IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 July 2024, accepted 2 August 2024, date of publication 7 August 2024, date of current version 16 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3439693

RESEARCH ARTICLE

Modeling Multi-Temporal Correlation of PV Stations Output via Improved Nonparametric Disaggregation Method in Power System Reliability Evaluation

FEI FAN[1], XIANGJIE ZHOU[1], LINGZHI ZHANG[1], ZHENJIANG HE[2], WEIXIANG XU[2], AND YUAN ZHAO[3], (Member, IEEE)

[1]Hunan Vocational College of Railway Technology, Zhuzhou 412006, China
[2]State Grid Zhuzhou Power Supply Company, Zhuzhou 412000, China
[3]State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400030, China

Corresponding author: Xiangjie Zhou (xiangjiez2024@163.com)

This work was supported in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 22C1117, in part by the Science and Technology Innovation Program of Hunan Province under Grant 2022RC1097, and in part by Hunan Provincial Natural Science Foundation of China under Grant 2023JJ50209 and Grant 2022JJ60075.

ABSTRACT The accurate modeling of multi-temporal correlation of photovoltaic stations output is important to achieve the precise power system reliability. However, the existing studies is mainly focused on the hourly temporal correlation between adjacent interval outputs of photovoltaic station. In this paper, a novel multi-temporal correlation modeling method for photovoltaic station output is proposed by using an improved nonparametric disaggregation (INPD) technique, and the multi-temporal characteristic representing the interactive correlation between the total daily output and hourly temporal outputs are deeply explored. Moreover, a three-stage sampling method for the INPD based power system reliability evaluation is further presented considering the multi-temporal correlation of photovoltaic stations output, which includes 1) the random sampling of daily photovoltaic station output; 2) the random disaggregation of daily output to the hourly temporal outputs and 3) the random sampling of power device. Finally, several numerical tests are conducted for a modified IEEE-RTS79 to validate the proposed method.

INDEX TERMS PV stations, improved nonparametric disaggregation, multi-temporal correlation, reliability assessment.

I. INTRODUCTION

Renewable energy sources has been increasingly developed in the past few decades across the world [1], [2]. Among all possible renewable energy sources, photovoltaic (PV) generation [3] is considered the most promising resource for power system despite its randomness, which is addressed by using energy storages [4]. Given that the random fluctuations of PV power has brought about an increasing challenge to the power system operational reliability, a rational and precise probability model for multi-temporal correlation of solar irradiance is crucial to achieve the accurate reliability evaluation of power system incorporating photovoltaic stations [5].

The solar irradiance mainly determined by the earth's rotation and climate factors, are not only random but also correlated at different moments. In the literature, a lot of probabilistic model have been developed to account for the random and correlated attributes of solar irradiance, and they are approximately classified into two categories. The first category is concentrated on the uncertainty modeling, and various probabilistic models for solar irradiance like the parameter fitting methods based on normal, lognormal, Weibull, extreme value (Type I) and Beta distributions [6], [7], [8], have been developed. The validity of such models

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

is often checked via goodness-of-fit, but unfortunately, there does not exist a universal parameter distribution which is effective for any application scenarios. To address the demerit of parametric distribution, a data-driven nonparametric kernel density estimation technique is employed to establish a PV probabilistic model without relying on any prior knowledge [9]. With the development of artificial intelligence(AI) algorithm, a novel convolution neural network framework [10] has been presented to model the uncertainty of the solar irradiance, and a chaos GA/PSO hybrid method was applied to demonstrate the superiority through all the simulation test results. The above models are useful but not applicable for the time-dependent probabilistic simulation analysis of power system, because only the uncertainty of solar irradiance at a certain moment or during a time interval (like a day, a month or a year) are considered. To conduct the time-dependent probabilistic simulation analysis, the temporal correlation besides the uncertainty are needed to be considered. Therefore, the second category is focused on the time correlation modeling. The auto-regressive moving average (ARMA) model of hourly solar irradiance is developed, which combines a deterministic model of solar radiation with stochastic simulations [11]. In addition, some studies have tried to combine the Beta distributions of hourly solar irradiance or hourly diffuse fraction k_h [12], [13], but fail to depict the time-series correlation of solar irradiance. The temporal variation rule of solar irradiance is described as a time function and a temporal dynamic probability distribution model of solar irradiance is built by superimposing the corresponding stochastic fluctuations [14], however a restriction of this method is that a fixed correlation matrix is used. To overcome this demerit, a conditional probability method was proposed to capture the randomness, chronology and the correlations among PV power outputs [15]. This improved method is successful for the random variables with temporal correlation, but ignores the summation constraint between the total daily solar irradiance and the hourly solar irradiance. Besides, the time series analysis for solar irradiance based on machine learning methods are studied, but its accuracy may be difficult to meet the requirement of solar irradiance prediction [16].

The summation correlation reflects the inherent relationship between the total solar irradiance during a time interval and the respective solar irradiance at each moment of this time interval, e.g., the relationship between the annual irradiance and monthly irradiance, the monthly irradiance and daily irradiance, or the daily irradiance and hourly irradiance. The semi-sine model that distributes the daily irradiance hour by hour is proposed in [17] to address the summation correlation, but cannot accurately reflect the stochastic fluctuation law. As similar as the semi-sine method, the disaggregation theory has been adopted to develop a parametric or nonparametric probability model in many fields, including hydrology stochastic simulation [18], [19], load modeling [20], and wind speed modeling [21]. Besides, a parametric disaggregation model is developed to model the summation relationship between the total daily irradiance to hourly irradiance [22], but encountering the demerit of parametric estimation. The taxonomy for similar studies on solar irradiance probabilistic modeling are shown in Table 1.

TABLE 1. Taxonomy table reviewing recent advances in probabilistic modeling of irradiance.

Taxonomy factors	Probabilistic model structure		
	Parametric	Nonparametric	AI algorithm
Uncertainty	[6], [7], [8]	[9]	[10]
Temporal correlation	[11],[12],[13],[14]	[15]	[16]
Summation correlation	[17], [22]	[18],[19],[20],[21]	-

This paper proposes a novel multi-temporal correlation modeling approach based on an improved nonparametric disaggregation (INPD) method for power system reliability assessment. This proposed method can capture the multi-temporal correlation between the total daily irradiance and hourly irradiance, and the generated chronological curve of PV station outputs is less biased from the actual curves. The main contributions of this paper are listed below,

1) A novel multi-temporal correlation analysis method concerning photovoltaic station output within an INPD framework is proposed. This approach allows for the identification of correlations between the total daily irradiance and hourly prediction values.

2) A refined three-stage sampling method for electrical power system reliability assessment considering the multi-temporal correlation of photovoltaic stations is developed.

The rest of this paper is organized as follows. The probability model of daily irradiance is proposed in Section II, and the INPD model for hourly irradiance is established in Section III. The three-stage stochastic sampling technique of hourly irradiance time series is presented in Section IV. The test results and practicality analyses are shown in Section V, followed by conclusions in Section VI.

