistical methods to calculate auto-correlation exponent is DFA method, namely $\gamma = 2 - \frac{2}{h}$ ($q = 2$) [13,71]. Recently B. Podobnik et. al. have demonstrated the relation between cross-correlation exponent, $\gamma_x$ and scaling exponent derived by equation (6) according to $\gamma_x = 2 - 2\lambda(q = 2)$ [49]. The magnitude of cross-correlation and scaling exponent derived by DFA and DCCA are reported in Table II.

To quantify the impact of El Niño (ENSO) phenomenon, we used the El Niño 3 index reported in [82] since 1950. Applying SVD method, the sinusoidal trends have been diminished, thereafter, cleaned data sets used as inputs for DCCA algorithm. Figure (13) indicates the results computed by DCCA method for the detrended data sets. There is a crossover time scale in the fluctuation function as a function of scale for all underlying rivers. The value of crossover equates to $s_x \sim 60$ months which is so-called the period of El Niño phenomenon. The value of scaling exponents and cross-correlation exponents reported in Table III, confirm that on $s \leq 60$ months, there exists a cross-correlation between ENSO phenomenon and rivers. For time scale larger than $s > 60$ months, this cross-correlation becomes ignorable [9,10]. We also compare the impact of ENSO phenomenon and sun activity on the streamflow of rivers. To this end, we should synchronize the El Niño, sunspot and river flow fluctuations. The value of scaling exponents and cross-correlation exponents for sunspot-river have also been reported in Table III. We find that the contribution of sun activity represented by sunspot is almost larger than ENSO phenomenon on the mentioned runoff rivers during mentioned period.

**IV. DISCUSSION AND CONCLUSION**

There are many motivations such as complexity of fluctuations in our environments which lead to examine them using robust and novel methods from complex systems and statistical physics point of views. Knowledge of natural variabilities are necessary to manage the energy resources and prevent disasters from social and economical point of views. In this paper, we analyzed three
important fluctuations in the nature, namely sun activity illustrated by sunspot numbers, El Niño phenomenon and streamflow of rivers. However, many climate indicators such as river flow fluctuations are affected by dominant seasonal trends, but recent researches have confirmed that, there are many variables causing the evolution of environmental conditions behave as a complex systems. Subsequently, applying the common methods in data analysis give incorrect or at least unreliable results. On the other hand to infer valuable results, the following necessary conditions should be satisfied: i) The length of underlying time series should be large enough. ii) The contribution of superimposed trends and noises on the recorded data must be small enough or at least distinguishable. Unfortunately above necessary conditions cannot be met in some practical measurements. To solve these problems, we rely on the robust methods in data analysis to explore the mutual effect of sunspot and runoff water fluctuations. Based on the mentioned motivation, we apply the most recent method, Multifractal Detrended Cross-Correlation Analysis (MF-DXA), to examine the cross-correlation and fractal properties of sunspot numbers and river flow fluctuations for some most famous and available rivers.

TABLE III. Values of the scaling and cross-correlation exponents of synchronized El Niño 3 index and river flow fluctuations as well as for sunspot and river in the same period, using the DCCA method.

| River      | $\lambda_{\text{ENSO}}$ | $\gamma_{\text{ENSO}}$ | $\lambda_{\text{sunspot}}$ | $\gamma_{\text{sunspot}}$ |
|------------|-------------------------|------------------------|-----------------------------|-----------------------------|
| Nolichucky | 0.89 ± 0.03             | 0.22 ± 0.06            | 0.94 ± 0.01                 | 0.12 ± 0.02                 |
| French Broad | 0.89 ± 0.03            | 0.22 ± 0.06            | 0.98 ± 0.02                 | 0.04 ± 0.02                 |
| Holston    | 0.85 ± 0.03             | 0.30 ± 0.06            | 0.90 ± 0.01                 | 0.20 ± 0.02                 |
| Daugava    | 0.74 ± 0.03             | 0.52 ± 0.06            | 0.77 ± 0.01                 | 0.46 ± 0.02                 |

FIG. 11. Fluctuation function given by DCCA method for detrended Holston, French Broad and Nolichucky rivers with their synchronized sunspot series in log-log scale versus time for $q = 2$. The maximum value of $s$ for each series equates to the integer part of half size of the corresponding river flow.

FIG. 12. Generalized Hurst exponent, $h(q)$ computed by MF-DFA (filled circle and square symbols) and $\lambda(q)$ derived by equation (6) in MF-DXA method (filled triangle symbol). Filled circle and square correspond to sunspot numbers and mentioned runoff river fluctuations, respectively.
scale, $s$. The first one corresponds to the annually period, the second is equal to the so-called solar activity, 11 years, while the last one may indicate the well-known Gleissberg period, but we must point out that this value with small uncertainty cannot be determined by this current data, because the size of data set is small. Here to eliminate those trends, we use the Singular Value Decomposition (SVD) method.

