Photovoltaic power impact analysis based on nonlinear correlation and time series model

Yang Zhen[1], Zhang Haiqing[2,*], Gong Cheng[1], Tang Xin[2], Yu Xi[3], Li Daiwei[2], Tang Dan[2]
[1.] Information Department, China Power Hua Chuang Electricity Technology Research Company LTD., Jiangsu 215125, Jiangsu, China
[2.] College of Software Engineering, Chengdu University of Information Technology, Chengdu 610225, China
[3.] College of Stirling, Chengdu University, Chengdu 610106, China
haiqing_zhang_zhq@163.com
*Corresponding author

2024 7th Asia Conference on Energy and Electrical Engineering (ACEEE) | 979-8-3503-5024-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ACEEE62329.2024.10652074

***Abstract*—The output power of photovoltaic power generation is a multi-variable and coupled nonlinear random process. Traditional correlation analysis methods are ineffective in detecting the nonlinear relationship between photovoltaic power generation and its related influencing factors. To solve this problem, a nonlinear correlation analysis algorithm is proposed based on shrinking and extending the time window of the Chatterjee correlation coefficient. The time window shrinking strategy utilizes time encapsulation windows to partition the entire time series into several subsequences. It then progressively validates the time window pairs within these subsequences. The time window extending strategy selects pairs of time series that meet nonlinear correlation relationships by traversing the minimum window and expanding to the left and right areas. Experimental results show that the proposed algorithm outperforms existing methods in terms of precision, recall, F-Score, and running time. At the same time, the analysis of Anhui Shijiahu photovoltaic power station shows that the proposed algorithm can analyze nonlinear relationships more accurately than existing methods.**

Keywords—Nonlinear correlation analysis; Time series; Photovoltaic power generation; Chatterjee correlation coefficient; Time window shrinking strategy; Time window extending strategy

I. INTRODUCTION

To enhance the reliability, safety, and stability of power grid operations, it is crucial to explore the factors that influence the prediction of photovoltaic power generation. Photovoltaic power generation is a multi-variable coupled nonlinear stochastic process, with main meteorological influencing factors including solar radiation, aerosols, sunshine duration, temperature, humidity, wind direction, wind speed, cloud cover, and so on. With the gradual expansion of the scale of photovoltaic power generation into the grid, the inherent intermittency and uncertainty of photovoltaic output has brought huge challenges to the reliable operation of the power grid. Therefore, domestic and foreign scholars have also carried out research on the meteorological factors related to photovoltaic power generation and analyze.

Yan Quanquan et al. [1] used the Pearson correlation coefficient to analyze the impact mechanism of typical meteorological factors on photovoltaic power generation in different seasons and the coupling relationship between multiple variables. Cao Yingli et al. [2] used the Pearson correlation analysis method to analyze the power generation of distributed photovoltaic power stations in Shenyang area and the daily solar radiation, sunshine hours, daily average ambient temperature, daily maximum temperature, daily minimum temperature, and daily temperature range. and the correlation between seven meteorological factors including daily average wind speed. Yisheng Cao et al. [3] used Spearman correlation coefficient to select an appropriate combination of meteorological features.

Although the above studies have drawn some useful conclusions, they have not taken into account the different regularities, periodicity and volatility of meteorological characteristics on time scales, resulting in fragmented selection of meteorological factors. Meanwhile, this paper analyzes the data of Photovoltaic Power Stations, it is found that there is a non-linear relationship between meteorological influencing factors and photovoltaic power output, so Pearson cannot be used to analyze correlation coefficient.

II. NONLINEAR CORRELATION ANALYSIS MEASUREMENT METHOD

A. Research status of nonlinear correlation measurement

The current classic correlation measurement methods include Pearson, Spearman and Kendall correlation coefficients. These correlation coefficients are very powerful for detecting linear and monotonic relationships. However, these three coefficients cannot effectively detect the relationship between nonlinear factors even if there is no noise at all. Therefore, in order to overcome the shortcomings of the classic correlation coefficient in nonlinear correlation detection, many research works have proposed the maximum correlation coefficient [4], correlation coefficient based on connection [5], and correlation coefficient based on pairwise distance [6]. However, the current correlation coefficient is designed based on the independence of the test and cannot well evaluate the strength of the correlation between variables [7]. Ideally, the correlation coefficient approaches its maximum value if and only if one variable looks more and more like a noiseless function of the other variable. Although the maximum correlation coefficient and especially the maximum information coefficient [8] have been widely used in nonlinear correlation analysis, MCC and MIC have been proven to draw wrong conclusions about non-linear correlations.

