IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 12, 2020, accepted April 22, 2020, date of publication April 28, 2020, date of current version May 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991109

Fractal Analysis of Power Grid Faults and Cross Correlation for the Faults and Meteorological Factors

TEJUN ZHOU[1,2], JIAZHENG LU[1], BO LI[1], AND YANJUN TAN[1], (Member, IEEE)
[1]State Key Laboratory of Disaster Prevention & Reduction for Power Grid Transmission and Distribution Equipment, Changsha 410007, China
[2]School of Electrical and Information Engineering, Changsha University of Science, Changsha 410007, China

Corresponding author: Tejun Zhou (zhoutejun1988@126.com)

This work was supported by the National Key Research and Development Project under Grant 2016YFC0800104.

ABSTRACT Power-grid faults pose a great threat to the economy and to social stability. In this paper, the detrended fluctuation analysis method is used to investigate the scaling properties of power grid faults and the correlation between faults and meteorological factors that are closely related to power grid faults. Multifractal detrended fluctuation analysis showed that the fault time series were multifractal. Further investigations revealed the origins of multifractality in power-grid fault time series, and the results showed that the temporal correlations in the data represent the distribution of the returns, which are a significant source of multifractal scaling. Then, cross correlations between the fault and the four meteorological factors were investigated using the detrended cross-correlation analysis method. The results showed that maximum wind speed has a considerable impact on the number of transmission system line faults per day, whereas daily precipitation, daily mean air temperature, and maximum wind speed have a considerable impact on the number of distribution system line faults per day.

INDEX TERMS Cross correlation, fault, meteorological factors, multifractality.

I. INTRODUCTION

Electric power transmission systems are a key element of national infrastructure, and blackouts of these systems have major direct and indirect consequences for the economy and national security. Major faults or blackouts of these power-grid systems have serious consequences. Extreme weather is partially responsible for the increase in power outages because an increase in the frequency and severity of extreme weather events [1]–[4], such as hurricanes [5], [6], floods [7], [8], wildfire [9] and winter storms [10], leads to damage to power systems.

Due to the special terrain and climatic conditions of Hunan Province, the percentage of faults caused by meteorological disasters in the Hunan power grid is greater than 90%. Individually, these faults can be attributed to specific extreme weather causes such as lightning strikes, ice storms, wildfires, or conductor galloping [8], [9]. In particular, ice and wildfire disasters last a long time and have widespread effects on the power grid. In 2008, an ice disaster in Hunan destroyed the power grid, and millions of people lost power for one month [10].

It is important to study the distribution law for power-grid faults for the design, operation, and maintenance of transmission lines. However, power-grid faults include several subsystems, such as the climate subsystem (faults caused by disasters) and human agency and lifestyle (faults caused by wildfires). Therefore, the power grid is a typically complex artificial system.

An exclusive focus on these individual causes can overlook the global dynamics of a complex system in which repeated major disruptions from a wide variety of sources are a virtual certainty. Duan *et al.* (2010) has calculated long-term correlations and probability distribution functions of faults in several power utilities and has revealed the self-organized criticality of fault time series in both transmission and distribution systems [11]. Dobson *et al.* (2004) stated that North American blackout data suggest that the frequency of large blackouts is governed by a power law [12]. Newman *et al.* (2011) suggested that blackout size distributions have a power-law

The associate editor coordinating the review of this manuscript and approving it for publication was Padmanabh Thakur.

form over much of their range and that the simulation of evolving network with cascading transmission line outages model shows that apparently sensible efforts to reduce the risk of smaller blackouts can sometimes increase the risk of large blackouts [13]. Five blackout models capturing various critical properties of power systems at different time scales are listed in [14]. Ren *et al.* (2008) analyzed the long-term effects of various policies such as the n-1 criterion and upgrading lines on the probability distribution of outage size [15].

The existing literature analyzes power law using R/S analysis and the detrended fluctuation analysis (DFA) method based on power-grid disturbance events. The indicators used to describe the seriousness of power-grid blackouts are loss of load [13], [16], [17], power outage interval [16], time of power outage [17], loss of power generation [17], number of line outages [18], and others.

However, no study has analyzed whether time series of power-grid faults have multifractal characteristics and revealed their origin. Second, the research just described was focused on time-series analysis of grid power outages and did not attend to the time series of meteorological factors that caused the power-grid disturbance events. In other words, the correlation between the fault time series and the meteorological-factor time series was not analyzed. Third, the differences in the impact of meteorological factors on different voltage grades on the power grid have not been analyzed. Fourth, there has been no study of how the time lags of meteorological factors affect the cross-correlation behavior between time series. If the above four problems can be effectively solved, power companies can assess power grid operation risks and then guide the operation and maintenance of transmission lines and improve the ability of the grid to resist extreme disasters. Furthermore, power companies can guide the design of transmission lines based on the global climate change trend and improve the level of differentiated design for transmission line bodies.

The paper is organized as follows. Section II describes the study area and data resources. The fractal and multifractal features of time series for power grid faults are discussed in Section III. The periodic uncertainty changes in cross-correlation exponents are examined in Section IV, and Section V concludes the paper.

II. STUDY AREA AND DATA RESOURCES

Hunan Province is situated between 24.62–30.13 degrees North latitude and 108.78–114.27 degrees East longitude. The province is tilted from south to north, with the east, south, and west sides surrounded by mountains and the central and northern parts relatively low-lying in a U-shaped basin. Mountains and hills occupy more than 80% of the province and plains less than 20%. About 4.3 percent of its surface area is higher than 1000 m above sea level, and large parts of it lie between 100 and 800 m above sea level. Hunan's climate is classified as humid subtropical, with short, cool, damp winters, very hot and humid summers, and plenty of rainfall. January temperatures typically range from -2 °C to 8 °C, whereas July temperatures typically range between 27 °C and 40 °C. The transition from spring into summer is a period of heavy rain and lightning strikes. Lightning activity in Hunan Province is mainly concentrated in April to September each year, during which 95.7% of total lightning strikes occur. Fifty percent of lightning strikes have a current amplitude greater than 24 kA, and 1% of strikes have a current amplitude greater than 130 kA. The risk of wildfires is particularly high every year in February to April. Sweeping tombs on Qingming Festival and land-clearing fires for plantations increase the potential sources of ignition. During the period of high wildfires, the number of wildfires can reach more than 1,000 per day. There are strong winds in the winter, with frequent ice disasters.

The Hunan power grid consists of 531 transmission lines of 110 kV and above (110, 220, 500, $\pm$500, and $\pm$800 kV) and 5217 distribution lines with voltage levels of 35 kV and 10 kV. By deriving fault data from the power production management system of the Hunan power grid, it was found that there were 2497 fault occurrences from January 2015 to June 2018. The frequency of faults was 1.54 times/day. The three main causes of faults were lightning strikes, wildfires, and ice storms, accounting for 46.8%, 16.04%, and 10.11% respectively [19].

The meteorological data came from the China Meteorological Data Service Center (http://data.cma.cn). The study used observed daily datasets closely related to power grid faults, which included average relative humidity, daily precipitation, daily mean air temperature, and maximum wind speed data from 37 ground meteorological stations in Hunan Province, including Xinhua, Sangzhi, and Huaihua. There were 81,030 Chinese records that covered the period from 2013 to 2018, with a continuous record over six years.

III. FRACTAL ANALYSIS OF POWER GRID FAULTS AND THEIR ORIGIN

A. SCALE EXPONENT CHARACTERISTICS OF FAULT TIME SERIES

The external manifestation of scale exponent characteristics is that self-organized criticality may govern the complex dynamics of power grid faults by proving that there is a long tail feature between frequency and event scale [17], [18], [20], [21], leading to autocorrelation between specific event scales. This paper further examines the hypothesis that various autocorrelation features may also reflect the dynamic characteristics of the time series at different time scales. The fluctuation of the number of transmission line faults has a seasonal, cyclical trend but exhibits random and other factors. If a sequence is to be studied in depth, it is important to decompose or remove its interference components. DFA is more effective than the traditional sequential analysis method at determining these features [22].

