


**Detrending-moving-average-based multivariate regression model for nonstationary series**Fang Wang <sup>\*</sup>*Key Laboratory of Intelligent Computing and Information Processing of Ministry of Education and Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan, 411105, China*Yuming Chen *Department of Mathematics, Wilfrid Laurier University, Waterloo, Ontario, Canada N2L 3C5*

(Received 23 February 2022; accepted 3 May 2022; published 18 May 2022)

Dependency between a response variable and the explanatory variables is a relationship of universal concern in various real-world problems. *Multivariate linear regression* (MLR) is a well-known method to focus on this issue. However, it is limited to dealing with stationary variables. In this work, we develop a MLR framework based on *detrending moving average* (DMA) analysis to reveal the actual dependency among variables with nonstationary measures. The DMA-based MLR can generate multiscale regression coefficients, which characterize different dependent behavior at different timescales. Artificial tests show that the DMA-MLR model can successfully resist the impact of trends on the studied series and produce more accurate regression coefficients with multiscale. Furthermore, some scale-dependent statistics are developed to deduce some important relationships in three typical DMA-based MLR models, which help us to deeply understand the DMA-MLR models in theory. The application of the proposed DMA-MLR framework to Beijing's air quality index system demonstrates that fine particulate matter with diameter  $\leq 2.5 \mu\text{m}$  ( $\text{PM}_{2.5}$ ) is the dominant pollutant affecting the air quality of Beijing in recent years.

DOI: [10.1103/PhysRevE.105.054129](https://doi.org/10.1103/PhysRevE.105.054129)**I. INTRODUCTION**

The rapid development of multivariate statistical technology has provided strong support for data analysis and decision making in various fields of information science. Modeling on multivariate nonstationary time series has been a research hotspot in recent decades [1–3]. An abundance of effective methods have been proposed to deal with nonstationary measures. Among them, *detrended fluctuation analysis* (DFA) [4] is undoubtedly quite popular since it is capable of exploring the long-range autocorrelations and multifractal features of time series in diverse fields [5–15]. As an important variant of DFA, *detrending moving average* (DMA) analysis [16] can also describe the long-term correlation in nonstationary systems, and it replaces DFA on many occasions. DMA and its extensions have also been widely used for various (multi)fractal analyses [17–24]. Synthetic tests have demonstrated that the two competitive methods have similar performances, while the DMA algorithm is computationally less demanding because it does not contain the process of polynomial fitting [25–28].

Essentially, these two methods dealing with nonstationary measures are similar. Both analyze the fluctuation characteristics of the residual obtained by splitting the profile and removing the local trends. That is to say, the scale-dependent residual is a special product of the detrended method family.

Naturally, the scale-dependent residual is translated into the language of variance and covariance by using its second-order moment and the second-order mixed moment, respectively, which makes it possible to realize an idea that connects the classical linear regression model and the DFA-DMA-based regression model [29,30]. Based on the *ordinary least squares* (OLS) framework and its expression with scale-dependent variances and covariances, Kristoufek [29,30] constructed two multiscale univariate linear regression frameworks using DFA-DMA and the bivariate generalization of *detrended cross-correlation analysis* (DXA) [31] as well as *detrended moving average cross-correlation analysis* (DMXA) [20]. They can allow us to interpret the dependence between two variables at different time periods. Hereafter, these methods have been improved and applied to describe the dependencies between the response variable and explanatory variables in various fields [32–40]. For example, Wang *et al.* [35] and Fan and Wang [38] extended the univariate DFA and DMA regression models to the bivariate case, which are applied to analyze the dependence of the *air quality index* (AQI) among three cities and three stock indices at different timescales, respectively.

With a different approach, namely, time series decomposition, Shen [32] proposed a DFA-based *multivariate linear regression* (MLR) model. Based on the proposed DFA-MLR, the author provided a *semipartial correlation coefficient* (SPCC) to assess the unique contribution of meteorological factors to the air pollution index. Inspired by the experience of Shen [32], in this work we develop an alternative but

<sup>\*</sup>Corresponding author: [popwang619@163.com](mailto:popwang619@163.com)

also complementary DMA-based MLR model. Based on this DMA-MLR, we further deeply explore the relationship between several scale-dependent parameters in MLR models. Specifically, we investigate three models, namely, a  $p$ -variable regression model, a  $p - 1$ -variable model established by the same explanatory variables except one, and a  $p - 1$ -variable model established by the former explanatory variables and taking one of them as the dependent variable. The relationships between the scale-dependent parameters (SPCC, partial correlation coefficient, coefficient of determination, tolerance of explanatory variables) are discussed for the proposed DMA-MLR. Through rigorous mathematical analyses and simulation experiments, we find that the relationships also hold for the multiscale case.

The paper is organized as follows: In Sec. II we first briefly review the DMA algorithm and then develop the DMA-based multivariate regression framework. In Sec. III based on the DMA-MLR model, a scale-dependent SPCC is proposed. In addition, we show that the three important relationships are connected by the DMA-based SPCC together with other scale-dependent statistics. Their derivation is given in the Appendix. In Sec. IV we launch a series of artificial experiments around the DMA-MLR. As a real-world application, the dependence between AQI and pollutant concentration series of Beijing is considered in Sec. V. The paper ends with a brief summary.

## II. METHODOLOGY

### A. Detrended moving average and cross-correlation analysis (DMA and DMXA) methodologies

Being important in the detrending method family for coping with nonstationary measures, the DMA and DMXA methodologies can well explore the long-term correlation of nonstationary series. They are well described in [16,18–20]. Here we briefly outline the variance and covariance procedure. For a given series  $\{x(t)\}$ ,  $t = 1, 2, \dots, N$ , we first calculate its profile series  $X(t) = \sum_{k=1}^t x(k)$  for  $t = 1, 2, \dots, N$ . Then the moving average function is defined by

$$\tilde{X}(t) = \frac{1}{s} \sum_{k=-\lfloor (s-1)\theta \rfloor}^{\lceil (s-1)(1-\theta) \rceil} X(t-k), \quad (1)$$

where  $s$  is the window size,  $\lfloor x \rfloor$  denotes the largest integer not greater than  $x$  while  $\lceil x \rceil$  represents the smallest integer not less than  $x$ , and  $\theta$  is the position parameter in the range of  $[0, 1]$  (specifically,  $\theta = 0, 0.5$ , and  $1$  correspond to forward, centered, and backward cases, respectively). For the sample points satisfying  $s - \lfloor (s-1)\theta \rfloor \leq i \leq N - \lfloor (s-1)\theta \rfloor$ , the residual series  $\epsilon(i)$  is obtained by removing the moving average function  $\tilde{X}(t)$  from the profile  $X(t)$ , i.e.,  $\epsilon(i) = X(i) - \tilde{X}(i)$ . Then we divide  $\epsilon$  into  $N_s = \lfloor N/s - 1 \rfloor$  disjoint segments with the equal length  $s$ . In the  $v$ th interval, the fluctuation function is defined by

$$f_x^2(s, v) = \frac{1}{s} \sum_{i=1}^s [X_v(i) - \tilde{X}_v(i)]^2. \quad (2)$$

The detrending variance can be calculated by averaging  $f_x^2(s, v)$  over all the segments

$$F_x^2(s) = \frac{1}{N_s} \sum_{v=1}^{N_s} f_x^2(s, v). \quad (3)$$

$F_x^2(s)$  is a critical product of the well-known DMA method, which describes the autocorrelation of the series  $\{x(t)\}$ .

In order to assess the long-term cross-correlation between the bivariate series  $\{x(t)\}$  and  $\{y(t)\}$ , one defines the bivariate fluctuation function in a similar way as in Eq. (2),

$$f_{xy}^2(s, v) = \frac{1}{s} \sum_{i=1}^s [X_v(i) - \tilde{X}_v(i)][Y_v(i) - \tilde{Y}_v(i)].$$

Then the main outcome of the DMXA method, the detrending covariance, is determined by

$$F_{xy}^2(s) = \frac{1}{N_s} \sum_{v=1}^{N_s} f_{xy}^2(s, v). \quad (4)$$

$F_x^2(s)$  and  $F_{xy}^2(s)$  are considered as scale-dependent variance and covariance, respectively.

Using the above scale-dependent variance and covariance, a DMA-based cross-correlation coefficient,

$$\rho_{DMXA}(s, x, y) = \frac{F_{xy}^2(s)}{\sqrt{F_x^2(s)F_y^2(s)}}, \quad (5)$$

was proposed to quantitatively characterize the degree of cross-correlation between the two bivariate series  $\{x(t)\}$  and  $\{y(t)\}$  [28].

### B. Description of the DMA-MLR framework

In this subsection, we extend the DMA-based linear regression model [30,38] to the multivariate case. Suppose that there are  $p$  explanatory variables,  $x_1, x_2, \dots, x_p$ , and one dependent variable  $y$ . The same length  $N$  samples are considered, i.e.,  $\{x_j(k)\}, \{y(k)\}$  ( $1 \leq k \leq N, 1 \leq j \leq p$ ). The time series may involve unknown trend or nonstationary measures. For the sake of simplicity, consider a centralized [i.e.,  $\sum_{k=1}^N x_j(k) = 0, \sum_{k=1}^N y(k) = 0$ ] multivariate linear regression (MLR) model, whose  $k$ th sample is

$$\text{Model I: } y(k) = \beta_1 x_1(k) + \beta_2 x_2(k) + \dots + \beta_p x_p(k) + \epsilon(k). \quad (6)$$

where the parameter  $\beta_j$  ( $1 \leq j \leq p$ ) is the regression coefficient and  $\epsilon = [\epsilon(1), \epsilon(2), \dots, \epsilon(N)]^T$  is the residual series, which may also contain potential trends.