II. CKDE OF DAILY SOLAR IRRADIANCE

Let I_t represents the solar irradiance on the t-th day, measured in W/m^2. Daily irradiance is random because of uncertain environmental factors like cloud changes. Besides, daily irradiance between two adjacent days is related to each other. In other words, the daily irradiance at the prior day affects that at the following day. To consider the randomness and correlation of daily irradiance simultaneously, this paper constructs the conditional probability density function of daily irradiance according to the conditional probability theory, which can be built as:

$$f(I_t\,|I_{t-1}) = \frac{f(I_t, I_{t-1})}{f(I_{t-1})} = \frac{f(\boldsymbol{I}_t)}{f(I_{t-1})} \tag{1}$$

where $f(I_t, I_{t-1})$ is the joint probability density function of I_t and I_{t-1}, and $f(I_{t-1})$ is the marginal probability density of I_{t-1}. Let vector $\boldsymbol{I}_t = [I_t, I_{t-1}]^{\mathrm{T}}$ denote the adjacent daily irradiance and n day samples are used, then $\boldsymbol{I}_{ti} = [I_{ti}, I_{ti-1}]^{\mathrm{T}}$ is the *i-th* sample of $\boldsymbol{I}_t$. Moreover, the sample sizes of I_{t-1} and $\boldsymbol{I}_t$ are n-1, respectively.

The bivariate kernel density estimation and the univariate kernel density estimation are used to evaluate $f(I_t, I_{t-1})$ and $f(I_{t-1})$, respectively. The Gaussian function is selected as the kernel function in this paper. In addition, let $\widehat{f}(I_{t-1})$ and $\widehat{f}(\boldsymbol{I}_t)$ denote the kernel estimation of $f(I_{t-1})$ and $f(\boldsymbol{I}_t)$, respectively. They are estimated as [23]:

$$\widehat{f}(I_{t-1}) = \frac{1}{(n-1)} \sum_{i=1}^{n-1} \frac{1}{(2\pi h_{t-1})^{1/2}} \cdot \exp\left(-\frac{(I_{t-1} - I_{ti-1})^2}{2h_{t-1}} \right) \quad (2)$$

$$\widehat{f}(\boldsymbol{I}_t) = \frac{1}{(n-1)} \sum_{i=1}^{n-1} \frac{1}{(2\pi)^{d_t/2} \det(\boldsymbol{H}_t)^{1/2}} \times \exp\left(-\frac{(\boldsymbol{I}_t - \boldsymbol{I}_{ti})^T \boldsymbol{H}_t^{-1} (\boldsymbol{I}_t - \boldsymbol{I}_{ti})}{2} \right) \quad (3)$$

where $d_t = 2$ denotes the dimension of vector $\boldsymbol{I}_t$, $det(\cdot)$ is the determinant operator, h_{t-1} is the bandwidth of the univariate kernel density estimation $\widehat{f}(I_{t-1})$, and $\boldsymbol{H}_t$ is a symmetric positive definite $d_t \times d_t$ bandwidth matrix of the bivariate kernel density estimation $\widehat{f}(\boldsymbol{I}_t)$.

According to the conditional kernel density estimation concept [24], formula (1) can be expressed as shown in (4):

$$\widehat{f}(I_t \,|I_{t-1}) = \frac{\widehat{f}(I_t, I_{t-1})}{\widehat{f}(I_{t-1})} = \frac{\widehat{f}(\boldsymbol{I}_t)}{\widehat{f}(I_{t-1})} \quad (4)$$

The expression for the conditional kernel density estimate of the daily irradiance can be derived by substituting (2) and (3) into (4), whose computational complexity lies in finding the bandwidth matrix $\boldsymbol{H}_t$, while the bandwidth h_{t-1} is only a special case of this matrix solution. Therefore, the bandwidth matrix $\boldsymbol{H}_t$ is the key to an accurate estimation of the kernel density estimation. Several optimization methods for choosing optimal values of the bandwidth matrix are available in the literature. This paper uses the asymptotic mean integrated squared error (AMISE) between estimated function $\widehat{f}(\boldsymbol{I}_t)$ and real function $f(\boldsymbol{I}_t)$, the AMISE optimal bandwidth matrix being obtained by a minimization of the following function:

$$\min AMISE = \frac{1}{n-1}(4\pi)^{-d_t/2} \det(\boldsymbol{H}_t)^{-1/2} + \frac{1}{4} d_t^2 \cdot \int tr^2 \left\{ \boldsymbol{H}_t f''(\boldsymbol{I}_t) \right\} d\boldsymbol{I}_t \quad (5)$$

where $tr\{\cdot\}$ is the trace operator, and $f''(\boldsymbol{I}_t)$ is the second-order partial derivatives of the unknown function $f(\boldsymbol{I}_t)$. If the bandwidth matrix is rewritten as [25]:

$$\boldsymbol{H}_t = \lambda_t^2 \boldsymbol{S}_t \quad (6)$$

Then (5) is minimized at optimal scale parameter λ_t, obtained by [25]:

$$\lambda_t = \left[(n-1) d_t (4\pi)^{d_t/2} R(f) \right]^{-1/(d_t+4)} \quad (7)$$

where $R(f) = \int tr^2 \left\{ \boldsymbol{S}_t \cdot f''(\boldsymbol{I}_t) \right\} d\boldsymbol{I}_t$, and $\boldsymbol{S}_t$ is the sample covariance matrix of the variable set $\boldsymbol{I}_t$, defined as:

$$\boldsymbol{S}_t = \begin{bmatrix} S_h & S_{hq} \\ S_{hq} & S_q \end{bmatrix} \quad (8)$$

Here S_h is the sample variance of I_t, S_q is the sample varian-ce of I_{t-1}, and S_{hq} is the sample covariance of I_t and I_{t-1}. Substituting (6) and (7) into (3), we have:

$$\widehat{f}(\boldsymbol{I}_t) = \frac{1}{(n-1)} \sum_{i=1}^{n-1} \frac{1}{(2\pi)^{d_t/2} \lambda_t^{d_t} \det(\boldsymbol{S}_t)^{1/2}} \cdot \exp\left(-\frac{(\boldsymbol{I}_t - \boldsymbol{I}_{ti})^T \boldsymbol{S}_t^{-1} (\boldsymbol{I}_t - \boldsymbol{I}_{ti})}{2\lambda_t^2} \right) \quad (9)$$

In the same way, if the bandwidth h_{t-1} is rewritten as:

$$h_{t-1} = \lambda_{t-1}^2 S_q \quad (10)$$

Then, optimal scale parameter λ_{t-1} can be determined using the above method as described in (5) - (7). Substituting (10) into (2), we have:

$$\widehat{f}(I_{t-1}) = \frac{1}{(n-1)} \sum_{j=1}^{n-1} \frac{1}{(2\pi)^{1/2} \lambda_{t-1} \det(S_q)^{1/2}} \times \exp\left(-\frac{(I_{t-1} - I_{tj-1})^2}{2\lambda_{t-1}^2 S_q} \right) \quad (11)$$

Therefore, the derivation of $\widehat{f}(I_t \,|I_{t-1})$ is as follows. Since the sample variance of I_{t-1} is generally non-zero, the elementary transformation can be applied to $\boldsymbol{S}_t$.

$$\boldsymbol{S}_t = \begin{bmatrix} 1 & S_{hq} S_q^{-1} \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} S_h - S_{hq}^2 S_q^{-1} & 0 \\ 0 & S_q \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ S_{hq} S_q^{-1} & 1 \end{bmatrix} \quad (12)$$