Suppose that there is a time series $\{x_k, k = 1, \ldots, N\}$, where k represents the time interval. The DFA method performs the following five steps:

Step 1: Calculate the accumulated deviations of sequence from the mean:

$$Y(i) = \sum_{k=1}^{i} (x_k - \langle x \rangle), \tag{1}$$

where $\langle x \rangle = \left(\sum_{k=1}^{N} x_k\right)/N$.

Step 2: The sequence Y is equally divided into N_s non-overlapping intervals of equal time length s, where $N_s = [N/s]$. The same is done for the reverse order of the sequence Y, giving $2N_s$ intervals of equal length for better use of the data.

Step 3: The data are fitted using the least squares method to obtain a local trend for each interval. The time series after removing local trends are recorded as $Y_s(i)$, which represents the difference between the original series and the fitted values $p_v(i)$,:

$$Y_s(i) = Y(i) - p_v(i). \tag{2}$$

Step 4: Calculate the variance after removing the trend for each interval. The order and reverse order sequences are calculated separately using Eqs. (3) and (4):

$$F_s^2(v) = \langle Y_s(i) \rangle = \frac{1}{s} \sum_{i=1}^{s} Y_s^2 [(v-1)\, s + i]$$
$$v = 1, 2, \ldots, N_s \tag{3}$$

$$F_s^2(v) = \frac{1}{s} \sum_{i=1}^{s} Y_s^2 [N - (v - N_s)\, s + i]$$
$$v = N_s + 1, N_s +, \ldots, 2N_s. \tag{4}$$

Step 5: Mean and square of the variances for all equal-length intervals are calculated to obtain the standard DFA fluctuation function:

$$F(\mathrm{s}) = \sqrt{\frac{1}{2N_s} \sum_{v=1}^{2N_s} \left[F^2(v, s)\right]}. \tag{5}$$

Obviously, as s becomes larger, the variance will also increase. If the sequence x is long-dependent, then $F(s)$ has a power-law relationship:

$$F(s) \sim s^{\lambda}. \tag{6}$$

where λ is the Hurst exponent.

In Figure 1, the parameter s represents the time scale and the function $F(s)r$ epresents the fluctuation magnitude. The parameter R denotes the correlation coefficients to the fitted line. Figure 1 shows significant autocorrelation in the number of transmission system line faults per day (TSLF) time series when the time scale is $>$ 8 days and $<$ 230 days and number of distribution system line faults per day (DSLF) time series when the time scale is $>$ 8 days and $<$ 530 days, respectively. This means that the current and future numbers of faults per day are affected by fault occurrences in the past. The positive correlation is in accordance with the scale exponent. The larger the scale exponent, the stronger the positive correlation will be. As shown in Figure 1, there are two crossover time

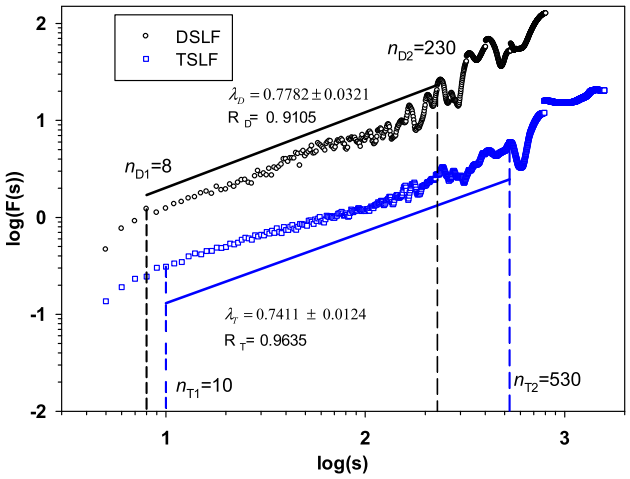

FIGURE 1. Scaling behaviors of fault time series.

TABLE 1. Hurst exponent and correlation to the fitted line.

Items compared	TSLF	DSLF
Hurst exponent	λ_T=0.7411	λ_D=0.7782
Correlation to the Fitted Line	R_T=0.9635	R_D=0.9105

scales $n_\times$ in the log-log plots of $F(s)$ versus s. These two crossovers divide $F(s)$ into three regions. The existence of these regions is due to the competition between noise and sinusoidal trend. For $s < n_1$ and $s > n_2$, the noise has the dominating effect. For $n_1 < s < n_2$, the sinusoidal trend dominates [23]–[25]. Therefore, the inflection point n_2 of the scale curve may correspond to a natural cycle taking 0.6 years or 1.5 years. The Hurst exponent and the correlation coefficient of the fluctuation function and fitted lines were calculated and are listed in Table 1.

B. MULTIFRACTAL CHARACTERISTICS OF FAULT TIME SERIES

The power-law relationship between the covariance function of the fault time series after eliminating the trend and the scale can only indicate the scale invariance of the series, which can be described by a scale exponent. When a sequence has a multifractal feature, the heterogeneity of the time distribution for power grid faults cannot be described. Therefore, several scale indices or exponential spectra are needed to fully describe its scale behavior.

An existing method called multifractal detrended fluctuation analysis (MFDFA) [26] can study the multifractal characteristics of non-stationary time series. Compared to DFA, MFDFA revealed a q-order fluctuation function $F_q(s)$ that varies with the parameter q introduced at the fourth step, as shown in Eqs. (7).

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} \left[F^2(v, s)\right]^{\frac{q}{2}} \right\}^{\frac{1}{q}}, \tag{7}$$

where N_s represent the number of sections into which the TSLF, DSLF, or meteorological time series is segmented and $N_s = [N/s]$. When $q = 2$, it is the standard DFA calculation formula, or in other words, DFA is a special case of MFDFA.

Log-log plots of parameter s versus function $F_q\ (s)$ make it possible to study the scale behavior of the fluctuation function. If the time series has autocorrelation features, then the fluctuation function $F_q\ (s)$ and the variable s satisfy the power-law relationship, which is defined as $F_q\ (s) \sim s^{h(q)}$, where $h(q)$ is the generalized Hurst exponent. In the case of $q < 0$, the magnitude of the small fluctuation deviation $F^2\ (v, s)$) determines the magnitude of $F_q\ (s)$. However, in the case $q > 0$, the magnitude of the large fluctuation deviation $F^2\ (v, s)$) determines the magnitude of $F_q\ (s)$.

Kantelhardt (2002) found that the relationship between the scaling exponents $\tau(q)$ and the generalized Hurst exponent $h(q)$ and the relationship between the multifractal spectrum $f(\alpha)$ and the generalized Hurst exponent $h(q)$ are as shown in Eqs. (8) and (9) respectively:

$$\tau(q) = qh\,(q) - 1, \tag{8}$$

When the time series has monofractal features, the graph of $\tau(q)$ is approximately linear with changing q. When the time series has multifractal features, $\tau(q)$ is nonlinear with changing q, and the stronger the non-linearity of $\tau(q)$, the greater will be the multifractal strength:

$$f(\alpha) = q\alpha - \tau\,(q)\,, \tag{9}$$

where α is the singular exponent, which can be used to describe the different degrees of singularity in each interval for transmission-line faults. The value of the multifractal spectrum $f(\alpha)$ reflects the fractal dimension with α. If the time series studied is monofractal, $f(\alpha)$ takes on a certain value. If the time series is multifractal, the $f(\alpha)$ curve takes on a unimodal bell shape [27].

The MFDFA method was used to study the multifractal characteristics of power grid faults. Then the generalized Hurst exponent curve, scaling exponent curve, multifractal spectrum curve, and fluctuation function curve were calculated, as shown in Figs. 2(a)–2(d).

Figure 2(a) shows the curves of $h(q)$ versus q for the fault time series in the Hunan power grid. Clearly, $h(q)$ decreased nonlinearly as q increased, indicating significant multifractal characteristics of the fault time series, which could not be characterized by monofractality.

From Fig. 2(b), it is clear that $\tau(q)$ increased nonlinearly as q increased, indicating significant multifractal characteristics of the fault time series.