According to Shen [32], the cumulative sum for Eq. (6) from 1 to  $t$  ( $1 \leq t \leq N$ ) is

$$\begin{aligned} \sum_{k=1}^t y(k) &= \beta_1 \sum_{k=1}^t x_1(k) + \beta_2 \sum_{k=1}^t x_2(k) \\ &+ \dots + \beta_p \sum_{k=1}^t x_p(k) + \sum_{k=1}^t \epsilon(k). \end{aligned} \quad (7)$$

Denote  $Y = [\sum_{k=1}^1 y(k), \sum_{k=1}^2 y(k), \dots, \sum_{k=1}^N y(k)]^T$ ,  $\beta = [\beta_1, \beta_2, \dots, \beta_p]^T$ , and  $E = [\sum_{k=1}^1 \epsilon(k), \sum_{k=1}^2 \epsilon(k), \dots, \sum_{k=1}^N \epsilon(k)]^T$ , which are column vectors of  $N$  dimension. Let

$$X = \begin{bmatrix} \sum_{k=1}^1 x_1(k) & \sum_{k=1}^1 x_2(k) & \cdots & \sum_{k=1}^1 x_p(k) \\ \sum_{k=1}^2 x_1(k) & \sum_{k=1}^2 x_2(k) & \cdots & \sum_{k=1}^2 x_p(k) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^N x_1(k) & \sum_{k=1}^N x_2(k) & \cdots & \sum_{k=1}^N x_p(k) \end{bmatrix}_{N \times p}.$$

Then the matrix form of Eq. (7) is

$$Y = X\beta + E.$$

Accordingly, the residuals series can be determined by  $E = Y - X\beta$ . Similar to the idea of OLS, the scale-based regression coefficients can be determined by minimizing the fluctuation of residuals for each scale, namely,

$$\begin{aligned} \min \text{Var}(E) &= F_\epsilon^2(s) = \frac{1}{N_s} \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s [E_v(t) - \tilde{E}_v(t)]^2 \\ &= \frac{1}{N_s} \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s \{[Y_v(t) - X_v(t)\beta] - [Y_v(t) - \tilde{X}_v(t)\beta]\}^2 \\ &= \frac{1}{N_s} \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s \{[Y_v(t) - \tilde{Y}_v(t)] - [X_v(t) - \tilde{X}_v(t)\beta]\}^2 \\ &= \frac{1}{N_s} \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s [Y'_v(t) - X'_v(t)\beta]^2, \end{aligned} \quad (8)$$

where  $E_v(t)$  and  $Y_v(t)$  denote the  $t$ th elements of the vectors  $E$  and  $Y$ , respectively, in segment  $v$ ;  $X_v(t)$  is the  $t$ th row of the matrix  $X$  in segment  $v$ ; and  $\tilde{E}_v(t)$ ,  $\tilde{Y}_v(t)$ , and  $\tilde{X}_v(t)$  are the detrending moving average functions of  $E(t)$ ,  $Y(t)$ , and  $X(t)$ , respectively, in the  $v$ th segment, which are calculated by Eq. (1). Suppose that  $Y' = Y - \tilde{Y}$  and  $X' = X - \tilde{X}$ . In light of OLS and Zhao and Shang [41],  $Y_v(t) - \tilde{X}_v(t)\beta = \tilde{Y}_v(t) - \tilde{X}_v(t)\beta$ .

Since the series  $Y'$  and the matrix  $X'$  have removed the trend components, their variances are not relevant to time  $t$ , but are relevant to the timescale  $s$ . Therefore, for every scale, the first-order condition can be conducted on Eq. (8) to obtain the minimum  $(\beta_1, \beta_2, \dots, \beta_p)$ . The normal equations can be obtained,

$$\frac{1}{N_s} \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s [Y'_v(t)X'_v(t, i) - X'_v(t, i)X'_v(t)\beta] = 0,$$

where  $X'(t, i)$  is the element of the  $t$ th row and  $i$ th column of the matrix  $X'$  ( $1 \leq i \leq p$ ). Note that  $(1/N_s) \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s Y'_v(t)X'_v(t, i)$  is the detrended covariance of the series  $\{y(t)\}$  and  $\{x_i(t)\}$ , and  $(1/N_s) \sum_{v=1}^{N_s} \frac{1}{s} \sum_{t=1}^s X'_v(t, j)X'_v(t, i)$  is the detrended covariance of  $\{x_j(t)\}$  and  $\{x_i(t)\}$  ( $1 \leq j, i \leq p$ ), which conform with Eqs. (3) and (4). We record column vectors  $\beta(s) = [\beta_1(s), \beta_2(s), \dots, \beta_p(s)]^T$  and

$T(s) = [F_{yx_1}^2(s), F_{yx_2}^2(s), \dots, F_{yx_p}^2(s)]^T$ . The detrended covariance matrix is

$$F(s) = \begin{bmatrix} F_{x_1}^2(s) & F_{x_1x_2}^2(s) & \cdots & F_{x_1x_p}^2(s) \\ F_{x_2x_1}^2(s) & F_{x_2}^2(s) & \cdots & F_{x_2x_p}^2(s) \\ \vdots & \vdots & \ddots & \vdots \\ F_{x_px_1}^2(s) & F_{x_px_2}^2(s) & \cdots & F_{x_p}^2(s) \end{bmatrix}_{p \times p}.$$

Thereupon, the normal equations of the DMA-based regression model can be rewritten in matrix form,

$$F(s)\beta(s) = T(s). \quad (9)$$

In the case where  $F(s)$  is invertible, the DMA-based regression coefficient  $\beta(s)$  can be estimated by

$$\hat{\beta}(s) = F^{-1}(s)T(s).$$

Using the DMA-based estimator  $\hat{\beta}(s)$ , the estimated scale-dependent residuals are obtained by

$$\hat{E}(s) = Y - X\hat{\beta}(s) - \langle Y - X\hat{\beta}(s) \rangle,$$

where the symbol  $\langle \cdot \rangle$  represents the average value. By using the estimated scale-dependent residuals, the variance of  $\hat{\beta}(s)$  at different scales can be calculated by the diagonal component of the matrix  $\frac{1}{N-p-1} F_\epsilon^2(s)F^{-1}(s)$ .

The corresponding scale-specific coefficient of determination  $R^2(s)$  is presented for assessing the explanatory ability of the independent variables to Model I at different timescales,

$$R^2(s) = 1 - \frac{F_\epsilon^2(s)}{F_y^2(s)}, \quad (10)$$

where  $F_y^2(s)$  and  $F_\epsilon^2(s)$  are the detrended variances of the dependent variable  $y$  and the residual  $\epsilon$ , respectively, computed by Eq. (3).

### III. IN-DEPTH ANALYSIS OF DMA-MLR

In this section, we further explore the statistical property of the DMA-MLR model. Besides Model I [see Eq. (6)], two other models of DMA-based MLR are also discussed [see Eqs. (12) and (15)]. In Sec. III A some critical scale-dependent statistical measures are developed, which help us to understand the relationship among these models. In Sec. III B we outline in detail the three important relationships connecting the three DMA-MLR models.

#### A. Scale-dependent statistics in the DMA-MLR model

In the MLR model, the dependent variable is always affected by different explanatory variables to varying degrees. To assess the relative importance of explanatory variables in determining the dependent variable  $y$ , a SPCC [42] is defined as the Pearson's correlation coefficient between the residual of one of the explanatory variables (e.g.,  $x_i$ ) on the other explanatory variables and the dependent variable  $y$ . The SPCC of  $y$  and  $x_i$  is defined as

$$\text{SPCC}(y, x_i) = \frac{\delta_i^T y}{\sqrt{\delta_i^T \delta_i} \sqrt{y^T y}}, \quad (11)$$

where  $y = [y(1), y(2), \dots, y(N)]^T$  is an observed value of the dependent variable  $y$ , and  $\delta_i = [\delta_i(1), \delta_i(2), \dots, \delta_i(N)]^T$  is an implementation value of the residual  $\delta$  of the variable  $x_i$  on the other explanatory variables, which is determined by the MLR model (denoted as Model II) below (also centralized for simplicity),

$$\text{Model II: } x_i(k) = \alpha_1 x_1(k) + \dots + \alpha_{i-1} x_{i-1}(k) + \alpha_{i+1} x_{i+1}(k) + \dots + \alpha_p x_p(k) + \delta_i(k), \quad (12)$$

where  $k = 1, 2, \dots, N$ .

The classical SPCC can deal only with stationary measures. If the dependent variable or explanatory variables possess trends or long-term correlation, the SPCC will give the false correlation since the assumption of steady covariance and variance is violated. What's more, the Pearson's correlation-based calculation makes SPCC not able to capture the correlation at different timescales. Here, to overcome these drawbacks, according to the DMA-DMXA language, we develop a DMA-based scale-dependent SPCC, which is defined as

$$\rho_{\text{DMSPCC}}(s, y, x_i) = \frac{F_{\delta_i y}^2(s)}{\sqrt{F_{\delta_i}^2(s) F_y^2(s)}}, \quad (13)$$

where  $F_{\delta_i}^2(s)$  and  $F_y^2(s)$  are the scale-dependent detrended variances computed with Eq. (3), and  $F_{\delta_i y}^2(s)$  is the scale-dependent detrended covariance computed by Eq. (4). The proposed  $\rho_{\text{DMSPCC}}$  can well assess the correlation at different scales between the dependent variable and one explanatory variable regardless of trends.

