mathematics

MDPI

Article

Multivariate Multifractal Detrending Moving Average Analysis of Air Pollutants

Milena Kojić [1], Petar Mitić [1,*], Marko Dimovski [2] and Jelena Minović [1]

[1] Institute of Economic Sciences, 11000 Belgrade, Serbia; milena.kojic@ien.bg.ac.rs (M.K.); jelena.minovic@ien.bg.ac.rs (J.M.)
[2] EVN AD, 1000 Skopje, North Macedonia; mdimovski16@gmail.com
* Correspondence: petar.mitic@ien.bg.ac.rs

Abstract: One of the most challenging endeavors of contemporary research is to describe and analyze the dynamic behavior of time series arising from real-world systems. To address the need for analyzing long-range correlations and multifractal properties of multivariate time series, we generalize the multifractal detrended moving average algorithm (MFDMA) to the multivariate case and propose a multivariate MFDMA algorithm (MV-MFDMA). The validity and performance of the proposed algorithm are tested by conducting numerical simulations on synthetic multivariate monofractal and multifractal time series. The MV-MFDMA algorithm is then utilized to analyze raw, seasonally adjusted, and remainder components of five air pollutant time series. Results from all three cases reveal multifractal properties with persistent long-range correlations.

Keywords: multifractal detrended moving average; multivariate analysis; air pollutants

Citation: Kojić, M.; Mitić, P.; Dimovski, M.; Minović, J. Multivariate Multifractal Detrending Moving Average Analysis of Air Pollutants. *Mathematics* **2021**, *9*, 711. https://dx.doi.org/10.3390/math9070711

Academic Editor: Raoul R. Nigmatullin

Received: 15 February 2021
Accepted: 18 March 2021
Published: 25 March 2021

1. Introduction

Research of complex systems has become increasingly important in both natural and social sciences. Time series derived from complex real-world systems exhibit nonlinear behavior that cannot be characterized by linear statistical models. Namely, time series from financial, environmental, and many other fields often manifest long-term memory and frequent large fluctuations that cannot be adequately explained with normal-like distributions. Conventional econometric models such as ARMA, GARCH, and EGARCH fail to capture and accommodate these properties and the nature of such series. The scaling law mostly used to describe these types of time series is a power law with a scaling exponent α that (at least) asymptotically describes the behavior of a quantity F as a function of a scale parameter s: $F(s) \sim s^{\alpha}$ [1]. The systems characterized by a scaling law typically represent fractals or multifractals, depending on whether they are described by one scaling exponent or by a multitude of scaling exponents [1,2].

Various methods have been developed to characterize the properties of fractals and multifractals. Detrended fluctuation analysis (DFA) is applied to time series data in various fields, but Peng et al. originally proposed it for identifying long-range dependence in DNA nucleotide sequences [3]. The DFA algorithm was extended to multifractal detrended fluctuation analysis (MFDFA) for describing the multifractal properties of time series [4]. Methods based on the moving average have also been developed. The first such method was introduced for estimating the Hurst exponent of self-affinity signals [5]. It was further extended to the detrending moving average (DMA) and the multifractal detrending moving average (MFDMA) by considering the second-order difference between the original time series and its moving average function [6]. The behavior of the multivariate time series has been widely studied [1,7–12]. The DFA and MFDFA have been recently extended to the multivariate cases (MVDFA and MV-MFDFA) to give insight into the multichannel data and auto-correlation behavior [1,12].

In this paper, inspired by Xiong and Shang [1] and Zhang et al. [12], we extend the method of MFDMA to a multivariate case (MV-MFDMA) to analyze the long-range correlations and multifractal properties of multivariate air pollutant time series data. A better understanding of the temporal and spatial variability of environmental time series is essential for modeling various phenomena. Algorithms that can handle multivariate time series have become exceedingly important, especially in dealing with environmental issues. Different environmental time series have been previously analyzed in the contexts of multifractality. For instance, Reference [13] used multifractal analysis on an air pollution index time series, while References [14,15] used six air pollutants for numerical experiments in their multifractal analysis. Fine particulate matter pollutants have been of interest for several research endeavors [16–18], while [19], for example, used a multifractal analysis on a time series from the European carbon futures markets. Multifractal analysis is often used for temperature analysis like [20] did for air temperature, or [21] for global methane concentrations and remotely-sensed temperature anomalies.

We chose data on air pollution as it directly impairs the environment, endangers entire ecosystems, causes biodiversity loses, and jeopardizes human health. Air pollution occurs when harmful or excessive amounts of gases, particles, and biological molecules are introduced into the atmosphere, whether they are of natural or anthropogenic origin. However, the consensus is that anthropogenic pollution is the main cause of most problems the environment faces today. Pollution is primarily caused by industrial activities, energy use, transport, and agricultural activities. Furthermore, some household activities, such as heating, can cause significant air pollution, although this is mostly relevant for developing and underdeveloped countries. Coordinated action and great willpower are needed to make positive changes because some pollutants, such as fine particulate matter and ground-level ozone continuously create significant health problems, while various emissions continue to damage the environment. Although air quality improvements are somewhat noticeable in recent years because of coordinated international actions and global, regional, and national policies, it is evident that a multidisciplinary approach is needed to advance this fundamentally important research area and create solutions to improve air quality.

Therefore, our paper has three objectives. We extend the MFDMA method to a multivariate case, suggest the MV-MFDMA algorithm, and then, conduct numerical experiments on multivariate processes to investigate the performance of the newly proposed MV-MFDMA. Finally, we apply the MV-MFDMA algorithm to explore the long-range correlations and multifractal properties of five air pollutant time series.

The remainder of the paper is organized as follows. We introduce the proposed MV-MFDMA algorithm in Section 2, while the validity of the MV-MFDMA algorithm is presented in Section 3. In Section 4, we outline the available data set and pre-processing analysis and provide the results obtained using air pollutant time series. Finally, the conclusions are drawn in Section 5.

2. Multivariate Multifractal Detrending Moving Average Analysis

In this section, we propose a multivariate multifractal detrended moving average algorithm (MV-MFDMA). Let $y = (y_{t,1}, y_{t,2}, \ldots, y_{t,p})$, $t = 1, 2, \ldots, N$ denote the p time series, where N is the number of observations in each time series. The newly constructed MV-MFDMA algorithm consists of the following steps.

Step 1: Calculate cumulative sums of each time series $i = 1, 2, \ldots, p$:

$$Y_{t,i} = \sum_{k=1}^{t} y_{k,i}. \tag{1}$$

Step 2: Calculate the moving average function of each time series $i = 1, 2, \ldots, p$ in a moving window of size n:

$$\tilde{Y}_{t,i} = \frac{1}{n} \sum_{k=-\lfloor (n-1)\theta \rfloor}^{\lceil (n-1)(1-\theta) \rceil} Y_{t-k,i}, \tag{2}$$

where $\lfloor \eta \rfloor$ is the largest integer not larger than η and $\lceil \eta \rceil$ is the smallest integer not smaller than η. The parameter $\theta \in [0, 1]$ specifies the position of the moving window. In general, the moving average function includes $\lceil (n-1)(1-\theta) \rceil$ data points in the past and $\lfloor (n-1)\theta \rfloor$ data points in the future. Here, we consider three different values of parameter $\theta = 0, 0.5, 1$. If $\theta = 0$, then the moving average function is calculated over all the past $(n-1)$ data points of the series, and hence, it refers to the backward moving average. For $\theta = 0.5$, the moving average function includes half past and half future data points in each window and refers to the central moving average. In the case of $\theta = 1$, the moving average function is calculated over all the future $(n-1)$ data points and refers to the forward moving average. For more details, see [6,22].

Step 3: Calculate the series residue by subtracting the moving average function $\tilde{Y}_{t,i}$ from $Y_{t,i}$:

$$e_{t,i} = Y_{t,i} - \tilde{Y}_{t,i}, \tag{3}$$

where t satisfies the criterion $n - \lfloor (n-1)\theta \rfloor \leq t \leq N - \lfloor (n-1)\theta \rfloor$.