Therefore, in order to solve the above problems, Sourav Chatterjee [7] proposed the Chatterjee correlation coefficient (CCC). CCC is a simple and interpretable correlation coefficient measure that can always accurately estimate the degree of dependence between variables. The CCC measure is 0 when and only when the variables are independent. In addition, the CCC measure is 1 when and only when one variable is dependent of another variable. Meanwhile, under the assumption of independence, the CCC coefficient has a simple asymptotic theory and is easy to calculate. Based on the above advantages of CCC coefficient, CCC coefficient is widely used in nonlinear correlation detection.

B. Chatterjee nonlinear correlation coefficient

Suppose (X, Y) is a pair of random variables and Y is not a constant. Suppose $(X_1, Y_1), \ldots, (X_n, Y_n)$ are independently and identically distributed, where $n \geq 2$, and X_i and Y_i do not have any equal values. Rearrange the data into $(X_{(1)}, Y_{(1)}) \ldots, (X_{(n)}, Y_{(n)})$, such that $X_{(1)} \leq \cdots \leq X_{(n)}$. Suppose r_i is the level of $Y_{(i)}$, that is, j is the number of $Y_{(j)} \leq Y_{(i)}$. Assuming that there is no autocorrelation between X_i, the CCC correlation coefficient [10] is defined as follows:

$$CCC(X,Y) := 1 - \frac{3\sum_{i=1}^{n-1}|r_{i+1} - r_i|}{n^2 - 1} \tag{1}$$

CCC(X,Y) is a consistent estimator of some correlation measure between random variables X and Y.

Assuming that there is autocorrelation between X_i, and the increasing sequence is selected by randomly and uniformly breaking the autocorrelation connection. Assume r_i is the level of $Y_{(i)}$, and define $l_{(i)}$ as the number of $Y_{(j)} \geq Y_{(i)}$, then the CCC correlation coefficient [10] is defined as follows:

$$CCC(X,Y) := 1 - \frac{n\sum_{i=1}^{n-1}|r_{i+1} - r_i|}{2\sum_{i=1}^{n} l_i(n - l_i)} \tag{2}$$

III. Nonlinear correlation analysis algorithm based on time window reduction and expansion

Photovoltaic power generation is a multi-variable coupled nonlinear random process. Two time series may be correlated over certain time intervals, but not throughout the entire time period. Therefore, when analyzing time series data, an important task is to evaluate the correlation between time series. Shuye Pan et al.[9] proposed a nonlinear correlation algorithm based on time window reduction and expansion. In this paper, the strategy based on time window reduction and expansion proposed by Shuye Pan et al. [9] is used to characterize the relationship between the full-field output power of photovoltaic power generation and related meteorological factors. Different from Shuye Pan et al.[9], in order to more accurately characterize the nonlinear correlation between factors, this work uses the CCC instead of the MIC.

A. Related parameter definitions

Definition 1 (Time Window) The time window $W=(s, l)$ is a segment over the entire continuous time period ($1 \leq s \leq n-l+1$). The mapping of time series $X=\{x_1, x_2, \ldots, x_n\}$ on the time window W is recorded as $X_W=\{x_s, \ldots, x_{s+l-1}\}$, or as X(*s*, *l*).

Definition 2 (Time Series Pair) Time window pair (X, Y)=($\{x_1, x_2, \ldots, x_n\}$,$\{y_1, y_2, \ldots, y_n\}$) is a time series within the same observation time period. The length of the time pair is: $n=|(X, Y)|$.