The semipartial correlation analysis [42] successfully solved the collinearity problem between explanatory variables in MLR. On the one hand, the proposed  $\rho_{\text{DMSPCC}}(s, y, x_i)$  reflects the unique contribution of the explanatory variable  $x_i$  to the multiple correlation coefficients of the MLR model at scale  $s$ , which is realized by eliminating the influence of the other explanatory variables. On the other hand, if the influence of the other explanatory variables on the dependent variable is also eliminated, then the "pure" relation between the studied explanatory variables and the dependent variable can be obtained, which is called the *partial correlation coefficient* (PCC). Recently a DMA-based PCC was proposed to assess the intrinsic correlation between multiple variables [43], denoted as  $\rho_{\text{DMPCC}}(s)$ .  $\rho_{\text{DMPCC}}$  is defined as the DMA-based cross-correlation coefficient of two residuals that are obtained by regression models established separately with two dependent variables affected by the same explanatory variables. Precisely,  $\rho_{\text{DMPCC}}(s)$  between variables  $y$  and  $x_i$  is defined by

$$\rho_{\text{DMPCC}}(s, y, x_i) = \frac{F_{\delta_i \eta_i}^2(s)}{\sqrt{F_{\delta_i}^2(s) F_{\eta_i}^2(s)}}, \quad (14)$$

where  $\delta_i$  and  $\eta_i = [\eta_i(1), \eta_i(2), \dots, \eta_i(N)]^T$  are, respectively, the scale-specific residuals of Model II [Eq. (12)] and Model III (also centralized) below,

$$\text{Model III: } y(k) = \gamma_1 x_1(k) + \dots + \gamma_{i-1} x_{i-1}(k) + \gamma_{i+1} x_{i+1}(k) + \dots + \gamma_p x_p(k) + \eta_i(k), \quad (15)$$

where  $k = 1, 2, \dots, N$ . Here, as before,  $F_{\delta_i}^2(s)$  and  $F_{\eta_i}^2(s)$  are computed with Eq. (3) while  $F_{\delta_i \eta_i}^2(s)$  is computed by Eq. (4).

## B. Important relationships in DMA-MLR models

In the MLR model, both SPCC and PCC are used to eliminate multicollinearity and assess the true correlation of explanatory variables and the dependent variable. In contrast, since SPCC reflects the unique contribution of explanatory variables to the multiple correlation coefficient, it is often used to screen the explanatory variables that have no significant impact on the dependent variable. In typical MLR models, SPCC is a bridge connecting the coefficient of determination of Model I [Eq. (6)] and that of Model III [Eq. (15)]. We find that the relationship also holds for the DMA-based regression framework, which is

$$R^2(s) = R_{yi}^2(s) + \rho_{\text{DMSPCC}}^2(s, y, x_i). \quad (16)$$

Here the coefficient of determination  $R^2(s)$  [Eq. (10)] and  $R_{yi}^2(s) [= 1 - F_{\eta_i}^2(s)/F_y^2(s)]$  measures the explanatory ability of all independent variables and the variables excluding  $x_i$  to Model I and Model III at scale  $s$ , respectively. Accordingly, for the calculated  $R^2(s)$ , if  $|\rho_{\text{DMSPCC}}(s, y, x_i)|$  is large, then  $R_{yi}^2(s)$  is small, which indicates that the influence of removing  $x_i$  on  $R^2(s)$  is great. This is explained as that the explanatory variable  $x_i$  is important for the dependent variable  $y$ . On the contrary, the larger  $R_{yi}^2(s)$  is, the less significant the contribution of  $x_i$  to  $y$  is. Moreover, the DMA-based PCC of  $y$  and  $x_i$  can be connected with the DMA-based SPCC at different scales by  $R_{yi}^2(s)$  through

$$\rho_{\text{DMPCC}}(s, y, x_i) = \frac{\rho_{\text{DMSPCC}}(s, y, x_i)}{\sqrt{1 - R_{yi}^2(s)}}. \quad (17)$$

According to Eq. (17), it is clear that  $\rho_{\text{DMPCC}}$  is not less than  $\rho_{\text{DMSPCC}}$  at any timescale. For the case where  $R_{yi}^2(s) = 0$ , that is, there is no relationship between the other explanatory variables and the dependent variable  $y$  at the given scale  $s$ ,  $\rho_{\text{DMPCC}}(s, y, x_i)$  is equal to  $\rho_{\text{DMSPCC}}(s, y, x_i)$  as expected.

In addition to connecting the DMA-MLR Models I and III, the DMA-based SPCC is also a bridge connecting Model I and Model II [Eq. (12)]. Noting that  $R_{xi}^2(s) [= 1 - F_{\delta_i}^2(s)/F_{x_i}^2(s)]$  is the scale-dependent coefficient of determination of Model II, we see that the SPCC between the variables  $x_i$  and  $y$  can also be calculated by

$$\rho_{\text{DMSPCC}}(s, y, x_i) = \beta_i^*(s) \sqrt{1 - R_{xi}^2(s)}, \quad (18)$$

where  $\beta_i^*(s) = \beta_i(s) \sqrt{F_{x_i}^2(s)/F_y^2(s)}$  is the standard DMA-based regression coefficient of Model I and  $\text{TOL}_i(s) = 1 - R_{xi}^2(s)$  is the tolerance of the explanatory variable  $x_i$  at the scale  $s$ . Obviously the smaller  $\text{TOL}_i(s)$  is, the more serious the collinearity of variable  $x_i$  by the other explanatory variables will be expected.

To sum up, Eq. (16) provides a quantitative description of the explanatory ability to the model by one of the explanatory variables; Eq. (17) shows the relationship between the DMA-based PCC and SPCC; Eq. (18) allows us to investigate the relationship between the partial regression coefficient and SPCC. These three important relationships help us to have

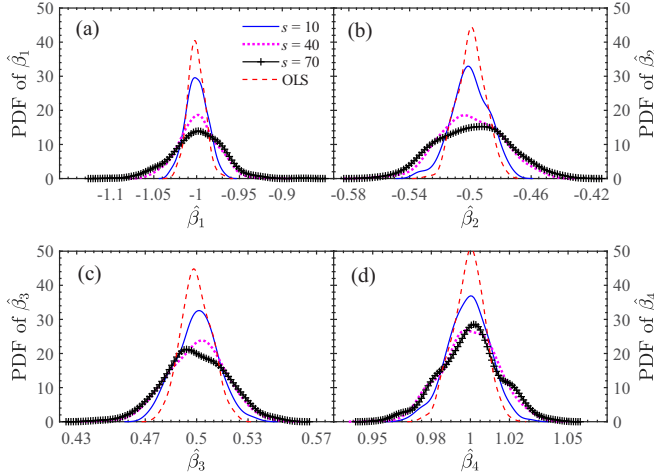


FIG. 1. PDFs of regression coefficients of the quaternary linear regression model ( $y = -x_1 - 0.5x_2 + 0.5x_3 + x_4 + \epsilon$ ) based on the DMA method and the traditional method. The four subplots are for the four regression coefficients in order. In each case, three scales  $s = 10, 40$ , and  $70$  are considered for the DMA-based regression model.

a deeper understanding of the DMA-based MLR models in theory. Meanwhile, they provide an easy way to calculate scale-dependent statistical measures in practical applications of DMA-MLR models. The proof is given in the Appendix.

#### IV. TEST AND DISCUSSION

##### A. Performance of DMA-based MLR

The DMA-DMXA is a method family having the ability to deal with nonstationary measures as the DFA-DCCA method family does, but the former is more efficient than the latter. However, research shows that there is a slight deviation in the results for different position types of moving average (as mentioned above, forward, centered, and backward correspond to  $\theta = 0, \theta = 0.5$ , and  $\theta = 1$ , respectively). The empirical result indicates that the best performance comes from the case of a centered moving average [18,19]. For the DMA-based binary regression model, Fan and Wang [38] compared the error results of different moving average schemes on two simulations and demonstrated that the centered case has the best result (see Fig. 1 there). Therefore, to avoid repetitive work, we utilize  $\theta = 0.5$  to do our test. In this subsection, we first test the validity of the proposed DMA-based multivariate regression framework and then examine the three relationships shown in Sec. III B.