Step 4: Divide the residue series $e_{t,i}$ into $N_n = \lfloor N/n - 1 \rfloor$ non-overlapping segments of equal length n. The segments are denoted by e_v such that $e_{v_{t,i}} = e_{l+t,i}$ for $1 \leq t \leq n$ with $l = (v-1)n$.

Step 5: Calculate the fluctuation variance $F^2(v, n)$ as a function of n for an arbitrary segment v:

$$F^2(v, n) = \frac{1}{n} \sum_{t=1}^{n} ||e_v||^2, \tag{4}$$

where $e_v = (e_{t,1}, e_{t,2}, \ldots, e_{t,p})$ and $|| \cdot ||$ stands for the Euclidean norm.

Step 6: Average over all local variances $F^2(v, n)$ to obtain the q^{th} order fluctuation function:

$$F_q^{MV-MFDMA}(n) = \left\{ \frac{1}{N_n} \sum_{v=1}^{N_n} [F^q(v, n)] \right\}^{\frac{1}{q}} \tag{5}$$

where $q \neq 0$ and $F^q(v, n) = (F^2(v, n))^{q/2}$ If $q = 0$,

$$F_0^{MV-MFDMA}(n) = \frac{1}{N_n} \sum_{v=1}^{N_n} \ln[F(v, n)]. \tag{6}$$

Step 7: Vary the values of segment size n to determine the power law relation between the function $F_q^{MV-MFDMA}(n)$ and the size scale n. If a time series exhibits multifractal properties, then $F_q^{MV-MFDMA}(n)$ for large values of n will follow a power law type of scaling relation, such as:

$$F_q^{MV-MFDMA}(n) \sim n^{h(q)}, \tag{7}$$

where $h(q)$ denotes the generalized Hurst exponent.

The generalized Hurst exponent $h(q)$ can be attained by the slope of the log-log plot of $F_q^{MV-MFDMA}(n)$ versus n through the method of ordinary least squares. The scaling exponent $h(q)$ when $q > 0$ traces the scaling behavior of the segments with larger fluctuations. Similarly, it shows the scaling behavior of segments with small fluctuations for $q < 0$. Usually, the large fluctuations are identified by a smaller scaling exponent

$h(q)$ for multifractal series. For monofractal time series, $h(q)$ is independent of q. If only short-range correlations or no correlations exist in the sequence, then the scaling exponent $h(2)$ equals 0.5. In this case, the time series display a random walk behavior. However, if there is long-range power law correlation, then $h \neq 0.5$. Furthermore, if $0.5 < h(2) < 1$, the long-range auto-correlations are persistent, which indicates that an increase is likely to be followed by another increase. If $0 < h(2) < 0.5$, we have long-range auto-correlations with anti-persistent behavior [3,4]. In this case, an increase is likely to be followed by a decrease.

Such a defined generalized Hurst exponent $h(q)$ is directly related to the multifractal scaling exponent $\tau(q)$: $\tau(q) = q \cdot h(q) - 1$. The monofractal time series are characterized by a linear form for the multifractal scaling exponent. The multifractality of the time series can be characterized by the singularity spectrum $f(\alpha)$ of the Holder exponent α. It is known that $\alpha = \frac{d\tau}{dq}$ and $f(\alpha) = q\alpha - \tau(q)$. The singularity spectrum $f(\alpha)$ indicates the dimension of the subset of the series that is characterized by α. The spectrum gives information about the relative dominance of various fractal exponents present in the series, while α characterizes the strength of the singularity. The wider the spectrum, the richer the multifractality behavior.

Note that if $i = 1$, then the MV-MFDMA reduces to the standard MFDMA. On the other hand, if $i \geq 2$, the MV-MFDMA investigates the multifractal features and long-range correlation properties of the multivariate process of dimension i as a whole.

3. Numerical Experiments on Synthetic Data Sets

The validity of the proposed MV-MFDMA algorithm is investigated in this section by performing numerical simulations on synthetic multivariate time series data. Synthetic data are modeled by the autoregressive fractionally integrated moving-average process (ARFIMA) following the procedure in [12]. Let us recall that the long-range correlations in stochastic variables can be modeled as:

$$z(t) = Z(d,t) + \epsilon(t), \tag{8}$$

where $\epsilon(t)$ follows a standard normal distribution, $d \in (-0.5, 0.5)$ is a memory parameter related to the Hurst exponent $h_z = 0.5 + d$, and $Z(d,t) = \sum_{n=1}^{\infty} a_n(d) z(t-n)$, $a_n(d) = d\Gamma(n-d)/(\Gamma(1-d)\Gamma(n+1))$ [12]. Then, the two-component ARFIMA process can be defined as:

$$x(t) = WX(d_1,t) + (1-W)Y(d_2,t)\epsilon_x(t) \tag{9}$$

$$y(t) = (1-W)X(d_1,t) + WY(d_2,t)\epsilon_y(t), \tag{10}$$

where $\epsilon_x(t)$ and $\epsilon_y(t)$ follow a standard normal distribution $N(0,1)$, $d_1, d_2 \in [0, 0.5]$ are the scaling parameters, and $W \in [0.5, 1]$ is a free parameter that controls the coupling strength between variables x and y [12]. If $W = 1$, then the variables x and y are fully decoupled and represent two separate ARFIMA processes. On the other hand, if W decreases from one to 0.5, then the correlations between variables x and y increase.

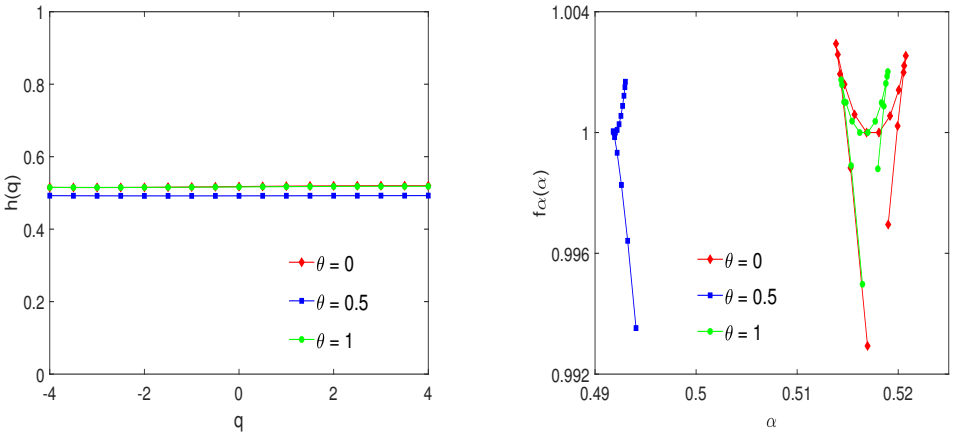
In the experiments, we considered multivariate monofractal time series, both independent and correlated. The length of each time series was $N = 2^{14}$. The results were averaged over 20 realizations of each type of test series.

3.1. Independent Multivariate Monofractal Series

We generated a trivariate uncorrelated monofractal time series with parameters corresponding to the scaling exponent $h = 0.5$ ($d = 0$) using the ARFIMA model (9) and (10). These series were considered as uncorrelated since the initial data channels were realizations of mutually independent uncorrelated monofractal series [12].