Let $A_t = S_h\text{-}S_{hq}^2 S_q^{-1}$, then the determinant of $\boldsymbol{S}_t$ is calculated as follows:

$$\det(\boldsymbol{S}_t) = \det(S_q) \det(A_t) \quad (13)$$

According to the matrix inverse formula and matrix algorithm rules, and substituting (9) and (11) into (4), the CKDE of daily irradiance can be derived as follows:

$$\widehat{f}(I_t \,|I_{t-1}) = \frac{\widehat{f}(\boldsymbol{I}_t)}{\widehat{f}(I_{t-1})} = \frac{\lambda_{t-1}}{(2\pi)^{1/2} \lambda_t^2 \det(A_t)^{1/2}} \times \sum_{i=1}^{n-1} \omega_{ti} \exp\left(-\frac{(I_t - B_{ti})^2}{2\lambda_t^2 A_t} \right) \quad (14)$$

where

$$B_{ti} = I_{ti} + S_{hq}S_q^{-1}(I_{t-1} - I_{ti-1}) \tag{15}$$

$$\omega_{ti} = \frac{C_{ti}}{\sum_{k=1}^{n-1} C_{tk}} \tag{16}$$

$$C_{ti} = \frac{\exp\left(-\frac{(I_{t-1}-I_{ti-1})^2}{2\lambda_t^2 S_q}\right)}{\sum_{j=1}^{n-1}\exp\left(-\frac{(I_{t-1}-I_{tj-1})^2}{2\lambda_{i-1}^2 S_q}\right)} \tag{17}$$

$$\sum_{i=1}^{n-1}\omega_{ti} = 1 \tag{18}$$

The conditional probability density estimate in (14) can be considered to be a mixture of n-1 univariate Gaussian functions (called kernels), each kernel having a mean equal to the i-th sample value B_{ti}, a variance equal to $\lambda_t^2 A_t$, and a weight equal to ω_{ti}, which depends directly on the distance between I_{t-1} and the i-th historical sample I_{ti-1}.

III. INPD OF HOURLY SOLAR IRRADIANCE

Let vector $\boldsymbol{r}_t = [r_{t,1}, r_{t,2}, \ldots, r_{t,d}]^T$ represent the hourly solar irradiance on the t-th day, and d denotes the number of hours in a day. n day samples are used and the i-th sample of hourly irradiance is $\boldsymbol{r}_i = [r_{i1}, r_{i2}, \ldots, r_{id}]^T$, $i = 1,2,\ldots, n$. Therefore, the additive correlation between daily irradiance and hourly irradiance can be defined by

$$I_t = r_{t,1} + r_{t,2} + \cdots + r_{t,d} \tag{19}$$

The hourly irradiance $\boldsymbol{r}_t$ is not only subject to the additive constraint with the daily irradiance I_t, but also influenced by the preceding day's hourly irradiance $\boldsymbol{r}_{t-1}$. According to the known $\boldsymbol{r}_{t-1}$ and I_t, the conditional probability density function $f(\boldsymbol{r}_t\,|r_{t-1}, I_t)$ can be built as follows:

$$f(\boldsymbol{r}_t\,|\boldsymbol{r}_{t-1}, I_t) = f(\boldsymbol{r}_t\,|\boldsymbol{V}_t) = \frac{f(\boldsymbol{r}_t, \boldsymbol{V}_t)}{f(\boldsymbol{V}_t)} \tag{20}$$

where $\boldsymbol{V}_t = (\boldsymbol{r}_{t-1}, I_t)^{\mathrm{T}}$, $f(\boldsymbol{r}_t, \boldsymbol{V}_t)$ is the $(2d+1)$-dimensional joint probability density function of $\boldsymbol{r}_t$ and $\boldsymbol{V}_t$, and $f(\boldsymbol{V}_t)$ is the $(d+1)$-dimensional marginal probability density of $\boldsymbol{V}_t$.

The conditional probability density in (20) can be specified through a linear transformation of the vector $\boldsymbol{r}_t$ into a new vector $\boldsymbol{Y}_t = (y_{t,1}, y_{t,2}, \ldots, y_{t,d})^{\mathrm{T}}$, defined as

$$\boldsymbol{Y}_t = \boldsymbol{R}\boldsymbol{r}_t \tag{21}$$

where $\boldsymbol{R}$ is a unit orthogonal matrix, i.e. $\boldsymbol{R}^T = R^{-1}$. Let $\boldsymbol{R}$= $(\boldsymbol{e}_1, \boldsymbol{e}_2,\ldots, \boldsymbol{e}_d)^{\mathrm{T}}$ and $\boldsymbol{e}_d = (1/\sqrt{d},1/\sqrt{d}, \ldots, 1/\sqrt{d})$, then row vector $\boldsymbol{e}_j(j = 1,2,\ldots, d\text{-}1)$ can be obtained by the Gram-Schmidt orthogonal transformation:

$$\boldsymbol{e'}_j = \boldsymbol{i}_j^T - \sum_{k=j+1}^{d}(\boldsymbol{e}_k \cdot \boldsymbol{i}_j)\boldsymbol{e}_k \tag{22}$$

$$\boldsymbol{e}_j = \boldsymbol{e'}_j / |\boldsymbol{e'}_j| j = d-1, d-2, \cdots, 1 \tag{23}$$

where $\boldsymbol{i}_1 = (1,0,\ldots,0)^T$, $\boldsymbol{i}_2 = (0,1,\ldots,0)^T$,…, $\boldsymbol{i}_d = (0,0,\ldots,1)^T$, and $|e_{j'}|$ is the 2-norm of $\boldsymbol{e}_{j'}$. According to (21), $y_{t,d} = (r_{t,1} + r_{t,2} + \ldots + r_{t,d})/\sqrt{d} = I_t/\sqrt{d}$. Let $I_{t'} = y_{t,d} = I_t/\sqrt{d}$, $\boldsymbol{U}_t = (y_{t,1}, y_{t,2}, \ldots, y_{t,d-1})^T$, thus $\boldsymbol{Y}_t = (y_{t,1}, y_{t,2}, \ldots, y_{t,d})^T = (\boldsymbol{U}_t^T, I_{t'})^T$, and we can transform (20) into

$$f(\boldsymbol{U}_t\,|\boldsymbol{U}_{t-1}, I'_t) = f(\boldsymbol{U}_t\,|\boldsymbol{V'}_t) = \frac{f(\boldsymbol{U}_t, \boldsymbol{V'}_t)}{f(\boldsymbol{V'}_t)} \tag{24}$$

where $\boldsymbol{V}_{t'} = (\boldsymbol{U}_{t-1}, I_{t'})^T$, and the dimension of $f(\boldsymbol{U}_t, \boldsymbol{V}_{t'})$ and $f(\boldsymbol{V}_{t'})$ is 2d-1 and d, respectively. According to multivariate kernel density estimation and the above bandwidth matrix method, $f(\boldsymbol{U}_t, \boldsymbol{V}_{t'})$ and $f(\boldsymbol{V}_{t'})$ can be estimated as

$$\begin{aligned}\widehat{f}(\boldsymbol{U}_t, \boldsymbol{V'}_t) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(2\pi\lambda_{\boldsymbol{UV'}}^2)^{(2d-1)/2}\det(\boldsymbol{S})^{1/2}} \\ &\times \exp\left\{-\begin{pmatrix}\boldsymbol{U}_t - \boldsymbol{U}_i \\ \boldsymbol{V'}_t - \boldsymbol{V'}_i\end{pmatrix}^T \boldsymbol{S}^{-1}\begin{pmatrix}\boldsymbol{U}_t - \boldsymbol{U}_i \\ \boldsymbol{V'}_t - \boldsymbol{V'}_i\end{pmatrix}\Big/ 2\lambda_{\boldsymbol{UV'}}^2\right\}\end{aligned} \tag{25}$$