We consider a quaternary linear regression model as  $y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \epsilon$ . The four explanatory variables  $x_i (i = 1, 2, 3, 4)$  are generated by four independent ARFIMA(0,  $d_i$ , 0) processes [44],  $x_i(t) = \sum_{n=1}^{\infty} \alpha_n(d_i)\xi_i(t-n)$ , where  $d_i$  is a fractional integration parameter with the range  $(-0.5, 0.5)$ ,  $\alpha_n(d_i) = \Gamma(n-d_i)/[\Gamma(-d_i)\Gamma(n+1)]$ ,  $\Gamma(\cdot)$  denotes the Gamma function, and  $\xi_i$  is an independent Gaussian noise. The residual  $\epsilon$  is the Gaussian white noise series. First, we investigate the distribution of the DMA-based regression coefficients. Suppose  $\beta_1 = -1$ ,  $\beta_2 = -0.5$ ,  $\beta_3 = 0.5$ , and  $\beta_4 = 1$ . The four ARFIMA series have the same fractional

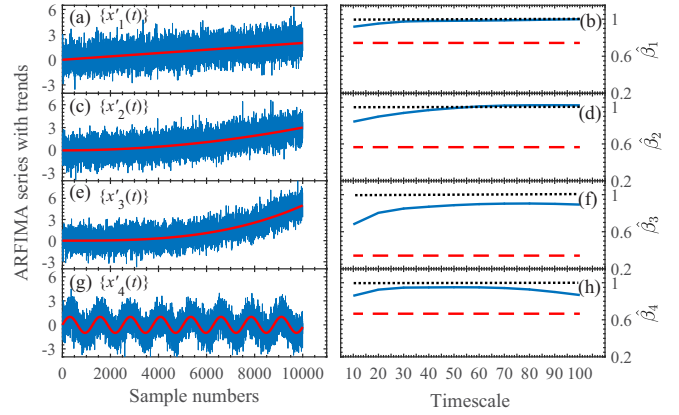


FIG. 2. Effect of the four types of trends on regression estimators. The left panel is for  $\{x_i(t)\}$  ( $i = 1, 2, 3, 4$ ) adding linear trend (a), quadratic trend (c), cubic trend (e), and sinusoidal trend (g) with the other variables unchanged. The right panel is for the corresponding  $\hat{\beta}_i$  estimated by DMA-based (solid line) and traditional OLS-based (dashed line) methods. Dotted line locates the expected result,  $\beta_i = 1$ . It demonstrates that the DMA-based estimators are close to the true values, but the estimations of the traditional OLS-based method deviate seriously.

integration parameter of  $d_i = 0.1$  but independently. The four explanatory variables' series  $\{x_i(t)\}$  have  $N = 10000$  sample points. We conduct 1000 independent simulations to estimate the regression coefficients. The probability density function (PDF) of the four regression coefficient estimators  $\hat{\beta}_i(s) - \beta_i(s)$  is shown in Figs. 1(a)–(d), respectively. Three timescales are selected for our consideration. As a comparison, the PDF obtained from the traditional MLR estimation (denoted as OLS) is also shown in the figure. There is no doubt that the PDFs are normal shape and are centered at the four designed coefficients  $-1, -0.5, 0.5$ , and  $1$ , for both methods. We find that the standard deviation of the DMA-based  $\hat{\beta}(s)$  increases with the increasing of the scale  $s$ . That the mean values of the simulations are centered at the true regression coefficients illustrates that the DMA-based regression estimators reflect the correct dependencies. Next, we wish to demonstrate the satisfactory trend resistance of the DMA-MLR. To do so, the above quaternary regression model is still employed. For comparison, this time, we set an identical coefficient  $\beta_i = 1$  ( $i = 1, 2, 3, 4$ ). The other settings are the same as those for the first test. Then we add four types of trends [denoted as  $\text{Tr}(t)$ ] to one of the explanatory variables [ $x'_i(t) = x_i(t) + \text{Tr}_i(t)$ ] and keep the others unchanged, namely,  $\text{Tr}_1(t) = 0.02t$ ,  $\text{Tr}_2(t) = 0.0003t^2$ ,  $\text{Tr}_3(t) = 0.000005t^3$ , and  $\text{Tr}_4(t) = \sin(0.5t)$ , where  $t = i/100$ ,  $i = 0, 1, 2, \dots, N-1$ . The four ARFIMA series embodied trends are shown in the left panels of Fig. 2.

As a comparison, we also show the traditional OLS-based regression estimators. Unfortunately, they deviate from the true values to varying degrees (see the dotted lines). This can be interpreted as that the traditional OLS-based regression estimators are greatly disturbed by the trends. However, owing to the fact that the DMA/DMXA has a strong ability to deal with nonstationary measures, the DMA-based  $\hat{\beta}(s)$  (the solid lines) is stably situated near 1 for available scales  $s$  regardless of the types of trends.

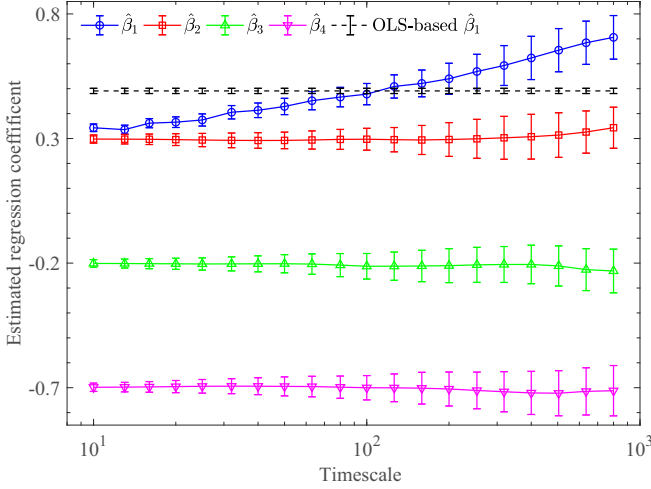


FIG. 3. Estimated DMA-based and OLS-based regression coefficients for the quaternary regression model  $y = 0.8x_1 + 0.3x_2 - 0.2x_3 - 0.7x_4 + \epsilon$ . Error bars indicate the standard deviations calculated from 100 independent realizations of the corresponding processes.

We now show the merit that DMA estimators can capture the scaling behavior of dependence at different timescales. The quaternary regression model  $y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \epsilon$  is also considered here. We set  $\beta_1 = 0.8, \beta_2 = 0.3, \beta_3 = -0.2$ , and  $\beta_4 = -0.7$  this time. The explanatory variable  $x_1$  is generated by a binomial multifractal series (BMFs) as  $x_1 = p^{n-n[k-1]}(1-p)^{n[k-1]}$ ,  $k = 1, 2, \dots, 2^n$ , where the parameter  $p \in (0, 0.5)$ ,  $n[k]$  denotes the number of digit 1 in the binary representation of the index  $k$ . Here we took  $p = 0.2$ . The  $x_2, x_3, x_4$  are independent Gaussian variables with 0 mean and 0.001 standard deviation. The error term  $\epsilon$  is also Gaussian noise with the same strength as  $x_2 \sim x_4$ . All the series are of length  $N = 2^{13}$ . Then we remove the values less than 0.000001 of the BMFs  $x_1$  so that only a few of the biggest values are left, and substitute Gaussian random numbers with 0 mean and 0.001 standard deviation in their places. Hence, we obtain a binomial cascade series embedded in random noise. Analysis of the dependence between the explained variable  $y$  and the four explanatory variables indicates that the estimated  $\hat{\beta}_2 \sim \hat{\beta}_4$  are unbiased at 0.3,  $-0.2$ , and  $-0.7$ , respectively, with a desirable errorbar for every timescale without suspense, as shown in Fig. 3. However, the performance of  $\hat{\beta}_1$  has changed significantly. The dependence between  $y$  and  $x_1$  is obviously less than the given value ( $\beta_1 = 0.8$ ) at the smaller scales contrary to the larger one. This is because in smaller scales, the dependency has been destroyed by the random noise while it remains unchanged in larger scales. But the traditional OLS-based estimators cannot capture this (see the black dashed line in Fig. 3).

**B. Test of DMA-based SPCC**

In this subsection, we focus on testing the performance of the DMA-based SPCC. The remarkable benefit of the proposed  $\rho_{\text{DMSPCC}}$  is that it can uncover the correct semipartial correlation regardless of trends. To verify this, the ARFIMA series are employed again. This time, three

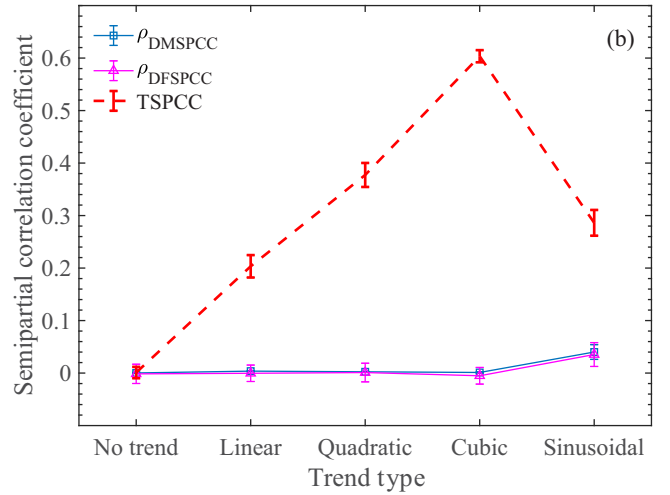
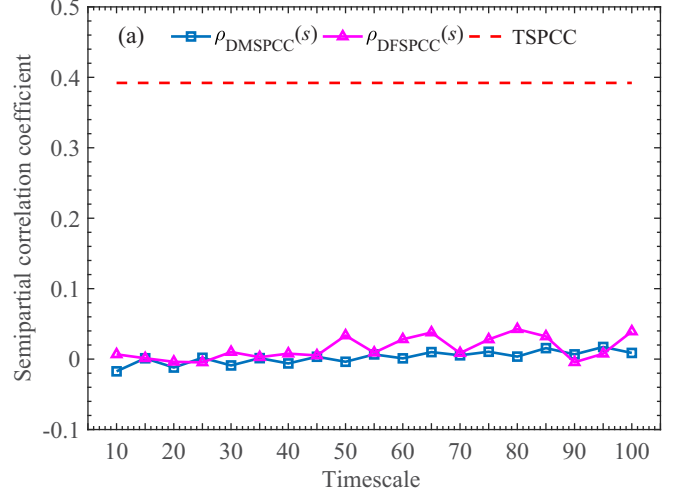