Figure 1 displays the scaling exponents $h(q)$ as a function of q and multifractal spectra $f(\alpha)$ obtained by applying the MV-MFDMA to trivariate uncorrelated monofractal series, for $\theta = 0, 0.5, 1$. As expected, regardless of θ, for trivariate uncorrelated monofractal series, the estimated scaling exponent $h(q)$ equals 0.5. This result implies the monofractal

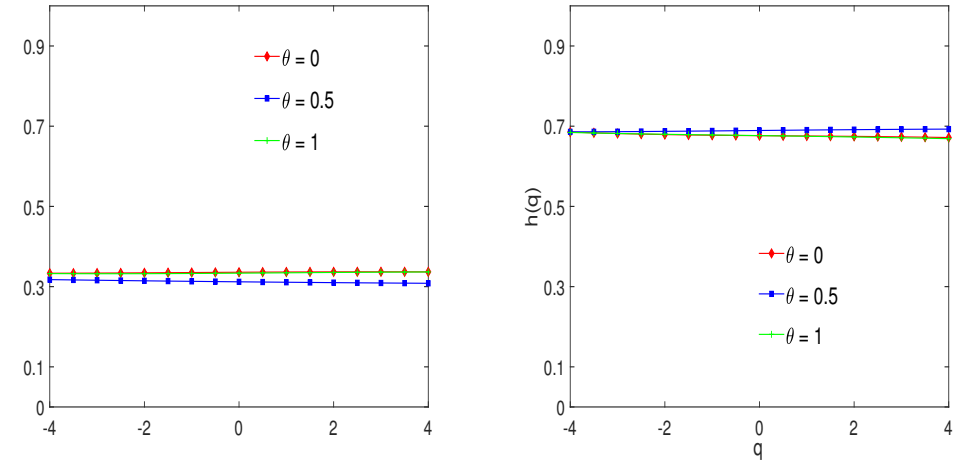
properties of multivariate series and holds regardless of the parameter θ. The multifractal spectra $f(\alpha)$ confirms the validity of the algorithm.



(**a**) The scaling exponents (**b**) The multifractal spectra

Figure 1. Uncorrelated monofractal series.

Next, we simulated trivariate anti-correlated and correlated monofractal time series with the following scaling exponents: $h = 0.3$ $(d = -0.2)$ and $h = 0.7$ $(d = 0.2)$; see Figure 2. The results show that $h(q)$ is independent of q and practicality equal to 0.3 and 0.7, respectively. These results hold regardless of the chosen parameter θ.

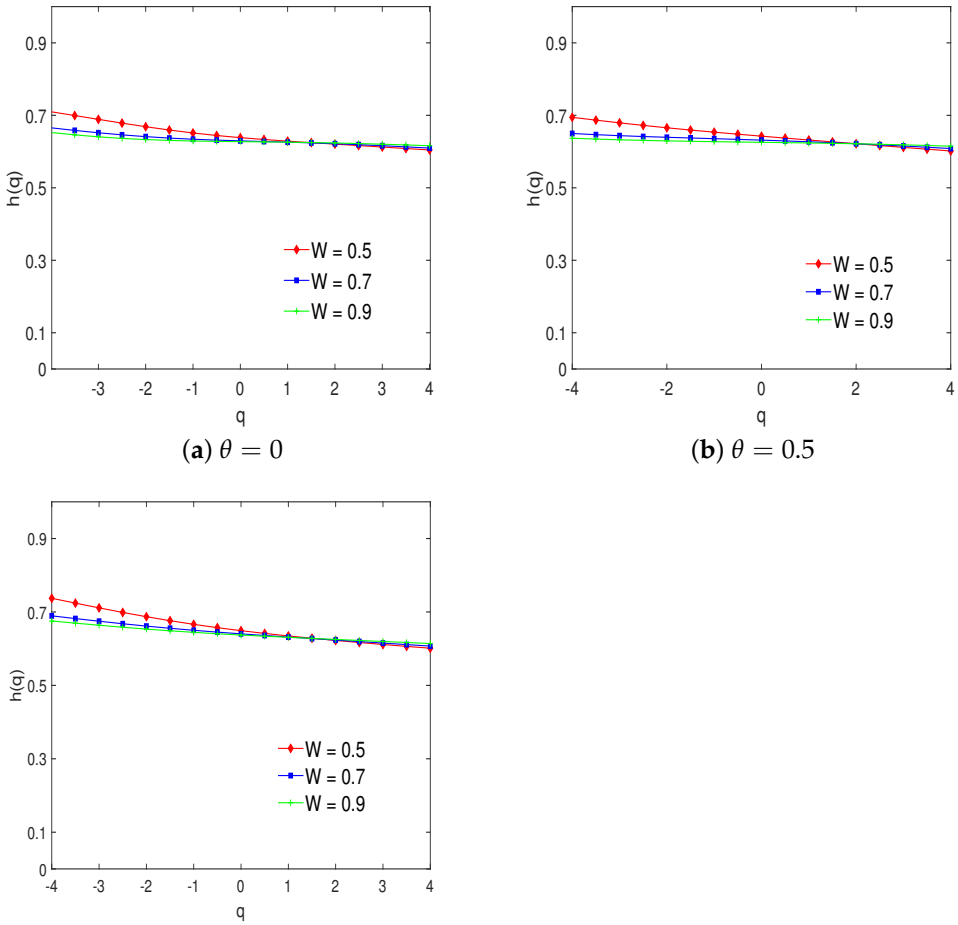


(**a**) Anti-correlated (**b**) Correlated

Figure 2. Independent monofractal series: the scaling exponents.

3.2. Correlated Bivariate Monofractal Series

The MV-MFDMA method was also applied and tested on the correlated bivariate time series simulated by the two-component ARFIMA model following the procedure in [12]. The following values of a free parameter controlling the coupling strength between time series were considered: $W = 0.5, 0.7, 0.9$, while the parameters of the model $d_1 = d_2 = 0.2$ corresponded to scaling coefficients $h_1 = h_2 = 0.7$. Figure 3 shows the average MV-MFDMA results of correlated bivariate time series with different correlation levels for $\theta = 0, 0.5, 1$. If we fix the coupling parameter W, the scaling exponents $h(q)$ do not depend on the value of q and are nearly equal to the value of 0.7. This result is in harmony with the scaling exponent of individual univariate series. Moreover, even if we change the coupling parameter W, the scaling exponents $h(q)$ remain the same for different values of q.



(**a**) $\theta = 0$ (**b**) $\theta = 0.5$

(**c**) $\theta = 1$

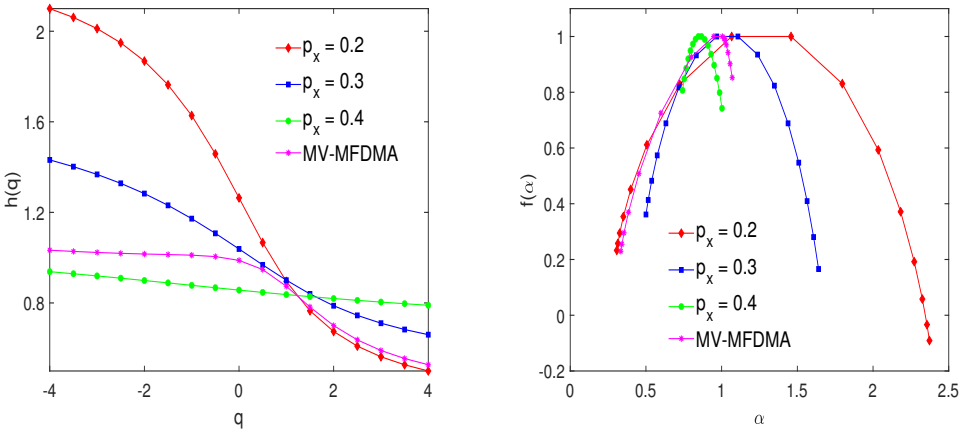
Figure 3. Correlated monofractal series: the scaling exponents.

3.3. Multivariate Multifractal Series

Finally, we applied the MV-MFDMA to multifractal time series simulated by the following binomial multifractal model:

$$x_k = p_x^{n(k-1)}(1-p_x)^{n_{max}-n(k-1)}, \tag{11}$$

where $p_x \in (0, 0.5)$ and $n(k)$ is the number of the digits equal to one in the binary representation of index k [12]. More precisely, we simulated trivariate time series x, y, z according to (11) and chose $p_x = 0.2$, $p_y = 0.3$ and $p_z = 0.4$, respectively. The proposed algorithm was applied to the generated series, and the obtained results are shown in Figure 4 in terms of scaling exponent $h(q)$ and multifractal spectra $f(\alpha)$ for $\theta = 0.5$. We also present the same results for individual time series.



(**a**) Scaling exponent (**b**) Multifractal spectra

Figure 4. Multifractal series.