Definition 3 (Time Window Pair) Assume that two time windows, X(*s*, *l*) and Y(*s'*, *l*), can form a time window pair (*X*(*s*, *l*), *Y*(*s'*, *l*)). The start times of the time window pairs are *s* and *s'* respectively, and the delay of the time relationship is defined as: τ=s'-s. A time window pair can be represented as a triplet $< s, l, \tau >$.

Definition 4 (Correlation of Time Window Pairs) For any set of time window pairs, X(s, l) and Y(s', l), if there is $CCC(X(s,l), Y(s',l)) \geq \theta$ (where, θ is the minimum threshold of nonlinear correlation), then X(s, l) and Y(s', l) are nonlinearly related. At the same time, it is defined that the nonlinear correlation pair is composed of the triplet CP=<*s*, *l*, *τ*>, where τ=s'-s is the time delay of the nonlinear correlation relationship. If $s_i+l_i \leqslant s_j$ and $s_i+\tau_i+l_i \leqslant s_j+\tau_i$ is exist, then the two related pairs CP_i=<s_i, l_i, τ_i> and CP_j=<s_j, l_j, τ_j> (assuming $s_i < s_j$) are not connected. Otherwise, there will be overlap in related pairs.

Definition 5 (Significance of Time-correlated Pairs) Nonlinear correlation is significant to CP_i, if there is a relationship CCC(CP_i) ⩾ CCC(CP_j) for any CP_j that has a connection intersection with CP_j.

B. Finding Strategies for Nonlinear Correlation Pairs

1) Nonlinear correlation analysis algorithm based on time window reduction

The strategy based on time window reduction is mainly to deal with the nonlinear correlation relationship of sparse distribution. For the pseudo code of the algorithm, see Algorithm 1: Nonlinear correlation search algorithm based on time window reduction strategy and CCC nonlinear correlation coefficient (referred to as CCC_NLCS).

TABLE I. CCC_NLCS Algorithm

Algorithm 1. CCC_NLCS Algorithm
Input: Time Series Pair of (*X*, *Y*)
Output: CP$_0$ candidate item set
1. **while** (*X*, *Y*) exist untraversed time pairs **do**
2. Initialize $CP \leftarrow < sp_i, w_e, 0 >$, CCC$_{best}$=0;
3. **for** $\tau \in [\tau_{\min}, \tau_{\max}]$ **do**
4. $EW \leftarrow (X_i, Y(sp_i + \tau, w_e))$
5. **if** CCC(EW)＞CCC$_{best}$ **then**
6. CCC$_{best}$=CCC(EW),
Update $CP \leftarrow < sp_i, w_e, \tau >$
7. **if** CCC$_{best}$＞θ **then**
8. CP.trim(miniL, θ)
9. Add CP to candidate item set CP$_0$

The purpose of Algorithm 1 is to find relevant time window pairs in the time series interval [l_{min}, l_{max}]. Set a larger time area range we to filter out candidate item sets that do not meet the requirements in the time series interval [l_{min}, l_{max}]. w_e is a time encapsulation window. The time window w_e is used to split the time series *X* into several subsequences, recorded as: $(X_1, X_2, X_i \cdots, X_{n/w_e})$, where $X_i = X(sp_i, w_e)$, $sp_i = (i-1) \bullet w_e + 1$.

First, the algorithm assumes that the time series pair (X, Y) has at least one nonlinearly related point pair in the entire time series. Algorithm 1 (step 2) initializes a time window pair <sp_i, w_e, 0> to be examined and initializes the nonlinear coefficient CCC to 0. Then, Algorithm 1 (steps 3-6) gradually traverses each time window pair divided by time window w_e. Calculating the nonlinear correlation coefficient $CCC(X_i, Y(sp_i+\tau, w_e))$ between the time window pairs X_i and $Y(sp_i+\tau, w_e)$ respectively. If all coefficient values in this time segment are less than the threshold θ, then there are no nonlinear correlation pairs in this area. Then the time series slides down to the next candidate time series X_{i+1} and time window w_e. Otherwise, if it exists $CCC(X_i, Y(sp_i+\tau, w_e)) \geq \theta$, selecting the point with the maximum correlation coefficient as the candidate $CP(sp_i, w_e, \tau)$, which means that on the time series X_i, there is a time window pair such that Xi and $Y(sp_i+\tau, w_e)$ is nonlinearly correlated with a delay of time τ.