FIG. 4. The performance of semipartial cross-correlation coefficient for ARFIMA series involving trends. (a) The multiscale SPCC obtained by DMA (blue squares) and DFA (violet triangles). The red dashed line denotes the traditional SPCC. (b) More cases with higher-order trends are shown. We take second-order polynomial fitting for the DFA method. The coefficient of each case is actually averaged over all the timescales in (a).

independent ARFIMA processes with identical  $d = 0.1$  are set as the explanatory variables  $x_1, x_2$ , and  $x_3$ .  $x_4 = x_1 + x_2 + x_3 + \delta$ , where  $\delta$  is Gaussian white noise. In this way,  $x_4$  will be explained well by the variables  $x_1, x_2$ , and  $x_3$ . Meanwhile, assume that  $y$  is another independent ARFIMA process with  $d = 0.4$ . Therefore,  $y$  and the residual of  $x_4$  on  $x_1, x_2$ , and  $x_3$  are unrelated in theory. Indeed, the simulation value of the SPCC calculated by Eq. (11) is 0.0028. However, if  $y$  and  $x_4$  are disturbed by trends, then the traditional SPCC won't work. To show this, we add quadratic trend (as in Sec. IV A) to  $x_4$  and  $y$ . As expected,  $\text{SPCC}(y, x_4) = 0.3921$ . The reason is obvious. This is because the series  $\{y(t)\}$  and  $\{x_4(t)\}$  are nonstationary and the traditional SPCC fails to deal with nonstationary measures. Fortunately, the DMA-based SPCC does this job well. Figure 4(a) illustrates that  $\rho_{\text{DMSPCC}}(y, x_4)$  always fluctuates slightly around 0 at every timescale. In this regard,  $\rho_{\text{DMSPCC}}$  can uncover the essential correlation between

TABLE I. Average coefficient of determination, SPCC, and PCC of the DMA-MLR framework over the given scales.

	$R_{yi}^2$	$\rho_{\text{DMSPCC}}$	$\rho_{\text{DMPCC}}$
$x_1$	0.9939	-0.0596	-0.8542
$x_2$	0.9907	0.0805	0.9104
$x_3$	0.9817	-0.1209	-0.9565
$x_4$	0.9283	0.2521	0.9893

two variables impacted by quadratic trends. As a matter of fact, it is immune to any trends. Figure 4(b) supports this, where the other three trends, namely, linear, cubic, and sinusoidal trends (set as in Sec. IV A), are used. As comparison, the traditional SPCC (the red dashed line) and DFA-based SPCC (denoted as  $\rho_{\text{DFSPCC}}$ , proposed by Shen [32]) are also shown in Fig. 4(b), where  $\rho_{\text{DMSPCC}}$  and  $\rho_{\text{DFSPCC}}$  are averaged over all the considered timescales. Compared with the unsatisfactory traditional SPCC, the fact that they are almost 0 demonstrates that both DMA-based and DFA-based SPCCs can well express the accurate result, and the performance of  $\rho_{\text{DMSPCC}}$  is slightly superior to that of  $\rho_{\text{DFSPCC}}$  due to the smaller fluctuations. It is worth noting that, in this test, we added the same trends to both series  $\{y(t)\}$  and  $\{x_i(t)\}$ . In fact, even if they are affected by different trends,  $\rho_{\text{DMSPCC}}$  is still satisfactory, which benefits from the merit of the DMA-DMXA method. Although it seems that both methods are slightly worse in removing the interference of sinusoidal trend [Fig. 4(b)], the immunity to trends is in general satisfactory.

Let us further study the DMA-based SPCC and its use in the DMA-MLR framework. SPCC is always used to evaluate the contribution of explanatory variables by eliminating collinearity in MLR. In this spirit, here a quaternary regression model with multicollinearity of explanatory variables is established. The four explanatory variables  $x_i$  ( $i = 1, 2, 3, 4$ ) and the dependent variable  $y$  are generated by binomial multifractal series (BMFs) with the parameter  $p = 0.1, 0.2, 0.3, 0.4$  and  $0.48$ , respectively. All the series have length  $N = 2^{13}$ . Figure 5(a) shows the DMXA coefficients [calculated by Eq. (5)] between every two explanatory variables. The big  $\rho_{\text{DMXA}}(s, x_i, x_j)$  value implies that there is a strong multicollinearity. The smaller  $\Delta p = |p_{x_i} - p_{x_j}|$  is, the stronger the collinearity is. By eliminating the collinearity effect,  $\rho_{\text{DMSPCC}}$ 's between  $y$  and  $x_i$  ( $i = 1, 2, 3, 4$ ) are shown in Fig. 5(b). As expected, the differences of the correlations between  $y$  and the four explanatory variables are significant. The explanatory  $x_2$  and  $x_4$  are positively correlated with  $y$  while  $x_1$  and  $x_3$  are negatively correlated with  $y$ . In contrast, the intrinsic correlation between  $y$  and  $x_4$  is the most significant due to the smallest  $\Delta p$ . This can also be confirmed by  $\rho_{\text{DMPCC}}(s, y, x_i)$  and  $R_{yi}^2(s)$  obtained from a ternary regression model (see Model III, by excluding  $x_i$ ). Table I lists the average  $R_{yi}^2$ ,  $\rho_{\text{DMPCC}}$ , and  $\rho_{\text{DMSPCC}}$  over scales  $s$  from 10 to 800. As expected,  $R_{y4}^2$  is the smallest, which implies that the model explanatory ability is greatly reduced by eliminating  $x_4$  (relative to the other variables). This is consistent with the fact that the values of  $\rho_{\text{DMSPCC}}(y, x_4)$  and  $\rho_{\text{DMPCC}}(y, x_4)$  are the largest.

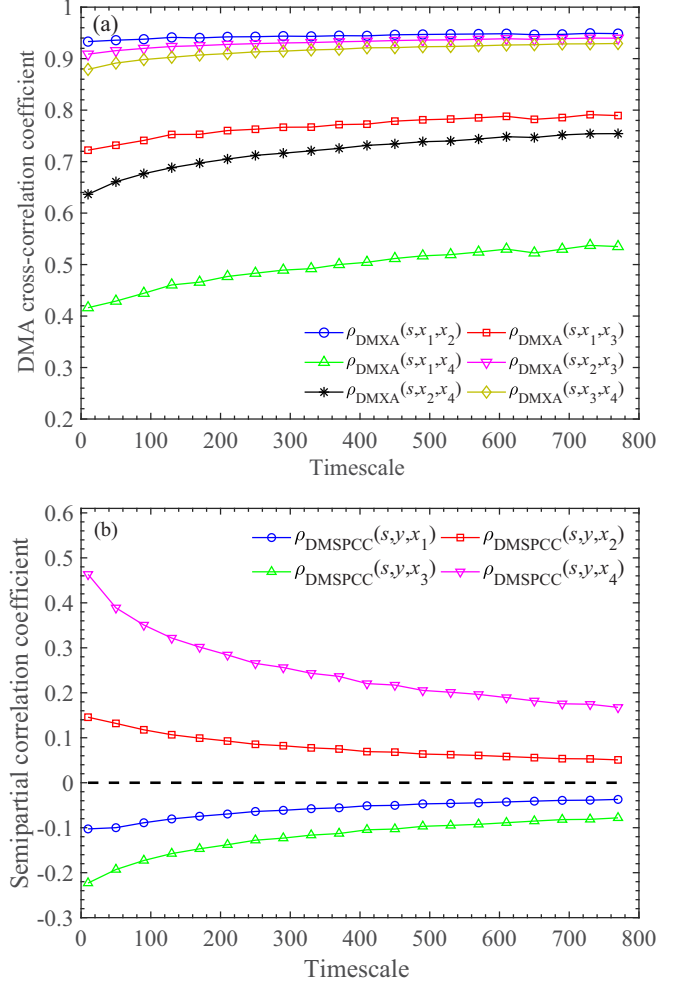


FIG. 5. DMA-MLR modeling for four explanatory variables  $\{x_i(t)\}$  ( $i = 1, 2, 3, 4$ ) and an explained variable  $\{y(t)\}$  generated by BMFs model. (a) DMXA coefficient of every two  $\{x_i(t)\}$ . (b) The DMA-based SPCC between series  $\{y(t)\}$  and each  $\{x_i(t)\}$ . The black dashed line represents zero.

In the above tests, although the Gaussian noise is used as residual in Models I–III, we should note that our model is also suitable for other distributed residuals.