The individual univariate time series manifest the expected multifractal characteristics, while it can be observed that the scaling exponents $h(q)$ of the considered trivariate multifractal series obtained by the new MV-MFDMA vary among the scaling exponents $h(q)$ of individual univariate series. For negative values of q, the scaling exponents $h(q)$ of the trivariate multifractal series are most harmonized to the scaling exponents of a univariate series for $p_z = 0.4$. On the other hand, for positive values of q, it is most adjusted to the scaling exponents of a univariate series for $p_x = 0.2$. This phenomenon was also revealed in [12]. The large fluctuations of univariate series with $p_x = 0.2$ have a major impact on the large fluctuation of multivariate series. Similarly, the major impact on the small fluctuations of multivariate series originates from the small fluctuations of univariate series with $p_z = 0.4$.

Based on the presented results, we can observe that the MV-MFDMA algorithm preserves the properties of the univariate MFDMA. Namely, it is known that algorithms based on the multifractal detrending moving average are sensitive to the sample size, the selection of the scaling range, the choice of q-orders, and the position of the moving window [23]. These algorithms require a larger sample size, i.e., a larger number of observations are needed to obtain a more accurate estimation of the multifractal spectrum. Then, the accuracy is also influenced by the selection of the appropriate scaling range. Equal spacing between scales is suggested, where the scaling range varies from 10 to $N/10$ [23]. The choice of the q-orders should include both positive and negative values. It is evident from the presented results that the algorithm depends on the position of the moving window. According to [24], the lowest accuracy is obtained when $\theta = 0.5$, while $\theta = 0$ and $\theta = 1$ provide higher accuracy.

A comparison of the proposed MV-MFDMA algorithm with other algorithms for detecting multifractal properties is given in Table 1.

Table 1. Comparison of the algorithms for detecting multifractality.

Method	Description	Comments
MFDFA [4]	• applies to univariate time series • detects a long-range correlation and multifractal properties • detrends the original series by removing its average	• requires a larger sample size for accurate estimates • non-overlapping segmentation can cause pseudo fluctuation errors
MFDMA [6]	• applies to univariate time series • detects a long-range correlation and multifractal properties • detrends the original series by removing the moving average function	• requires a larger sample size for accurate estimates • the backward MFDMA outperforms the MFDFA
MVDFA [1]	• applies to multivariate time series • represents a generalization of the DFA • describes the autocorrelations' behavior	• preserves the characteristics of the univariate DFA • characterized by one scaling exponent • requires an equal length of time series
MV-MFDFA [12]	• applies to multivariate time series • represents a generalization of the MFDFA • detects a long-range correlation and the multifractal properties of multichannel data directly	• preserves the characteristics of the univariate MFDFA • requires an equal length of time series
MV-MFDMA Section 2	• applies to multivariate time series • represents a generalization of the MFDMA • detects a long-range correlation and the multifractal properties of multichannel data directly	• preserves the characteristics of the univariate MFDMA • requires an equal length of time series • represents an alternative of both the MV-MFDFA and the MFDMA

4. Air Pollutant Time Series via MV-MFDMA

In this section, we present the data and analyze the multifractal properties of multivariate air pollutant time series. We performed the MV-MFDMA algorithm on raw series, seasonally adjusted series, and remainder components and then compared the results.

4.1. Data

Data for this paper were acquired from [25]. For the purposes of testing the MV-MFDMA, we used datasets of five air pollutants presented in Table 2. To obtain uniform series in terms of length, values were used only on days when all five variables were measured.

Table 2. Variables.

Variable	Time-Span	Unit	Frequency	State	Monitor Site
Carbon Monoxide (CO)	2010–2019	ppm	daily	California	60010009
Nitrogen Dioxide (NO_2)	2010–2019	ppb	daily	California	60010009
Ozone (O_3)	2010–2019	ppm	daily	California	60010009
Particulate Matter ($PM_{2.5}$)	2010–2019	$\mu g/m^3$	daily	California	60010009
Sulfur Dioxide (SO_2)	2010–2019	ppb	daily	California	60010011

Table 3 provides the descriptive statistics of the air pollutant time series. Positive values of skewness indicate that the distributions of all air pollutants are positively skewed. Kurtosis is greater than three for all air pollutants, which indicates that the homogeneity of the distribution in all cases is leptokurtic in relation to the normal distribution. The Jarque–Bera test rejects the null hypothesis of normality in air pollutants' series distribution, which is further corroborated by the fact that kurtosis is larger than three, and there is no zero skewness in any of the five cases. The statistics of the augmented Dickey–Fuller (ADF) unit root test based on both the Akaike (AIC) and Schwarz information criterion (SIC) reject the null hypothesis of a unit root at the 1% significance level. Additionally, the Philips–Perron (PP) test also rejects the null hypothesis of a unit root at the same significance level. Both the ADF and PP tests were done for both the trend and intercept and just the intercept cases. The *t*-statistics for all six cases of unit root tests are also presented in Table 3 and indicate that the air pollutants series are stationary.

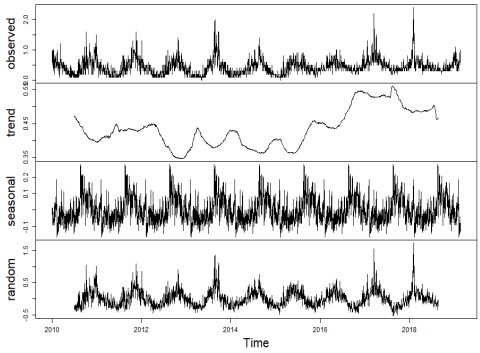
Table 3. Descriptive statistics: raw data series.

		CO	NO_2	O_3	$PM_{2.5}$	SO_2
Mean		0.4375	21.4214	0.0280	8.8429	1.9949
Std. Dev.		0.2508	11.0571	0.0091	7.7520	2.7298
Skewness		1.4251	0.6122	0.1371	8.5209	7.9297
Kurtosis		7.4013	3.4522	4.2429	131.7860	141.7981
Jarque–Bera		3825.237 *	237.0335 *	225.3758 *	2,347,909 *	2,715,223 *
Observations		3339	3339	3339	3339	3339
ADF (AIC)	intercept	−4.4779 *	−4.8280 *	−6.6122 *	−12.1969 *	−7.4589 *
	trend and intercept	−4.5956 *	−5.0117 *	−6.6436 *	−12.1979 *	−8.6055 *
ADF (SIC)	intercept	−5.9964 *	−5.8778 *	−8.0346 *	−16.5856 *	−10.4407 *
	trend and intercept	−6.1243 *	−6.0941 *	−8.0991 *	−16.5850 *	−11.4677 *
PP	intercept	−27.8682 *	−34.4109 *	−40.4884 *	−19.5129 *	−52.3557 *
	trend and intercept	−28.07811 *	−35.2493 *	−40.7289 *	−19.5138 *	−50.2368 *

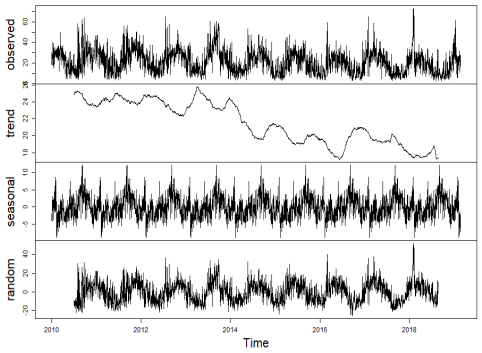
Note: ADF—augmented Dickey–Fuller unit root test; PP—Philips–Perron unit root test; AIC—Akaike information criterion; SIC—Schwarz information criterion; *—stat.sig.< 0.01.