2) Nonlinear correlation analysis algorithm based on time window expansion

When the distribution of nonlinear correlation pairs is relatively dense, if the large window indentation method is still used, then two different correlation pairs that meet the conditions may appear in an envelope window w_e. According to Algorithm 1, since only one correlation pair is retained in an envelope window w_e, some correlation pairs will be ignored. The window expansion strategy is the opposite of the window reduction strategy, which use a small window to find the candidate item set. The pseudo code of the algorithm is shown in Algorithm 2: Nonlinear correlation search algorithm based on time window expansion strategy and CCC nonlinear correlation coefficient (CCC_NLCE for short).

TABLE II. CCC_NLCE ALGORITHM

Algorithm2. CCC_NLCE Algorithm
Input：Time Series Pair of (*X*, *Y*)
Output：CP_0 candidate item set
1. **while** (*X*, *Y*) exist untraversed time pairs **do**
2. Initialize $CP \leftarrow < sp_i, L_{min}, 0 >$， CCC_{best}=0;
3. Adopt the steps 3-6 to get CCC_{best}，UpdateCP。
4. if $CCC_{best} \geqslant \theta$ **then**
5. $CP_1 \leftarrow < sp, L_{\min} + L_{\max}, \tau >$,
6. $CP_2 \leftarrow < sp - L_{\max}, L_{\min} + L_{\max}, \tau >$,
7. $CP_1.trim(\min iL, \theta)$
8. $CP_2.trim(\min iL, \theta)$
9. **if** $CCC(CP_1) \geqslant CCC(CP_2)$ **then**
10. add CP1 to CP0，update $sp \leftarrow sp + \lvert CP_1 \rvert$,
11. else
12. add CP2 to CP0，update $sp \leftarrow sp + \lvert CP_2 \rvert$
13. else
14. $sp \leftarrow sp + L_{min}$

First, Algorithm 2 uses the minimum window as L_{min} to expand the traversal of the time series X (step 2), and note that the current time window is $X(sp, L_{min})$. Calculate the correlation coefficient between the time series sequence Time window $X(sp+L_{min}, L_{min})$ (step 14 of the algorithm). Otherwise, record the time window pair $<sp, L_{min}, \tau>$ when the CCC value is maximum.

Different from the time window reduction strategy, the relevant pairs currently filtered out by the time window expansion strategy may be a segment of the entire time series. So, at this point, in order to find the complete set of candidate items, the algorithm expands the time window from both ends. For the time series pair $X(sp, L_{min})$ and $Y(sp+\tau, L_{min})$, the window expansion method is as follows:

$$CP_1 =< sp, L_{\min} + L_{\max}, \tau > \quad (3)$$

$$CP_2 =< sp - L_{\max}, L_{\min} + L_{\max}, \tau > \quad (4)$$

Then, in steps 7 and 8, in order to obtain a more optimized set of candidate items, an optimization strategy is used to optimize CP_1 and CP_2. Compare the values of the nonlinear correlation coefficients generated by CP_1 and CP_2 respectively, and add the window pair that generates a larger value to the candidate item set CP_0 (steps 10 and 12).

IV. PERFORMANCE ANALYSIS OF NONLINEAR CORRELATION ALGORITHM

First, the nonlinear correlation data set synthesized by Shuye Pan et al. [9] is used to verify the effectiveness of the algorithm. The synthetic correlation relationship consists of nonlinear relationships such as logarithm, exponential, square, fourth power, etc. The way to form a nonlinear correlation pair is: 1). Randomly select a time starting point and a custom nonlinear relationship type; 2). The nonlinear correlation pair CP (set parameter value range is $l \in [400,850] \wedge \tau \in [-150,150]$) is generated based on time Point t. This data set gives 40 pairs of nonlinear correlation windows.