As shown in Fig. 2(c), the multifractal spectrum has a long right tail and left tail, reflecting that the multifractal structure of time series is sensitive to the local fluctuations with large magnitudes and to those with small magnitudes, respectively. The width of multifractal spectrum ($\Delta\alpha$L+$\Delta\alpha$R) is 1.186. These results demonstrated that the fault time series in the

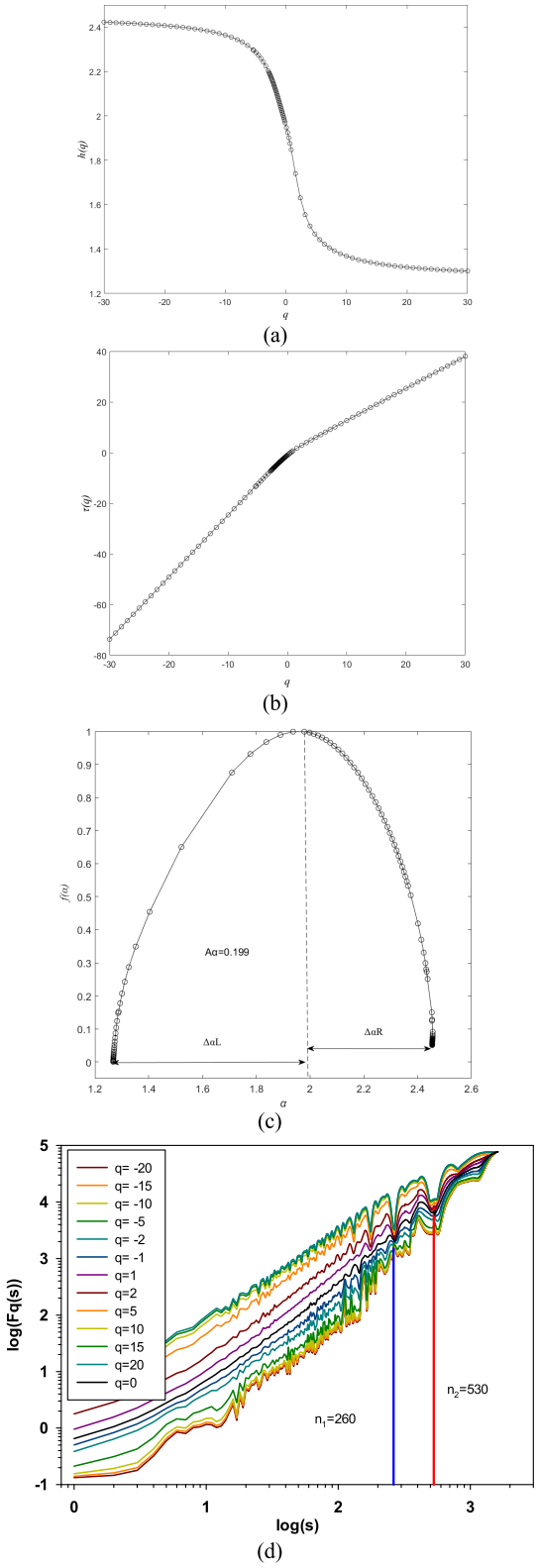

FIGURE 2. Scale-free features of fault time series based on the MFDFA results: (a) generalized Hurst exponent for power grid fault time series; (b) $\tau(q)$ for fault time series; (c) multi-spectrum parameter $f(\alpha)$ and singular exponent α changes in the fault time series; (d) fluctuations in function $F_q\ (s)$ and s changes in the fault time series.

Hunan power grid was characterized by multifractality. The shape of the multifractal spectrum $f(\alpha)$–α curve is truncated on the right. Then the asymmetry parameter Aα was introduced to quantify the asymmetric [30]. The value of Aα means that the faults of Hunan power grid are sensitive to the large magnitude of local fluctuations [27]–[29].

As q takes on different values, the fluctuation function curve varies over a distinctly linear interval. Figure 2(d) shows that in the scale-free range between 8 days and $N/4$, i.e., 404 days, when the time scale is short, the logarithmic value of the fluctuation function changes greatly when the order q is negative, but the numerical change is small when the number q is positive, indicating that a small fluctuation deviation plays a determining role in power-grid fault series in the short time. When the time scale is long, the logarithmic value of the fluctuation function changes little when the order q is negative, but the numerical change is much larger when the number q is positive, indicating that a large fluctuation deviation plays a determining role in power grid fault series over the long term.

Ninety percent of transmission line faults in Hunan Province are caused by lighting, wildfires, storms, ice storms, and similar causes [19]. The above causes are closely related to meteorological factors. However, meteorological factors have a lower probability of change in Hunan Province over a short period of time, during which the number of faults remained similar. The fluctuation deviation was small over the short period of time studied. As time scales become longer, the variability of meteorological factors may become significant and dominant, causing the number of faults to vary greatly and leading to large differences and fluctuations.

C. ORIGINS OF MULTIFRACTALITY

It is generally accepted that there are two possible sources of multifractal scaling in time-series data [26]. It can be predominantly due to (1) long-term correlations of small and large fluctuations or to (2) the fact that the data are drawn from a heavy-tailed probability distribution. Both these influences can individually be removed from the data to reveal what impact they have on the multifractality of the time series.

A simple way to check whether correlations in the data produce any scaling is to shuffle the data using the method as suggested by Kantelhardt *et al.* [26]. Shuffling removes time correlations, and any scaling that remains must be due to the probability distribution from which the data are drawn. The distribution of the values is not affected by reordering the series. Any individual shuffle may still contain some correlations, and therefore, to be sure to completely rid the data of all correlations, the fault data were shuffled for a number of times equal to twenty times the length of the time series, each random permutation beginning with a new random number generator seed in MATLAB (https://jp.mathworks.com/help/matlab/ref/rng.html). Phase randomization of time series can weaken the non-Gaussian nature of a distribution. Hence, a Fourier transform of the time series was taken, after which the phases were randomized,

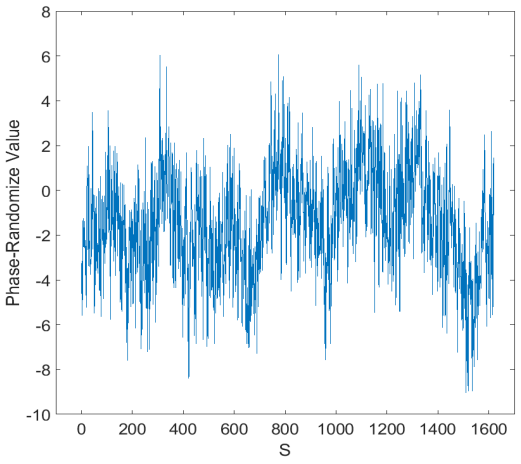

FIGURE 3. Phase-randomized surrogates of fault time series based on Fourier transform.

and an inverse Fourier transform was performed. The result was a surrogate time series for the power grid fault time series, as shown in Fig. 3.

(1) As shown in Figs. 4(a)–4(c), the generalized Hurst exponents of the surrogate time series, $h_{\mathrm{surr}}(q)$, and of the shuffled time series, $h_{\mathrm{shuf}}(q)$, decreased with increasing q (Fig. 4(a)). $\tau(q)$ of the shuffled sequence and the surrogate sequence for the Hunan power-grid fault time series were nonlinear with q. $\tau(q)$ of the shuffled sequence and the surrogate sequence for the Hunan power-grid fault time series were nonlinear with q, but their nonlinearities were obviously weaker than that of the original sequence. (Fig. 4(b)). The $f(\alpha)$ curves of the shuffled sequence and the surrogate sequence are unimodal bell-shaped images. Figure 4(c) shows that the multifractal spectrum widths of the shuffled sequence, $f_{\mathrm{shuf}}(\alpha)$, and the surrogate sequence, $f_{\mathrm{surr}}(\alpha)$, are significantly smaller than that of the original sequence, $f_{\mathrm{origin}}(\alpha)$, indicating that rearrangement and phase randomization of the sequence both weakened the multifractal time-series characteristics. These showed that the long-range correlation and the fat-tailed distribution of the fault time series for transmission lines in the Hunan power grid both caused multiple fractal features.