The state of California was chosen because it is usually singled out as a leader in air pollution among other U.S. states. Even though all air pollutants are high in California, the O_3 levels particularly stand out. Since seasonal trends can sometimes influence the multifractal behavior of time series, it is preferable to perform seasonal decomposition; see for example [15,26,27]. It is interesting to note that a common approach for analyzing trendless fluctuations is absent. Nigmatullin and Vorobev [28] stated that in most cases, authors use traditional methods, such as the Fourier method, the wavelet method, Yulmenteyev's method, and Timashev's method or some additional processing algorithms that are also based on conventional methods containing some model assumptions and treatment methods associated with continuous mathematics. Nigmatullin, Lino, and Maione [29] pointed out that these sets of methods solve some specific tasks, but cannot be viewed as universal. To address this gap, a universal "platform" for treating various types of different trendless sequences has been proposed (see [28,29] for more details). For the purposes of this paper, we adjusted the considered time series by performing seasonal

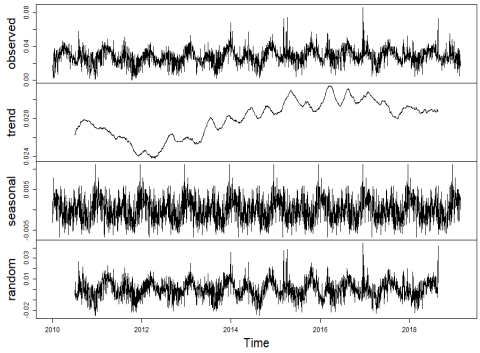
and trend decomposition using LOESS (STL) decomposition [30]. Through this algorithm, each time series is additively decomposed into deterministic trends, seasonal components, and stochastic remainder components. Removing seasonal components from the raw time series does not significantly influence the descriptive statistics results. All the descriptive statistics and unit root tests conclusions from the raw data are valid for the seasonally adjusted data as well (see Appendix A). Figure 5 shows that the seasonal components of all time series are characterized by the annual oscillation. Figure 5 also demonstrates, among others, the evolution trend of five air pollutant concentration series. The $PM_{2.5}$ and SO_2 series are more prone to extreme spikes, whereas NO_2, CO, and O_3 concentrations have somewhat different trends in winter and summer, indicating a seasonal inclination. On the other hand, the trend components are characterized by a very small range of variability and do not show a large temporal evolution. The range of variability is the largest for series NO_2 and $PM_{2.5}$. The remainder components of all time series do not display discernible patterns, and there are small fluctuations around zero.



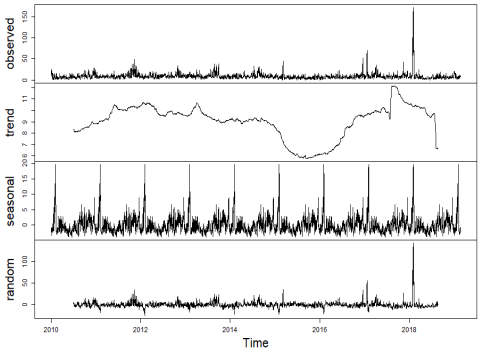
(**a**) CO



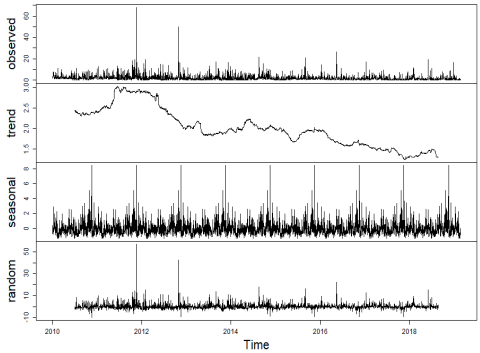
(**b**) NO_2



(**c**) O_3



(**d**) $PM_{2.5}$



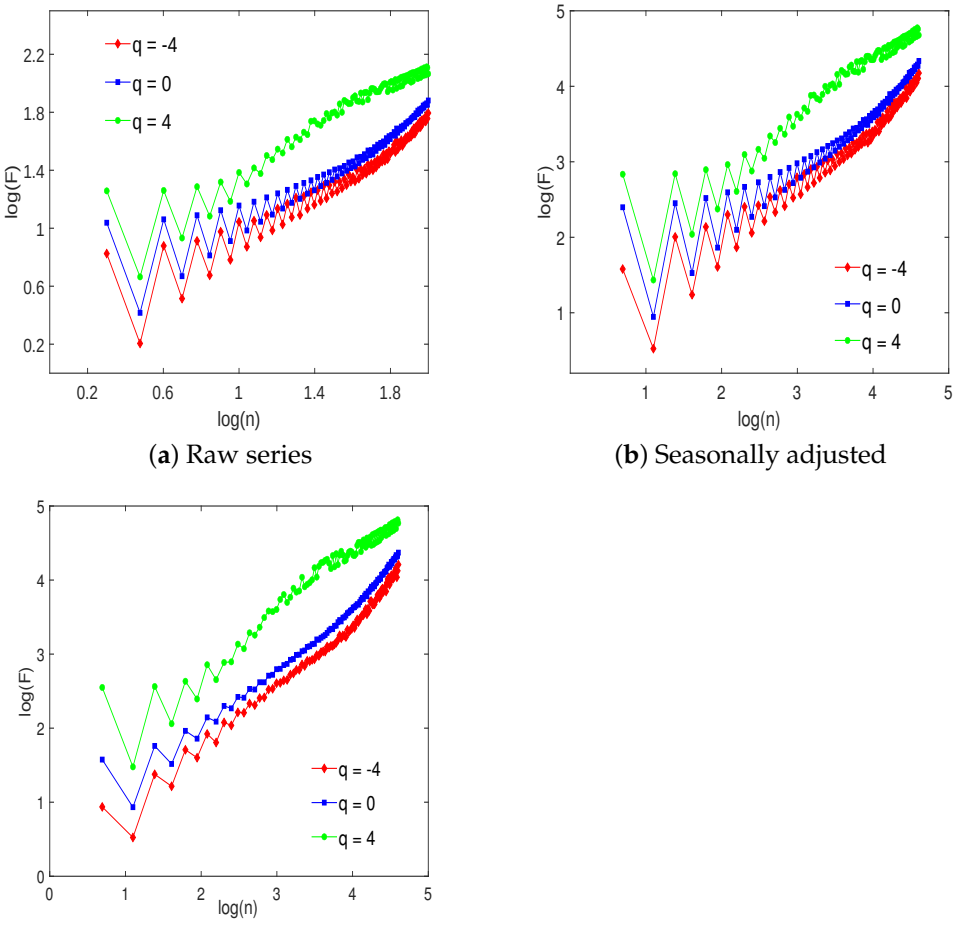
(**e**) SO_2

Figure 5. STL decomposition of air pollutant time series.

4.2. Results

The MV-MFDMA presented in Section 2 was applied to the above-mentioned raw and seasonally adjusted air pollutant time series obtained upon removal of the seasonal components, as well as on the remainder components of the series. Applying multifractal analysis on the remainder components ensures the identification of the dynamic characteristics of air pollutants' inner fluctuations and enhances the robustness of the results. We set the range of time scale n to be $10 < n < N/10$, where N is the length of each time series.

Figure 6 illustrates the behavior of the MV-MFDMA fluctuation function versus time in the logarithmic scale for all three case: raw, seasonally adjusted, and remainder. The plots are presented for $q = -4, 0, 4$ and fixed $\theta = 0.5$. All curves are approximately linear under large scales, indicating the power law behavior and the presence of multifractality. The MV-MFDMA fluctuation functions have quite similar behavior in all three cases, while we can note that fluctuations are the smallest for the remainder components. Fluctuations are observed regardless of the q values for small $log(s)$ values. That is, the results show crossover time scales, i.e., different scaling laws and scaling exponents for time scales $n < n^*$ (where $n^* \approx 3.32$ for raw and $n^* \approx 12$ for the other two cases). When $n > n^*$, we find that the MV-MFDMA fluctuation functions nicely respect the scaling relation, i.e., scale well with the scale size. Additionally, although not displayed, the MV-MFDMA is tested for $\theta = 0, 1$. The results confirm the power law behavior and demonstrate the absence of large fluctuations.