The algorithm uses precision, recall and F-score to verify the effectiveness of the algorithm. Based on the results of the dataset, all-time series pairs are classified into two types: correlated and uncorrelated. There are four situations for each pair of time series: true positive (the time series pair is calculated as correlated and marked as related in the nonlinear relationship, denoted as TP), false positive (the time series pair is calculated as correlated and is marked as uncorrelated in the nonlinear relationship, marked as FP), False Negative (time series pair calculated as uncorrelated and marked as correlated in a nonlinear relationship, marked as FN), True Negative (time series pair calculated as uncorrelated and Marked as irrelevant in a nonlinear relationship , recorded as TN).

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F\text{-}score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

For the analysis of this data set, the experimental results are shown in Table III. From the analysis of algorithm performance, the performance of the time window reduction and expansion algorithm improved by using CCC nonlinear correlation coefficient is higher than that of Shuye Pan et al. [9]. In particular, CCC_NLCE achieved the highest performance in terms of precision, F-score and time performance, and CCC_NLCS achieved the best performance in recall rate.

TABLE III. ALGORITHM PERFORMANCE ANALYSIS OF IMPROVED TIME WINDOW SHRINKING AND EXPANDING ALGORITHM

	Precision	Recall	F-score	Running Time
MI_NLC$_S$	0.9849	0.8444	0.9092	138.968
CCC_NLC$_S$	**0.9929**	***0.8532***	**0.9178**	**112.754**
MI_NLC$_E$	0.9362	0.7642	0.8415	102.283
CCC_NLC$_E$	***0.9994***	**0.8504**	***0.9189***	***86.652***

In order to further analyze the correlation between time series pairs obtained by the improved time window reduction and expansion algorithm, this study has conducted the normalized operation of the correlation coefficients in terms of MI_NLCS [9] (the correlation coefficient obtained by the time window reduction algorithm based on MI coefficient), MI_NLCE [9] (the correlation coefficient obtained by the time window expansion algorithm based on the MI coefficient), CCC_NLCS, and CCC_NLCE. The results are shown in Figures 1 and 2. Figures 1 and 2 show the first 36 points extracted by all algorithms. It can be seen from the graph that the values of MI_NLCS, MI_NLCE, CCC_NLCS, CCC_NLCE and the given correlation coefficient are relatively close. As can be seen from the points in Figure 1 and Figure 2, MI_NLCS, MI_NLCE, CCC_NLCS, CCC_NLCE and the given correlation coefficient are all in the same direction. This also explains why the performance Precision, Recall, and F-score of all algorithms in Table III are above 80%.

V. ANALYSIS OF INFLUENCING FACTORS OF PHOTOVOLTAIC POWER GENERATION BASED ON NONLINEAR CORRELATION ALGORITHM

In order to verify the effectiveness of the proposed algorithm, the algorithms were used to analyze influencing factors of Anhui Shijiahu Photovoltaic Power Station.