$$\begin{aligned}\widehat{f}(\boldsymbol{V'}_t) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(2\pi\lambda_{\boldsymbol{V'}}^2)^{d/2}\det(\boldsymbol{S}_{\boldsymbol{V'}})^{1/2}} \\ &\times \exp\left[-\frac{(\boldsymbol{V'}_t - \boldsymbol{V'}_i)^T \boldsymbol{S}_{\boldsymbol{V'}}^{-1}(\boldsymbol{V'}_t - \boldsymbol{V'}_i)}{2\lambda_{\boldsymbol{V'}}^2}\right]\end{aligned} \tag{26}$$

and

$$\boldsymbol{S} = \begin{bmatrix}\boldsymbol{S}_{\boldsymbol{U}} & \boldsymbol{S}_{\boldsymbol{UV'}} \\ \boldsymbol{S}_{\boldsymbol{UV'}}^T & \boldsymbol{S}_{\boldsymbol{V'}}\end{bmatrix} \tag{27}$$

where $\lambda_{\boldsymbol{UV'}}$ is the optimal scale parameter for $\widehat{f}(\boldsymbol{U}_t, \boldsymbol{V'}_t)$, $\lambda_{\boldsymbol{V'}}$ is the optimal scale parameter for $\widehat{f}(\boldsymbol{V'}_t)$,and the i-th sample $\boldsymbol{V}_{i'} = (\boldsymbol{U}_{i-1}, I_{ti'})^T$ and $\boldsymbol{U}_i$ (i = 1,2,…, n) can be generated from the linear transformation of $\boldsymbol{r}_i$. Besides, $\boldsymbol{S}$ is the sample covariance matrix of $(\boldsymbol{U}_t, \boldsymbol{V}_{t'})$, $\boldsymbol{S}_{\boldsymbol{U}}$ is the $(d\text{-}1)\times(d\text{-}1)$ sample covariance matrix of $\boldsymbol{U}_t$, $\boldsymbol{S}_{\boldsymbol{UV'}}$ is the $(d\text{-}1)\times d$ covariance vector for $\boldsymbol{U}_t$ and $\boldsymbol{V}_{t'}$, $\boldsymbol{S}_{\boldsymbol{V'}}$ is the $d\times d$ sample covariance matrix of $\boldsymbol{V}_{t'}$.

Similarly, since the sample covariance matrix of $\boldsymbol{V}_{t'}$ is generally non-zero, the elementary transformation can be applied to $\boldsymbol{S}$.

$$\begin{aligned}\boldsymbol{S} &= \begin{pmatrix}\boldsymbol{E}_{d-1} & \boldsymbol{S}_{\boldsymbol{UV'}}\boldsymbol{S}_{\boldsymbol{V'}}^{-1} \\ 0 & \boldsymbol{E}_d\end{pmatrix} \times \begin{pmatrix}\boldsymbol{S}_{\boldsymbol{U}} - \boldsymbol{S}_{\boldsymbol{UV'}}\boldsymbol{S}_{\boldsymbol{V'}}^{-1}\boldsymbol{S}_{\boldsymbol{UV'}}^T & 0 \\ 0 & \boldsymbol{S}_{\boldsymbol{V'}}\end{pmatrix} \\ &\times \begin{pmatrix}\boldsymbol{E}_{d-1} & 0 \\ \boldsymbol{S}_{\boldsymbol{V'}}^{-1}\boldsymbol{S}_{\boldsymbol{UV'}}^T & \boldsymbol{E}_d\end{pmatrix}\end{aligned} \tag{28}$$

where $\boldsymbol{E}_{d-1}$ is the $(d\text{-}1)$-order identity matrix, $\boldsymbol{E}_d$ is the d-order identity matrix.

Let $\boldsymbol{A} = \boldsymbol{S}_{\boldsymbol{U}} - \boldsymbol{S}_{\boldsymbol{UV'}}\boldsymbol{S}_{\boldsymbol{V'}}^{-1}\boldsymbol{S}_{\boldsymbol{UV'}}^T$, thus $\boldsymbol{A}$ is a symmetric matrix, i.e. $\boldsymbol{A}^T = \boldsymbol{A}$, and the determinant of $\boldsymbol{S}$ can be calculated as follows:

$$\det(\boldsymbol{S}) = \det(\boldsymbol{S}_{\boldsymbol{V'}})\det(\boldsymbol{A}) \tag{29}$$

According to the block matrix inversion formula and matrix algorithm rules, the INPD of hourly irradiance can be derived as follows:

$$\hat{f}(\boldsymbol{U}_t|\,\boldsymbol{V'}_t)=\frac{\hat{f}(\boldsymbol{U}_t,\boldsymbol{V'}_t)}{\hat{f}(\boldsymbol{V'}_t)}=\frac{\lambda_{\boldsymbol{V'}}^{d}}{(2\pi)^{(d-1)/2}\,\lambda_{\boldsymbol{UV'}}^{2d-1}\det(\boldsymbol{A})^{1/2}}\times\sum_{i=1}^{n}\omega_i\exp\left(-\frac{(\boldsymbol{U}_t-\boldsymbol{B}_i)^T\boldsymbol{A}^{-1}(\boldsymbol{U}_t-\boldsymbol{B}_i)}{2\lambda_{\boldsymbol{UV'}}^{2}}\right)\tag{30}$$

where

$$\boldsymbol{B}_i=\boldsymbol{U}_i+\boldsymbol{S}_{\boldsymbol{UV'}}\boldsymbol{S}_{\boldsymbol{V'}}^{-1}\left(\boldsymbol{V'}_t-\boldsymbol{V'}_i\right)\tag{31}$$

$$\omega_i=\frac{C_i}{\sum_{k=1}^{n}C_k}\tag{32}$$

$$C_i=\frac{\exp\left(-\frac{(\boldsymbol{V'}_t-\boldsymbol{V'}_i)^T\boldsymbol{S}_{\boldsymbol{V'}}^{-1}(\boldsymbol{V'}_t-\boldsymbol{V'}_i)}{2\lambda_{\boldsymbol{UV'}}^{2}}\right)}{\sum_{j=1}^{n}\exp\left\{-\left(\boldsymbol{V'}_t-\boldsymbol{V'}_j\right)^T\boldsymbol{S}_{\boldsymbol{V'}}^{-1}\left(\boldsymbol{V'}_t-\boldsymbol{V'}_j\right)\Big/2\lambda_{\boldsymbol{V'}}^{2}\right\}}\tag{33}$$

$$\sum_{i=1}^{n}\omega_i=1\tag{34}$$

The $\widehat{f}(\boldsymbol{U}_t\,|\boldsymbol{V'}_t)$ in (30) can be seen as a weighted sum of n kernels each with mean vector $\boldsymbol{B}_i$ and covariance matrix $\lambda_{\boldsymbol{UV'}}^2\boldsymbol{A}$. Equation (32) and (33) show that the weight ω_i which only relies on the conditioning $\boldsymbol{V}_t\prime$ depends directly on the distance between $\boldsymbol{V}_t\prime$ and the sample $\boldsymbol{V}_i\prime$. Utilizing this feature the sample can be easily generated by the following three-stage sampling method.