### C. Examination of the important relationships

Now we would like to verify the three important relationships in DMA-based MLR models. To this end, the above BMFs are employed. Besides the above quaternary regression model (denoted as Model A, a specific realization of Model I), two other kinds of MLRs are established. One is the ternary regression model family by using one of  $\{x_i(t)\}$  as the dependent variable and the other three as explanatory variables (denoted as Model B, a specific realization of Model II). The other is also the ternary regression model family but by using  $\{y(t)\}$  as the dependent variable and three of  $\{x_i(t)\}$  as explanatory variables (denoted as Model C, a specific realization of Model III). Figures 6–8 show the simulation results. In Fig. 6  $R^2(s)$  (the black dashed line) is the coefficient of determination of Model A and  $R_{yi}^2(s)$  is obtained from

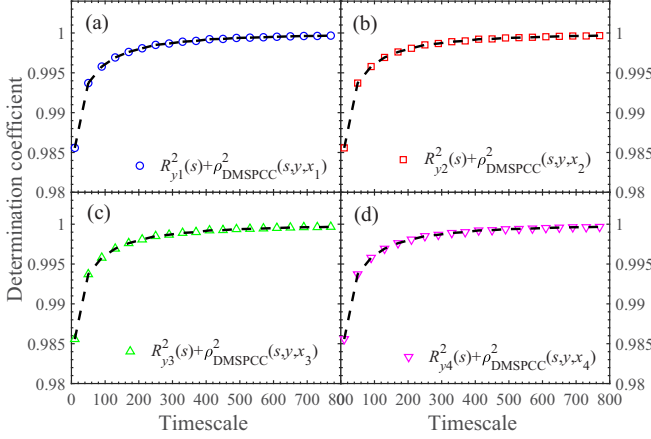


FIG. 6. Examination of the relationship between  $R^2(s)$  and  $R_{yi}^2(s)$ , which is the coefficient of determination of the quaternary and ternary linear regression models based on the DMA framework. Subplots (a), (b), (c), and (d) are for Model C established by three  $x_i$  with the exclusion of  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$ , respectively. The dashed line in the four subplots is  $R^2(s)$  of Model A (all of them are identical). The hollow symbols denote the quantities determined by the right side of Eq. (16). The perfect agreement suggests that the relationship (16) holds.

Model C. In each subplot, the colored hollow symbol denotes the sum of  $R_{yi}^2(s)$  and the corresponding  $\rho_{DMSPCC}^2(s, y, x_i)$ . Figure 7 shows the DMA-based PCC between  $y$  and  $x_i$ , which is calculated by Eq. (14) (hollow symbols) and Eq. (17) (dashed lines). The examination of the relationship between the DMA-based SPCC and DMA-based regression coefficient of Model A is shown in Fig. 8. The hollow symbol stands for  $\rho_{DMSPCC}(s, y, x_i)$  calculated by Eq. (13), and the dashed line represents that determined by Eq. (18). Seen from these figures, the perfect agreement demonstrates clearly that the proposed relationships hold for the DMA-based MLR.

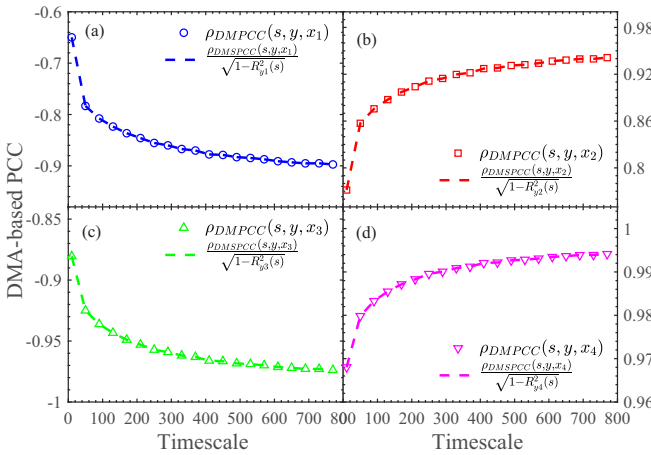


FIG. 7. Examination of the relationship between the DMA-based PCC and SPCC. The four subplots correspond to DMA-based PCCs between the series  $\{y(t)\}$  and one of  $\{x_i(t)\}$ . The hollow symbols and the dashed line are the quantities determined by the left-hand side and right-hand side of Eq. (17), respectively. The perfect agreement suggests that the relationship (17) holds.

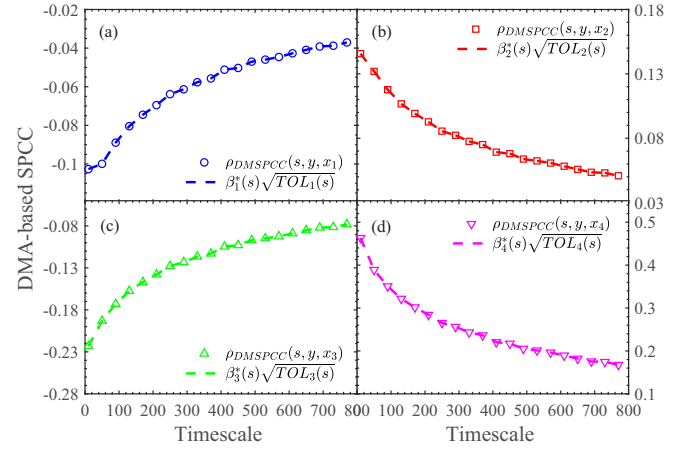


FIG. 8. Examination of the relationship between the DMA-based SPCC and DMA-based regression coefficient. The four subplots are the DMA-based SPCC between series  $\{y(t)\}$  and one of  $\{x_i(t)\}$ . The hollow symbols and the dashed line are the quantities determined by the left-hand side and right-hand side of Eq. (18), respectively. The perfect coincidence suggests that the relationship (18) holds.

## V. APPLICATION TO BEIJING'S AIR QUALITY INDEX SYSTEM

Nowadays, air pollution has become a global problem related to public health [45], especially in cities and regions with rapid industrialization. In our previous studies, we have investigated the interactions of air pollution in adjacent cities [35,46], the difference in dynamic structures of AQI between cities in northern and southern China [47], and correlation between major air pollutants in Beijing [14,15]. In this work, we focus on the impact of the four typical air pollutants on AQI of Beijing using the proposed DMA-based MLR framework. From the six factors involved in the AQI system, we select the four most serious factors affecting the air quality in Beijing, namely, two particulate matters (with diameter  $\leq 2.5 \mu\text{m}$  denoted as  $\text{PM}_{2.5}$  and diameter  $\leq 10 \mu\text{m}$  denoted as  $\text{PM}_{10}$ ) and the two conventional pollutants [carbon monoxide (CO) and nitrogen dioxide ( $\text{NO}_2$ )]. The concentrations of the four pollutants are regarded as the explanatory variables and AQI is regarded as the dependent variable, where a quaternary linear regression model is established. The AQI series and the four series of concentrations of pollutants were recorded daily from January 1, 2014, to December 31, 2019. We first centralize and standardize these series to remove their dimensions. Then we investigate the dependency between Beijing's AQI and the four pollutants at different timescales. For the whole daily data of 2014–2019 (there are 2191 observations for each series), we set timescale  $s$  from 7 days to 364 days with step size 7 days. The estimated scale-dependent regression coefficient together with the SPCC are shown in Fig. 9. Since the data are standardized, the regression coefficient provides the information that can be used to compare the dependence of AQI on the four explanatory variables. Obviously, the two particulate matters, especially  $\text{PM}_{2.5}$ , have a great impact on AQI at all the timescales. Meanwhile CO has almost no effect on AQI due to the fact that its corresponding  $\hat{\beta}(s)$  is close to 0. In addition, with the increase of timescale,  $\text{NO}_2$  has



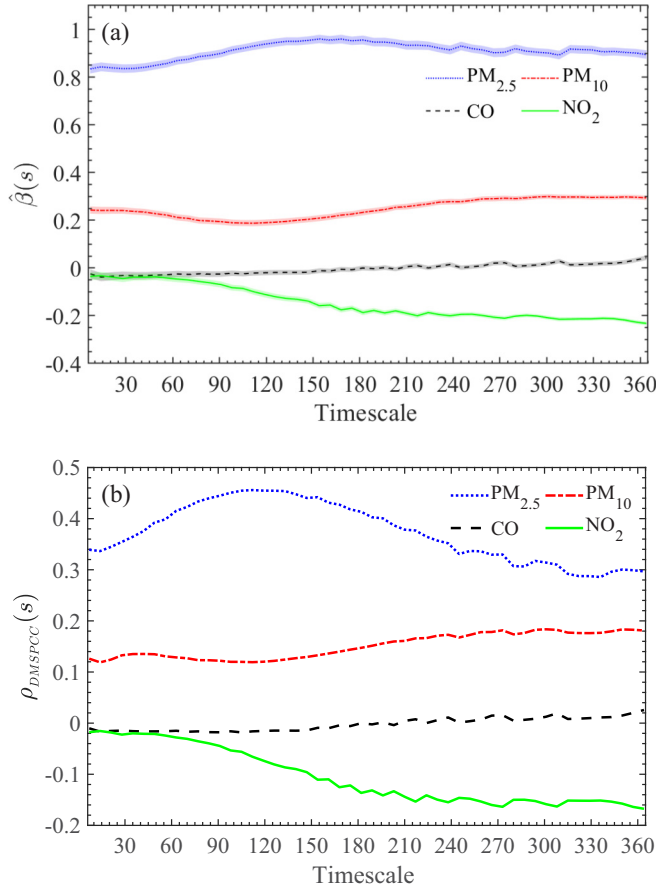


FIG. 9. The quaternary linear regression model of Beijing’s AQI for the years 2014–2019. Panels (a) and (b) are the estimated scale-dependent regression coefficient and scale-dependent SPCC based on the DMA-MLR framework, respectively. The shaded zones around  $\hat{\beta}(s)$  denote the 95% confidence intervals.

a reverse effect on AQI. The above outcomes can also be confirmed by  $\rho_{\text{DMSPPC}}(s)$  shown in Fig. 9(b). The  $\rho_{\text{DMSPPC}}(s)$  between  $\text{PM}_{2.5}$  and AQI is undoubtedly significantly greater than those of the other three. Its contribution to AQI gradually strengthens with the increasing timescale until about 120 days and then gradually weakens. The fact that  $\hat{\beta}(s)$  together with  $\rho_{\text{DMSPPC}}(s)$  presents changeable values at different timescales manifests that the scale-dependent regression model provides richer information on the dependency between those studied variables.