(2) The extent of the decrease was less for $h_{\mathrm{surr}}(q)$ than for $h_{\mathrm{shuf}}(q)$, and the extents of decrease for $h_{\mathrm{surr}}(q)$ and $h_{\mathrm{shuf}}(q)$ were less than that of $h_{\mathrm{origin}}(q)$ (Fig. 4(a)). The nonlinearity of the surrogate data $\tau_{\mathrm{surr}}(q)$ was weaker than that of the rearranged data $\tau_{\mathrm{shuf}}(q)$ (Fig. 4(b)). Moreover, the width of the multifractal spectrum, $f_{\mathrm{shuf}}(\alpha)$, of the shuffled sequence ($\Delta\alpha \mathrm{L}_{\mathrm{shuf}} + \Delta\alpha \mathrm{R}_{\mathrm{shuf}} = 0.836$) was significantly larger than the width of the multifractal spectrum, $f_{\mathrm{surr}}(\alpha)$, for the surrogate sequence ($\Delta\alpha \mathrm{L}_{\mathrm{surr}} + \Delta\alpha \mathrm{R}_{\mathrm{surr}} = 0.717$) (Fig. 4(c)). As seen in upper panel of Figs. 4(c), the scaling exponents of the shuffled sequence (slope of $F_{\mathrm{q}}(s)$) slightly depend on q for $q < 0$ while weakly for $q > 0$. The scaling exponents of the surrogate sequence (slope of $F_{\mathrm{q}}(s)$) visibly depend on q for $q < 0$ while very weakly for $q > 0$. The opposite is true for scaling exponents of the origin sequence (Fig. 2(d)).

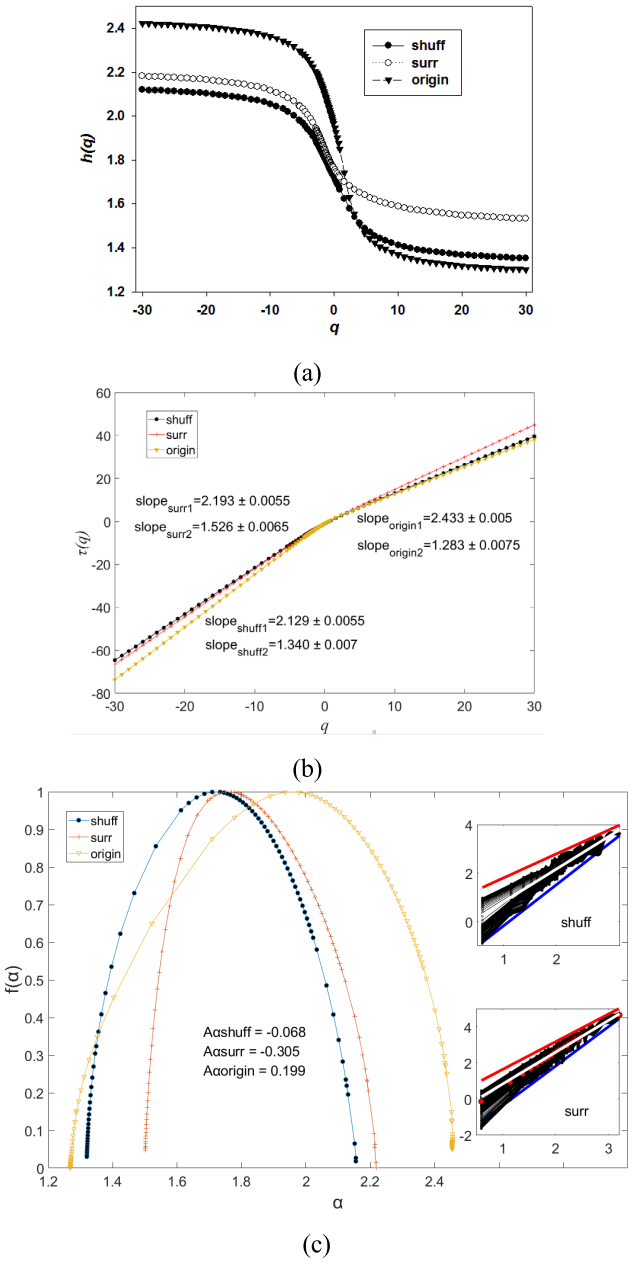

FIGURE 4. Comparison of scale-free features for shuffled, surrogate, and original fault time series: (a) generalized Hurst exponent for shuffled, surrogate, and original fault time series; (b) $\tau(q)$ for shuffled, surrogate, and original fault time series; (c) changes in the multispectral parameters $f(\alpha)$ and the singular exponent α for shuffled, surrogate, and original fault time series.

This correlates with orientations of asymmetries seen in left panel of Figs. 4(c) [30]. These indicate that the fat-tailed probability distribution plays an important role in the multiscale changes in transmission-line faults in the Hunan power grid and that the effect of long memory on the multiscale changes of transmission-line faults in the Hunan power grid is relatively weak, which is consistent with the result that heavy-tailed pdf is responsible for apparent multifractality for short time series [31], [32].

IV. ANALYSIS OF THE EFFECT OF METEOROLOGICAL FACTORS ON FAULTS

A. DETRENDED CROSS-CORRELATION ANALYSIS (DCCA) BETWEEN FAULTS AND METEOROLOGICAL FACTORS

Podobnik and Stanley proposed the DCCA method [33], which describes the power-law correlation between two non-stationary time series. The principle of the method is to minimize the impact of external trends on the cross-correlation calculations by eliminating trend covariances. It is a generalization of the DFA method. The long-range cross-correlation characteristics of the two sets of time series depend not only on their own set's past values, but also on the historical values of the other set's variable [34]. If the correlation between the two time series is strong, it indicates that the consistency of wave direction between the two time series is strong.

The research results in [35] indicate that average relative humidity, average daily precipitation, mean temperature, and extreme wind-speed data have fractal characteristics on the time scale ranging between 25 and 366 days. Therefore, the time-scale maximum for analyzing time-series correlation is 366 days when investigating the cross correlation between meteorological factor time series and TSLF/DSLF.

Suppose there are two time series $\{x_k, k = 1, \ldots, N\}$ and $\{x'_k, k = 1, \ldots, N\}$, where k represents the time interval. This method performs the following six steps:

Step 1: Calculate the accumulated deviation of sequences from the mean:

$$Y(i) = \sum_{k=1}^{i} (x_k - \langle x \rangle), \tag{10}$$

$$Y'(i) = \sum_{k=1}^{i} (x'_k - \langle x' \rangle), \tag{11}$$

where $\langle x \rangle = \left(\sum_{k=1}^{N} x_k\right)/N$, $\langle x' \rangle = \left(\sum_{k=1}^{N} x'_k\right)/N$.

Step 2: The sequences Y and $Y'(i)$ are equally divided into N_s non-overlapping intervals of equal time length s, where $N_s = [N/s]$.

Step 3: The data are fitted using the least squares method to obtain a local trend for each interval. The time series after removing local trends are recorded as $Y_s(i)$ and $Y'_s(i)$, which represent the difference between the original series and the fitted values $p_v(i)$ and $p'_v(i)$:

$$Y_s(i) = Y(i) - p_v(i), Y'_s(i) = Y'(i) - p'_v(i). \tag{12}$$

Step 4: Calculate the variance after removing the trend for each interval using Eq. (13):

$$f^2_{DCCA}(s, v) = \frac{1}{s} \sum_{k=1}^{s} (Y_k - p_v(i))(Y'_k - p'_v(i)). \tag{13}$$

Step 5: Calculate the mean and square of the variances for all equal length intervals to obtain the detrended cross-correlation magnitude $F_{DCCA}(s)$, which is defined in Eq. (14):

$$F_{DCCA}(s) = \sqrt{\frac{1}{2m} \sum_{v=1}^{2m} [f^2(v, s)]} \tag{14}$$

Step 6: Develop the $F_{DCCA}(s)$ to s relationship in double logarithmic coordinates. Then fit a curve to $F_{DCCA}(s)$ and s using the least squares method, so that the scaling exponent λ can be obtained:

$$F_{DCCA}(s) \sim s^{\lambda}. \tag{15}$$

The scaling exponent λ is a parameter to evaluate the degree of correlation between the time series x_k and x'_k.