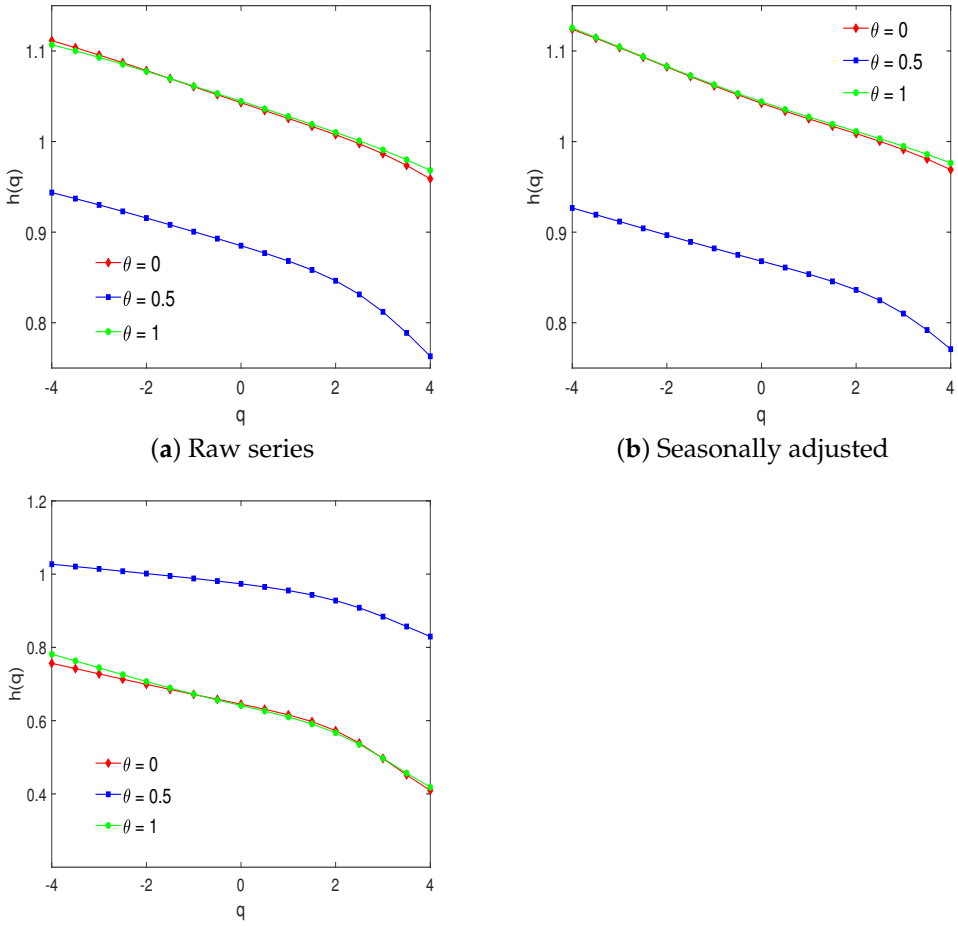
(**a**) Raw series (**b**) Seasonally adjusted

(**c**) Remainder

Figure 6. Power law dependence of the fluctuation functions with respect to the scale (log-log plots).

Figure 7 shows the values of the generalized Hurst exponents $h(q)$ versus q, where the variation of q belongs to the interval $[-4, 4]$ for raw, seasonally adjusted, and remainder

series. In all cases for $\theta = 0, 0.5, 1$, we observe the dependence of MV-MFDMA $h(q)$ on q. The MV-MFDMA $h(q)$ monotonically decreases when the value of q increases. This result implies the multifractality of the examined time series. It is interesting to note that raw and seasonally adjusted series display almost the same behavior of the generalized Hurst exponents $h(q)$ for all values of parameter θ. On the contrary, the values of the generalized Hurst exponents $h(q)$ for the remainder series are above 0.5 for $\theta = 0.5$, which indicates a strong long-term persistence. However, in the cases of the forward and backward moving average, the values of $h(q)$ are lower.



(**a**) Raw series (**b**) Seasonally adjusted

(**c**) Remainder

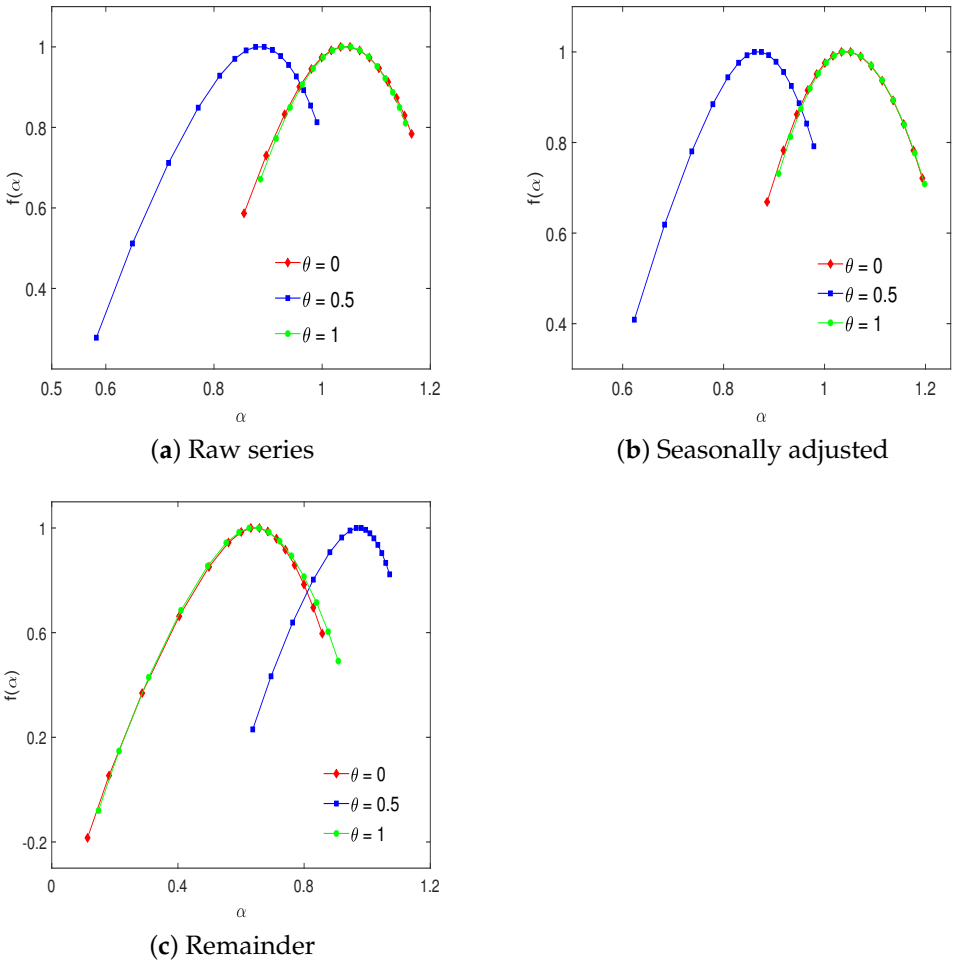
Figure 7. Generalized Hurst exponents.

Figure 8 displays the multifractal spectra of $\alpha \sim f(\alpha)$ via the Legendre transform. The shapes of the inverted U parabola of the spectra confirm the fact of the multifractality in the time series. The smallest range of the multifractality is observed when $\theta = 0.5$. The results hold for all raw, seasonally adjusted, and remainder series.

With the aim of quantifying the degree of the multifractality of the multivariate system, we use the following measures (12) and (13) [12]:

$$\Delta h = \max h(q) - \min h(q) \tag{12}$$

$$\Delta \alpha = \alpha_{max} - \alpha_{min}. \tag{13}$$



(**a**) Raw series (**b**) Seasonally adjusted

(**c**) Remainder

Figure 8. The multifractal spectra.

The larger the measures of Δh and $\Delta\alpha$, the richer and stronger the degree of multifractality of the time series is. Tables 4–6 present the scaling exponent $h(2)$ and the multifractality degrees Δh and $\Delta\alpha$ for the individual and multivariate air pollutants for all levels $\theta = 0, 0.5, 1$ and for raw, seasonally adjusted, and remainder series, respectively.