IV. PROPOSED THREE STAGE SAMPLING METHOD

On the basis of the stochastic sampling technique of the multivariate normal distribution [26], a three-stage sampling method of the solar irradiance time series is developed using the proposed conditional probability model. Stage 1 involves the initialization of sample parameters, while stage 2 focuses on generating random samples of daily irradiance using the above CKDE model. Lastly, in stage 3, hourly irradiance random samples are generated using the above INPD model. The specific procedure for sampling is as follows:

Stage 1: 1) Generate the samples $\boldsymbol{I}_{ti}$ and $\boldsymbol{r}_i$ from the measured data of solar irradiance. 2) Use the linear transformation in (21) to obtain the samples $\boldsymbol{U}_i$ and $\boldsymbol{V}_i\prime=(\boldsymbol{U}_{i-1},\,I_{ti}\prime\,)^T$. 3) Compute $\boldsymbol{S}_t$, $\boldsymbol{S}$, λ_t, λ_{t-1}, $\lambda_{\boldsymbol{UV}\prime}$ and $\lambda_{\boldsymbol{V}\prime}$. 4) Generate A_t and $\boldsymbol{A}$ from $\boldsymbol{S}_t$ and $\boldsymbol{S}$, then use Cholesky decomposition to get $A_t=L_tL_t^T$ and $\boldsymbol{A}=LL^T$.

Stage 2: 1) Let $t=2$, and initialize I_{t-1}. 2) Given I_{t-1}, calculate the weight $\omega_{ti}(i=1,2,\ldots,n\text{-}1)$ associated with each sample I_{ti-1} by (16), and divide [0,1] into n-1 subintervals where the length of the i-th subinterval is ω_{ti}. 3) Generate a uniformly distributed sample x over [0,1], and select the responsible kernel by this means: if the subinterval k contains x, i.e. $\sum_{i=1}^{k-1}\omega_{ti}\le x\le\sum_{i=1}^{k}\omega_{ti}$, then the k-th kernel function with mean B_{tk} and variance $\lambda_t^2A_t$ is selected. 4) Generate a standard normally distributed random number V_k, then calculate the conditioned I_t by $I_t=B_{tk}+\lambda_tL_tV_k$ under the known conditioning I_{t-1}. If $t\le 365$ day, assign I_t to I_{t-1} as a new conditioning value and set $t=t+1$ to step 2); otherwise, proceed to stage 3.

Stage 3: 1) Let $t=2$, and initialize $\boldsymbol{r}_{t-1}$ by I_{t-1}. 2) Given $\boldsymbol{r}_{t-1}$ and I_t, i.e. $\boldsymbol{V}_t\prime$, calculate the weight $\omega_i(i=1,2,\ldots,n)$ associated with each sample $\boldsymbol{V}_i\prime$ by (32), and divide [0,1] into n subintervals where the length of the i-th subinterval is ω_i. 3) Generate a uniformly distributed sample x over [0,1], and select the responsible kernel by this means: if the subinterval k contains x, i.e. $\sum_{i=1}^{k-1}\omega_i\le x\le\sum_{i=1}^{k}\omega_i$, then the k-th kernel function with mean vector $\boldsymbol{B}_k$ and covariance matrix $\lambda_{\boldsymbol{UV'}}^2\boldsymbol{A}$ is selected. 4) Generate a $(d\text{-}1)\times1$ standard normally distributed random vector $\boldsymbol{Q}$, and calculate the conditioned $\boldsymbol{U}_t$ by $\boldsymbol{U}_t=B_k+\lambda_{\boldsymbol{UV}\prime}\boldsymbol{LQ}$ under the known conditioning $\boldsymbol{V}_t\prime$, thus $\boldsymbol{r}_t=R^TY_t$ is achieved by $\boldsymbol{Y}_t=(\boldsymbol{U}_t^T,\,I_t\prime\,)^T$. If $t\le365$ day, let $t=t+1$ and go back to step 2); otherwise, exit.

V. CASE STUDIES

The hourly irradiance data of the Ashland and UO Solar Awning installation in Eugene (called Eugene), provided by the Solar Radiation Monitoring Laboratory at Oregon University in the US [27] from 2020 to 2022, are used to test the practicality of the model proposed here, where the above two PV stations are marked by Pv1 and Pv2. Then two 40 MW Pv1 and Pv2 are added to bus 15 and bus 20 in the IEEE-RTS79 [28] for power system with PV stations reliability evaluation. Finally, the power output of a PV station is calculated using the relationship between the power output of the PV station and solar irradiance, which is given by the solar irradiance-to-energy conversion function ignoring the effect of the size or type of the PV modules on the power out [29].

In the case, the unit solar irradiance in the standard environment is set to 1000 W/m^2, the specific solar irradiance is taken as 150 W/m^2, and the forced outage rate and repair time are assumed to be 0.12 and 60 hours, respectively.

A. PRACTICALITY ASSESSMENT OF THE MODEL

short sequence method [30] is adopted to test the practicability of the model, which generates 900 sets of irradiance sequences simulated using the nonrarametric model described earlier. The length of each simulation set was chosen to match the length of the measured data, ensuring that the variability of sample statistics across these realizations is representative of the sampling variability of the historical data. Both models (CKDE of daily irradiance and INPD of hourly irradiance) compared with the ARMA model mentioned above [11] were tested for their ability to reproduce the following statistics of the historic data: mean; standard deviation (S_{td}); coefficient of variation (C_v); lag 1 autocorrelation coefficient ($R1$); and cross correlation between hourly irradiance. Besides, the probability-probability (P-P) plot and

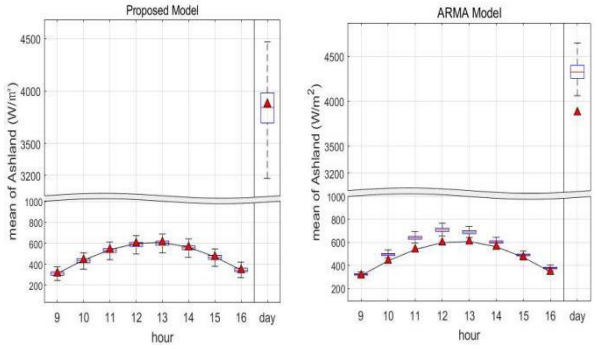


FIGURE 1. Box plot of the mean using hourly INPD and daily CKDE model compared with the ARMA model at the Pv1 site.

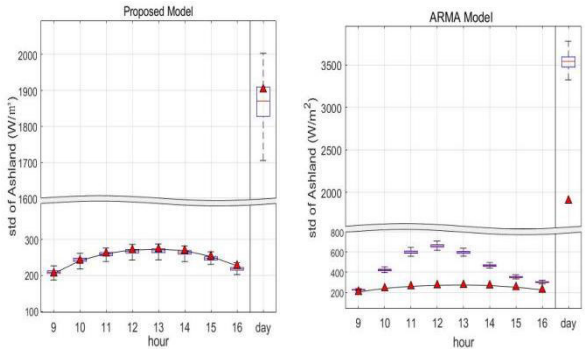


FIGURE 2. Box plot of the S_{td} using hourly INPD and daily CKDE model compared with the ARMA model at the Pv1 site.