Results given by DFA method for original runoff rivers and sunspot numbers, demonstrated that all underlying runoff river behave as the anti-persistent long-range correlated series for time scale $s < 12$ months at $1\sigma$ confidence interval (Figure (6)). On the other hand, the value of Hurst exponents for series without sinusoidal trends exhibit runoff rivers and sunspot numbers are long-range correlated process (see Table II).

MF-DXA analysis for non-detrended underlying data sets implied universal scaling exponent of fluctuation function for $q = 2$ for all underlying rivers flow at intermediate time scales, $[12 - 24] \leq s \leq 130$ months, namely $\lambda = 1.17 \pm 0.04$ (see Figure (8)). This finding has recently been reported in Ref. [81] from power spectrum analysis. On the other hand there is a similar crossover in cross-correlation fluctuation functions at $s_x \sim 130$ months which is known as cycle of solar activity, for all studied rivers versus sunspot [13].

Data without sinusoidal trends were used as the inputs for DCCA and its generalized MF-DXA methods. Due to the scaling behavior of fluctuations function (equation (6)) versus time scale, $s$, we concluded that there exists a significant cross-correlation between sun activity and waster runoff river fluctuations. This cross-correlation confirms that the influence of other reasons on the river flow fluctuations such as human activity, drainage network morphology, land use patterns and topography may be ignored or at least cannot be distinguished form major effect which here is sun activity [81]. It must point out that, we know that the amount of radiation given off by the Sun (solar irradiance) is greatest when there are lots of sunspots. As discussed in detail in Ref. [9], the higher value of galactic cosmic ray, the higher the cloud cover which increases water resource of river, so we expect streamflow to be almost cross-correlated with solar activity. Other explanation are also investigated in [83,84].

The cross-correlation exponent, $\gamma_x$, of underlying data set in the presence of non-stationarities and trends have been determined by DCCA method. According to the relation between scaling exponent, $\lambda(q = 2)$, and cross-correlation exponent, $\gamma_x$, namely $\gamma_x = 2 - 2\lambda(q = 2)$, one find that for long-range cross-correlated signals, the value of $\gamma_x$ goes to small value demonstrating the cross-correlation function decreases more slowly and statistically two underlying data sets tend their present situation in their cross-correlation behavior. According to the values reported in Table II and Figures (10) and (11), there exists a long-range cross-correlation between rivers flow fluctuations used in this paper and sun activity indicated by sunspot numbers.

Weak $q$-dependency of the generalized Hurst, $h(q)$, and scaling, $\lambda(q)$, exponents, demonstrates that the original underlying series and also cross-correlation between sunspot numbers and each river behave as almost multifractal processes, demonstrating another universal characteristics. The value of $\lambda(q)$ for $q > 0$ remains almost constant, this indicates that the rare events in the cross-correlation function behave in similar way. The similarity in behavior of fluctuation function, $F_q(s)$ in the MF-DXA method demonstrates the various type of fluctuations in different time scales have the same fractal features. Based on this result, this may be useful to extend the fractal properties of cross-correlation function derived by using short time scales to larger one without relying on the large time series. In addition, the rare events in streamflow and sunspot cross-correlation analysis, are not affected by local characteristics.

According to our results, the empirical relation between $\lambda(q = 2)$ and Hurst exponent, $\lambda(q = 2) \approx [h_{sun}(q = 2) + h_{river}(q = 2)]/2$, has also been confirmed.

Since the value of cross-correlation of Daugava river and sunspot numbers is less than other rivers, one can conclude that, beside the sun as a main resource of energy in the nature, the river’s geographical situation and the source of rainfall for that river may have reasonable impact on runoff water fluctuation. Recently, even an opposite behavior between solar activity and rainfall fluctuation in equatorial east Africa has been noticed in the
recent report of the East Africa [85] explained in Ref. [9]. In addition, the value of $\lambda$ for French Broad and Nolichucky rivers are larger than two remained runoff rivers. Figure (9) also indicated, the power spectrum of French Broad and Nolichucky rivers behave in similar way around the well-known solar activity period. We also found a sharp crossover in the fluctuation functions derived by MF-DXA method for mentioned rivers in Figure (7). Since, the geographical positions of French Broad and Nolichucky rivers are close together, subsequently as expressed before, the impact of sun activity represented by sunspot, depends on the geographical properties of rivers.

One of the most important phenomenon which can affect the runoff rivers is El Niño (ENSO). To compare the contribution of ENSO phenomenon and sun activity, we used El Niño index 3. The results derived by DCCA have been indicated in Figure (13) and also reported in Table III. There exists a crossover time scale in the log-log plot of fluctuation function versus scale. The value of this crossover time scale is $s_\times \sim 60$ months regarding to well-known ENSO period. ENSO phenomenon is cross-correlated by streamflow fluctuations for $s \leq 60$ months, while for larger than this period, one can ignore this cross-correlation behavior. By comparison, the effect of ENSO phenomenon and sun activity on the river flows, we found that the contribution of sunspot is almost larger than ENSO index since 1950.