This study used the DCCA method to study the cross correlation between TSLF/DSLF and the meteorological factors that influence the two [36], [37], [38].

To perform a quantitative analysis of the long-range cross correlation between the time series for meteorological factors and TSLF/DSLF, an indicator of the inter-correlation significance level is needed. The relationship $F^2_{f+g} - (F^2_f + F^2_g) \approx 0$ should hold true when two time series, $f(i)$, $g(i)$, have no cross correlation [40]. Therefore, this study used $|(F^2_{f+g} - F^2_f - F^2_g)/F^2_{f+g}|$ as the indicator of the inter-correlation significance level.

In Figs. 5(a), 5(b), and 5(c), the cross correlation did not satisfy the power-law property. Although each individual time series (TSLF, daily average precipitation, daily average temperature, and daily average humidity) has long-range correlation characteristics, the TSLF and the daily average precipitation, daily average temperature, and daily average humidity time series have weak cross correlation, with no long-term memory behavior. The long-range correlations between the time series for TSLF and maximum wind speed are stronger than those between the time series for TSLF and daily mean air temperature, average relative humidity, or daily precipitation. The value of the scaling exponent λ is 1.05 (Fig. 5(d))

In Figs. 6(a), 6(b), and 6(c), the inter-correlation significance levels between the time series for TSLF and daily mean air temperature, average relative humidity, and daily precipitation remained between 0 and 0.1. Most inter-correlation significance levels between the time series for TSLF and for maximum wind speed remained between 0.15 and 0.30, which is consistent with the linear fitting confidence results shown in Fig. 5(d).

One of the reasons for these results was the transmission system line to withstand the high relative humidity, high daily precipitation and high daily mean air temperature, while unable to withstand disasters related to extreme winds. As far as the Hunan power grid faults data is concerned, strong winds cause transmission-line faults in several ways. The main mechanism is windage yaw discharge of the transmission line due to the phase distance under the lower parameter. Strong winds also raise up foreign materials, such as plastic belts and kites lying on the ground, causing discharge of transmission lines. It may be that transmission lines become iced up in winter and that strong wind causes the ice-covered line to gallop. Wind-based faults and gallops are directly related to the magnitude of the extreme wind speed. The other reason is that the change trends of the TSLF are not consistent

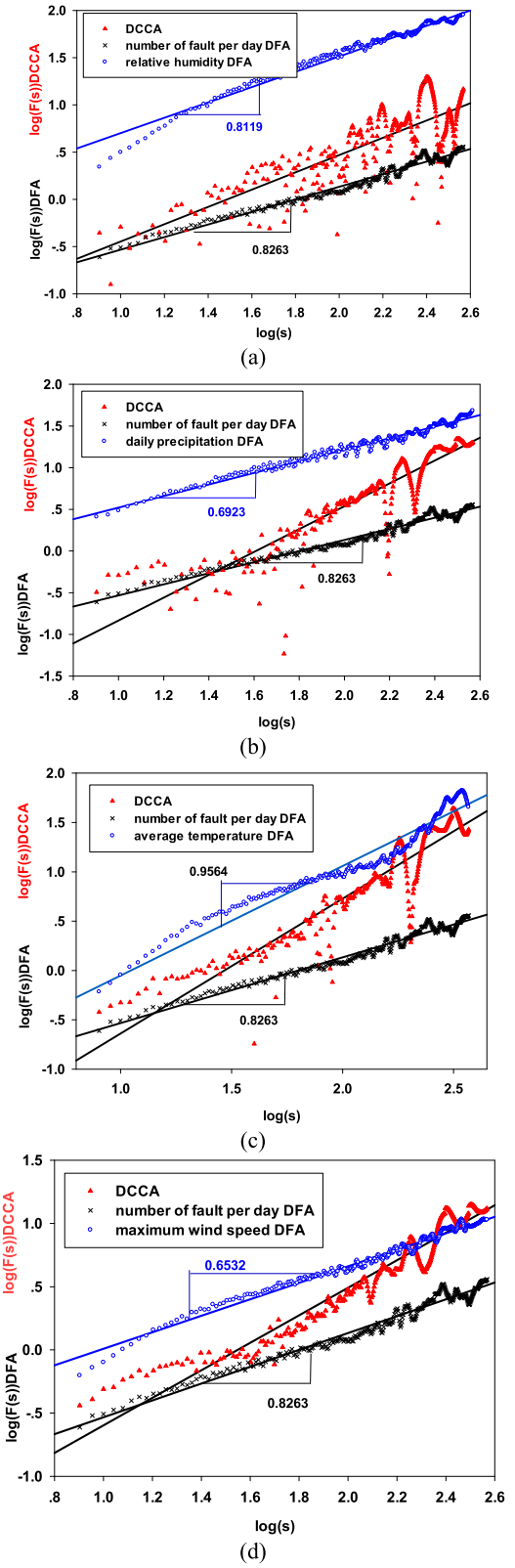

FIGURE 5. Cross-correlation analysis of TSLF and meteorological factor time series based on DCCA: (a) average relative humidity and TSLF time series; (b) daily precipitation and TSLF time series; (c) daily mean air temperature and TSLF time series; and (d) maximum wind speed and TSLF time series.

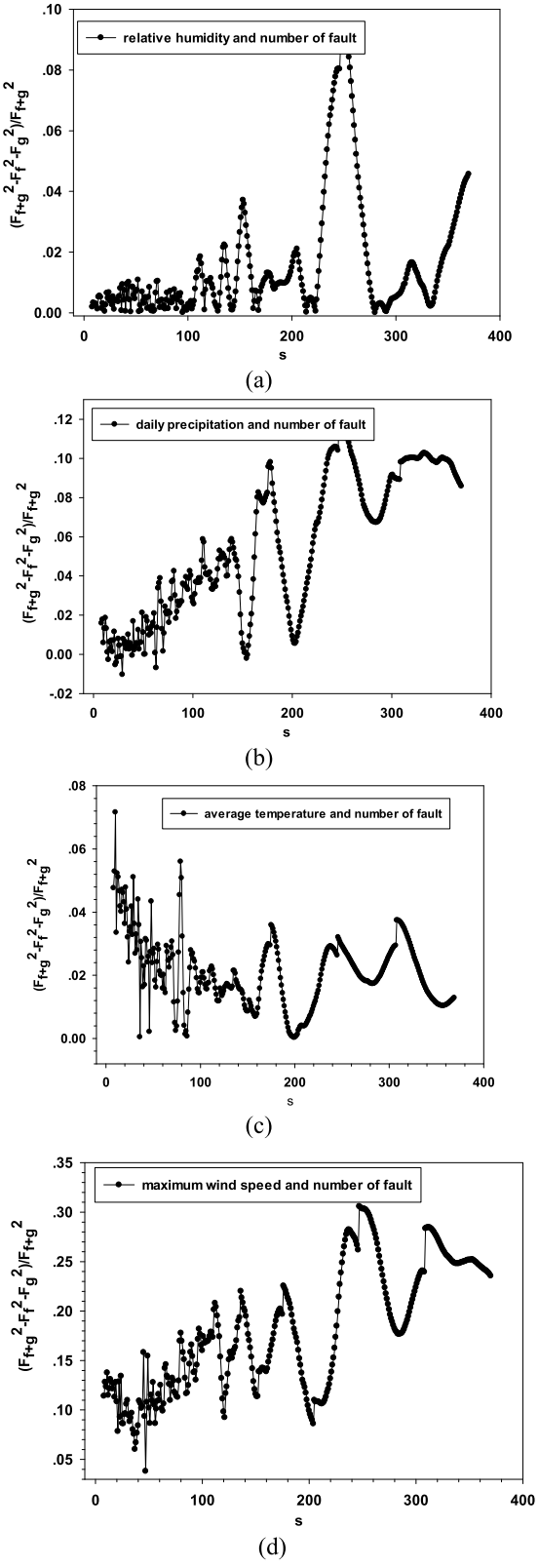

FIGURE 6. Significance levels of the correlation coefficients between TSLF and meteorological data time series: (a) TSLF and average relative humidity; (b) TSLF and daily precipitation; (c) TSLF and daily mean air temperature; and (d) TSLF and maximum wind speed.

with that of average relative humidity, daily precipitation and daily mean air temperature. Although many faults, such as ice-flashing trips, are caused by ice disasters, the factors affecting ice disasters include precipitation, average temperature, and humidity. The mechanism of these multiple factors acting on the grid may have multiple aspects that may at times oppose each other. Extremely high and low temperatures can cause the load to increase, causing transmission-line trips. Low temperatures can also cause the lines to become covered with ice and the towers to break.