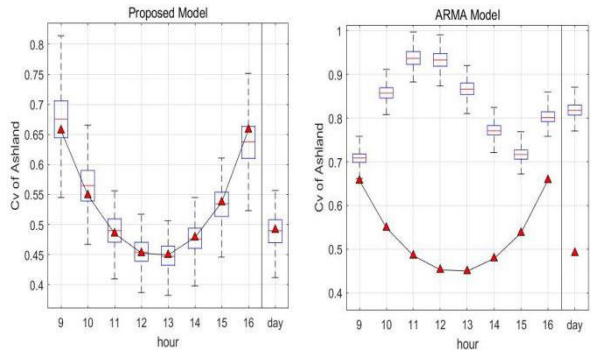


FIGURE 3. Box plot of the *Cv* using hourly INPD and daily CKDE model compared with the ARMA model at the Pv1 site.

marginal density distribution curves will be used to test the goodness-of-fit of the proposed model.

As used here, the simulated statistics of daily and hourly solar irradiance for the proposed and compared models at the Pv1 site are represented using box plots, which consist of a box that represents 25% to 75% percentiles, a red line in the center of the box that represents the median, and whiskers that extend to the 5% and 95% percentiles. Similarly, The observed statistics of daily and hourly solar irradiance at the Pv1 site are denoted by the broken line and triangle labels. Finally, the results are shown in Figures 1-4.

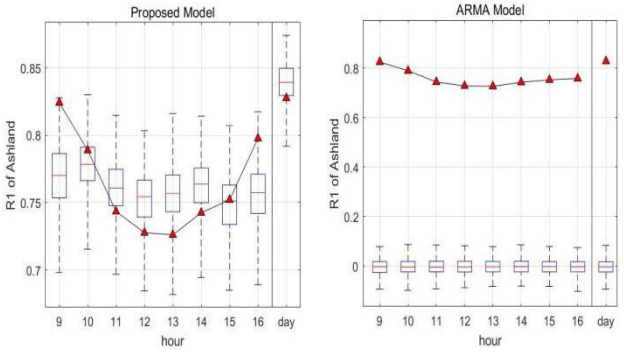


FIGURE 4. Box plot of the $R1$ using hourly INPD and daily CKDE model compared with the ARMA model at the Pv1 site.

The hourly and daily mean, S_{td} and C_v of the observed and simulated irradiance are illustrated in Figures 1-3. As one can seen from figures, the statistics of the measured data for the proposed model fall within the range of the boxes, thus all statistics were reproduced well by the proposed hourly and daily irradiance model, which shows that differences between data and model can be ascribed to sampling variability. But, the historical statistics for ARMA model is outside the range of the boxes, which indicates the parametric model does not reproduce.

Figure 4 illustrates the $R1$ of the historical and simulated samples from the proposed and compared models. It shows that the proposed model reproduces this statistic well. This is an important result as it denotes that longer-term dependence is being properly represented. Besides, the aggregate daily irradiance $R1$ for the observed and simulated samples is presented from the CKDE model in the Figure. Note how well the daily lag 1 correlation is simulated. This is also an important result as the model has not been designed to ensure the proper representation of dependence at the daily timescale. However, the ARMA model is a linear model, in which the auto-correlation of the simulated samples can not be reproduced due to the inherent nonlinear characteristic and time varying of solar irradiance.

To test the simulation effectiveness of the proposed model on modeling the correlation between hourly irradiance, the correlation coefficient matrices and its relative errors of hourly irradiance at the Pv1 and Pv2 sites are calculated and analyzed. The correlation coefficients shown in the upper triangle of Tables 2 and 3 are calculated by using the observed irradiance. The lower triangle of Tables 2 and 3 denote the correlation coefficients calculated by using the simulated irradiance. The percentages in the brackets in Tables 2 and 3 represent the errors of the indices. It can be seen that the correlation coefficients calculated for the simulated sequences of adjacent hourly irradiance from Pv1 and Pv2 sites exhibit a relatively small error compared to the results obtained from measured data, controlled within 5%. Conversely, the errors of correlation coefficients for non-adjacent hourly irradiance are relatively larger but still maintained

TABLE 2. Correlation coefficient error matrix of hourly irradiance in Pv1.

hour	9	10	11	12	13	14	15	16
9	1.0000	0.9583	0.9061	0.8735	0.8499	0.8418	0.8337	0.8447
10	0.9512(0.74%)	1.0000	0.9479	0.9091	0.8826	0.8636	0.8474	0.8438
11	0.8903(1.73%)	0.9431(0.51%)	1.0000	0.9455	0.9082	0.8782	0.8482	0.8347
12	0.8463(3.11%)	0.8938(1.68%)	0.9364(0.95%)	1.0000	0.9435	0.9031	0.8631	0.8455
13	0.8279(2.58%)	0.8678(1.67%)	0.8974(1.18%)	0.9362(0.77%)	1.0000	0.9437	0.9002	0.8779
14	0.8203(2.54%)	0.8457(2.07%)	0.8611(1.95%)	0.8906(1.37%)	0.9349(0.93%)	1.0000	0.9425	0.9072
15	0.8168(2.02%)	0.8327(1.72%)	0.8351(1.55%)	0.8463(1.93%)	0.8836(1.84)	0.9336(0.94%)	1.0000	0.9459
16	0.8234(2.51%)	0.8221(2.58%)	0.8172(2.08%)	0.8246(2.46%)	0.8559(2.51%)	0.8901(1.88%)	0.9377(0.87%)	1.0000

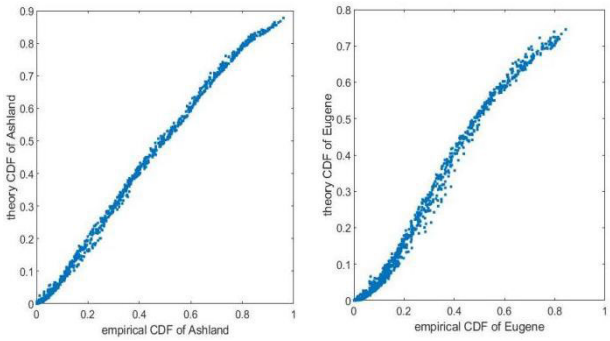


FIGURE 5. P-P plots of INPD model at the Pv1 and Pv2 sites.

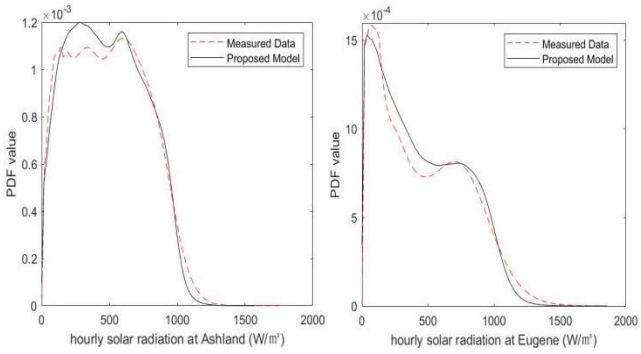


FIGURE 6. Probability density curves of hourly irradiance at the Pv1 and Pv2 sites.

within a reasonable range of 10%. The results verify that the proposed method effectively considers temporal correlations between adjacent and nonadjacent hourly irradiance as well as the additive correlation, resulting in improved simulation accuracy.