There are many methods to predict the sun activity [35,36], so based on our current results which is demonstrated the cross-correlation between detrended sunspots and water runoff, one may use those results to predict river flow fluctuations. Finally, one must point out that it should be interesting to extend the present analysis to a various set of runoff water fluctuations and other sun activity indicators such as solar irradiance, galactic cosmic rays flux and geomagnetic activity to find whether the values for the main parameters used in this analysis have a more universal validity of the solar influence on the streamflow fluctuations assigned as a climate variable in this paper.

Acknowledgements This research has been financially supported by Shahid Beheshti University research deputy affairs under grant No. 600/405.

[1] E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, H.-J. Schellnhuber, Phys. Rev. Lett. 81, 729 (1998).
[2] S.M. Barbosa, M.J. Fernandes and M.E. Silva, Physica A 371 (2006) 72573.
[3] K. Fraedrich, U. Luksch, R. Blender, Phys. Rev. E 70 (2004) 037301.
[4] W. B. White, J. Lean, D. R. Cayan, and M. D. Dettinger, J. Geophys. Res. 102, 3255 (1997).
[5] W. B. White, Journal of Geophysical Research (Oceans) 111, 9020 (2006).
[6] W. B. White and Z. Liu, Journal of Geophysical Research (Oceans) 113, 1002 (2008).
[7] N. D. Marsh and H. Svensmark, Physical Review Letters 85, 5004 (2000).
[8] B. Sun and R. S. Bradley, Journal of Geophysical Research (Atmospheres) 107, 4211 (2002).
[9] P. J. D. Mauas, E. Flamenco, A. P. Buccino, Phys. Rev. Lett. 101, 168501 (2008).
[10] Pablo J.D. Mauas, Andrea P. Buccino and Eduardo Flamenco, Journal of Atmospheres and Solar-Terrestrial Physics, in press, arXiv:1003.0414.
[11] V. Livina, Y. Ashkenazy, P. Braun, R. Monetti, A. Bunde and S. Havlin, Physical Review E, 67 (2003) 042101.
[12] V. Livina et al., Physica A, 330 (2003) 283-290.
[13] M. Sadegh Movahed and Eva1ds Hermanis, Physica A 387, 915 (2008).
[14] Sergio Arianos and Anna Carbone, JSTAT, P03037 (2009).
[15] Q. Zhang et al., Physics A, 388 (2009), 927-934.
[16] R.U. Murdock, J.S. Gulliver, J. Water Res. Pl-Asce. 119 (1993) 473.
[17] C.N. Kroll, R.M. Vogel, J. Hydrol. Eng. 7 (2003) 137.
[18] H.E. Hurst, Transact. Am. Soc. Civil Eng. 116 770 (1951).
[19] Y. Tessier, S. Lovejoy, P. Hubert, D. Schertzer, S. Pecnknold, J. Geophys. Res. Atmos. 101 (1996) 26 427.
[20] G. Pandey, S. Lovejoy, D. Schertzer, J. Hydrol. 208 (1998) 62.
[21] J.D. Pelletier, D.L. Turcotte, J. Hydrol. 203 (1997) (14) 198.
[22] E. Koscielny-Bunde, J.W. Kantelhardt, P. Braun, A. Bunde, S. Havlin, Proceedings of Conference Hydrofractals 2003, arXiv:physics/0305078.
[23] J.W. Kantelhardt, D. Rybski, S.A. Zschiegner, P. Braun, E. Koscielny-Bunde, V. Livina, S. Havlin, A. Bunde, Physica A 330 (2003) doi:10.1016/j.physa.2003.08.019 [these proceedings].
[24] V. Ashkenazy et al., Phys. Rev. L., 86(9) (2001) 1900-1903.
[25] V. Ashkenazy et al., Physica A, 323 (2003) 19-41.
[26] N. Loboda, Geophysical Research Abstracts, 8(2006) 00797.
[27] A. Veronig, M. Messerotti and A. Hanslmeier, 2000 Astronomy and Astrophysics, 357, 337-350
[28] V. I. Abramenko, 2005 Solar Physics, 228, 29-42
[29] V. I. Zhukov, arXiv:astro-ph/0304456.
[30] M. Temmer, A. Veronig, 2002 J. Rybák and A. Hanslmeier, Proc. 10th. European Solar Physics Meeting.
[31] M. Temmer, A. Veronig and A. Hanslmeier, Astronomy and Astrophysics, 390, 707-715 (2002).
[32] R. S. Bogart, 1982 Solar Physics, 76, 155
[33] M. Temmer and et. al., Astronomy and Astrophysics, 392-2, 699-712 (2002).
[34] M. Temmer, A. Veronig and A. Hanslmeier, 2003 Solar Physics, 512, 111-129