When the results in Figs. 7 and 8 are combined, the significance levels of the correlation coefficient between distributed line faults and average relative humidity are below 10%, which means that they are only weakly correlated. The significance levels of the correlation coefficients between distributed line faults and daily precipitation, between distributed line faults and daily mean air temperature, and between distributed line faults and maximum wind speed are greater than 20%, which shows that they are all strongly correlated.

The ability to withstand extreme weather disasters may be the major reason for the correlation difference of TSLF and DSLF. As for the transmission lines of the Hunan power grid, the structure of the power grid with voltage levels of 110 kV and above is relatively strong, and therefore it is unusual that the transmission lines become overloaded due to excessively high or low temperatures, which tend to increase the electrical load sharply. Furthermore, the Hunan power grid belongs to the state-owned electric utility monopoly of China. The construction, operation, and maintenance standards of the backbone network (voltage levels of 110 kV and above) are very high. The clearance distances to the ground of the transmission lines are high enough to protect against floods. High average relative humidity has limited impact on the operation of the power grid. Therefore, the long-range correlation between the time series of temperature, average relative humidity, and daily precipitation and the time series of transmission-line faults in the 110 kV and above network is not obvious. However, for the distribution lines of the Hunan power grid, the structure of the distribution-line grid is still relatively weak due to insufficient investment. Its ability to withstand natural disasters is still very inadequate. Extremely hot weather has increased the temperature-related load, and distribution lines are becoming more heavily loaded in summer and winter. Moreover, China has a huge population, but inadequate land. Many distribution lines and substations are not fully equipped for floods and geological disasters (which mainly refers to landslides in Hunan). Once heavy rains occur, they are likely to cause large-scale blackouts. Similarly, high winds can cause faults in the distribution system network or even break ice-covered towers. Therefore, unlike TSLF, the long-range correlations between the time series of temperature, daily precipitation, and extreme wind speed and DSLF are obvious.

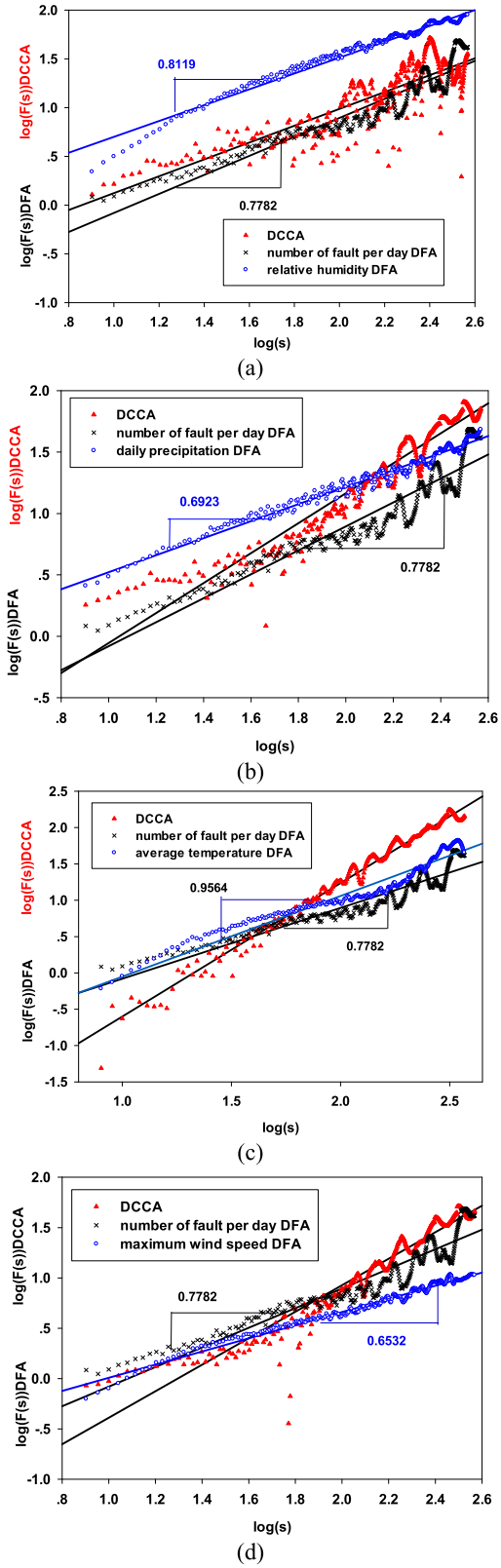

FIGURE 7. Cross-correlation analysis of DSLF and meteorological factor time series based on DCCA: (a) average relative humidity and DSLF time series; (b) daily precipitation and DSLF time series; (c) daily mean air temperature and DSLF time series; and (d) maximum wind speed and DSLF time series.

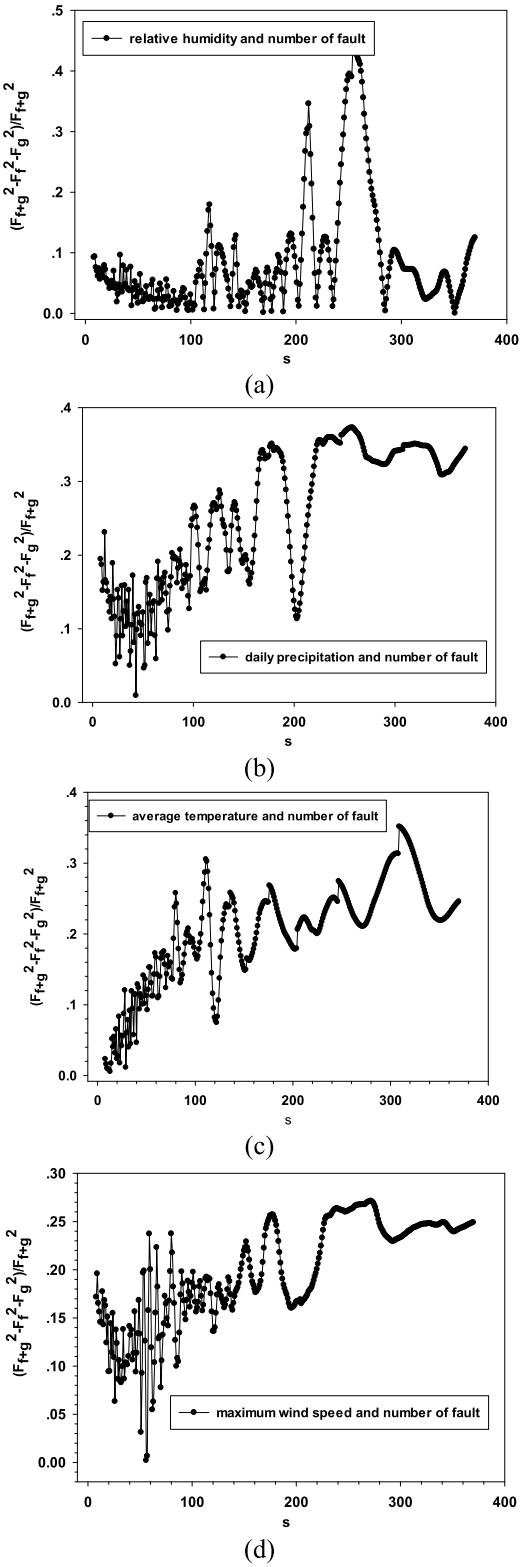

FIGURE 8. Significance levels of the correlation coefficients between DSLF and meteorological data time series: (a) DSLF and average relative humidity; (b) DSLF and daily precipitation; (c) DSLF and daily mean air temperature; and (d) DSLF and maximum wind speed.