Probability-Probability (called PP) plot will be used to test the goodness-of-fit of multidimensional variables. In the unit square shown in Figure 5, the abscissa represents the cumulative empirical distribution values of hourly irradiance obtained by the measured data, and the ordinate denotes the cumulative distribution function values of the simulated samples obtained using the proposed model. It can be seen that the scatter sequence is basically located on the diagonal, which indicates that the proposed model has sufficient fitting accuracy for the measured data and can better maintain the probability distribution characteristics of the observed samples.

Furthermore, the probability density distributions of hourly and daily irradiance for the two PV sites estimated using the actual measured data and the data sampled by the proposed models are shown in Figures 6-7. It can be seen that the probability distribution curves of hourly and daily irradiance calculated by the proposed method can well track the probability density distribution curves of the measured sample data, which once again indicates that the proposed model has relatively high simulation accuracy.

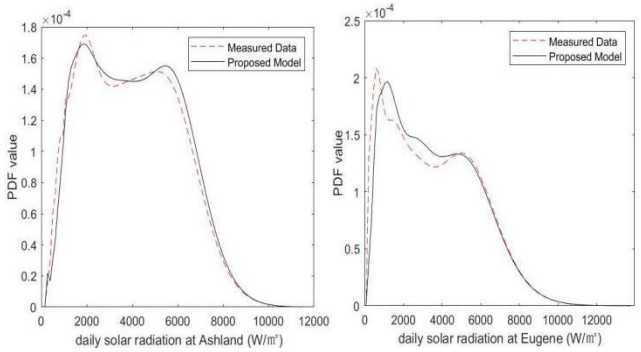


FIGURE 7. Probability density curves of daily irradiance at the Pv1 and Pv2 sites.

B. RELIABILITY ANALYSIS FOR POWER SYSTEM WITH PV STATIONS

The IEEE-RTS79 with Pv1 and Pv2 is used to assess the impact of PV station on power system reliability. Table 3 presents, for three cases, the reliability indices loss of load probability (LOLP) and expected energy not supplied (EENS) obtained from sequential Monte Carlo simulation [31]. Case *a* shows the reliability indices before adding the two PV stations, case *b* shows the reliability indices of 2 × 40 MW PV stations added, and case *c* shows the reliability indices with the same capacity of two traditional

TABLE 3. Correlation coefficient error matrix of hourly irradiance in Pv2.

hour	9	10	11	12	13	14	15	16
9	1.0000	0.8733	0.7977	0.7314	0.6816	0.6513	0.6363	0.6404
10	0.8736(0.04%)	1.0000	0.9118	0.8309	0.7628	0.7281	0.6984	0.6888
11	0.7969(0.09%)	0.9128(0.11%)	1.0000	0.9065	0.8294	0.7852	0.7523	0.7235
12	0.7299(0.19)	0.8324(0.18%)	0.9036(0.32%)	1.0000	0.8958	0.8195	0.7816	0.7382
13	0.6775(0.61%)	0.7652(0.32%)	0.8275(0.23%)	0.8997(0.44%)	1.0000	0.9086	0.8435	0.7646
14	0.6511(0.04%)	0.7282(0.04%)	0.7801(0.63%)	0.8204(0.11%)	0.9023(0.69%)	1.0000	0.9119	0.8361
15	0.6124(3.76)	0.6751(3.32%)	0.7252(3.62%)	0.7585(2.95%)	0.8161(3.24%)	0.8963(1.71%)	1.0000	0.9178
16	0.5769(9.92%)	0.6213(9.79%)	0.6504(10.10%)	0.6661(9.76%)	0.6902(9.72%)	0.7851(6.11%)	0.8893(3.11%)	1.0000

TABLE 4. Reliability indices in different cases.

Reliability indices	case a	case b	case c
LOLP	0.001 112	0.000 922	0.000 607
EENS (MWh/yr)	1226.25	976.36	608.32

TABLE 5. Reliability indices for different capacities.

Reliability indices	Added capacity of PV stations (MW)				
	2×40	2×80	2×120	2×160	2×200
LOLP	0.000922	0.000787	0.000727	0.000701	0.000625
EENS (MWh/yr)	976.36	821.62	801.38	767.29	683.96

units. Besides, the convergence criterion for all cases is $\beta_{\text{EENS}} \leq 5\%$.

As shown in Table 4, the LOLP index obtained from case *b* results in a 17.08% decrease from 0.001112 to 0.000922, and a 20.37% decrease in EENS indicator, indicating that PV station can partially meet the energy demand of load and improve system reliability. In case *c*, the LOLP and EENS decreased by 45.41% and 50.39% respectively, due to the stable output of traditional units, in contrast to the random fluctuations observed in PV stations, which do not generate electricity during nighttime. Therefore, traditional unit contributes more to the system reliability than the PV station at the same capacity.

Furthermore, the reliability indicators with different capacities of two added PV stations obtained using the proposed method are shown in Table 5. The reliability indicators decrease with the increase of capacities of added PV stations. However, when the capacity of the PV station increases to 200 MW, both LOLP and EENS remain larger than the reliability index obtained by the same capacities of added traditional units. This further means that the randomness of PV station has a great impact on the system reliability.

VI. CONCLUSION

Compared with existing studies that focus on the self-correlation or hourly correlation characteristic of PV station output, this paper presents a novel solar radiation generation model that successfully retains the multi-temporal correlation of photovoltaic station outputs between daily irradiance amount and hourly prediction value. And the hierarchical global-to-local relations of photovoltaic station output can also be guaranteed.

To verify its goodness, the proposed method compared with ARMA model were tested for their ability to reproduce the four statistics. Three tests, including correlation coefficient matrices, P-P plot and probability distribution were then performed for the practicality analysis. The test results indicate that the simulated statistics for the proposed model can be reproduced better than ARMA model, and that the proposed model can also be suitable to the PV station output with distinct patterns in uncertainty and multi-temporal correlation. Finally, the proposed three-stage sampling method was used to assess the reliability of power system with PV stations, which indicates that the proposed method can provide a more accurate and reliable reliability index of the power system.

REFERENCES

[1] C. Zhu, M. Wang, M. Guo, J. Deng, Q. Du, W. Wei, Y. Zhang, and S. S. A. Talesh, "Optimizing solar-driven multi-generation systems: A cascade heat recovery approach for power, cooling, and freshwater production," *Appl. Thermal Eng.*, vol. 240, Mar. 2024, Art. no. 122214.

[2] S. Su, Y. Hu, L. He, K. Yamashita, and S. Wang, "An assessment procedure of distribution network reliability considering photovoltaic power integration," *IEEE Access*, vol. 7, pp. 60171–60185, 2019.

[3] Y. Ju, W. Liu, Z. Zhang, and R. Zhang, "Distributed three-phase power flow for AC/DC hybrid networked microgrids considering converter limiting constraints," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 1691–1708, May 2022.

[4] C. Zhu, Y. Zhang, M. Wang, J. Deng, Y. Cai, W. Wei, and M. Guo, "Optimization, validation and analyses of a hybrid PV-battery-diesel power system using enhanced electromagnetic field optimization algorithm and ε-constraint," *Energy Rep.*, vol. 11, pp. 5335–5349, Jun. 2024.

[5] H. Abunima and J. Teh, "Reliability modeling of PV systems based on time-varying failure rates," *IEEE Access*, vol. 8, pp. 14367–14376, 2020.

[6] R. Aguiar and M. Collares-Pereira, "TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation," *Sol. Energy*, vol. 49, no. 3, pp. 167–174, Sep. 1992.