To further investigate whether the particulate matters are the dominant air pollutants in Beijing in recent years, we will assess the changes of the impact of the four pollutants on AQI in these six natural years. To do so, we average the model parameters over the timescales from 7 days to 56 days, which is shown in Fig. 10, where the changes in the six years can be easily obtained from the bar plots. Figures 10(a) and 10(b) show the two coefficients of SPCC and PCC based on the DMA-based MLR framework, respectively. It is easily observed that the two coefficients of AQI and the two particulate matters are positive for the years 2014–2019, while the other two pollutants are negatively correlated with AQI in most of the years.  $\rho_{\text{DMSPPC}}$  and  $\rho_{\text{DMPCC}}$  also illustrate that the intrinsic

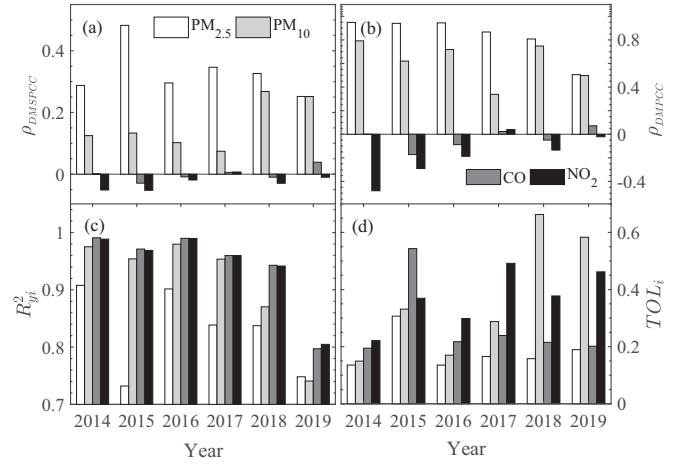


FIG. 10. Annual indicators of the quaternary linear regression model of Beijing’s AQI. (a–d) The averaged SPCC, averaged PCC, averaged coefficient of determination, and averaged tolerance over timescales of 7 to 56 days based on the DMA-MLR framework, respectively.

correlation between AQI and  $\text{PM}_{2.5}$  is the most significant one, which indicates that  $\text{PM}_{2.5}$  is the main pollutant affecting Beijing’s air quality. It was the most serious one in 2015. Besides the two correlation coefficients, the bars of coefficient of determination ( $R^2_{yi}$ ) and tolerance of each pollutants ( $\text{TOL}_i$ ) in the six years are also shown in Fig. 10. The large  $R^2_{yi}$  of the pollutants  $\text{PM}_{10}$ , CO, and  $\text{NO}_2$  confirm that relative to  $\text{PM}_{2.5}$ , the three pollutants are less significant in the contribution to Beijing’s AQI. However, in 2019 the contributions of these four pollutants to AQI were becoming more balanced. The indicator  $\text{TOL}_i$  shown in Fig. 10(d) exhibits the degree of multicollinearity among the series of concentrations of the four pollutants. The  $\text{TOL}_i$  of  $\text{PM}_{2.5}$  to the other pollutants is the smallest in the six years, which implies that the collinearity between  $\text{PM}_{2.5}$  and other three pollutants is always the most serious one.  $\text{PM}_{10}$  and  $\text{NO}_2$  express relatively weak collinearity to the other factors in 2017–2019. The findings uncover that the AQI of Beijing can be mainly reflected by  $\text{PM}_{2.5}$  as  $\text{PM}_{2.5}$  is closely related to the other pollutants.

VI. CONCLUSION

In this work, on the one hand, we extend the DMA-based linear regression model [30,38] to the multivariable case. Artificial tests show that DMA-MLR can not only accurately estimate the regression coefficients but also successfully resist the influence of trends on the variables while the traditional MLR can not. Furthermore, the generated multiscale statistics such as regression coefficient, coefficient of determination, and SPCC can provide richer information than the traditional linear regression analysis can. On the other hand, based on DMA-MLR, we develop a SPCC (denoted as  $\rho_{\text{DMSPPC}}$ ), which can well assess the contribution of explanatory variables to the dependent variable at different scales. Revolving around  $\rho_{\text{DMSPPC}}$ , three important relationships in the DMA-MLR framework are deduced, which help us better understand the relationships among the three typical DMA-MLR models.

The findings from applying these scale-dependent statistics to Beijing’s AQI system provide evidence of the primary contribution of PM<sub>2.5</sub> to Beijing’s air pollution in recent years. Finally, an important point we should note is that in most fields, it is necessary to probe the scaling dependence behavior between the multiple data at different fluctuations. Hence, the  $q$ -order fluctuations [5,18,19] may be considered in our proposed DMA-based MLR framework.

**ACKNOWLEDGMENT**

This research is supported partially by the National Natural Science Foundation of China (Grant No. 61973111) and the Natural Science Foundation of Hunan Province (CN) (Grant No. 2020JJ4377).

**APPENDIX: PROOF OF THE THREE RELATIONSHIPS**

The goal here is to deduce the three important relationships, Eqs. (16)–(18), for the DMA-based MLR.

For Models II and III, we take  $x_i = x_1$  as an example and repeat the process of calculating the regression coefficient as in Model I. The matrix form of the cumulative Models I–III is

$$\begin{cases} Y = X\beta + E, \\ X_1 = X_2\alpha + \Delta, \\ Y = X_2\gamma + H, \end{cases}$$

where  $E = [\sum_{k=1}^1 \epsilon(k), \sum_{k=1}^2 \epsilon(k), \dots, \sum_{k=1}^N \epsilon(k)]^T$ ,  $\Delta = [\sum_{k=1}^1 \delta_1(k), \sum_{k=1}^2 \delta_1(k), \dots, \sum_{k=1}^N \delta_1(k)]^T$ , and  $H = [\sum_{k=1}^1 \eta_1(k), \sum_{k=1}^2 \eta_1(k), \dots, \sum_{k=1}^N \eta_1(k)]^T$ . Set  $X = [X_1, X_2]$ , where  $X_1$  is the first column of the matrix  $X$  with  $N$  elements and  $X_2$  is the last  $p - 1$  columns of  $X$  with size  $N \times (p - 1)$ . Decompose  $T(s)$  into  $[T_1(s), T_2(s)]^T$ , where  $T_1(s) = F_{yx_1}^2(s)$  and  $T_2(s) = [F_{yx_2}^2(s), F_{yx_3}^2(s), \dots, F_{yx_p}^2(s)]^T$  with size  $(p - 1) \times 1$ . Partition the matrix  $F(s)$  into

$$F(s) = \begin{bmatrix} F_1(s) & F_{12}(s)^T \\ F_{12}(s) & F_{22}(s) \end{bmatrix},$$

where  $F_1(s) = F_{x_1}^2(s)$ ,  $F_{12}(s) = [F_{x_1x_2}^2(s), F_{x_1x_3}^2(s), \dots, F_{x_1x_p}^2(s)]^T$  with size  $(p - 1) \times 1$ , and  $F_{22}(s)$  is a symmetric matrix with size  $(p - 1) \times (p - 1)$  obtained by deleting first row and first column from the matrix  $F(s)$ . Let

$$\beta_{22}(s) = \begin{bmatrix} \beta_2(s) \\ \beta_3(s) \\ \vdots \\ \beta_p(s) \end{bmatrix}, \alpha(s) = \begin{bmatrix} \alpha_2(s) \\ \alpha_3(s) \\ \vdots \\ \alpha_p(s) \end{bmatrix}, \gamma(s) = \begin{bmatrix} \gamma_2(s) \\ \gamma_3(s) \\ \vdots \\ \gamma_p(s) \end{bmatrix}.$$

Then  $\beta(s) = [\beta_1(s), \beta_{22}(s)]^T$ . One can easily obtain the relationship among these regression coefficients as

$$\alpha(s) = -\frac{1}{\beta_1(s)}[\beta_{22}(s) - \gamma(s)]. \tag{A1}$$

With the partitions of  $F(s)$  and  $T(s)$ , the normal equation (9) can be rewritten as

$$\begin{cases} F_1(s)\beta_1(s) + F_{12}(s)^T\beta_{22}(s) = T_1(s), \\ F_{12}(s)\beta_1(s) + F_{22}(s)\beta_{22}(s) = T_2(s). \end{cases} \tag{A2}$$