B. DCCA ANALYSIS BASED ON TIME DELAY (TLDCCA)

The previous subsection described the dynamics of the cross correlation of power grid faults and meteorological data at a single point in time. The long-range correlations between the time series of temperature, daily precipitation, and extreme wind speed and DSLF are obvious, and the long-range correlation between the time series for TSLF and maximum wind speed is strong.

Therefore, this subsection will describe the application of DCCA to cross correlate these four pair datasets based on a time lag, τ, as shown in Table 2.

TABLE 2. Pair datasets for cross correlation based on a time lag.

Num	Time Series 1	Time Series 2
1	TSLF	Maximum wind speed time series
2	DSLF	Maximum wind speed
3	DSLF	Daily precipitation
4	DSLF	Daily mean air temperature

The covariance of DCCA based on a time delay [40], [41] between time series x_k and x_k' is defined as:

$$f^2_{DCCA}(v_1, s) = \frac{1}{s}\sum_{i=1}^{s}(Y_{\mathrm{k}} - p_{v1}(i))(Y'_{k+\tau} - p'_{v1,\tau}(i)), \quad (16)$$

where τ is the time delay, which was varied between -600 days and -1 day by a step of 1 day.

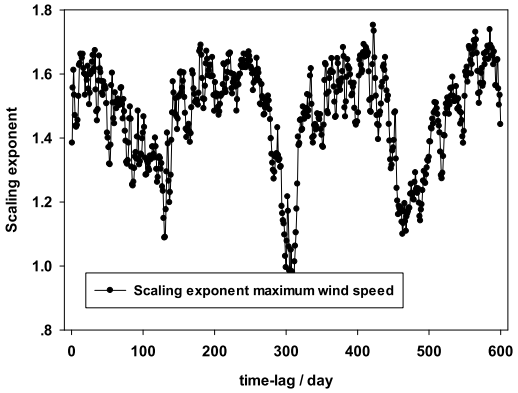

FIGURE 9. Scaling exponents of DCCA based on time lag for TSLF and maximum wind speed.

As shown in Figs. 9 and 10, for any time lag, the scaling exponents between the transmission-line faults (TSLF and DSLF) and the meteorological factor time series range from 0.8 to 1.9. This shows that there is a strong cross correlation between the above four sets of time series. The scaling exponent changes periodically with the time lag τ, with a period of about 180 days. The reason may be that the affecting factors have a semi-annual impact on the Hunan power grid, which is consistent with actual meteorological events. However, the scaling exponents of DCCA based on time lag τ for

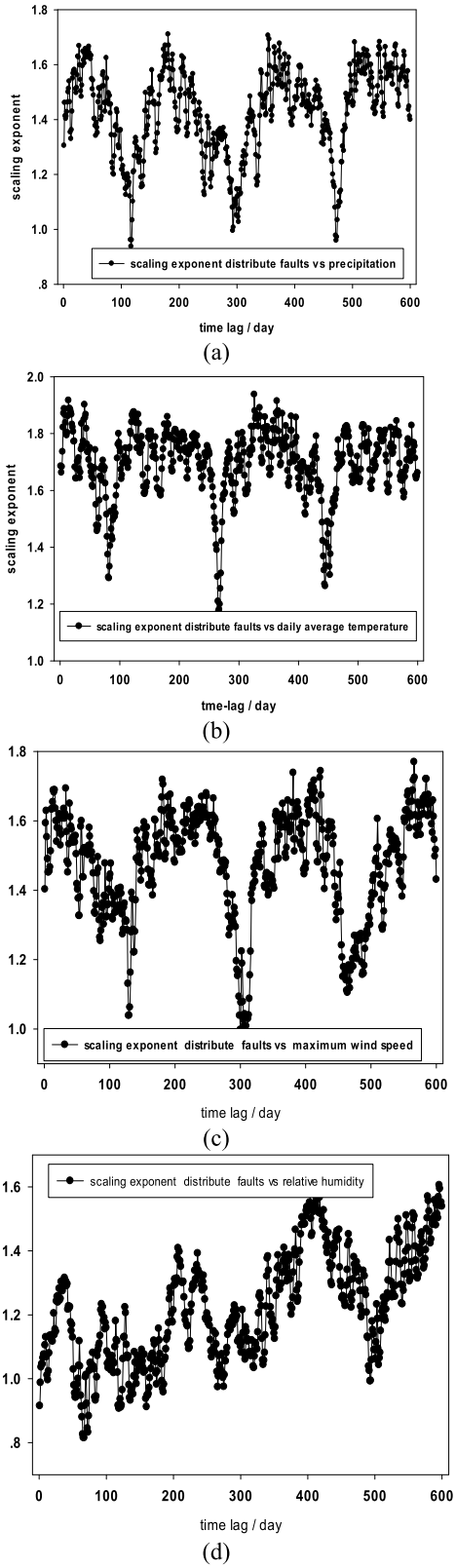

FIGURE 10. Scaling exponents of DCCA based on time lag for DSLF and meteorological factors: (a) DSLF and daily precipitation; (b) DSLF and daily mean air temperature; (c) DSLF and maximum wind speed; and (d) DSLF and average relative humidity.

DSLF and average relative humidity are not periodic, which is consistent with the results in the previous subsection.

V. CONCLUSION

Most of the faults were caused by meteorological disasters in the Hunan power grid, and the time series for power-grid faults have non-stationary characteristics. The fractal features of faults were studied by using DFA, MFDFA, DCCA, and TLDCCA methods.

1) The DFA method was used to study the time series for power grid faults. The time series had stable positive correlations over shorter time scales (<230 days for TSLF and <530 days for DSLF).

2) The power grid fault time series were analyzed using the MFDFA method and were found to have multiple fractal components. Further investigations revealed that the multifractal features of the fault time series of transmission lines in the Hunan power grid are due to long-range correlation of the grid and its fat-tailed distribution. The results showed that the fat-tailed distribution's contribution to multiscale behavior is greater than that of the long-range correlation to multiscale behavior. Given enough data over several years, it is possible to forecast how large the average 10- or 50-year faults will be.

3) The DCCA and TLDCCA methods were used to study the cross correlation between the time series for TSLF, DSLF and meteorological factors. The results shown that TSLF and maximum wind speed are strongly correlated. Unlike TSLF, the long-range correlations between DSLF and temperature, daily precipitation, and extreme wind speed are obvious. One interesting finding was that the cross-correlation exponents have periodic uncertainty changes with time lag. The affecting factors have a semi-annual impact on the Hunan power grid. With the trend of global climate change, the results of this study can help assess the risk of large-scale grid failures in the future.

ACKNOWLEDGMENT

The authors would like to thank International Science Editing (http://www.internationalscienceediting.com) for editing this manuscript.

AUTHOR CONTRIBUTIONS

Jiazheng Lu designed conception,
Bo Li and Tejun Zhou analyzed the data,
Yanjun Tan and Tejun Zhou wrote the manuscript.

The authors declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

REFERENCES

[1] D. Luo, Y. Xia, Y. Zeng, C. Li, B. Zhou, H. Yu, and Q. Wu, "Evaluation method of distribution network resilience focusing on critical loads," *IEEE Access*, vol. 6, pp. 61633–61639, 2018.

[2] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3747–3757, Sep. 2017.

[3] S. Mukherjee, R. Nateghi, and M. Hastak, "A multi-hazard approach to assess severe weather-induced major power outage risks in the U.S.," *Rel. Eng. Syst. Saf.*, vol. 175, pp. 283–305, Jul. 2018.

[4] P. Dehghanian, B. Zhang, T. Dokic, and M. Kezunovic, "Predictive risk analytics for weather-resilient operation of electric power systems," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 3–15, Jan. 2019.

[5] Y. Yang, W. Tang, Y. Liu, Y. Xin, and Q. Wu, "Quantitative resilience assessment for power transmission systems under typhoon weather," *IEEE Access*, vol. 6, pp. 40747–40756, 2018.

[6] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, "Predicting hurricane power outages to support storm response planning," *IEEE Access*, vol. 2, pp. 1364–1373, 2014.

[7] M. A. Mohamed, T. Chen, W. Su, and T. Jin, "Proactive resilience of power systems against natural disasters: A literature review," *IEEE Access*, vol. 7, pp. 163778–163795, 2019.