[7] O. M. Kam, S. Noël, H. Ramenah, P. Kasser, and C. Tanougast, "Comparative Weibull distribution methods for reliable global solar irradiance assessment in France areas," *Renew. Energy*, vol. 165, pp. 194–210, Mar. 2021.

[8] E. S. Oda, A. M. A. E. Hamed, A. Ali, A. A. Elbaset, M. A. E. Sattar, and M. Ebeed, "Stochastic optimal planning of distribution system considering integrated photovoltaic-based DG and DSTATCOM under uncertainties of loads and solar irradiance," *IEEE Access*, vol. 9, pp. 26541–26555, 2021.
[9] M. Wahbah, B. Mohandes, T. H. M. El-Fouly, and M. S. El Moursi, "Unbiased cross-validation kernel density estimation for wind and PV probabilistic modelling," *Energy Convers. Manage.*, vol. 266, Aug. 2022, Art. no. 115811.
[10] N. Dong, J.-F. Chang, A.-G. Wu, and Z.-K. Gao, "A novel convolutional neural network framework based solar irradiance prediction method," *Int. J. Electr. Power Energy Syst.*, vol. 114, Jan. 2020, Art. no. 105411.
[11] M. David, F. Ramahatana, P. J. Trombe, and P. Lauret, "Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models," *Sol. Energy*, vol. 133, pp. 55–72, Aug. 2016.
[12] D. K. Khatod, V. Pant, and J. Sharma, "Analytical approach for well-being assessment of small autonomous power systems with solar and wind energy sources," *IEEE Trans. Energy Convers.*, vol. 25, no. 2, pp. 535–545, Jun. 2010.
[13] V. A. Graham, K. G. T. Hollands, and T. E. Unny, "A time series model for K_t with application to global synthetic weather generation," *Sol. Energy*, vol. 40, no. 2, pp. 83–92, 1988.
[14] Y. Kun, C. Yijia, C. Xingying, G. Chuangxin, and Z. Hua, "Dynamic probability power flow of district grid containing distributed generation," *Proc. CSEE*, vol. 31, no. 1, pp. 20–25, 2011.
[15] Z. Ren, W. Yan, X. Zhao, W. Li, and J. Yu, "Chronological probability model of photovoltaic generation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1077–1088, May 2014.
[16] L.-G. Chen, H.-D. Chiang, N. Dong, and R.-P. Liu, "Group-based chaos genetic algorithm and non-linear ensemble of neural networks for short-term load forecasting," *IET Gener. Transmiss. Distrib.*, vol. 10, no. 6, pp. 1440–1447, 2016.
[17] T. N. Goh and K. J. Tan, "Stochastic modeling and forecasting of solar radiation data," *Sol. Energy*, vol. 19, no. 6, pp. 755–757, 1977.
[18] A. Sharma and R. O'Neill, "A nonparametric approach for representing interannual dependence in monthly streamflow sequences," *Water Resour. Res.*, vol. 38, no. 7, pp. 189–206, Jul. 2002.
[19] D. G. Tarboton, A. Sharma, and U. Lall, "Disaggregation procedures for stochastic hydrology based on nonparametric density estimation," *Water Resour. Res.*, vol. 34, no. 1, pp. 107–119, Jan. 1998.
[20] Z. Yuan, F. Fei, and W. Jie, "Nonparametric disaggregation load model in power system reliability evaluation incorporating the additive correlation," *Proc. CSEE*, vol. 35, no. 23, pp. 6039–6047, 2015.
[21] Y. Zhao, J. Kuang, K. Xie, W. Li, and J. Yu, "Dimension reduction based non-parametric disaggregation for dependence modeling in composite system reliability evaluation," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 159–168, Jan. 2021.
[22] V. A. Graham and K. G. T. Hollands, "A method to generate synthetic hourly solar radiation globally," *Sol. Energy*, vol. 44, no. 6, pp. 333–341, 1990.
[23] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis*. New York, NY, USA: Theory and Practice, 2006.
[24] J. Fan, "Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems," *Biometrika*, vol. 83, no. 1, pp. 189–206, Mar. 1996.
[25] M. Kristan, A. Leonardis, and D. Skočaj, "Multivariate online kernel density estimation with Gaussian kernels," *Pattern Recognit.*, vol. 44, nos. 10–11, pp. 2630–2642, Oct. 2011.
[26] W. Li and R. Billinton, "Effect of bus load uncertainty and correlation in composite system adequacy evaluation," *IEEE Trans. Power Syst.*, vol. 6, no. 4, pp. 1522–1529, Nov. 1991.
[27] *The Hourly PV Power Data*. [Online]. Available: http://solardata.uoregon.edu/SelectArchival.html
[28] P. Subcommittee, "IEEE reliability test system," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 6, pp. 2047–2054, Nov. 1979.
[29] J. Park, W. Liang, J. Choi, A. A. El-Keib, M. Shahidehpour, and R. Billinton, "A probabilistic reliability evaluation of a power system including solar/photovoltaic cell generator," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2009, pp. 1–6.
[30] P. Ai, C. Xiong, K. Li, Y. Song, S. Gong, and Z. Yue, "Effect of data characteristics inconsistency on medium and long-term runoff forecasting by machine learning," *IEEE Access*, vol. 11, pp. 11601–11612, 2023.
[31] R. Billinton and L. Wenyuan, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*. New York, NY, USA: Plenum, 1994.

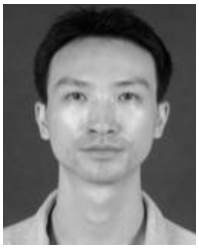
FEI FAN received the B.S. degree in electrical engineering from Northeastern University, Shenyang, China, in 2012, and the M.S. degree in electrical engineering from Chongqing University, Chongqing, China, in 2015. He is currently a Lecturer with Hunan Vocational College of Railway Technology, China. His research interests include power system planning and reliability evaluation.

XIANGJIE ZHOU received the B.S. degree in energy engineering from Shijiazhuang Tiedao University, Shijiazhuang, China, in 2008, and the M.S. degree in energy engineering from Beihang University, Beijing, China, in 2011. He is currently a Professor with Hunan Vocational College of Railway Technology, China. His research interests include photovoltaic systems, power quality management, and power electronics.

LINGZHI ZHANG received the B.S. degree in literature from Jinan University, Guangzhou, China, in 2003, and the M.S. degree in education from Southwest Jiaotong University, Chengdu, China, in 2015. He is currently a Professor with Hunan Vocational College of Railway Technology, China. His research interests include dynamic load panoramic characteristics of rail transit, power supply operation, and maintenance technology.

ZHENJIANG HE received the B.S. degree in civil engineering from Xiangtan University, Xiangtan, China, in 2008. He is currently a Senior Engineer with State Grid Zhuzhou Power Supply Company, China. His research interests include the operational analysis of transmission lines, smart grid technology, and new energy grid connection technology.

WEIXIANG XU received the B.S. degree in electrical engineering and the M.S. degree in business administration from Three Gorges University, Yichang, China, in 2015 and 2023, respectively. He is currently an Engineer with State Grid Zhuzhou Power Supply Company, China. His research interests include transmission line operation and smart grid technology.



YUAN ZHAO (Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2004. He is currently a Professor with the Electrical Engineering College, Chongqing University. His research interests include power system planning and reliability evaluation.