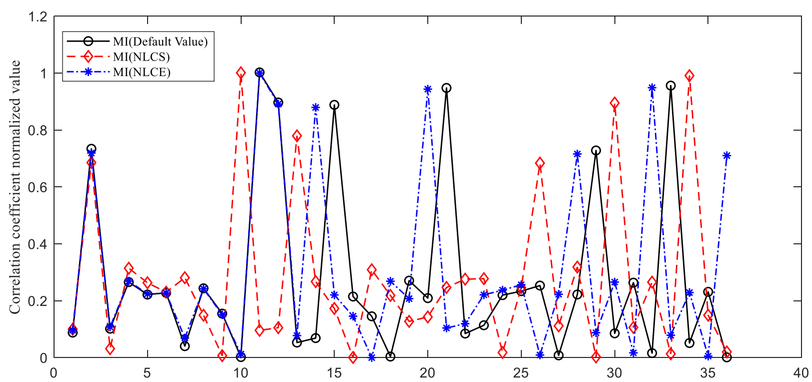

Fig. 1. Correlation Analysis of Time Series Pairs Based on MI Coefficient

The data analyzed by Anhui Shijiahu Photovoltaic Power Station is 242,784 pieces of data from 2021/01/01 to 2023/04/23. MI_NLCS, MI_NLCE, CCC_NLCS, and CCC_NLCE are used to evaluate the nonlinear correlation between various influencing factors and photovoltaic power generation. The results are shown in Figure 3. Observing the conclusions drawn from MI_NLCS, MI_NLCE, CCC_NLCS and CCC_NLCE, it can be concluded that there is a strong correlation between the ‘site irradiation intensity’, ‘tilt irradiation intensity’, ‘horizontal scattered radiation intensity’ and the full field power of photovoltaic power generation. The difference is that CCC_NLCS and CCC_NLCE believe that 'site temperature', 'ambient temperature' and 'ambient humidity' have a weak influence relationship. Meanwhile, ‘wind direction’ and ‘wind speed’ have the weakest influence. However, MI_NLCS and MI_NLCE believe that 'site temperature', 'ambient temperature' and 'ambient humidity' are very weak. Meanwhile, it is believed that there is no relationship between 'wind direction' and 'wind speed' with the overall power. In order to more accurately evaluate the correctness of the conclusions obtained, the author of uses a scatter plot to plot the relationship between all influencing factors and the full-field power. It can be seen that 'station temperature', 'ambient temperature' and 'ambient humidity' do have a certain correlation with the whole-field power, while the relationship between 'wind direction' and 'wind speed' and the whole-field power is indeed weak. Therefore, the conclusions drawn by CCC_NLCS and CCC_NLCE are more credible.

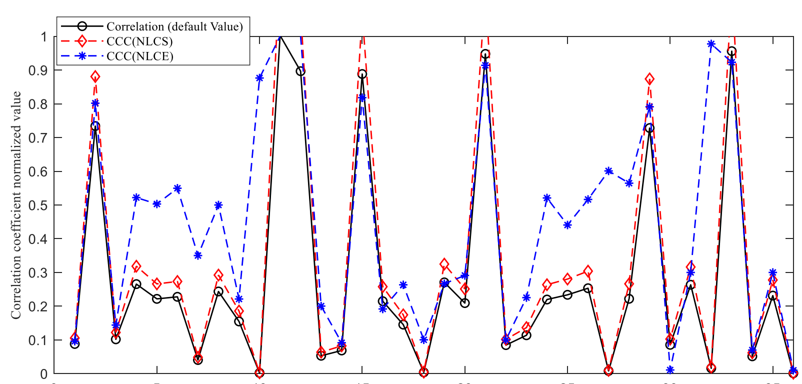

Fig. 2. Correlation Analysis of Time Series Pairs Based on CCC Coefficient

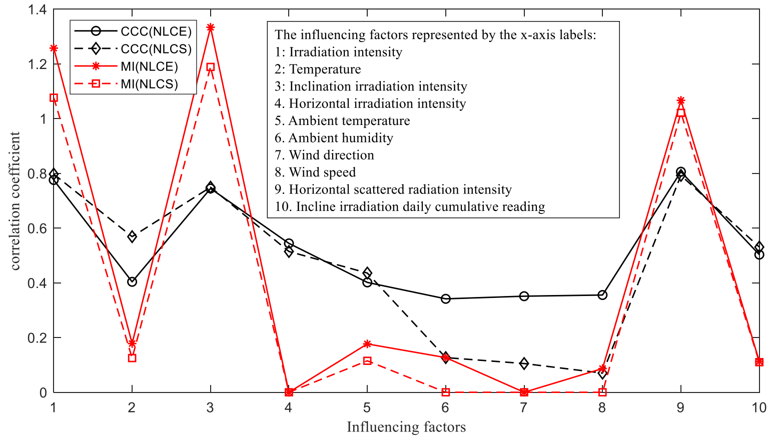

Fig. 3. Analysis of influencing factors of photovoltaic power generation (Shijiahu Photovoltaic Power Station, Anhui)

VI. Conclusion

Photovoltaic power generation is currently the fastest growing and most potential renewable energy source. In order to improve the reliability, security and stability of power grid operation, it is crucial to study the related influencing factors of photovoltaic power generation. Traditional correlation coefficients all describe linear relationships and cannot perform a good analysis of the nonlinear influencing factors. This study proposes a nonlinear correlation analysis algorithm.