[8] Y. Di, J. Lu, X. Xu, T. Feng, and L. Li, "A response characteristics study of widespread power grid icing to el nino," *Math. Problems Eng.*, vol. 2019, pp. 1–7, Mar. 2019.

[9] J. Lu, J. Guo, Z. Jian, and X. Xu, "Research and application of efficient risk analysis method for electric power grid multiple faults under widespread wildfire disasters," *Int. Trans. Electr. Energy Syst.*, vol. 29, no. 9, p. e12055, Sep. 2019.

[10] Q. Chen, X. Yin, D. You, H. Hou, G. Tong, B. Wang, and H. Liu, "Review on blackout process in China southern area main power grid in 2008 snow disaster," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2009, pp. 1–8.

[11] D. Xianzhong and S. Sheng, "Self-organized criticality in time series of power systems fault, its mechanism, and potential application," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1857–1864, Nov. 2010.

[12] I. Dobson, "Where is the edge for cascading failure?: Challenges and opportunities for quantifying blackout risk," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2007, pp. 1–8.

[13] D. E. Newman, B. A. Carreras, V. E. Lynch, and I. Dobson, "Exploring complex systems aspects of blackout risk and mitigation," *IEEE Trans. Rel.*, vol. 60, no. 1, pp. 134–143, Mar. 2011.

[14] S. Mei, A. Xue, and X. Zhang, "On power system blackout modeling and analysis based on self-organized criticality," *Sci. China E, Technol. Sci.*, vol. 51, no. 2, pp. 209–219, Feb. 2008.

[15] H. Ren, I. Dobson, and B. A. Carreras, "Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1217–1225, Aug. 2008.

[16] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, "Evidence for self-organized criticality in a time series of electric power system blackouts," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1733–1740, Sep. 2004.

[17] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, "Initial evidence for self-organized criticality in electric power system blackouts," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2000, pp. 1–6.

[18] S. Su, Y. Li, and X. Duan, "Self-organized criticality of power system faults and its application in adaptation to extreme climate," *Chin. Sci. Bull.*, vol. 54, no. 7, pp. 1251–1259, Apr. 2009.

[19] L. Jiazheng, Z. Tejun, and W. Chuanping, "Fault statistics of 220 kV and above power transmission line in province-level power grid," (in Chinese), *High Voltage Eng.*, vol. 42, no. 1, pp. 200–207, 2016.

[20] S. Mei, Y. Ni, G. Wang, and S. Wu, "A study of self-organized criticality of power system under cascading failures based on AC-OPF with voltage stability margin," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1719–1726, Nov. 2008.

[21] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, "Evidence for self-organized criticality in electric power system blackouts," in *Proc. 34th Annu. Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, Jan. 2001, pp. 705–709.

[22] T. C. P. Bak and K. Wiesenfeld, "Self-organized criticality: An explanation of the $1/f$ noise," *Phys. Rev. Lett.*, vol. 59, pp. 381–384, Jul. 1987.

[23] M. S. Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar, and M. R. R. Tabar, "Multifractal detrended fluctuation analysis of sunspot time series," *J. Stat. Mech., Theory Exp.*, vol. 2006, no. 2, 2006, Art. no. P02003.

[24] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, "Effect of trends on detrended fluctuation analysis," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 1, Jun. 2001, Art. no. 011114.

[25] C. G. Tzanis, I. Koutsogiannis, K. Philippopoulos, and N. Kalamaras, "Multifractal detrended cross-correlation analysis of global methane and temperature," *Remote Sens.*, vol. 12, no. 3, p. 557, 2020.

[26] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, "Multifractal detrended fluctuation analysis of nonstationary time series," *Phys. A, Stat. Mech. Appl.*, vol. 316, nos. 1–4, pp. 87–114, Feb. 2002.

[27] Y. Wu, Y. He, M. Wu, C. Lu, S. Gao, and Y. Xu, "Multifractality and cross-correlation analysis of streamflow and sediment fluctuation at the apex of the pearl river delta," *Sci. Rep.*, vol. 8, no. 1, pp. 1–11, Dec. 2018.
[28] N. Kalamaras, K. Philippopoulos, D. Deligiorgi, C. G. Tzanis, and G. Karvounis, "Multifractal scaling properties of daily air temperature time series," *Chaos, Solitons Fractals*, vol. 98, pp. 38–43, May 2017.
[29] E. A. F. Ihlen, "Introduction to multifractal detrended fluctuation analysis in MATLAB," *Frontiers Physiol.*, vol. 3, p. 141, Jun. 2012.
[30] S. Drożdż and P. Oświęcimka, "Detecting and interpreting distortions in hierarchical organization of complex time series," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 91, no. 3, Mar. 2015, Art. no. 030902.
[31] R. J. Buonocore, T. Aste, and T. Di Matteo, "Measuring multiscaling in financial time-series," *Chaos, Solitons Fractals*, vol. 88, pp. 38–47, Jul. 2016.
[32] S. Drożdż, J. Kwapień, P. Oświecimka, and R. Rak, "Quantitative features of multifractal subtleties in time series," *Europhys. Lett.*, vol. 88, no. 6, p. 60003, Dec. 2009.
[33] B. Podobnik and H. E. Stanley, "Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series," *Phys. Rev. Lett.*, vol. 100, Feb. 2008, Art. no. 084102.
[34] L. Kristoufek, "Measuring correlations between non-stationary series with DCCA coefficient," *Phys. A, Stat. Mech. Appl.*, vol. 402, pp. 291–298, May 2014.
[35] J. Lu, T. Zhou, B. Li, and C. Wu, "Scale analysis and correlation study of wildfire and the meteorological factors that influence it," *Math. Problems Eng.*, vol. 2018, pp. 1–10, Jun. 2018.
[36] R. Billinton and G. Singh, "Application of adverse and extreme adverse weather: Modeling in transmission and distribution system reliability evaluation," *Proc. Inst. Elect. Eng., Gen. Trans., Distrib.*, vol. 153, no. 1, pp. 115–120, Jan. 2006.
[37] Å. J. Holmgren and S. Molin, "Using disturbance data to assess vulnerability of electric power delivery systems," *J. Infrastruct. Syst.*, vol. 12, no. 4, pp. 243–251, Dec. 2006.
[38] S. Sheng, D. Xianzhong, and W. L. Chan, "Probability distribution of fault in distribution system," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1521–1522, Aug. 2008.
[39] N. Xu, "The fractal and chaotic analysis of time series," Ph.D. dissertation, Beijing Jiao Tong Univ., Beijing, China, (in Chinese), 2010.
[40] C. Shen, "Analysis of detrended time-lagged cross-correlation between two nonstationary time series," *Phys. Lett. A*, vol. 379, no. 7, pp. 680–687, Mar. 2015.
[41] A. Lin, P. Shang, and X. Zhao, "The cross-correlations of stock markets based on DCCA and time-delay DCCA," *Nonlinear Dyn.*, vol. 67, no. 1, pp. 425–435, Jan. 2012.

TEJUN ZHOU received the master's degree from the Huazhong University of Science and Technology (HUST), Wuhan, China. He is currently pursuing the Ph.D. degree with the School of Electrical and Information Engineering, Changsha University of Science.

He is currently an Engineer at the State Key Laboratory of Disaster Prevention & Reduction for Power Grid Transmission and Distribution Equipment. His research interests are data mining and power system disaster prevention.

JIAZHENG LU received the Ph.D. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1995. He is currently the Director of the State Key Laboratory of Disaster Prevention & Reduction for Power Grid Transmission and Distribution Equipment. He is also a Professor with the College of Electrical and Electronics Engineering, HUST. His research interests include power system disaster prevention, high voltage technique, and power system analysis.

BO LI received the bachelor's degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1995. He is currently a Professor of engineering at the State Key Laboratory of Disaster Prevention & Reduction for Power Grid Transmission and Distribution Equipment. His research interest is disaster prevention and reduction for power grid.

YANJUN TAN (Member, IEEE) received the master's degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2008. He is currently a Professor of engineering at the State Key Laboratory of Disaster Prevention & Reduction for Power Grid Transmission and Distribution Equipment. His research interest is prevention and control of ice disaster in power grid.