The similar normal equations are obtained for Models II and III:

$$F_{22}(s)\alpha(s) = F_{12}(s), \quad F_{22}(s)\gamma(s) = T_2(s). \tag{A3}$$

Using Eqs. (A2) and (A3), we can simplify  $F_\epsilon^2(s)$ ,  $F_\delta^2(s)$ , and  $F_\eta^2(s)$  into

$$\begin{cases} F_\epsilon^2(s) = F_y^2(s) - \beta_1(s)T_1(s) - \beta_{22}(s)^T T_2(s), \\ F_\delta^2(s) = F_1(s) - \alpha(s)^T F_{12}(s), \\ F_\eta^2(s) = F_y^2(s) - \gamma(s)^T T_2(s). \end{cases}$$

Similarly, the covariances  $F_{\delta y}^2(s)$  and  $F_{\delta \eta}^2(s)$  can be calculated by

$$\begin{cases} F_{\delta y}^2(s) = T_1(s) - \alpha(s)^T T_2(s), \\ F_{\delta \eta}^2(s) = T_1(s) - \gamma(s)^T F_{12}(s) = F_{\delta y}^2(s). \end{cases} \tag{A4}$$

Now we are ready to deduce the three relationships. With Eq. (A2), for given  $s$ ,  $F_{\delta y}^2(s)$  can be further transformed into (for simplicity, in the sequel, we omit  $s$ )

$$\begin{aligned} F_{\delta y}^2 &= \beta_1 F_1 + \beta_{22}^T F_{12} - \beta_1 \alpha^T F_{12} - \beta_{22} F_{22} \alpha \\ &= \beta_1 (F_1 - \alpha^T F_{12}) + \beta_{22}^T (F_{12} - F_{22} \alpha) \\ &= \beta_1 (F_1 - \alpha^T F_{12}) = \beta_1 F_\delta^2. \end{aligned}$$

According to Eqs. (A1), (A2), and (A3), we have

$$\begin{aligned} R^2 - R_{y1}^2 &= \left(1 - \frac{F_\epsilon^2}{F_y^2}\right) - \left(1 - \frac{F_\eta^2}{F_y^2}\right) \\ &= \frac{1}{F_y^2} (F_\eta^2 - F_\epsilon^2) \\ &= \frac{1}{F_y^2} [\beta_1 T_1 + (\beta_{22}^T - \gamma^T) T_2] \\ &= \frac{\beta_1}{F_y^2} (T_1 - \alpha^T T_2) = \frac{\beta_1^2}{F_y^2} (F_1 - \alpha^T F_{12}) \\ &= \frac{\beta_1^2 F_\delta^2}{F_y^2} = \frac{(F_{\delta y}^2)^2}{F_\delta^2 F_y^2} = \rho_{\text{DMSPCC}}^2(y, x_1). \end{aligned}$$

Therefore, Eq. (16) holds. In addition, according to Eq. (A4),

$$\begin{aligned} \frac{\rho_{\text{DMSPCC}}(y, x_1)}{\rho_{\text{DMPCC}}(y, x_1)} &= \frac{F_{\delta y}^2}{\sqrt{F_\delta^2 F_y^2}} \frac{\sqrt{F_\delta^2 F_\eta^2}}{F_\delta^2} \\ &= \frac{F_{\delta y}^2 \sqrt{F_\eta^2}}{F_\delta^2 \sqrt{F_y^2}} \\ &= \sqrt{1 - R_{y1}^2} \end{aligned}$$

and

$$\frac{\rho_{\text{DMSPCC}}(y, x_1)}{\beta_1^*} = \frac{F_{\delta y}^2}{\sqrt{F_\delta^2 F_y^2}} \frac{\sqrt{F_y^2}}{\beta_1 \sqrt{F_{x_1}^2}} = \frac{\sqrt{F_\delta^2}}{\sqrt{F_{x_1}^2}} = \sqrt{1 - R_{x_1}^2}.$$

Therefore, Eqs. (17) and (18) hold.

- [1] S. Jiang, B. G. Li, Z. G. Yu, F. Wang, V. Anh, and Y. Zhou, *Chaos* **30**, 023134 (2020).
- [2] V. A. Reisen, A. M. Sgrancio, C. Lévy-Leduc, P. Bondon, E. Z. Monte, H. H. A. Cotta, and F. A. Ziegelmann, *Appl. Math. Comput.* **346**, 842 (2019).
- [3] S. Okuno, K. Aihara, and Y. Hirata, *Chaos* **29**, 033128 (2019).
- [4] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys. Rev. E* **49**, 1685 (1994).
- [5] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, *Physica A* **316**, 87 (2002).
- [6] W. X. Zhou, *Phys. Rev. E* **77**, 066211 (2008).
- [7] F. Wang, Q. J. Fan, and H. E. Stanley, *Phys. Rev. E* **93**, 042213 (2016).
- [8] M. Höll, K. Kiyono, and H. Kantz, *Phys. Rev. E* **99**, 033305 (2019).
- [9] R. Gebarowski, P. Oświecimka, M. Wątopek, and S. Drożdż, *Nonlinear Dyn.* **98**, 2349 (2019).
- [10] Z. Q. Jiang, W. J. Xie, W. X. Zhou, and D. Sornette, *Rep. Prog. Phys.* **82**, 125901 (2019).
- [11] S. Drożdż, P. Oświecimka, A. Kulig, J. Kwapien, K. Bazarnik, I. G. Gradzińska, J. Rybicki, and M. Stanuszek, *Inf. Sci.* **331**, 32 (2016).
- [12] F. Wang, *Chaos* **26**, 063109 (2016).
- [13] A. N. Pavlov, O. N. Pavlova, O. V. Semyachkina-Glushkovskaya, and J. Kurths, *Eur. Phys. J. Plus* **136**, 1 (2021).
- [14] F. Wang and Q. Fan, *Commun. Nonlinear Sci. Numer. Simul.* **94**, 105579 (2021).
- [15] F. Wang, J. Xu, and Q. Fan, *Commun. Nonlinear Sci. Numer. Simul.* **99**, 105781 (2021).
- [16] E. Alessio, A. Carbone, G. Castelli, and V. Frappietros, *Eur. Phys. J. B* **27**, 197 (2002).
- [17] S. Arianos and A. Carbone, *Physica A* **382**, 9 (2007).
- [18] G. F. Gu and W. X. Zhou, *Phys. Rev. E* **82**, 011136 (2010).
- [19] Z. Q. Jiang and W. X. Zhou, *Phys. Rev. E* **84**, 016106 (2011).
- [20] L. Y. He and S. P. Chen, *Physica A* **390**, 3806 (2011).
- [21] F. Wang, L. Wang, and R. B. Zou, *Chaos* **24**, 033127 (2014).
- [22] Y. Tsujimoto, Y. Miki, S. Shimatani, and K. Kiyono, *Phys. Rev. E* **93**, 053304 (2016).
- [23] M. Kojić, P. Mitić, M. Dimovski, and J. Minović, *Mathematics* **9**, 711 (2021).
- [24] F. X. Zhou, S. Wang, G. S. Han, J. S. Jiang, and Z. G. Yu, *Chaos* **30**, 053113 (2020).
- [25] D. Grech and Z. Mazur, *Acta. Phys. Pol. B* **36**, 2403 (2005).
- [26] L. Xu, P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley, *Phys. Rev. E* **71**, 051101 (2005).
- [27] Y. H. Shao, G. F. Gu, Z. Q. Jiang, W. X. Zhou, and D. Sornette, *Sci. Rep.* **2**, 835 (2012).
- [28] L. Kristoufek, *Physica A* **406**, 169 (2014).
- [29] L. Kristoufek, *Phys. Rev. E* **91**, 022802 (2015).
- [30] L. Kristoufek, *Acta Phys. Pol. A* **129**, 908 (2016).
- [31] B. Podobnik and H. E. Stanley, *Phys. Rev. Lett.* **100**, 084102 (2008).
- [32] C. Shen, *Phys. Lett. A* **379**, 2962 (2015).
- [33] P. Ferreira and L. Kristoufek, *Physica A* **486**, 554 (2017).
- [34] L. Kristoufek, *Chaos Solitons Fractals* **110**, 69 (2018).
- [35] F. Wang, L. Wang, and Y. M. Chen, *Sci. Rep.* **8**, 7475 (2018).
- [36] A. D. Likens, P. G. Amazeen, S. G. West, and C. T. Gibbons, *Physica A* **532**, 121580 (2019).
- [37] O. Tilfani, P. Ferreira, and Y. El Boukfaoui, *Physica A* **532**, 121758 (2019).
- [38] Q. Fan and F. Wang, *Phys. Rev. E* **102**, 012218 (2020).
- [39] P. Ferreira and L. Kristoufek, *Physica A* **545**, 123803 (2020).
- [40] O. Tilfani, L. Kristoufek, P. Ferreira, and M. Y. El Boukfaouia, *Physica A* **588**, 126530 (2022).
- [41] X. Zhao and P. Shang, *Nonlinear Dyn.* **84**, 1033 (2016).
- [42] E. Walker, *Technometrics* **31**, 117 (1989).
- [43] N. Zhang, A. Lin, and P. Yang, *Physica A* **542**, 122960 (2020).
- [44] J. Hosking, *Biometrika* **68**, 165 (1981).
- [45] M. Ives, <https://www.nytimes.com/2016/09/28/world/air-pollution-smog-who.html>, Sept. 27 (2016).
- [46] F. Wang, L. Wang, and Y. Chen, *Sci. Rep.* **7**, 10109 (2017).
- [47] F. Wang, W. C. Zhao, and S. Jiang, *Nonlinear Dyn.* **99**, 1451 (2020).