Table 4. The scaling exponent $h(2)$, multifractality degrees Δh, and $\Delta\alpha$: raw series.

		CO	NO_2	O_3	$PM_{2.5}$	SO_2	Multiv.
$\theta = 0$		0.9980	1.0164	1.0190	0.9823	0.9785	1.0076
$\theta = 0.5$	$h(2)$	0.9295	0.8741	0.7073	0.7071	0.6414	0.8463
$\theta = 1$		0.9937	1.0126	1.0190	0.9775	0.9837	1.0101
$\theta = 0$		0.2204	0.1544	0.0559	0.3006	0.2016	0.1525
$\theta = 0.5$	Δh	0.2074	0.1658	0.0301	0.2875	0.4126	0.1807
$\theta = 1$		0.2126	0.1501	0.0559	0.2844	0.3074	0.1387
$\theta = 0$		0.3888	0.2678	0.1078	0.6490	0.5710	0.3099
$\theta = 0.5$	$\Delta\alpha$	0.3854	0.3003	0.0650	0.6950	0.7838	0.4080
$\theta = 1$		0.3667	0.2573	0.1078	0.6045	0.6098	0.2681

Table 5. The scaling exponent $h(2)$, multifractality degrees Δh, and $\Delta\alpha$: seasonally adjusted series.

		CO	**NO_2**	**O_3**	**$PM_{2.5}$**	**SO_2**	**Multiv.**
$\theta = 0$		1.1277	1.2088	1.0583	1.0980	1.1991	1.1140
$\theta = 0.5$	$h(2)$	1.1333	0.9998	0.7086	0.8579	0.7928	0.9192
$\theta = 1$		1.1399	1.1881	1.0608	1.1425	1.2301	1.1149
$\theta = 0$		0.1758	0.2485	0.0522	0.2778	0.2814	0.1551
$\theta = 0.5$	Δh	0.2988	0.1793	0.0190	0.3510	0.2923	0.1559
$\theta = 1$		0.2002	0.2162	0.0570	0.3208	0.3142	0.1490
$\theta = 0$		0.3051	0.4907	0.1006	0.6269	0.5462	0.3076
$\theta = 0.5$	$\Delta\alpha$	0.6336	0.3357	0.0385	0.8267	0.5706	0.3557
$\theta = 1$		0.3834	0.4052	0.1150	0.7193	0.6135	0.2891

Table 6. The scaling exponent $h(2)$, multifractality degrees Δh, and $\Delta\alpha$: remainder component.

		CO	**NO_2**	**O_3**	**$PM_{2.5}$**	**SO_2**	**Multiv.**
$\theta = 0$		0.6209	0.6123	0.5481	0.3940	0.4522	0.5727
$\theta = 0.5$	$h(2)$	1.0275	0.9819	0.8858	0.7224	0.6188	0.9279
$\theta = 1$		0.6206	0.6065	0.5603	0.3878	0.4435	0.5669
$\theta = 0$		0.4764	0.3253	0.3109	0.6585	0.5222	0.3471
$\theta = 0.5$	Δh	0.1945	0.1448	0.1793	0.4061	0.3391	0.1974
$\theta = 1$		0.4954	0.3624	0.3160	0.6304	0.3737	0.3632
$\theta = 0$		0.8184	0.6024	0.5106	1.1022	1.0444	0.7440
$\theta = 0.5$	$\Delta\alpha$	0.3740	0.2705	0.3500	0.7443	0.6627	0.4341
$\theta = 1$		0.8240	0.6388	0.5368	1.0571	0.7054	0.7603

Raw and seasonally adjusted series have similar characteristics of multifractal parameters. For $\theta = 0$ and $\theta = 1$, the values of $h(2)$ are approximately around one, while in the case of $\theta = 0.5$, the values are a bit lower, namely $h(2) \approx 0.85$ and $h(2) \approx 0.9$ for raw and seasonally adjusted series, respectively. This holds for both individual and multivariate time series. The results imply that both individual and multivariate series show persistent long-range correlation. Analogous to $1/f$, noise can be deducted in the case of the scaling exponent $h(2)$ of multivariate time series less than one. In the case of the remainder components, the fluctuations are also characterized by a long-term persistence. However, the values of $h(2)$ suggest that the remainder components are characterized by a weak persistence suggesting stable trends in the next period. In summary, multivariate series of five air pollutants show consistent, positive autocorrelation behavior. In all three cases, an interesting fact is that all values of Δh and $\Delta\alpha$ are significantly larger than zero. On the other hand, the values for multivariate air pollutants range around those values for individual time series and are lower than the maximum values of the individuals in each case. Both the Δh and $\Delta\alpha$ values in the remainder components are higher than the values of both raw and seasonally adjusted series. This finding reveals that the considered air pollutants can be characterized by multifractality behavior. Furthermore, since the relationship always exists among the individuals, it possibly makes the multifractal level weaker. As a result, the Δh and $\Delta\alpha$ values for multivariate time series are smaller than the maximum and average of the Δh and $\Delta\alpha$ values of individuals.

5. Conclusions

The multivariate approach and analysis may have a key role in cases when a high degree of uncertainty and coupling underlying dynamical mechanics is present. Hence, in this paper, we introduced the multivariate multifractal detrended moving average analysis (MV-MFDMA) as a multivariate generalization of the MFDMA method. The proposed method provides a way to analyze and investigate fractal dynamics information in multi-

variate time series data sets generated by complex systems. The validity of the proposed method was investigated by conducting numerical simulations on synthetic monofractal and multifractal time series. The correlation properties of multivariate monofractal time series are in line with the univariate case, and the fractal behavior of correlated bivariate time series is independent of its correlation level. In the case of the multivariate multifractal series, the multivariate system exhibits multifractal properties. The corresponding multifractal degree decreases with the decreasing of the multifractal degree of individual series. According to our simulation results, the MV-MFDMA represents a reliable technique for measuring the long-term correlations of non-stationary multivariate time series.

The MV-MFDMA was also utilized to investigate the multifractal properties of air pollutant time series where the individual air pollutants were considered as different variables from the same system. We analyzed raw and seasonally adjusted time series of air pollutants obtained upon removing the annual oscillations. We found that air pollutants exhibited multifractal auto-correlation behavior, even after removing the seasonal pattern. In both cases, the air pollutant time series data exhibited multifractal properties, with persistent long-range correlations. In order to capture the dynamic characteristics of the inner fluctuations of the air pollutants, we also applied the algorithm to the stochastic remainder components of the time series. The results confirmed that multivariate time series of air pollutants possessed multifractal properties and exhibited long-term persistence such that an increase (decrease) in the previous period will be followed by an increase (decrease) in the next period.

The findings of this paper provide an elevated understanding of the evolutionary activity and temporal links of five air pollutants in the state of California. In general, it is necessary to achieve a better comprehension of the associations between different air pollution time series to manage the environmental air quality based on evidence. Researchers from any scientific field that operate with multivariate frequent time series are encouraged to apply the proposed algorithm since it provides an improvement in detecting long-range correlations and multifractal properties of multichannel data. However, the proposed algorithm has certain limitations. It applies only to frequent and large-dimensional multivariate time series with an equal length N. Moreover, when applying the proposed algorithm, it is suggested to use an equal spacing between scales, where the scaling range varies from 10 to $N/10$, while the q-orders should include both positive and negative values. Finally, in some cases, seasonality can influence the selected time series, so if that is the case, the series should be firstly depersonalized before applying the algorithm. Future research papers can use the proposed MV-MFDMA algorithm, not only the environmental quality ones, but on the widest range of different real-world complex system time series, for example on financial, economic, meteorological, and health care time series. Therefore, directions for future research should go towards studying new time series when a better comprehension of their interconnections is necessary.

Author Contributions: Conceptualization, M.K. and P.M.; methodology M.K.; software, M.D.; validation, M.D.; writing, all; visualization, M.D.; supervision, J.M. All authors read and agreed to the published version of the manuscript.