The time window reduction strategy algorithm (CCC_NLCS) divides the entire time series pair into several subsequences through the preset time encapsulation window, and gradually traverses whether the time window pairs in the subsequences satisfy the nonlinear correlation relationship, thereby achieving extraction based on Candidate non-linear correlation of time series for the purpose of sequence pairs. The time window expansion strategy algorithm (CCC_NLCE) traverses the set minimum window and expands to the left and right areas of the time series to select time series pairs that satisfy the nonlinear correlation relationship.

Experimental results show that the proposed CCC_NLCS and CCC_NLCE algorithms have significant improvements in precision, recall, F-Score, running time and other indicators compared with existing algorithms. In order to further verify the effectiveness of the proposed algorithm, nonlinear analysis was conducted on the photovoltaic power-related influencing factors of Anhui Shijiahu Photovoltaic Power Station. It can be concluded that the proposed nonlinear correlation algorithm is better than the existing method. In future research work, more photovoltaic power plant data will be used to verify the effectiveness of the algorithm.

Acknowledgment

This research is supported by the Sichuan Province Science and Technology Program (No. 2022ZDZX0001, No. 2021YFH0107, No. 2022NSFSC0571, No. 2022YFS0544, No.2023NSFSC0778), and supported by Science and Technology Innovation Capability Enhancement Program of Chengdu University of Information Technology (No.KYQN202223).

References

[1] Y. Quanquan, L. I. Fen, W. Lijuan, S. Jing, Q. Jialin, and Z. Jinbin, "Diagnostic Analysis of the Impact of Meteorological Factors on Photovoltaic Power Generation in Different Seasons," *Journal of Shanghai University of Electric Power*, Accessed: May 22, 2024. [Online]. Available: http://en.cnki.com.cn/Article_en/CJFDTOTAL-DYXY201603012.htm

[2] C. Ying-Li, F. Shi-Qi, W. Yang, Y. U. Bing-Xin, Z. Huan-Cheng, and X. U. Tong-Yu, "Diagnostic Analysis of Distributed Photovoltaic Power Characteristics and the Impact of Meteorological Factors," *Journal of Shenyang Agricultural University*, 2018, Accessed: May 22, 2024. [Online]. Available: http://en.cnki.com.cn/Article_en/CJFDTotal-SYNY201803018.htm

[3] Y. Cao, G. Liu, D. Luo, D. P. Bavirisetti, and G. Xiao, "Multi-timescale photovoltaic power forecasting using an improved Stacking ensemble algorithm based LSTM-Informer model," *Energy*, vol. 283, p. 128669, 2023.

[4] J. Yang *et al.*, "Understanding land surface temperature impact factors based on local climate zones," *Sustainable Cities and Society*, vol. 69, p. 102818, 2021.

[5] K. Zhang, "BET on Independence," *Journal of the American Statistical Association*, vol. 114, no. 528, pp. 1620–1637, Oct. 2019, doi: 10.1080/01621459.2018.1537921.

[6] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," 2007, Accessed: Jan. 22, 2024. [Online]. Available: https://projecteuclid.org/journals/annals-of-statistics/volume-35/issue-6/Measuring-and-testing-dependence-by-correlation-of-distances/10.1214/009053607000000505.short

[7] S. Chatterjee, "A New Coefficient of Correlation," *Journal of the American Statistical Association*, vol. 116, no. 536, pp. 2009–2022, Oct. 2021, doi: 10.1080/01621459.2020.1758115.

[8] D. N. Reshef *et al.*, "Detecting Novel Associations in Large Data Sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, Dec. 2011, doi: 10.1126/science.1205438.

[9] S. Pan, P. Wang, C. Wang, W. Wang, and J. Wang, "NLC: search correlated window pairs on long time series," *Proc. VLDB Endow.*, vol. 15, no. 7, pp. 1363–1375, Mar. 2022, doi: 10.14778/3523210.3523215.