Funding: This research has been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Collected data are available from the authors.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Appendix A

Table A1 presents descriptive statistics of seasonally adjusted series.

Table A1. Descriptive statistics: seasonally adjusted series.

		CO	NO_2	O_3	$PM_{2.5}$	SO_2
Mean		0.4381	21.4366	0.0279	8.8198	1.9964
Std. Dev.		0.2422	10.7698	0.0088	7.4037	2.6018
Skewness		1.2054	0.3761	0.1193	6.5486	6.0940
Kurtosis		6.0841	3.0578	3.7585	96.2279	102.8836
Jarque–Bera		2131.856 *	79.1695 *	87.9690 *	1,233,060 *	1,408,679 *
Observations		3339	3339	3339	3339	3339
ADF (AIC)	intercept	−4.3080 *	−4.6241 *	−6.5685 *	−12.3825 *	−7.2438 *
	trend and intercept	−4.4298 *	−4.7864 *	−6.5767 *	−12.3867 *	−8.5387 *
ADF (SIC)	intercept	−5.8546 *	−5.7474 *	−7.9224 *	−16.3571 *	−10.2605 *
	trend and intercept	−5.9854 *	−5.9493 *	−7.9653 *	16.3589 *	−11.4403 *
PP	intercept	−26.7990 *	−33.2027 *	−39.6205 *	−19.1607 *	−51.9978 *
	trend and intercept	−26.9846 *	−34.1516 *	−39.8951 *	−19.1657 *	−49.5006 *

Note: ADF—augmented Dickey–Fuller unit root test; PP–Philips–Perron unit root test; AIC—Akaike information criterion; SIC—Schwarz information criterion; *—stat. sig. < 0.01.

References

1. Xiong, H.; Shang P. Detrended fluctuation analysis of multivariate time series. *Commun. Nonlinear Sci.* **2017**, *42*, 12–21. [CrossRef]
2. Ihlen, E.A.F. Introduction to multifractal detrended fluctuation analysis in Matlab. *Front. Physiol.* **2012**, *3*, 141. [CrossRef]
3. Peng, C.K.; Buldyrev, S.V.; Havlin, S.; Simons, M.; Stanley, H.E.; Goldberger, A.L. Mosaic organization of DNA nucleotides. *Phys. Rev. E* **1994**, *49*, 1685–1689. [CrossRef] [PubMed]
4. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A* **2002**, *316*, 87–114. [CrossRef]
5. Vandewalle, N.; Ausloos, M. Crossing of two mobile averages: A method for measuring the roughness exponent. *Phys. Rev. E* **1998**, *58*, 6832–6834. [CrossRef]
6. Gu, G.-F.; Zhou, W.-X. Detrending moving average algorithm for multifractals. *Phys. Rev. E* **2010**, *82*, 011136. [CrossRef] [PubMed]
7. Flores-Marquez, E.L.; Ramírez-Rojas, A.; Telesca, L. Multifractal detrended fluctuation analysis of earthquake magnitude series of Mexican South Pacific Region. *Appl. Math. Comput.* **2015**, *265*, 1106–1114. [CrossRef]
8. Wang, Y.; Wu, C. Futures Markets: New Evidence from Multifractal Detrending Moving Average Analysis. *Comput. Econ.* **2013**, *42*, 393–414. [CrossRef]
9. Mali, P. Multifractal detrended moving average analysis of global temperature records. *J. Stat. Mech. Theory Exp.* **2016**, *2016*, 013201. [CrossRef]
10. Ahmed, M.U.; Mandic, D.P. Multivariate multi-scale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918. [CrossRef]
11. Cao, L.; Mees, A.; Judd, K. Dynamics from multivariate time series. *Physica D* **1998**, *121*, 75–88. [CrossRef]
12. Zhang, X.; Zeng, M.; Meng, Q. Multivariate multifractal detrended fluctuation analysis of 3D wind field signals. *Physica A* **2018**, *490*, 513–523. [CrossRef]
13. Manimaran, P.; Narayana, A.C. Multifractal detrended cross-correlation analysis on air pollutants of University of Hyderabad Campus, India. *Physica A* **2018**, *502*, 228–235. [CrossRef]
14. Fan, Q.; Liu, S.; Wang, K. Multiscale multifractal detrended fluctuation analysis of multivariate time series. *Physica A* **2019**, *532*, 121864. [CrossRef]
15. Wang, F.; Fan, Q. Coupling correlation detrended analysis for multiple nonstationary series. *Commun. Nonlinear Sci.* **2021**, *94*, 105579. [CrossRef]
16. Zhang, C.; Ni, Z.; Ni, L.; Li, J.; Zhou, L. Asymmetric multifractal detrending moving average analysis in time series of $PM_{2.5}$ concentration. *Physica A* **2016**, *457*, 322–330. [CrossRef]
17. Zhang, C.; Wang, X.; Chen, S.; Zou, L.; Zhang, X.; Tang, C. A study on daily $PM_{2.5}$ concentrations in Hong Kong using the EMD-based MFDFA method. *Physica A* **2019**, *530*, 121182. [CrossRef]
18. Zhang, C.; Wang, X.; Chen, S.; Zou, L.; Tang, C. Coupling detrended fluctuation analysis of the relationship between $PM_{2.5}$ concentration and weather elements. *Physica A* **2019**, *531*, 121757. [CrossRef]
19. Zou, S.; Zhang, T. Multifractal detrended cross-correlation analysis of the relation between price and volume in European carbon futures markets. *Physica A* **2020**, *537*, 122310. [CrossRef]
20. Da Silva, H.S.; Silva, J.R.S.; Stosic, T. Multifractal analysis of air temperature in Brazil. *Physica A* **2020**, *549*, 124333. [CrossRef]

21. Tzanis, C.G.; Koutsogiannis, I.; Philippopoulos, K.; Kalamaras, N. Multifractal detrended cross-correlation analysis of global methane and temperature. *Remote Sens.* **2020**, *12*, 557. [CrossRef]
22. Wang, Y; Wu, C.; Pan, Z. Multifractal detrending moving average analysis on the US Dollar exchange rates. *Physica A* **2011**, *390*, 3512–3523. [CrossRef]
23. Xi, C.P.; Zhang, S.-N.; Xiong, G; Zhao, H.-C. A comparative study of multifractal detrended fluctuation analysis and multifractal detrended moving average algorithm to estimate the multifractal spectrum. *Acta Phys. Sin.* **2015**, 64, 136403.
24. Lai, S.; Wan, L.; Zeng, Z.: Comparative Study of Sliding Window Multifractal Detrended Fluctuation Analysis and Multifractal Moving Average Algorithm. *J. Phys. Conf. Ser.* **2019**, 1345, 042086. [CrossRef]
25. United States Environmental Protection Agency: Outdoor Air Quality Data. Available online: https://www.epa.gov/outdoor-air-quality-data/download-daily-data (accessed on 15 August 2020).
26. Kalamaras, N.; Tzanis, C.G.; Deligiorgi, D.; Philippopoulos, K.; Koutsogiannis, I. Distribution of air temperature multifractal characteristics over Greece. *Atmosphere* **2019**, *10*, 45. [CrossRef]
27. Krzyszczak, J.; Baranowski, P.; Zubik, M.; Hoffmann, H. Temporal scale influence on multifractal properties of agro-meteorological time series. *Agric. Forest Meteorol.* **2017**, 239, 223–235. [CrossRef]
28. Nigmatullin, R.; Vorobev, A. The "Universal" Set of Quantitative Parameters for Reading of the Trendless Sequences. *Fluct. Noise Lett.* **2019**, *18*, 1950023. [CrossRef]
29. Nigmatullin, R.; Lino, P.; Maione, G. Reduction of Trendless Sequences of Data by Universal Parameters. In *New Digital Signal Processing Methods*; Springer: Cham, Switzerland, 2020; pp. 409–429.
30. Cleveland, R.B.; Cleveland,W.S.; McRae, J.E.; Terpenning, I. STL: A seasonal-trend decomposition procedure based on loess. *J. Off. Stat.* **1990**, *6*, 3–73.